THE TERRITORIAL DIMENSION IN SDG INDICATORS: GEOSPATIAL DATA ANALYSIS AND ITS INTEGRATION WITH STATISTICAL DATA

UN-GGIM: EUROPE | WORKING GROUP ON DATA INTEGRATION

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL | Edited by Statistics Portugal

UN-GGIM EUROPE
UNITED NATIONS COMMITTEE OF EXPERTS ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT

THE TERRITORIAL DIMENSION IN SDG INDICATORS: GEOSPATIAL DATA ANALYSIS AND ITS INTEGRATION WITH STATISTICAL DATA

UN-GGIM: EUROPE | WORKING GROUP ON DATA INTEGRATION

UN-GGIM EUROPE

UNITED NATIONS COMMITTEE OF EXPERTS ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT

This work is published under the steering of the UN-GGIM: Europe Executive Committee. The opinions expressed and arguments employed herein do not necessarily reflect the official views of all UN-GGIM: Europe member countries.
This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Chair of the UN-GGIM: Europe Executive Committee
Tomaž Petek (Director General of the Surveying and Mapping Authority of the Republic of Slovenia).

Title
The territorial dimension in SDG indicators:
geospatial data analysis and its integration with statistical data

Editor
Instituto Nacional de Estatística, I.P.
Av. António José de Almeida
1000-043 Lisboa - Portugal
Telefone: 21 842 61 00
Fax: 21 845 40 84

President of Statistics Portugal Board
Francisco Lima

Please cite this publication as:

UN-GGIM: Europe (2019), *The territorial dimension in SDG indicators: geospatial data analysis and its integration with statistical data*, Instituto Nacional de Estatística, Lisboa.

Author
UN-GGIM: Europe - Working Group Data Integration, led by Pier-Giorgio Zaccheddu (Federal Agency for Cartography and Geodesy, Germany)

Coordination
Francisco Vala (Statistics Portugal), Pier-Giorgio Zaccheddu (Federal Agency for Cartography and Geodesy, Germany)

Rapporteurs
Cátia Nunes (Statistics Portugal), Jerker Moström (Statistics Sweden), Fabio Volpe (e-GEOS Italy), Ingrid Kaminger (Statistics Austria) supported by Sabine Afflerbach-Thom, Jeanette Kretz (Federal Agency for Cartography and Geodesy, Germany)

Design and Edition
Statistics Portugal

ISBN 978-989-25-0485-8

Monograph

Content

FOREWORD FROM

Chair of UN-GGIM: Europe Executive Committee

Statistical information coupled with geospatial information is indeed a powerful tool

The United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) was established with the objective of setting the agenda for global geospatial information development and very critically for promoting the benefits of geospatial information for addressing national policy and key global challenges. In Europe the Regional Committee, UN-GGIM: Europe, aims to ensure that the National Mapping and Cadastral Authorities (NMCAs) and National Statistical Institutes (NSIs) in the European UN Member-States, the European Institutions and associated bodies work together to contribute to a more effective management and availability of geospatial information in Europe.

UN-GGIM: Europe is proud to present this work embarked by the Working Group on Data Integration, which promotes one of the primary goals of the Regional Committee – that of exploring how the integration of statistical and geospatial and other information can meet and satisfy user needs and requirements.

The Sustainable Development Goals (SDGs) are the blueprint to achieve a better and more sustainable future for all. By addressing the global challenges, including those related to poverty, inequality, climate, environmental degradation, prosperity, and peace and justice, the Goals intend to leave no one behind. The monitoring framework for the 2030 Agenda on Sustainable Development includes a number of indicators and a raft of information is required to measure and monitor the progress of these indicators.

While statistical information has long been a traditional source for information, the 2030 Agenda specifically calls for the contribution made by a wide range of data, including earth observation and geospatial information to the monitoring framework of the SDG indicators. However, the contribution of geospatial data goes far beyond simply location variables that allow the disaggregation, aggregation, mapping and visualisation of information. Geospatial analysis and models, similarly to statistical analysis, provide methods of interpreting the data based on the application of specific techniques and data processing models. Statistical information coupled with geospatial information is indeed a powerful tool, greater than the sum of their component parts.

This report illustrates how the Working Group went about addressing the territorial dimension of the Sustainable Development Goals indicators by focusing on the contribution of geospatial data analysis and its integration with statistical data based on global, European and national perspectives. Focusing on the background and experiences of European and national initiatives addressing the SDGs the Working Group on Data Integration also took into account the global level, by linking with the activities of the Inter-Agency Expert Group on Sustainable Development Goals Indicators Working Group on Geospatial Information (IAEG-SDG WG GI).

The Working Group's work focused on four indicators, and their efforts and outputs, presented in this report, highlight the benefits of collaboration between National Statistical Institutes and National Mapping and Cadastral Agencies within countries and across European and Global institutions.

This collaboration is fundamental to ensuring that the 2030 Agenda is achieved and that indeed no one is left behind.

Tomaž Petek

Chair of UN-GGIM: Europe Executive Committee

FOREWORD FROM

President of Statistics Portugal Board

Data that counts for people's everyday life

Statistics Portugal has been actively taking part in UN-GGIM: Europe activities since the beginning, in 2014, as we recognise this initiative as a leading forum for addressing geospatial information management issues, including from a statistical production perspective.

Space (like time) is an essential component of statistical production. To capture this fundamental data dimension, the use of geospatial information to properly address the location element present in all phases of statistical production is essential – from the design phase, to efficiently collect the raw data, up to the dissemination stage, to structure statistical results and to allow a territorial visual perception of data.

Geospatial information plays a major role in statistical production transformation, including data integration from different types of sources – from both public and private administrative information to big data and earth observations – by allowing accurate data linkage and (spatial) data matching.

This report on data integration to address Sustainable Development Goals indicators in a territorial perspective is an excellent showcase to highlight the potential of statistical and geospatial data combination and geospatial data analysis. This combination allows producing new indicators – as the case of tier II and III indicators, where territorial assets are imbedded in statistical information – and more relevant statistics – as the example of tier I indicator, with a higher spatial detail than the one obtained with traditional data sources. Moreover, this report confirms the advantages of cooperation between statistical and geospatial communities, as well as geospatial data producers and statistical analysts, to enhance innovation in statistical production. The Portuguese experience on Land Use Efficiency indicator and on the whole new Land Use and Land Cover statistical project – LCLUStats – based on a close cooperation with the Directorate-General for Territory (the Portuguese National Mapping and Cadastral Agency) is an example on how bridging different types of expertise can deliver better statistics.

Statistics Portugal is committed in keeping pace with the 2030 Agenda, by disseminating the official statistics already available for monitoring the Sustainable Development Goals – through a dedicated webpage in our Portal and a yearly analytical publication – and, at the same time, by following up innovative projects to better address the challenging dimensions that the motto leaving no one behind entails.

The territorial dimension for SDG monitoring is a challenging issue to address by UN-GGIM. Therefore, further work of the European Data Integration group on this issue is very much welcomed.

Data integration is on the verge of moving from a stovepipe model of statistical production – centred on economic, social and environmental specific issues – to a horizontal and flexible model of production that enables a faster and higher quality reply to emerging cross cutting issues, including greater spatial granularity: data that counts for people's everyday life.

Aligned with this reasoning, Statistics Portugal is developing a project NDI – the National Data Infrastructure – aiming to make a more intensive and integrated use of administrative data and other types of data by means of data vault modelling in a ultra-safe environment, stimulating 'on progress statistics' dissemination. It is expected that this project on data integration and analysis will allow adding value to corporate and public agencies information, thus better serving society's information needs, while also reducing redundancy on data collection initiatives across public administration and response burden in traditional official statistical surveys.

Geocoding and spatial analysis will, in this vein, be crucial for this Data Integration Vision for the National Statistical System and to increase the possibilities of combining and analysing data according to different and more flexible territorial arrangements to better cope with data needs for policy making, research and official statistical production.

Francisco Lima
President of Statistics Portugal Board

Management Summary

1. The aim of this report is to address the territorial dimension of the Sustainable Development Goals indicators by focusing on the contribution of geospatial data analysis and its integration with statistical data based on a global, European and national perspective, by taking into account, at the global level, the activities of the Inter-Agency Expert Group on Sustainable Development Goals Indicators Working Group on Geospatial Information (IAEG-SDG WG GI), and also the background and experiences of European and national initiatives addressing the SDGs from a geospatial perspective.

2. In order to address this goal, the working group members provided structured comments on a wider reference list of SDG indicators identified as directly or indirectly benefiting from geospatial information and from its integration with statistical data based on a template for indicator analysis that included: **i)** the current reporting situation both at global and national levels with metadata analysis; **ii)** gap analysis on the methodology and geospatial data integration suggested for the indicator; and **iii)** identification of corresponding EU SDG indicators and of specific national indicators.

3. Based on the number of contributions and inputs provided by the working group members, the maturity of the indicators operationalization, also reflected in the tier level classification, the possibility of calculation at European level and at national level, and the policy relevance in the European context, four SDG indicators were selected:

11.2.1 | Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities

11.3.1 | Ratio of land consumption rate to population growth rate

11.7.1 | Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities

15.1.1 | Forest area as a proportion of total land area

4. The outcomes and findings of the analysis carried out on the territorial dimension of the selected SDG indicators have allowed to agree on the following set of recommendations to enhance the contribution of geospatial data analysis and its integration with statistical data to address the SDG indicators, which can be directly linked to Global Statistical Geospatial Framework principles [see Figure 3]:

✓ Harmonize relevant geospatial data themes in Europe, such as Buildings, Addresses, Land Use and Land Cover data as well as Cadastral data, according to the UN-GGIM: Europe Core Data recommendations.

✓ Implement Cadastral and Land Cover data as key national authoritative data for the operationalization of SDG indicators measurement and encourage European institutions to financially support the Member-States on the implementation of this type of data and its regular update.

✓ Use geospatial layers generated from Earth Observation data with a stable and validated methodology at global (e.g. Global Human Settlement Layer) and European level (e.g.,Copernicus High Resolution Layers, CORINE) to enable data comparability across countries.

✓ Create capacity building initiatives for National Statistical Institutes to take full advantage of Earth Observation based data to produce new statistical indicators and to increase territorial disaggregation of traditional indicators already reported by NSIs.

✓ Define and implement National Spatial Data Infrastructures having in mind the requirements for statistical production to meet the needs of the Sustainable Development Goals monitoring framework and to improve the modernisation of official statistics.

✓ Implement consistent and stable sub-national spatial units based on different geographical levels of detail (including grid systems), and its correspondent models of codification, to produce and disseminate coherent and comparable statistical data and indicators over time.

✓ Develop and use population grids and other grid-based statistics as a way to increase statistical and geospatial data integration, including geospatial data processing analysis to calculate relevant indicators for the Sustainable Development Goals monitoring framework.

✓ Adopt harmonised and comparable concepts, definitions and classifications and build consensus among Geospatial Agencies and National Statistical Institutes in common thematic and technical domains within statistical and geospatial communities.

✓ Ensure availability and accessibility of processing workflows, including open formats of programming codes, allowing the automatic or semi-automatic extraction of information from satellite images, the development of algorithms for indicator calculation and territorial classifications (e.g. ESS Degree of urbanization) and of its associated metadata, as a way to improve reporting harmonization and comparability of data.

- ✓ Develop initiatives that promote availability, accessibility and usability of geospatial data by making use of standard metadata and quality reference frameworks aligned with the requisites of the Generic Statistical Business Process Model and Metadata Reporting Standards for statistical production.

- ✓ Increase the collaboration with researchers and data providers to take full advantage of the available data and processing infrastructures and also for tuning operational workflows and regular computation of SDG indicators.

- ✓ Increase cooperation between National Statistical Institutes and Geospatial Agencies for the calculation of SDG indicators, to better address the territorial dimension of SDG indicators and to promote the relevance of geographical data in institutional national forums for SDG reporting.

5. In general, it is fundamental to increase collaboration between National Statistical Institutes and Geospatial Agencies (comprising the National Mapping and Cadastral Agencies) within countries and across European and Global institutions. This could be a way to improve processes and methodologies, to harmonize concepts, definitions and procedures, to develop relevant new statistical indicators and to assure consistent points of view in international forums. Action plans of cooperation in these domains can be defined and implemented with the aim of modernizing the statistical production chain and of promoting better data integration.

6. The specific analysis of the four selected SDG indicators on a global, European and national perspective regarding the contribution of geospatial data analysis and its integration with statistical data resulted on the following findings, as detailed in sections 3 and 4 of this report. The analytical outcomes should not be understood as exclusively related to one indicator, as they may be also relevant for other SDG indicators. These findings constitute a contribution, from a European perspective, to the work being developed by the Inter-Agency Expert Group on Sustainable Development Goals Indicators Working Group on Geospatial Information (IAEG-SDG WG GI).

7. The contribution of geospatial data to the monitoring framework of the SDG indicators goes further than having location variables that allow the disaggregation, aggregation and mapping of information. Geospatial analysis and models, similarly to statistical analysis, consists of a way of interpreting the data based on the application of specific techniques and data processing models. Measuring accessibility has a strong spatial character, since it is intrinsically associated to the physical distance to a place and the selected SDG tier II indicator **11.2.1 *Proportion of population that has convenient access to public transport by sex, age and persons with disabilities*** shows that geospatial data and modelling is at the core of this indicator. The detailed analysis on this indicator can be summarised in the following outcomes and findings:

Concepts	The Global Human Settlement Layer (GHSL) definitions could be used at the global level and, at the European level, a common definition of urban areas should be considered based on available common territorial typologies, such as the Degree of Urbanization (DEGURBA). The global metadata would benefit from a clear definition of stops and of terms such as comfortable environment and frequent service.
Data sources	Point-based population data provides a more flexible and easier way to calculate the indicator as no disaggregation procedure is needed. Public transport data on stops can entail different levels of detail on available routes, number of services and timetable information whether the indicator is being computed at the global, European or national level. A more general approach on the selection of public transport stops can be followed for global and European comparability and at the national level a more refined and differentiating approach can be applied when detailed data on public transport is available.
Computation and algorithm	Network distances, such as the shortest path, can provide more accurate results than Euclidian distances, but it requires quality data on the road network, including pedestrian walks to account for walking access. The proposed global computation takes into account the places of residence as reference points, but other reference points could be considered, such as schools, workplaces or markets.
Challenges regarding the use of geospatial data	▪ Availability of comparable point-based data on public transport data, especially including information on routes, frequencies and timetables. ▪ Availability of road network data for computation of distances, including for walking distances. ▪ Complying with the proposed global disaggregation by disability.

8. The geospatial analysis combining land cover and population data provides the possibility of deriving new metrics that are relevant to grasp important dimensions on human settlement planning and management. The SDG tier II indicator 11.3.1 ***Ratio of land consumption rate to population growth rate*** is a very straightforward example on this type of data combination. The detailed analysis on this indicator can be summarised in the following outcomes and findings:

Concepts	Global metadata should be more precise on the definition and use of urban area/city and built-up area as they refer to two different concepts. Urban area/city operationalization would benefit from an international statistical definition to deal with issues of comparability. The EU definitions as presented in the TERCET regulation can be taken as a reference for a worldwide concept following the discussions in UN Statistical Division for a global definition of cities as proposed by the European Commission, OECD and partners. The concept of built-up area should be used as a metric to capture artificial land and the expansion of land consumption over time and it is, therefore, a clear distinct concept from the one relating to urban area/city.
Data sources	GHSL should be considered as a ready to use product and/or European Space Agency Land Cover products, but special attention should also be given to stimulate European remote sensing derived products initiatives worldwide, such as Copernicus Imperviousness High Resolution Layer and CORINE Land Cover. National data sources can provide more detailed and high quality geospatial data for greater territorial disaggregation of the indicator.
Computation and algorithm	The algorithm on Land Use Efficiency (LUE) as proposed by the Joint Research Centre (JRC) should be considered as it deals with issues of zero population growth and periodicity of the available information by recommending a normalization of the results for a 10 year reference. JRC developed a toolbox in open code format which provides a good way to increase indicator harmonization and comparability.
Challenges regarding the use of geospatial data	▪ Accessibility, periodicity, timeliness and methodological stability of the data sources in order to measure progress over time. ▪ Identification of precise land cover components or categories to derive built-up areas based on different geospatial data products, including at the national level. ▪ Spatial resolution of input data and its impact on the quality of the territorial disaggregation of statistical outputs.

9. The monitoring framework for the 2030 Agenda on Sustainable Development includes a number of more challenging indicators due to the lack of data availability and existing established methodology. This is the case of the selected SDG former tier III (now tier II) indicator **11.7.1 *Average share of the built-up area of cities that is open space for public use by sex, age and persons with disabilities***, for which land use and cadastral data obtained using different geospatial based products can play a significant contribution for its operationalization. The detailed analysis on this indicator can be summarised in the following outcomes and findings:

Concepts	The concepts of urban area/city and built-up area should be defined in line with other indicators of Goal 11. In this vein, the concept of built-up area should be measured as artificial land and available common and harmonised territorial typologies (e.g. EU TERCET) should be used to capture the urban dimension. The global metadata should benefit from a more detailed description on the different dimensions that are supposed to be captured by the concept of open space for public use, especially, in order to achieve comparability across countries at the global level.
Data sources	GHSL provides global coverage to measure built-up areas and other remote sensing derived products, such as Copernicus Imperviousness High Resolution Layers and CORINE Land Cover provide ready to use comparable data at the European level. National data sources might have more detailed information and are able to better address the conceptual definitions of the indicator, namely the differentiation between public and private open space.
Computation and algorithm	Cadastral data can provide better data coverage for a more detailed territorial disaggregation and a more consistent and stable classification for measurements over time. The global metadata should include a reference on the substitutability of field data/non-geospatial inventory data for geospatial information as a valuable approach to increase comparability across countries.
Challenges regarding the use of geospatial data	▪ Definition and availability of data sources in order to measure the concept of public open space ensuring international data comparability. ▪ Availability of data sources with ownership (public vs. private). ▪ Combination of different geospatial data sources in order to grasp the necessary definitions to calculate the indicator: urban delimitation, built-up areas and open space for public use.

10. Earth observation data is a relevant source of information to monitor progress towards the SDGs, as this type of data can provide an objective and consistent view of the earth for different periods in time, at different scales and ensuring a coherent basis for comparability between different countries. The selected SDG tier I indicator **15.1.1 *Forest area as a proportion of total land area*** is a very good example to showcase the potential of earth observation data to increase the scope of territorial disaggregation of statistical indicators. The detailed analysis on this indicator can be summarised in the following outcomes and findings:

Concepts	Common technical specifications for targets to be managed by the indicator, in particular forest and inland waters should be better defined. A shared forest definition according to FAO standards that could be captured by geospatial layers would provide a way to increase data comparability at the global level.
Data sources	Global (ESA Land cover) and European (e.g. Copernicus High Resolution Layers, CORINE Land Cover) geospatial layers, based on remote sensing techniques, could be used for this indicator allowing a more detailed territorial segmentation of the indicator. At the European level, geospatial data layers are generated on a regular basis and the situation is the same for satellite data, since EU Sentinel satellites are planned to remain in orbit for several years.
Computation and algorithm	Vector or raster format geospatial layers should be used in order to allow proper data aggregation/disaggregation. Different levels of technical algorithm specifications can be considered for a single country depending on its objectives, but common, harmonized specifications and its availability in open formats provide a way to increase data comparability across countries.
Challenges regarding the use of geospatial data	▪ Stability of the geospatial data sources and of its methodology regarding processing workflows with a high degree of automation. ▪ Clear definition of who is responsible for the generation of the output product quality at the global level. ▪ Capacity building in using earth observation data and derived products to produce statistical indicators.

Background and acknowledgments

The United Nations initiative on Global Geospatial Information Management (UN-GGIM), established in July 2011, aims at playing a leading role in setting the agenda for global geospatial information development and promoting the benefits of geospatial information for addressing national policy and key global challenges.

The regional committee UN-GGIM: Europe, officially established on 1 October 2014, aims to ensure that the National Mapping and Cadastral Authorities (NMCA) and National Statistical Institutes (NSI) in the European UN Member-States, the European Institutions and associated bodies work together to contribute to a more effective management and availability of geospatial information in Europe, and its integration with other information, based on user needs and requirements. Three working groups have been established under the scope of this initiative, one focusing on Core Data, other dealing with Data Integration issues and one addressing Geodetic Reference Frames.

The Working Group (WG) on Data Integration is chaired by Germany (Pier-Giorgio Zaccheddu from the Federal Agency for Cartography and Geodesy – BKG being the technical leader). Under the scope of the work plan defined for the period between 2014 and 2017 three deliverables were produced:

1. Definition of the priority user needs for combinations of data[1]
2. Recommendations for methods implementing the prioritised combinations of data[2]
3. Recommendations about how to manage side-effects induced by data combinations[3]

UN-GGIM: Europe WG on Data Integration activities have been taking into account the background of the Sustainable Development Goals (SDG) monitoring framework. The work plan defined for the 2017-2019 period placed the focus of activities on the contribution of geospatial data and analysis, particularly when integrated with statistical data, for the purpose of addressing SDG indicators and its monitoring framework, by comprising supportive tasks to the global process and specific tasks with a European perspective. In addition, the 2017-2019 work plan for UN-GGIM: Europe foresaw a two-way

[1] UN-GGIM: Europe Working Group B1 report on Definition of the priority user needs for combinations of data.
[2] UN-GGIM: Europe Working Group B2 report on Recommendations for methods implementing the prioritised combinations of data.
[3] UN-GGIM: Europe Working Group B3 report on Recommendations about how to manage side-effects induced by data combinations.

interaction with the Inter-Agency and Expert Group on Sustainable Development Goals (IAEG-SDG) Working Group on Geospatial Information (WG GI)[4] and established that this interaction should be guaranteed by the WG on Data Integration[5].

UN-GGIM: Europe activities have been supported by Eurostat. The 2018 European Statistical Programme[6], under the strategic objective of increasing use of spatial information combined with social, territorial, economic and environmental statistical information, clearly states that Eurostat actively supports the work of this initiative, especially on increasing the cooperation and synergies between statistical and geospatial communities. Additionally, the work developed under the UN-GGIM: Europe initiative is carried out under a close collaboration with the United Nations Economic Commission for Europe (UNECE)[7], namely through the identification of the strategic areas of collaboration.

The following tasks were accepted by the UN-GGIM: Europe at its 4th Plenary Session on 7-8 June 2017 and re-confirmed at the 5th Plenary Session on 6-7 June 2018:

1. Draft an outreach paper aiming at responsible ministries but also at stakeholders that are responsible for coordinating the political agenda across ministries. The paper shall make use of the findings/recommendations of the Work Group deliverables already published.

2. Select and analyse Sustainable Development global, regional and national indicators reflecting "data integration" aspects and cross-cutting issues regarding the integration of geospatial and statistical data based on a global, European and national perspective.

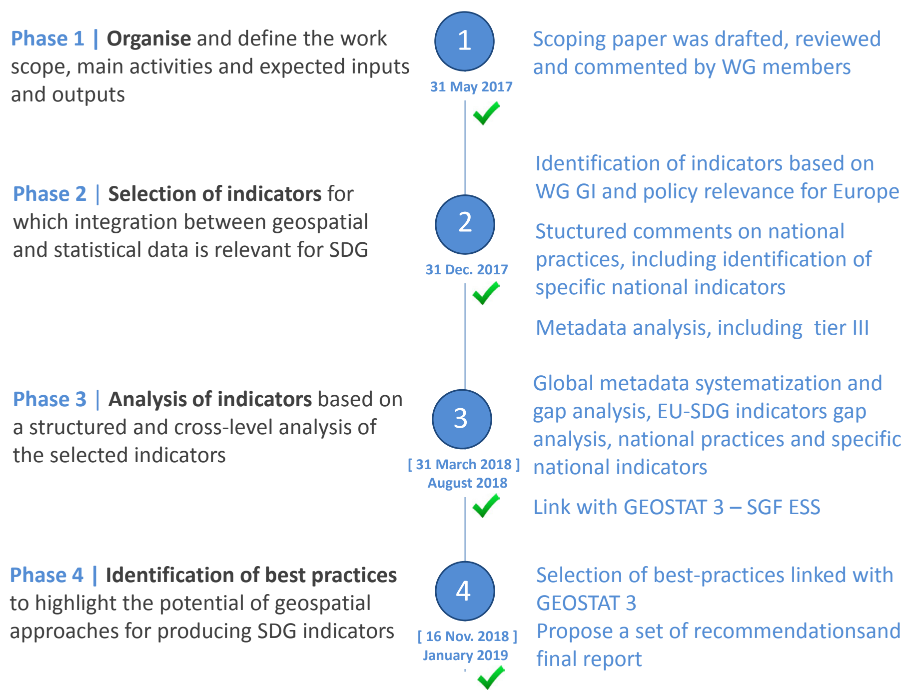
[4] In the third meeting of the IAEG-SDG held in Mexico City, Mexico (from 30 March to 1 April 2016), the work plan and next steps agreed included the creation of three working groups on SDMX, Geospatial information, and Interlinkages.
[5] UN-GGIM: Europe nominated two participants to contribute to the IAEG-SDGs WG GI. On behalf of the Executive Committee of UN-GGIM: Europe, Mr. Pier-Giorgio Zaccheddu (BKG, Germany), Technical Leader of WG on Data Integration, was nominated to represent UN-GGIM: Europe in this Working Group. UN-GGIM: Europe's second representative to the WG GI is Mr. Fabio Volpe (Geo Content Innovation at e-GEOS, Italy).
[6] The Commission adopted the work programme 2018 for the implementation of the European statistical programme 2013-2020 on 19 December 2017. The European Statistical Programme 2018 defines the strategic priorities for statistical production.
[7] UNECE (2018) *Improving collaboration between United Nations Economic Commission for Europe and the United Nations Global Geospatial Information Management Europe on geospatial data and statistics*, 66th plenary session of the Conference of European Statisticians, Geneva, 18-20 June.

Two subgroups have been established to address these tasks. Task 1 was led by Eurostat (Ekkehard Petri) and task 2 by the National Statistical Institute (NSI) of Portugal (Francisco Vala). The activities undertaken in both tasks relied on the participation of WG members from 19 countries, representing either geospatial agencies or statistical offices. Besides Eurostat, the WG also benefited from the participation of other European institutions, namely the European Environment Agency (EEA) and the Joint Research Centre (JRC). In addition, the WG has been involved with the GEOSTAT 3 project, BKG being the focal point, by contributing to the development of the European version of the Global Statistical Geospatial Framework (GSGF) and on the contribution of geospatial information for SDG indicators.

The present report aims at presenting the results of the work developed by the subgroup dealing with task 2. This subgroup focused on the contribution of geospatial data analysis and its integration with statistical data, for the purpose of addressing specific SDG indicators and established a line of work based on four main phases. Figure 1 shows a brief outline of the work phases developed and a more detailed description is presented next.

Figure 1 I

Outline of the work phases and actions achieved



Phase 1 I Organise the work scope by defining the main activities to be carried out and the expected inputs and outputs:

Actions:

- Research and collection of information with the objective of elaborating a scoping paper reflecting on the work to be carried out, including the main phases and actions.

Outputs:

- Scoping paper document – *The territorial dimension in SDG indicators: the contribution of geospatial data and analysis and its combination with statistical data.*

Phase 2 I Select specific indicators for which integration between geospatial and statistical data is relevant for SDG monitoring based on a territorial perspective:

Actions:

- Identification of the SDG indicators to be analysed based on the group of indicators selected by the IAEG-SDG WG GI and indicators addressing issues on accessibility, environment and urban/rural segmentation, based on policy relevance for the European context and relevant European and national initiatives.
- Structured comments by subgroup 2 members based on a template for indicator analysis which comprised a set of fields to describe: i) the current reporting situation both at global and national level with metadata analysis, including also tier III indicators; ii) gap analysis on the methodology and geospatial data integration suggested for the indicator; iii) identification of corresponding EU SDG indicators and of specific national indicators.

Outputs:

- Documents compiling all the structured comments received on the indicator analysis.
- Selection of the SDG indicators, and corresponding EU SDG indicators and specific national indicators.

Phase 3 I Analysis of indicators based on a structured and cross-level (global, European and national) analysis of the selected SDG indicators:

Actions:

- Analysis of the selected SDG indicators focusing on i) global metadata systematization and gap analysis; ii) gap analysis for the corresponding EU SDG indicators and iii) systematization of national practices and of specific national indicators.
- Link with GEOSTAT 3 work to maximise the understanding of how the new Global Statistical Geospatial Framework for the European Statistical System will help countries to integrate statistical and geospatial data sources.

Outputs:

- Consolidated detailed analysis documents for each selected SDG indicators based on a harmonised template comprising global metadata systematization, national practices and corresponding EU SDG indicator systematization – Gap analysis.
- Brief discussion papers for each selected SDG indicator.

Phase 4 I Identification of best practices to highlight the potential of geospatial approaches resulting from the integration of geospatial and statistical data for producing SDG indicators:

Actions:

- Selection of best practices on the contribution of geospatial data and analysis to the selected SDG indications, taking into consideration the work of GEOSTAT 3.
- Propose a set of recommendations and elaboration of the final report.

Outputs:

- Final report

To fulfil its task, the subgroup relied on the active participation and contributions of members from 13 countries, 4 from geospatial agencies and 9 from statistical offices.

Indicators coordinators were nominated to summarize and discuss countries and Eurostat contributions for the *detailed analysis* of the four selected indicators and to produce a *Brief discussion paper* for each indicator: 11.2.1 – Statistics Austria; 11.3.1 – Statistics Portugal; 11.7.1 – Statistics Sweden; 15.1.1 – e-GEOS (Italy) [Annex II]. A full list of those who have contributed can be found at the end of this report [Annex III].

We take this opportunity to thank all the members who have provided inputs and comments regarding the selected indicators and who have actively participated in the WG meetings and discussions.

The UN-GGIM: Europe Executive Committee approved this report in April 2019 and uploaded it to the UN-GGIM website.

1 INTRODUCTION

On 25 September, 2015 the United Nations General Assembly adopted the resolution Transforming our world: the 2030 Agenda for Sustainable Development[8] which defines, at the global level, 17 Sustainable Development Goals [Figure 2] and 169 associated targets. These goals and targets were defined to be monitored according to a global indicator framework to be defined and implemented by the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDG).

The global indicator framework was adopted by the General Assembly on 6 July 2017[9], encompassing around 232 indicators[10] to monitor the progress towards Sustainable Development and emphasizing the importance of geographical disaggregation of the indicators, along with sex, age, income, race, ethnicity, migratory status, and disability in order to cope with the motto of *leaving no one behind*.

Figure 2 |

The 17 Sustainable Development Goals

[8] Resolution adopted by the General Assembly on 25 September, 2015 on Transforming our world: the 2030 Agenda for Sustainable Development (A/RES/70/1).

[9] Resolution adopted by the General Assembly on 6 July 2017 on Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development (A/RES/71/313).

[10] The list includes 232 indicators, but the total number of indicators listed in the global indicator framework of SDG indicators corresponds to 244, since nine indicators repeat under two or three different targets - *List of global Sustainable Development Goal indicators* as agreed on the 48th session of the UN Statistical Commission, March 2017.

At the global level, the IAEG-SDG has defined a three tier system of classification of indicators regarding data availability and established methodology:

i) a first tier of indicators for which an established methodology exists and data are already widely available (*tier I*);

ii) a second tier for which a methodology has been established but for which data are not easily available (*tier II*);

iii) a third tier for which an internationally agreed methodology has not yet been developed (*tier III*).

Therefore, geospatial information and its integration with statistical data can provide relevant contributions to address these gaps.

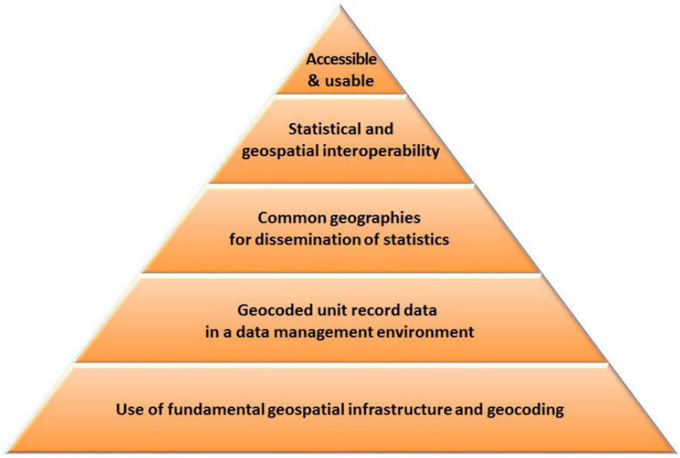
The first report of the WG on Data Integration on the *Priority user needs for combinations of data*, provided a number of use cases addressing relevant information for the SDGs and recommended that Member-States should initiate a process to increase the number of national, authoritative geospatial datasets (addresses and others) meeting stakeholders requirements (like statistics) and to promote the use of geospatial workflows and technology, as a key to advance on the integration of geospatial and thematic information.

At the global level, it is also important to take into consideration the principles presented in the United Nations Expert Group on Integration of Statistical and Geospatial Information (EG-ISGI) proposal for a *Global Statistical Geospatial Framework*[11] (GSGF). This proposal describes how to better achieve this integration in an effective and consistent way, thus facilitating a consistent production and approach for integration of geo-statistical information. The EG-ISGI has defined five guiding principles that are considered essential for integrating geospatial and statistical information [Figure 3]. The principles were adopted at the 6th session of the United Nations Committee of Experts on Global Geospatial Information Management and were endorsed at the 48th session of the United Nations Statistical Commission.

[11] United Nations Expert Group on the Integration of Statistical and Geospatial Information's proposal for a Global Statistical Geospatial Framework - 47th Session of the United Nations Statistical Commission and a Global Consultation.

Figure 3 |

The five principles of the Global Statistical Geospatial Framework



Source: Australian Bureau of Statistics (ABS)/ UN-GGIM.

The 232 indicators defined at the global level attempt to capture a worldwide progress towards Sustainable Development, but countries around the world are in different stages of economic, social and environmental development. So, global indicators are not equally relevant to every country, in the same way and at the same point in time. Therefore, countries and regions are encouraged by the United Nations to develop specific indicators to address the specificities of their national or regional circumstances to monitor the progress towards Sustainable Development.

The European Union responded with a Commission Communication *Next steps for a sustainable European future*[12] that called for a detailed regular monitoring of the SDGs in an EU context from 2017 onwards and the development of a reference indicator framework for this purpose. Eurostat, in close cooperation with other EU Directorates-General, has established an EU SDG indicator set to measure progress towards the SDGs in an EU context. The definition of the EU SDG indicator set was based on specific principles, namely the indicator set is structured along the 17 global SDGs to allow monitoring EU policies in the perspective of the United Nations 2030 Agenda, and it includes a maximum of 100 indicators. The indicator selection was based on a three-step approach, involving:

i) assessment of the indicator policy relevance;

ii) admissibility criteria (based on the Code of Practice of European Statistics);

iii) quality rating based on Euro-SDMX Metadata Structure (ESMS) and the ESS quality framework.

[12] Commission Communication on *Next steps for a sustainable European future* (COM(2016) 739 adopted on 22 November 2016).

[13] The latest publication corresponds to the 2018 edition: Eurostat (2018) *Sustainable development in the European Union: Monitoring report on the progress towards the SDGs on an EU context*. Luxembourg: Publications Office of the EU.

Since its establishment, an annual overview has been published with the aim of monitoring progress towards the SDGs in the EU context[13].

At the European level, there is also on-going work on establishing the GSGF in Europe, including recommendations on its implementation and testing its usefulness for SDG monitoring. This is a task that has been undertaken by the GEOSTAT 3 project. The recommendations from the GEOSTAT 2 project on establishing a geocoding infrastructure for statistics based on geocoded address registers is one of the key contributions to the development and the establishment of the GSGF in Europe. Systematic geocoding of administrative, statistical and other types of data using national geocoding infrastructures will geo-enable massive amounts of existing data that are directly or indirectly relevant to address SDG indicators.

Additionally, the European Union's INSPIRE Directive to create a spatial data infrastructure for the purpose of environmental policies and the Copernicus Programme [see Box 1] to develop information services based on satellite Earth Observation and in situ data, provide a particular relevant regional context to manage a wider integration of geospatial and statistical data to include a territorial dimension within the SDG indicators.

At the national level, some countries have been addressing the 2030 agenda by defining and implementing national strategies to monitor their national progress towards the SDGs.

Besides taking into consideration the goals, targets and indicators defined at the global level, national strategies on Sustainable Development have also defined specific national targets and corresponding indicators to monitor them. This way, countries are able to monitor progress having their specific national context as a reference, by defining targets that are in line with their specific policy agenda and the corresponding indicators to measure them.

In Europe, some countries have defined national strategies and national implementation plans on Sustainable Development. This is the case, for example, of Ireland (Sustainable Development Goals National Implementation Plan 2018-2020), Germany (German Sustainable Development Strategy), Sweden (Sweden's Action Plan for the 2030 Agenda) and Switzerland (Sustainable Development Strategy 2016-2019).

More specifically, in the case of Ireland, the national implementation plan for 2018-2020 aligns the 17 SDGs with the eight strategic national themes and principles for achieving Sustainable Development in Ireland[15], and provides the conditions in which Ireland expects to fulfil its plan. This includes, regarding data and indicators, the establishment of an online SDG platform for exploring, downloading and combining available data, including geospatial data. This platform results from a collaboration between the NMCA and the NSI with a focus on visualising indicators using geographic information technologies and, thus, providing map layers of combined statistical and geographical information.

[15] Defined in the Irish background strategy on sustainable development *Our Sustainable Future – the Framework for Sustainable Development*.

Aim of the report

Taking into consideration this broad context on the Sustainable Development Goals monitoring across different levels of implementation, the aim of this report is to address the territorial dimension of SDG indicators by focusing on the contribution of geospatial data analysis and its integration with statistical data based on a global, European and national perspective, by taking into account, at the global level, the activities of the IAEG-SDG WG GI, and also the background and experiences of European and national initiatives addressing the SDGs from a geospatial perspective.

2 FROM GEOSPATIAL DATA TO STATISTICAL INDICATORS: CHALLENGES TO ADDRESS THE SDGs

The need and relevance of geospatial data to address the SDGs

How geospatial information can be used in the production of statistical indicators is an important step forward to address the challenges of computing SDG indicators and to monitor the 2030 Agenda for Sustainable Development in a territorial perspective, *leaving no one behind*.

The *in-depth review on geospatial information services based on official statistics* prepared by the ONS[16] and discussed in the 64th UN Conference of European Statisticians (Paris, 27-29 April 2016) states that geospatial data and spatial analysis can play a very important role and make several contributions to the indicator framework supporting SDG monitoring.

[16] *In-depth review of developing geospatial information services based on official statistics*, Note by the UK Office for National Statistics.

Specifically, its contribution can be considered at three levels:

i) by contributing directly to the proposed framework, not only in terms of sources of information to increase data availability and spatial disaggregation, but also in terms of methods and analysis to produce indicators resulting from the integration of geospatial and statistical information;

ii) by promoting common accepted standards and frameworks to guarantee comparability;

iii) by encouraging innovation and modernization.

Recognising the relevance of geospatial data and the possibilities of location-based variables to support the SDG monitoring for a more detailed picture of the progress in and across countries, a Working Group on Geospatial Information, reporting to the IAEG-SDG (IAEG-SDG WG GI) was created in March 2016, with the main purpose of guaranteeing that a statistical and geographic location is reflected in the global indicator framework. One of the first tasks tackled by this working group was to review the global indicators and metadata according to a geospatial lens, which resulted in the identification of a set of SDG indicators that directly or indirectly benefit from geospatial information, the so called *Short list*[17].

The resolution adopted by the Economic and Social Council (ECOSOC) on 27 March 2016 on *Strengthening institutional arrangements on geospatial information*[18] also came to stress the need of having a coordinated and coherent global geospatial information management regarding data collection, data sharing, data dissemination and capacity building for the implementation of the 2030 Agenda for Sustainable Development. More recently, the *Cape Town Global Action Plan for Sustainable Development Data*[19] issued on 15 January 2017, as a result of the first UN World Data Forum, adopted by the United Nations Statistical Commission in March 2017, recognizes that the implementation and monitoring of the 2030 Agenda for Sustainable Development requires the collection, analysis and availability of an unprecedented amount of data and indicators at global, regional, national and sub-national levels. In order to address this requirement, this action plan stresses the relevance of National Statistical Systems and of their ability to respond to the increasing need of data, including at different levels of territorial disaggregation.

Besides other objectives, this action plan aims at promoting the modernization of statistical standards through the implementation of standardised structures for the exchange of data and metadata on the social, economic and environmental dimensions of Sustainable Development and the integration

[17] IAEG-SDG WG GI Short list results of the analysis of the Global Indicator Framework with a “geographic location” lens.

[18] Resolution adopted by the Economic and Social Council on 27 July 2016 on *Strengthening institutional arrangements on geospatial information management* (E/RES/2016/27).

[19] The Cape Town Global Action for Sustainable Development Data was prepared by the High-level Group for Partnership, Coordination and Capacity Building for Statistics for the 2030 Agenda for Sustainable Development.

of geospatial data into statistical production programmes at all levels and through the integration of modern geospatial information management systems within mainstream management statistical production, including in terms of metadata and technological infrastructures.

In August 2018, the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM), working together with a wide range of experts, UN Member-States and the World Bank, adopted a new strategic framework for the geospatial community with the purpose of helping countries strengthen their management of geospatial information. The *Integrated Geospatial Information Framework and its Overarching Strategy*[20] states that geospatial information management is crucial to implement the SDGs, stressing that accessible and quality geospatial information is a core element to meet data needs to monitor progress towards this global agenda.

Therefore, the potential of geospatial information to advance the 2030 Agenda has been put forward by the official governance model for its implementation – IAEG-SDG WG GI – and by other relevant international references. Additionally, the earth observation community, such as the Group on Earth Observations (GEO[21]) and its specific initiative on Earth Observations in Service of the 2030 Agenda supports the efforts of integrating earth observations and geospatial information into national development and monitoring frameworks for the SDGs.

At the European level, information services based on earth observation and in situ data have been developed and made available by the Copernicus Programme [Box 1], providing harmonized relevant data themes for SDG monitoring.

[20] *Integrated Geospatial Information Framework: A Strategic Guide to Develop and Strengthen National Geospatial Information Management*, adopted in the 8th session of the UN-GGIM Committee of Experts held in New York, 1-3 August 2018.
[21] The Ministerial Declaration endorsed at the last GEO Summit (Mexico City, November 2015) stressed the importance of leveraging Earth observations to support the implementation, monitoring and evaluation of the SDGs and called for an initiative in this field.

Box 1 |

Copernicus Programme

Copernicus is the EU Observation Programme, providing information services based on satellite Earth Observation and in situ (non-space) data, implemented in partnership with the Member-States, the European Space Agency (ESA), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), EU Agencies and Mercator Ocean.

Satellite data are collected by a family of dedicated, EU-owned satellites, the Sentinel constellation, providing regular information over Europe and the rest of the world. Actually, the constellation satellites in orbit are:

- Sentinel-1A and 1B providing radar images, therefore all-weather, day and night images
- Sentinel-2A and 2B providing multispectral high resolution optical images
- Sentinel 3A and 3B providing high-accuracy optical, radar and altimetry data
- Sentinel-5P providing atmospheric measurements with high spatio-temporal resolution

The Copernicus Services transform data provided by this constellation and in situ data into value-added information and geospatial layers by processing and analysing the data, integrating it with other sources and validating the results.

The free, full and open data policy adopted for the Copernicus programme foresees access available to all users for the Sentinel data products, through the Sentinel data access hub. In addition to the satellite data download services, the Sentinel Data Products are available in the Copernicus Data and Information Access Service (DIAS) cloud environments. Each DIAS provides processing resources, tools and complimentary data sources at commercial conditions to further facilitate the access to Sentinel data.

Therefore, for the generation of customized geospatial layers and information, end users can download satellite data and take care of their processing, or use the DIAS platforms data and functionalities. Otherwise, users can access to the six thematic streams of Copernicus services, providing reliable and up-to-date information products in six thematic areas:

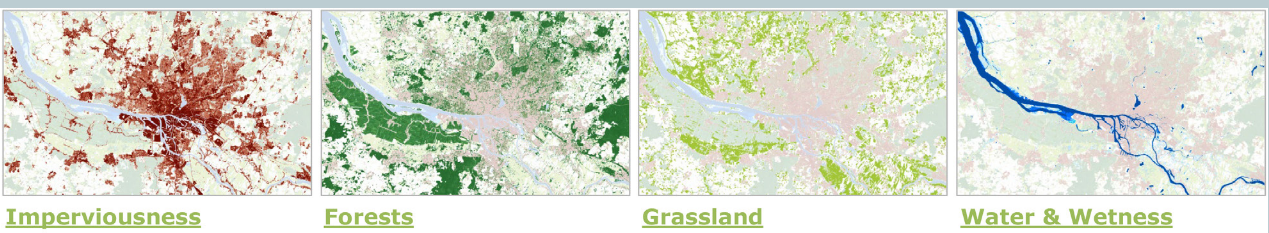
- Atmosphere Monitoring
- Marine Environment Monitoring
- Land Monitoring
- Climate Change
- Emergency Management
- Security

Of relevant interest is the Copernicus Land Monitoring Service (CLMS) providing geographical information on land cover and its changes, land use, vegetation state, water cycle and earth surface energy variables, supporting applications in many domains. CLMS consists of four main components:

- The systematic monitoring of biophysical parameters produces mainly a series of qualified bio-geophysical products on the status and evolution of the land surface.
- Land cover and land use mapping produces land cover classifications at various level of detail, both within a Pan-European and global context. At the Pan-European level, these are complemented by detailed layers on land cover characteristics, such as imperviousness, forests, grassland, water and wetness [Figure 4].
- Thematic hot-spot mapping aims to provide tailored and more detailed information on specific areas of interest.
- Imagery and reference data provide satellite image mosaic in high and very high resolutions and reference datasets (such as hydrography and elevation).

Figure 4 |

Copernicus Pan-European High Resolution layers from the Land Monitoring Services



Challenges of data integration

The integration of geospatial and statistical information with the purpose of producing relevant statistical indicators for the Sustainable Development Agenda does not come without a certain number of challenges. One first and relevant challenge is related to the fact that geospatial and statistical data come from two different domains and communities. Therefore, there are conceptual, methodological and technical differences that need to be understood in the context of the statistical production process model and its Generic Statistical Business Process Model (GSBPM) and the standard references regarding metadata and quality, such as the Single Integrated Metadata Structure (SIMS) and its underlying reporting structures as well as the ESS Quality and Reference Metadata Standards (ESQRS).

An additional challenge is related to increasing the understanding and knowledge of statistics on the value and use of geospatial information not only for greater spatial disaggregation, but also to produce new and relevant statistical indicators to grasp emerging dimensions of the social and economic reality, and as a way to support the modernization, flexibility and efficiency of the statistical production process model.

One crucial dimension to address SDG indicators in a territorial perspective is to have common regional and territorial frames to structure the statistical outputs within a country in a comparable way to depict internal differentiation of progress towards sustainable goals. The ESS established such a territorial framework in the early 70s – the Nomenclature of Territorial Units for Statistics (NUTS) – which became legally binding for European Union countries in 2003. Eurostat manages the implementation of this regional framework for EU countries, for the European Free Trade Association (EFTA) countries and accessing countries.

The Organisation for Economic Co-operation and Development (OECD) also has a similar regional structure for its member countries which is coherent with the ESS structure for the countries covered by the NUTS classification. More recently, the ESS has established a common set of territorial typologies (TERCET) to better grasp the asymmetries of development within countries and OECD also applies some of these typologies [Box 2].

Box 2 |

The European Statistical System territorial framework for data integration and comparability: NUTS and TERCET

At the beginning of the 1970s, Eurostat set up the NUTS classification as a single, coherent system for dividing up the EU's territory in order to produce regional statistics for the Community. For around thirty years, implementation and updating of the NUTS classification was managed under a series of "gentlemen's agreements" between the Member-States and Eurostat. The Commission Regulation (EC) No 1059/2003, was adopted in May 2003 and entered into force in July 2003, providing a legal status to the NUTS classification.

The NUTS classification is a three level hierarchical system for dividing up the economic territory of the EU member countries:

NUTS 1 major socio-economic regions

NUTS 2 basic regions for the application of regional policies

NUTS 3 small regions for specific diagnoses

This hierarchical system is managed according to three principles:

Principle 1 *POPULATION THRESHOLDS*

The NUTS regulation defines minimum and maximum population thresholds for the size of the NUTS regions: NUTS 1 between a minimum of 3 and a maximum of 7 million inhabitants; NUTS 2 between 800 thousand and 3 million; NUTS 3 between 150 thousand and 800 thousand.

Principle 2 *NUTS FAVOURS ADMINISTRATIVE DIVISIONS*

For practical reasons the NUTS classification generally mirrors the territorial administrative division of the Member-States. This supports the availability of data and the implementation capacity of policy.

Principle 3 *STABILITY*

The NUTS classification can be amended, but generally not more frequently than every three years.

Additionally, 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 3, Local Administrative Units (LAU) and grid cells and also details the definition and purpose of the LAU and establishes a system of statistical grids for statistical purposes.

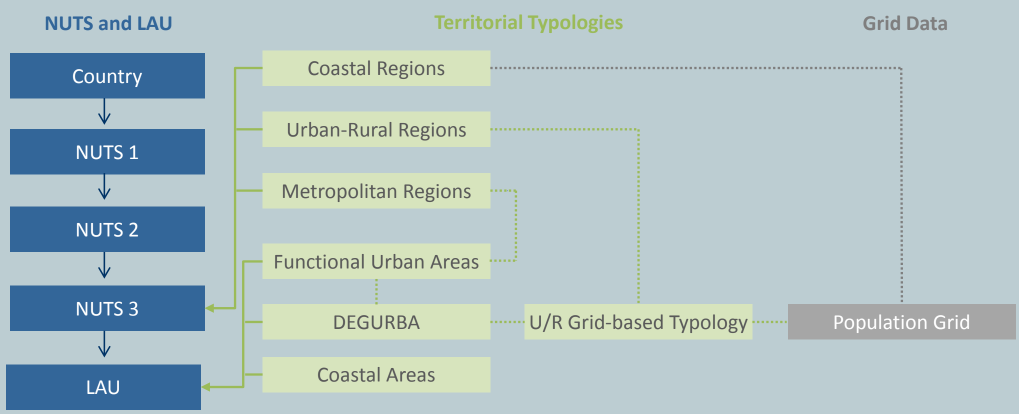
The **grid-based typology** (1 km^2) defines 'urban centres', 'urban clusters' and 'rural grid cells'.

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' based on the grid typology; 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 distinguishes between 'Coastal regions' and 'Non-coastal regions'.

Figure 5 |

EU Territorial architecture for data integration and comparability



Source: Based on Eurostat, NUTS – Nomenclature of Territorial Units for Statistics.

For the sake of data integration and comparability at global level it is extremely important to adopt common territorial frameworks worldwide, at regional and local levels – including grid systems and territorial typologies – by taking the example of the EU/OECD work towards comparable sub-national spatial units for statistical analysis.

Increasing the articulation between geospatial and statistical communities is an important step in this context. The work being developed under the UN-GGIM: Europe initiative and, at the EU level, the work of the European Forum for Geography and Statistics (EFGS) and the GEOSTAT projects are as well extremely relevant for this purpose. Specifically, the GEOSTAT 2 project included as one its objective the promotion of the application of spatial statistics and the integration of geospatial information into the statistical production chain, within the framework of the GSBPM and, more recently, the work carried out by GEOSTAT 3 to provide an implementation guide for the GSGF in Europe is expected to contribute to the harmonization of methods for the integration of statistical and geospatial information within the European Statistical System (ESS) and to its modernization, increased efficiency and flexibility regarding statistical outputs, and to provide a better foundation for collaboration between statistical and geospatial communities.

On this point in particular, it is important to highlight the results of the questionnaire poll to promote data integration for policy outreach, conducted as part of the activities of the UN-GGIM: Europe Work Group on Data Integration and its subgroup on the policy outreach paper[22], which point out that 76% of the countries already have an active agreement between the NSI and the NMCA and, in addition, 35% reported having a national action plan for their cooperation. Furthermore, and despite 53% of the countries having reported that they have a National Geospatial Information Management, only 29% countries indicated that it covers aspects of data integration with statistics.

[22] This questionnaire was addressed to both NSI and NMCA for a coordinated reply. A total of 28 European countries completed the questionnaire, namely Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Lithuania, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom and Ukraine.

The selection process of the SDG indicators

For the process of selecting the SDG indicators to be analysed, the starting point was to consider the IAEG-SDG WG GI selected Short list of 15 indicators for which geospatial information and its integration with statistical data can contribute directly. Additionally, this Short list includes 9 indicators for which geospatial information can indirectly support its production[23] [Table 1].

An analysis carried out by the Group on Earth Observations (GEO)[24], with the purpose of highlighting the potential role of earth observations to support the global indicator framework, identified a group of 29 indicators that can be directly or indirectly supported by earth observations. Table 1 shows that 14 out of the 29 indicators identified by the GEO group corresponded to the same indicators identified by the IAEG-SDG WG GI, and that 11 match the 15 indicators identified by the IAEG-SDG WG GI for which geospatial information has a direct contribution.

At the European level, Eurostat addressed the relevance of the spatial dimension in SDG indicators and identified a group of 78 out of the 232 indicators[25]. Although, the spatial dimension seems to be more associated with some of the Goals (Goals 6, 11 and 15), Eurostat's systematization showed that the spatial dimension can be identified in all of the 17 SDGs. The spatial dimension is associated to different types of data breakdown, namely the importance of the Degree of urbanization (DEGURBA) to capture urban/rural segmentation [see Box 2 and Box 6].

[23] This list of SDG indicators directly or indirectly benefiting from geospatial indicators was the result of the IAEG-SDG WG GI 2nd meeting that took place in Mexico City between 12-14 December 2016 in the premises of Instituto Nacional de Estadística y Geografía (INEGI).

[24] *Earth Observations on support of the 2030 Agenda for Sustainable Development.*

[25] Jortray, M. *Spatial statistics for Sustainable Development – challenges for UN-GGIM: Europe.* UN-GGIM: Europe Plenary Meeting, 7 October2015, Belgrade, Serbia.

Table 1 |

SDG and EU SDG indicators directly or indirectly benefiting from geospatial information

Indicators identified by the IAEG-SDG WG GI for which geospatial information has a direct contribution	**Tier**	**GEO group**	**Urban Audit**	**EU SDG indicators**
2.4.1 | Proportion of agricultural area under productive and sustainable agriculture	II	✓		Area under organic farming [partial]
6.3.2 | Proportion of bodies of water with good ambient water quality	II	✓		
6.5.2 | Proportion of transboundary basin area with an operational arrangement for water cooperation	I			
6.6.1 | Change in the extent of water-related ecosystems over time	I	✓		
9.1.1 | Proportion of the rural population who live within 2 km of an all season road	II	✓		
9.c.1 | Proportion of population covered by a mobile network, by technology	I			
11.2.1 | Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities	II	✓	✓	Difficulty in accessing public transport [similar]
11.3.1 | Ratio of land consumption rate to population growth rate	II	✓	✓	Artificial land cover per capita [similar]
11.7.1 | Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities	II	✓	✓	
14.2.1 | Proportion of national exclusive economic zones managed using ecosystem-based approaches	III			
14.5.1 | Coverage of protected areas in relation to marine areas	I	✓		Surface of marine sites designated under NATURA 2000 [similar]
15.1.1 | Forest area as a proportion of total land area	I	✓		Forest area as a proportion of total land area [integral]
15.1.2 | Proportion of important sites for terrestrial and freshwater biodiversity that are covered by protected areas, by ecosystem type	I			Surface of terrestrial sites designated under NATURA 2000 [similar]
15.3.1 | Proportion of land that is degraded over total land area	II	✓		
15.4.1 | Coverage by protected areas of important sites for mountain biodiversity	I	✓		Surface of terrestrial sites designated under NATURA 2000 [similar]
Indicators identified by the IAEG-SDG WG GI for which geospatial information has a indirect contribution	**Tier**	**GEO group**	**Urban Audit**	**EU SDG**
1.1.1 | Proportion of population below the international poverty line, by sex, age, employment status and geographical location (urban/rural)	I			
1.4.2 | Proportion of total adult population with secure tenure rights to land, with legally recognized documentation and who perceive their rights to land as secure, by sex and by type of tenure	II	✓		
4.5.1 | Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict affected, as data become available) for all	I/II/III			
5.4.1 | Proportion of time spent on unpaid domestic and care work, by sex, age and location	II			
5.a.1 | Proportion of total agricultural population with ownership or secure rights over agricultural land	II	✓		
5.a.2 | Proportion of countries where the legal framework (including customary law) guarantees women's equal rights to land ownership and/or control	II			
5.2.2 | Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner in the previous 12 months, by age and place of occurrence	II			
11.7.2 | Proportion of persons victim of physical or sexual harassment, by sex, age, disability status and place of occurrence, in the previous 12 months	III			
15.4.2 | Mountain Green Cover Index	I	✓		
Additional indicators identified by the Group on Earth Observations (GEO)	**Tier**	**GEO group**	**Urban Audit**	**EU SDG**
3.9.1 | Mortality rate attributed to household and ambient air pollution	I	✓		
6.3.1| Proportion of wastewater safely treated	II	✓		
6.4.2 | Level of water stress: freshwater withdrawal as a proportion of available freshwater resources	I	✓		
6.5.1 | Degree of integrated water resources management implementation (0-100)	II	✓		
7.1.1 | Proportion of population with access to electricity	I	✓		
9.4.1 | CO2 emission per unit of value added	I	✓		
11.1.1 | Proportion of urban population living in slums, informal settlements or inadequate housing	I	✓		
11.6.2 | Annual mean levels of fine particulate matter (e.g. PM2.5 and PM10) in cities (population weighted)	I	✓		Exposure to air pollution by particulate matter [integral]
12.a.1 | Amount of support to developing countries on research and development for sustainable consumption and production and environmentally sound technologies	III	✓		
13.1.1 | Number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population	II	✓		
14.3.1 | Average marine acidity (pH) measured at agreed suite of representative sampling stations	II	✓		
14.4.1 | Proportion of fish stocks within biologically sustainable levels	I	✓		
17.6.1 | Number of science and/or technology cooperation agreements and programmes between countries, by type of cooperation	III			
17.18.1 | Proportion of sustainable development indicators produced at the national level with full disaggregation when relevant to the target, in accordance with the Fundamental Principles of Official Statistics	III	✓		
				Estimated soil erosion by water

Note: The table provides the tier classification updated as of 13 February 2019.

Additionally, in the context of the EU Urban Audit project[26], Eurostat analysed with Member-States possible contributions to the project of Urban Audit to the UN SDG Agenda, by mainly focusing on the scope of the indicators associated to Goal 11 - *Make cities and human settlements inclusive, safe, resilient and sustainable*. This assessment was taken into account for the new list of variables to be considered within the Urban Audit project and 7 SDG indicators were identified as relevant for monitoring the state of the European cities, with 3 indicators identified as benefiting from the use of geospatial information for calculation [see Table 1].

The work being carried out by Eurostat on the EU SDG indicator set to monitor EU policies in the perspective of the UN 2030 Agenda was also taken into consideration. The full list of EU SDG indicators is composed by 100 indicators and the analysis of the list referenced to April 2017 identified that 31 indicators included disaggregation at NUTS 2 level, 34 according to DEGURBA and 14 by both NUTS 2 and DEGURBA, and that a sub-group of EU SDG indicators seems to potentially benefit from geospatial information and of its integration with statistical data. These indicators have an integral, similar or partial correspondence with the global SDG indicator, as presented in Table 1.

Therefore, the background for the final reference list for selecting the global SDG indicators considered all the 15 indicators identified by the IAEG-SDG WG GI for which geospatial information has a direct contribution and two additional indicators selected by the GEO, namely indicators: 11.1.1 *Proportion of urban population living in slums, informal settlements or inadequate housing*, and 11.6.2 *Annual mean levels of fine particulate matter (e.g. PM2.5 and PM10) in cities*.

Taking into consideration, the tier level classification available at the time of the selection of indicators (September 2017), the final reference list for selection included 17 SDG indicators – 6 tier I, 4 tier II and 7 tier III. Nine EU SDG indicators were also identified as potentially benefiting from geospatial information and of its integration with statistical data. Additionally, the identification of indicators also considered the possibility of including national indicators defined within the context of national SDG monitoring strategies.

[26] The Urban Audit is a data collection project of sub-national statistics on the quality of life in European cities, based on established territorial typologies (cities, greater cities, functional urban areas) for a specific number of variables covering different thematic areas. The data collection exercise is undertaken jointly by the National Statistical Institutes, the Directorate-General for Regional and Urban Policy and Eurostat.

The working group members provided structured comments based on a template for 26 indicators analysis which included:

i) the current reporting situation both at global and national levels with metadata analysis;

ii) gap analysis on the methodology and geospatial data integration suggested for the indicator;

iii) identification of corresponding EU SDG indicators and of specific national indicators.

Based on the number of contributions and inputs provided by the working group members, the maturity of the indicators' operationalization, also reflected in the tier level classification, the possibility of calculation at European level and at national level and the policy relevance in the European context, four SDG indicators were selected for an in-depth analysis [Figure 6].

Figure 6 |
Selected SDG indicators

11 SUSTAINABLE CITIES AND COMMUNITIES

11.2.1

tier II indicator

Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities

11 SUSTAINABLE CITIES AND COMMUNITIES

11.3.1

tier II indicator

Ratio of land consumption rate to population growth rate

11 SUSTAINABLE CITIES AND COMMUNITIES

11.7.1

tier III indicator

[moved to tier II as of 27 November 2018]

Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities

15 LIFE ON LAND

15.1.1

tier I indicator

Forest area as a proportion of total land area

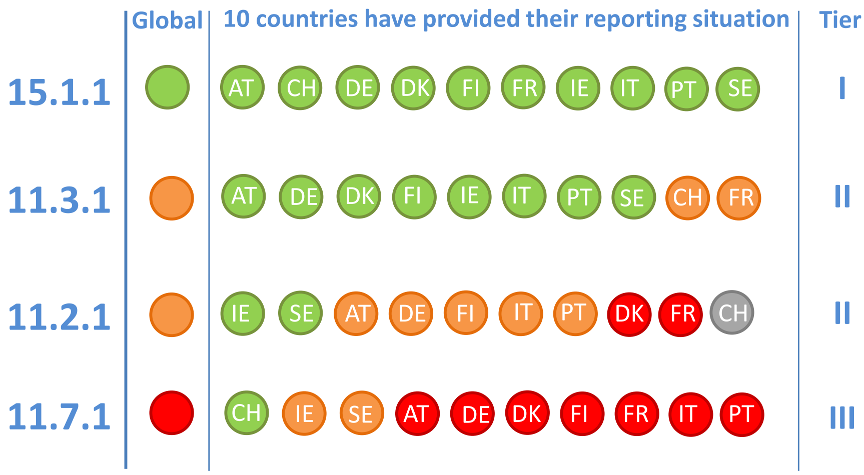
National report assessment of the selected SDG indicators

To complement the detailed analysis at global, European and national level carried out on the four selected SDG indicators, and as a way to provide a state of the art on the current official national reporting situation, the assessment classification model proposed by Sweden in the context of the IAEG-SDG WG GI analysis of the SDG indications under a geospatial lens was considered, which includes a colour evaluation: ***Green*** meaning it is possible to report or already being reported; ***Orange*** meaning it is possible to develop, data integration needed or changes to current surveys; and ***Red*** meaning it is very difficult to report, no current survey, no available method; and ***Grey*** not relevant or global data enough.

Figure 7 presents the national official reporting assessments regarding the four selected SDG indicators and shows that for the tier **I** indicator 15.1.1 *Forest area as a proportion of total land area* all of the 10 countries that have provided an assessment indicated that it is possible to report this indicator. Concerning the tier **II** indicators 11.3.1 *Ratio of land consumption rate to population growth rate* and 11.2.1 *Proportion of population that has convenient access to public transport* the number of countries reporting the green colour decreases, especially regarding indicator 11.2.1 *Proportion of population that has convenient access to public transport.* For the tier **III** indicator 11.7.1 *Average share of the built-up area of cities that is open space for public use* only Switzerland reported this indicator.

Figure 7 I

National official report assessment on the four selected SDG indicators



Note: 11.2.1 Sweden - Disaggregation for disabled persons not possible yet; 11.2.1 Switzerland - A national level indicator on the autonomous utilization of public transport by persons with disabilities has been defined; 11.3.1 Portugal – based on the Land used efficiency as proposed by JRC; 11.7.1 Sweden - A national complementary indicator on the share of public green areas is already available as official statistics;

The following section present the outcome of the analysis carried out for the four selected SDG indicators based on a global, European and national perspective. Each point is structured around the selected SDG indicators that best showcases the use of specific data types and methods regarding geospatial and statistical data integration.

Therefore, the analytical outcomes should not be understood as exclusively related to one indicator, as they can also contribute to other SDG indicators.

3 THE CONTRIBUTION OF GEOSPATIAL AND STATISTICAL DATA INTEGRATION TO DERIVE SDG INDICATORS

The contribution of geospatial data to the monitoring framework of SDG indicators goes further than having location variables that allow the disaggregation, aggregation and mapping of information. Geospatial analysis and models, similarly to statistical analysis, consists of a way of interpreting the data based on the application of specific techniques and data processing models.

Measuring accessibility using spatial modelling and analysis – the case of indicator 11.2.1 *Proportion of population that has convenient access to public transport*

Make cities and human settlements inclusive, safe, resilient and sustainable

Target 11.2 I By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons

Measuring accessibility has a strong spatial character, since it is intrinsically associated to a physical distance to the place and the selected SDG tier II indicator 11.2.1 *Proportion of population that has convenient access to public transport* shows that geospatial data and modelling is at the heart and nature of this indicator.

Analysis at the Global level

At the global level, UN-Habitat (United Nations Human Settlements Programme) and other partners are leading the efforts to compile data for this indicator. Conceptually, the indicator requires two spatial components: the location of the population and the location of the public transport stations. As the indicator is part of Goal 11, the delimitation of cities and human settlements is a prerequisite, as well. The definition of convenient access is defined by criteria such as distance (e.g. maximum 500 m), accessibility for special-need customers, frequency of service, safe and comfortable station environment. The indicator should be disaggregated by age, sex and persons with disabilities and should be reported in a two to five year interval based on availability of new data.

The global suggested method to calculate the proportion of population that has convenient access to public transport is based on four steps:

i) spatial analysis to delimit the built-up area of the urban agglomeration (see indicator 11.3.1 in the next point of the report);

ii) inventory of the public transport stops in the city and calculation of the service area (various options on the type of distance and the actual distance considered convenient as approximately 500 m);

iii) calculation of urban areas with access to public transport and identification of the population served;

iv) calculation of the proportion of the population with convenient access to public means of transport out of the total population of the city.

When available, information can be disaggregated by various demographic variables as well as variables based on transport frequency and accessibility. The temporal measurement on avalability of public transport is left out completely for global comparison, but countries that can additionally capture this component are encouraged to collect and report this information as part of the disaggregation.

The basic methodology is described using population data on the level of Census enumeration districts which is most likely to be available on a worldwide basis. The recommended data source, however, is the location of dwelling units as GIS data including the number of residents per dwelling unit. Optionally, the data source can rely on Census or household surveys that collect information on the proportion of households that declare having convenient access to public means of transport (e.g. EU Survey on Living Conditions 2012 module) and possibly also collect information about the quality of the service.

This indicator is categorized under tier II, meaning the indicator is conceptually clear and an established methodology exists, but data is not easily available. No internationally agreed methodology exists for measuring convenience and service quality of public transport. Moreover, data is not harmonized and comparable at the global level. Obtaining this data will require collecting it at municipal/city level with serious limitations in some areas, such as data on mass transit and on transport infrastructure.

In addition, an open-source software platform for measuring accessibility, the Open Trip Planner Analyst (OTPA) accessibility tool, is proposed at the global level to be available to government officials and all urban transport practitioners. This tool was developed by the World Bank in conjunction with Conveyal, leveraging the power of the OTPA engine and open standardized data to model block-level accessibility. The added value of the tool – free and user friendly – is its ability to easily calculate the accessibility for various settings and transport scenarios and will ensure a more uniform and standard format for reporting on this indicator.

The analysis of the WG members regarding the metadata on this indicator has pointed out:

- At the conceptual level:

At the global level, the definitions underlying the GHSL – Global Human Settlement Layer could be used and at the European level [see Box 5], a common definition of urban areas should be considered based on available common territorial typologies, such as the Degree of urbanization (DEGURBA, see Box 6).

The indicator should reflect the urban dimension in the name as its methodological scope begins with the delimitation of urban agglomeration (e.g. proportion of urban population with convenient access to public transport). On the other hand, measuring access to transport is a relevant indicator to an overall assessment of the national situation and, therefore, it could also be reported for the whole country, including segmentation for rural areas. The definition of stops could also be a challenge as, typically, large stations have several entrances and all entrances should be included as access points to public transport, when available. Moreover, the meaning of terms such as ‘safe and comfortable environment’ and ‘frequent service’ would benefit from further theoretical clarification and clear conceptual definition.

- At the data source level and geospatial processing level:

Availability of data on road networks, including pedestrian walks, as well as on public transport spots and corresponding timetables might also be a challenge, especially in order to guarantee global and European comparability. Information might also not be easily available regarding population with disabilities, but another perspective would be to have information on stops accessible for people with disabilities, as anyone can be temporarily injured or have the need to push a wheel chair or a baby stroller (e.g. proportion of population with access to public transport stops accessible for people with disabilities).

The methodology to compute the general indicator measuring the proportion of population only takes into account the place of residence as reference points. Other reference points could be considered as major population concentration hubs, such as schools, workplaces or markets. This indicator would then measure the proportion of schools, workplaces and markets with convenient access to public transport. Schools, workplaces and markets represent, however, a different object type, and the indicator would measure the proportion of schools, workplaces or markets with convenient access to public transport.

In terms of calculation, the shortest network distance can provide more accurate results than the Euclidian distance, but it requires quality data on the road network, including walkways and bicycle lanes. In addition, if population data is accurately assigned to point-location by means of geocoding using authoritative address, buildings or dwelling registers, the computation will not only be more accurate, but also more flexible and easier to conduct as no proxy data or disaggregation procedure will be needed. This should be the preferred approach in Europe as in the European context a growing number of countries are implementing point-based geocoding infrastructures allowing them to calculate very accurate figures.

Analysis at the European level

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, a similar indicator on the *Distribution of population by level of difficulty in accessing public transport* based on data from the EU Survey on Income and Living Condition (EU-SILC) ad hoc module of 2012. The indicator measures the share of population reporting i) very low; ii) low; iii) very high or iv) high level of difficulty in accessing public transport. This indicator reflects people's perception and is neither based on spatial analysis nor does it have a clear correspondence to the global indicator.

A further European approach was developed by the Directorate-General for Regional Development (DG REGIO) in the working paper *Measuring access to public transport in European cities* (Poelman and Dijkstra, 2015) using geospatial data and public timetable information. DG REGIO study measured networks distances (instead of Euclidean distance) for 29 EU cities [see Box 3], defined using the EU Urban Audit city definition of urban centres (high-density clusters from the grid based DEGURBA).

The study considered:

i) gridded population with grid cell sizes 1 km^2 or smaller or house blocks (building blocks corresponding to polygons of the Copernicus Urban Atlas layer (2006) and in areas with no data, population was estimated to 100 m^2 grid, downscaling from the EU 2006 population grid), and

ii) information on transport stops and road network including footpath information relied on TomTom MultiNet and national data sources on public transport were also used. Transport stops and timetable information is not available for all transport networks and for all Member-States, which is an obstacle for calculating the indicator at EU level. Google also provided transport timetables available, but with coverage gaps.

Analysis at the National level

At the national level, from the national practices collected (from Austria, France, Ireland, Sweden, and Switzerland), it was possible to identify that this indicator has been calculated for the national context. The results from the EU-SILC ad hoc module 2012 have been used in the context of evaluating access to public transports in the case of Austria, whereas in the case of Switzerland the reference source to evaluate accessibility for the SDG is the Swiss Health Survey. National cases have identified the NSI as the agency responsible for the indicator and in the case of France and Switzerland together with the National Geographic Institute and the Swiss Federal Office for Spatial Planning, respectively. In the case of Sweden, the evaluation of this indicator also includes the transport authority. As for all SDG 11 goals, the definition of the geographic delimitation of urban areas is necessary. Countries used either the EU definition or followed their national methodology. All five countries have data sources with population data geocoded to address or building point location, which can be used for the spatial analysis proximity to the stops as well as for the demographic disaggregation of the indicator by age and sex.

It was pointed out that the indicator would require reliable data on public transport and its stops, and possible data sources were mentioned. In France data on public transport stops exists, but it is provided by the local or city authorities. So the main difficulty would be to gather the information at national level from various data source providers. Sweden is a good example, where data on public transport stops (coordinates and traffic frequency) is available as open data in GTFS format (Google General Transit Feed Specification) and is provided jointly by the public transport service (www.trafiklab.se).

In the context of the GEOSTAT 3 testing of this indicator, the case of Estonia also relied on national public transport provided by the public transport register as open data with descriptions, timetables and locations of stops on domestic public transport routes that are fit to Google GTFS data model. The testing of the global indicator considered the

Box 3 I

Measuring access to public transport in European cities (DG REGIO)

DG REGIO calculation using geospatial information, population distribution and public transport stops timetable took into account:

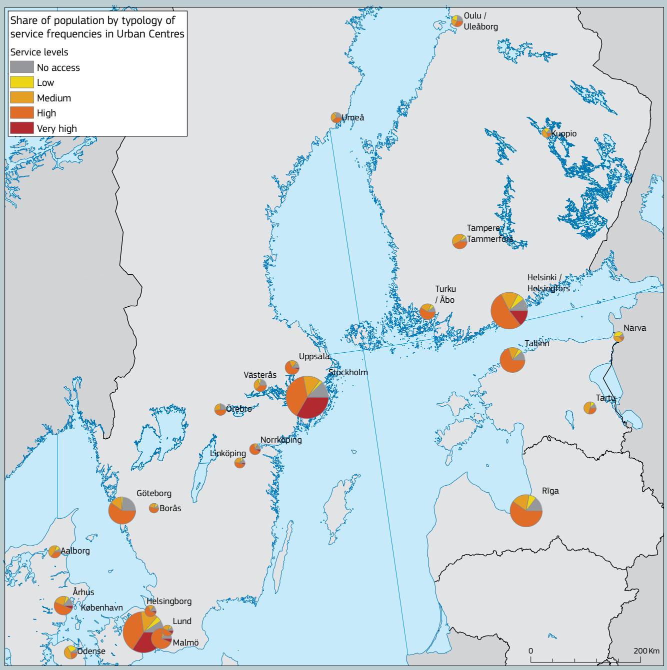
- Walking distances (service areas around each stop) using the street network (accounting for obstacles such as rivers, steep slopes, highways and railroads). It was assumed that people would be willing to walk five minutes (417m) to a bus/tram stop or 10 minutes (833m) to train/metro.

- Time table data on departures on a normal weekday (6:00 – 20:00) by considering five groups based on access and departure frequency [see Figure 8]: i) no access: people cannot easily walk to a public transport stop, it takes more than 5 minutes to reach a bus or tram stop and more than 10 minutes to reach a metro or train station; ii) low access: people can easily walk to a public transport stop with less than four departures an hour; iii) medium: people can easily walk to a public transport stop with between 4-10 departures an hour; iv) high: people can easily walk to a bus or tram stop with more than 10 departures an hour or people can easily walk to a metro or train station with more than 10 departures an hour (not both); v) very high: people can easily walk to a bus or tram stop with more than 10 departures an hour and a metro or train station with more than 10 departures an hour.

- Each of the service area polygons was characterized by the sum of the hourly average number of departures available at the stop around which it is created. The study assumed that the stop with the most frequent departures is the most probable choice. The service areas within each of the groups of transport modes were intersected and, in case of overlapping areas, the maximum value of the hourly average number of departures was attributed.

Figure 8 I

Access to public transport in urban centres in Denmark, Sweden, Finland, Estonia and Latvia



Source: Poelman, H. and Dijkstra, L. (2015). *Measuring access to public transport in European cities*. Regional Working Paper, DG REGIO.

selection of stops with at least one trip between 6:00 am and 8:00 pm on a specific day (Wednesday) and population data geocoded to address point location. In addition, and based on the assessment that the global approach is too general to provide a differentiation on mobility needs that are relevant for mobility/transport policy, Estonia also developed a national approach that differentiates access by frequency and urban settlement, namely for:

i) urban areas at least 6 trips per hour (very good access) and at least 2 trips per hour (good access) between 6:00 am and 8:00 pm on a working day for a service area of 400 m;

ii) rural areas at least 3 trips per day (at least 1 trip in range 6:00 am-9:00 am, 3:00 pm-6:00 pm and 6:00 pm-8:00 pm), and at least 6 trips per day (between 6:00 am-8:00 pm on a working day) for a service area of 1 000 m.

In the case of Austria the data on transport stops and timetables for different transport modes is gathered by the organisation VAO (Verkehrsauskunft Österreich) and provided to various routing applications for intermodal door to door routing. However the access for statistical purposes has not been clarified yet. More recently, transport projects carried out by the Austrian Conference on Spatial Planning and the Austrian Ministry for Transport, with the collaboration of Austrotech, have produced two relevant products, namely: a 100 m^2 grid matrix with travel time of the best available intermodal route based on criteria such as type of transport, time and frequency of service, number of changes; and a 100 m^2 grid with a public transport quality grading system, providing for each cell information on how well each grid cell is served by public transport (including aspects such as type of transport, distance to stop, timetable information). Products like these could be a useful tool to measure accessibility based on data at a very detailed geographical level.

In terms of geographical information, the calculation of this indicator for the Irish context relies on data from the National Transport Authority, which is made available as open data (data.gov.ie) on a regular basis. The data for public transport networks and stations, along with road network, includes coordinates along with extensive information about routes, trips and traffic frequency for each stop. Data is provided through an API under open data license in GTFS format. Information relating to identifying stops with "frequent service" during peak or off-peak travel times can be done by using the timetable information connected to each stop.

In the case of Switzerland and Sweden, national indicators have been proposed to address the global target 11.2 on *access to safe, affordable, accessible and sustainable transport systems for all*. Switzerland has defined at the national level the indicator *Autonomous utilization of public transport by persons with disabilities*, with the purpose of measuring the percentage of seriously handicapped people between 15 and 64 years old living in private households which can use public transport autonomously (without aid by a third person) and without difficulties. The indicator is based on the results of the Swiss Health Survey and, therefore, represents a subjective self-evaluation of the persons questioned. In addition, Switzerland has proposed another indicator, the *Average distance to the next public transport stop*, but aiming mainly to support target 9.1 concerning the quality of infrastructures. This indicator relies on:

i) point-based population (population and household statistics are geocoded up to the building level and updated annually);

ii) national road network as a product of the large-scale topographical landscape model produced and maintained by the Swiss Federal Office for Topography Swisstopo (Swiss NMCA), revised and updated at a six-year periodicity, i.e., every year one sixth of the national territory of Switzerland is updated;

iii) public transport stops from the Federal Office of Transport, combined with further analysis of frequencies from Federal Office for Spatial Development (data available in the Swiss geographical portal).

In the case of Sweden, the Swedish National Board of Housing, Building and Planning proposed to measure accessibility to public transport based on the number of dwellings and new dwellings developed in proximity of public transport stops, which has been calculated based on geospatial data on households and public transport stops [Box 4].

Box 4 |

Sweden national complementary indicator on Housing in proximity of public transports

As a national complementary indicator, the Swedish National Board of Housing, Building and Planning proposed to measure the number of dwellings and new dwellings developed in proximity of public transport stops. The rationale for this indicator is to follow up the sustainability of urban planning; assuming that housing close to public transports will require less need for cars. The indicator will be updated annually in order to follow the trend of new housing in proximity of public transports.

The steps for calculating the indicator are the following:

Step 1 public transport data, select only those stops that match the desired frequency of departures.

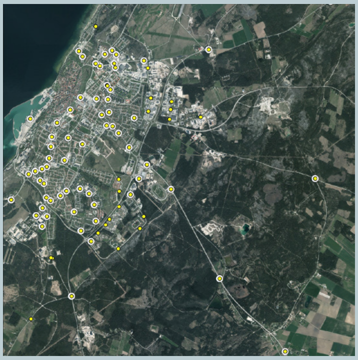


image showing all public transport stops (yellow dots) and transport stops considered “frequently trafficked” (yellow dots surrounded by a bigger white dot)

Step 2 create service areas around each public stop: Buffers with varying sizes (400, 1 000 and 2 000 meters).

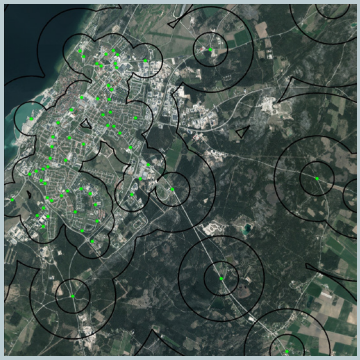


image showing frequently trafficked public transport stops with service areas

Step 3 conduct a point-in-polygon operation to find out which dwellings are within the range of the service areas (both in total and dwellings in buildings completed during the reference year of interest).

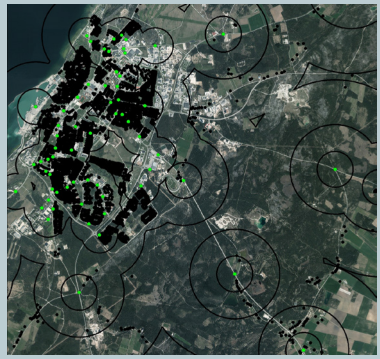


image showing buildings with registered dwellings together with service areas of public transport stops

Step 4 conduct a point-in-polygon operation also on population data geocoded to the level of address locations to find out how many people live within the range of the service areas.

Step 5 use the total figure for dwellings and population by county and municipality to calculate a share.

Step 6 publish the information in the Statistical database from which the National Board of Housing, Building and Planning can retrieve data either by means of searches or by means of machine-readable data served through an API.

The results show that, on national level, 78% of all dwellings are located within 400 m from a “frequently trafficked” public transport stop and 90% of the dwellings were located within 1 000 m.

Among the new dwellings (completed throughout the year of 2015) some 83% were located within 400 m from a frequently trafficked public transport stop.

Source: Lantmäteriet, Trafiklab and Statistics Sweden.

Deriving new metrics integrating land cover and population data – the case of indicator 11.3.1 *Ratio of land consumption rate to population growth rate*

Make cities and human settlements inclusive, safe, resilient and sustainable

Target 11.3 | By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries

The geospatial analysis combining land cover and population data provides the possibility of deriving new metrics that are relevant to grasp important dimensions on human settlement planning and management. The SDG tier II indicator 11.3.1 *Ratio of land consumption rate to population growth rate* is a very straightforward example on this type of data combination.

Analysis at the Global level

At the global level, UN-Habitat and other partners such as the institutions involved in the Global Human Settlement Layer (GHSL) team and ESRI will support various components for reporting on this indicator. Conceptually, the indicator requires defining the following two components:

i) 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;

ii) 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. 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 a comparable scale. At the global level, the suggested method to calculate the indicator proposes to first estimate the land consumption rate, and then to estimate the population growth rate and finally to compute the ratio of land consumption rate to population growth rate.

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) [Box 5] technology open framework is proposed for global open spatial baseline data production (built-up and population grids).

Box 5 I

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 Directorate-General for Regional Development (DG REGIO) of the European Commission, together with the international partnership of the GEO Human Planet Initiative.

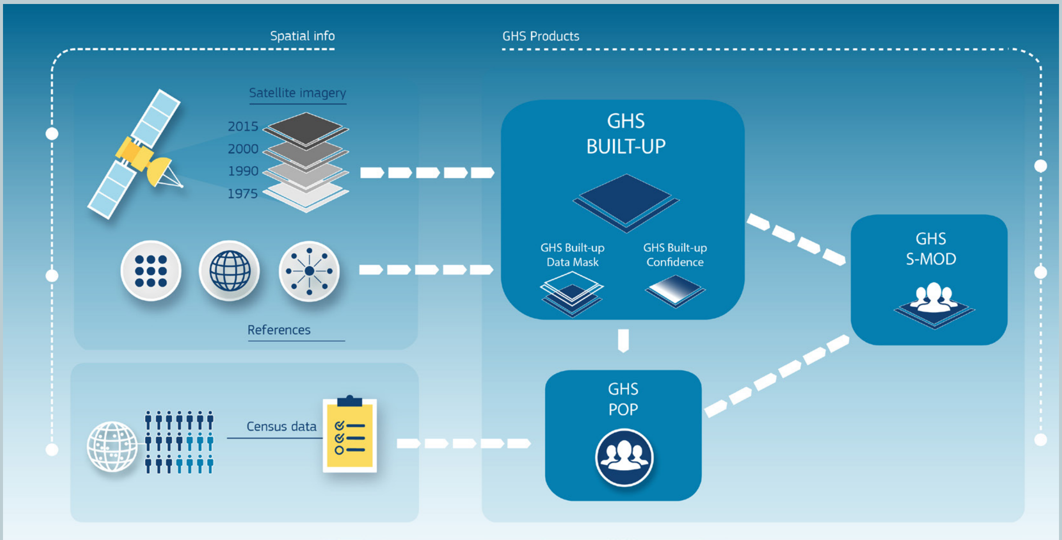
This tool includes a review of previous efforts to map settlements at various scales and using different datasets in order to produce new global spatial information, evidence-based analytics and knowledge for describing the human presence on the planet. The paradigm underlying the GHSL is the design 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 9].

GHS 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. GHS built-up were produced for a 250 x 250 m and a 38 x 38 m 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 250 m^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 and the population grid cell value represents the numbers of inhabitants. The GHS 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 ESS DEGURBA framework.

Figure 9 I

Global Human Settlement – general methodology



Source: European Commission, Global Human Settlement.

The analysis 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 regarding the concepts of 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, which 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 way 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.

In this context, urban area/city operationalization should be based on an international statistical definition to deal with issues of comparability. Having in mind the TERCET regulation, we can refer either to the concept of 'city' or the 'functional urban area' or the 'urban area' from DEGURBA, which correspond to 'cities' and 'towns and suburbs'. The grid-based typology concepts can also be used with advantages in terms of detail and comparability between countries. The EU definitions can be taken as a reference to a worldwide definition following the discussions in the context of Habitat III and, more recently, in the UN Statistical Division based on the approach proposed by the European Commission, OECD, World Bank and partners [Box 6]. Additionally, countries may also have national classifications. The indicator could also account for the entire territory and not be limited to an urban definition.

- At the data source and geospatial processing 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. Particular specifications are needed for geospatial processing, namely the spatial resolution for input and output of geospatial data processing needs to be better identified. 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 (20 m) and CORINE Land Cover Map (CLC) could be possible data sources for built-up areas identification, but in the case of 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).

Box 6 I

ESS Territorial typologies relevant for urban delimitation - the grid-based typology, DEGURBA and FUA - and the Global DEGURBA approach

At the European level, the Regulation (EU) 2017/2391 of the European Parliament and of the Council of 12 December 2017, 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 typologies to identify urban at grid and LAU levels are particularly relevant for Goal 11 monitoring but also to structure other indicators according to urban and rural segmentation:

The grid-based typology (1 km²) defines [Figure 10]:

- 'Urban centres' or 'High density clusters': Contiguous (without diagonals) 1 km² grid cells within the 'urban cluster' with a density of at least 1 500 inhabitants/km² and a minimum of 50 000 inhabitants in the cluster (gaps in the cluster are filled).
- 'Urban clusters' or 'Moderate density clusters': Contiguous (including diagonals) 1 km² grid cells with a density of at least 300 inhabitants per km², and a minimum of 5 000 inhabitants in the cluster.
- 'Rural grid cells' or 'Low density grid cells': 1 km² grid cells with density below 300 inhabitants/km² and other cells outside urban clusters.

The Degree of urbanization (DEGURBA) identifies at LAU level [Figure 11]:

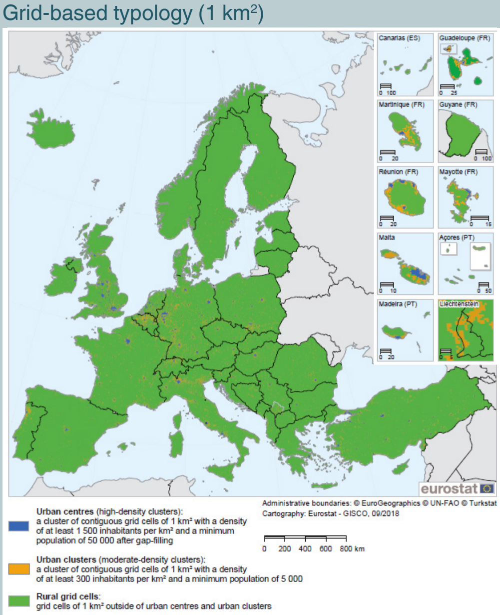
- 'Cities' or 'Densely populated areas': LAU level territorial units where at least 50% of the population live in 'urban centres'.
- 'Towns and suburbs' or 'Intermediate density areas': LAU level territorial units where less than 50 % of the population lives in 'rural grid cells' and less than 50 % lives in 'urban centres'.
- 'Urban areas': 'Cities' and 'Towns and suburbs'.
- 'Rural areas' or 'Thinly populated areas': LAU level territorial units where at least 50% of the population live in rural grid cells.

The Functional urban areas (FUA), identifies at LAU level:

- 'Cities' plus their 'Commuting zones' defined as LAU level territorial units from which at least 15% of the employed population commute to the city, whereby enclaves are included and exclaves are excluded.

Figure 10 I

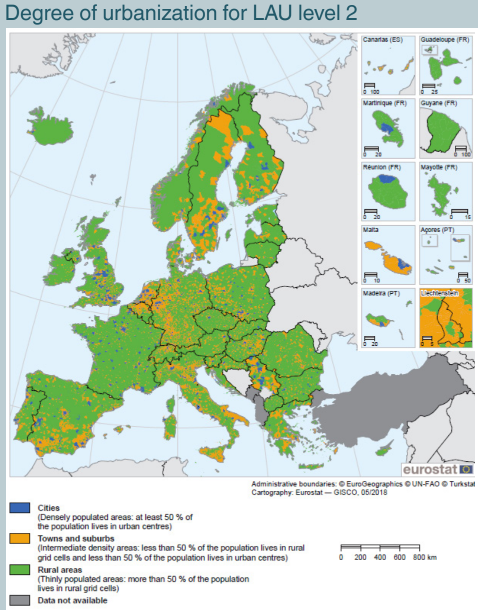
Grid-based typology (1 km²)



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

Figure 11 I

Degree of urbanization for LAU level 2



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

These typologies have been discussed in the UN Expert Group Meeting on Statistical Methodology for Delineating Cities and Rural Areas and in particular a **refined DEGURBA definition** intended to be applicable worldwide to report data to UN Statistical Division and SDG monitoring. This Global DEGURBA details the three classes into six allowing the spatial identification of 'Cities', 'Towns', 'Suburbs', 'Villages', 'Dispersed rural areas' and 'Mostly uninhabited areas' at both grid and local levels.

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 the CCI Land Cover products of the ESA Climate Change Initiative (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 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 disasters 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) using the Global Human Settlement Layer, which can be adapted to other input data. 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)[27]. A script that can be installed in the toolbox of Quantum GIS (QGIS) has been made available.

Analysis at the European level

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: *Artificial land cover per capita and Change in artificial land cover per year*[28]. These two EU SDG indicators are based on the LUCAS Survey, which corresponds to a harmonised in situ land cover and land use data collection over EU's territory, based on a standardised methodology in terms of sampling plan, classifications, data collection and statistical estimators. 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 aiming to study the most appropriate 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

[27] 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.
[28] The indicator Change in artificial land cover per year has been excluded from the EU SDG indicator set in the 2018 revision.

have been put forward by other departments of the Commission. The Copernicus Imperviousness HRL could also 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 in three year cycles. It has a minimum mapping unit (MMU) of 20 m (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 (25 ha). Nevertheless, the next generation of CLC, CLC+, starting with 2018 reference year, 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, size and shape of buildings and parcels, 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 Annex III theme building data is not yet available in all Member-States.

Analysis at the National level

At the national level, from the cases analysed (Finland, Ireland, Italy and Portugal) it is possible to identify that this indicator has not always been calculated, disseminated or reported by countries. National cases have mainly 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).

Countries use 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. On the other hand, data for land consumption is not produced so regularly. Other data sources, besides national thematic layers, could be used to compute this indicator at national level, namely Italy has referred data from Sentinel-2 satellite images and European thematic layers, such as Copernicus Imperviousness HRL, CORINE Land Cover.

In this vein, in the context of the GEOSTAT 3 testing, the exercise developed by Sweden on the computation of this indicator considered a comparative approach on the delimitation of urban agglomeration by using different data sources, namely the national classification of urban areas (Swedish localities), the European grid-based typology to define urban centres and urban clusters, and the built-up maps from the GHSL. The testing also included the computation of results using the formula defined in the global metadata and the LUE proxy indicator proposed by JRC. The results showed that national localities occupy a much larger space than based on the other datasets. Swedish national data takes into account a number of other spatial factors, besides the population density, such as connectivity, barriers and land use, hence offering a more precise and detailed representation of the urban outline than the grid based clustering. The European grid cluster is much more generalized, as it consists of large grids, only taking in to account the population size of each grid and the GHSL only captures built-up in the sense of impervious land.

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 and, among others, including gardens. Way objects in the PRIME2 database also represent all drivable and walkable roads and paths from motorways down to sidewalks [see Box 10].

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 Mainland Portugal for four reference years - COS 1995, COS 2007, COS 2010 and COS 2015 - and correspond to polygonal maps that represent homogenous land use/cover units. COS is based on a vector data model with a reference MMU of ha 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' [Box 7].

Box 7 |

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

In December 2018, Statistics Portugal published a new set of Land Use and Land Cover Statistics based on the Land Use and Land Cover Map (COS) produced by the Directorate-General for Territory (the Portuguese NMCA) that includes the calculation at municipality level of the global SDG 11.3.1 indicator based on the Land Use Efficiency (LUE) formula as proposed by the JRC [Figure 12] and of the corresponding EU SDG indicator defined by Eurostat to monitor Goal 11 at the EU level [Figure 13].

For both indicators, the megaclass 1 of COS nomenclature "artificial land", excluding the class 133 corresponding to "areas under construction" was used and 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 LUE indicator was calculated using data from COS 2010 and COS 2015 and data from annual resident population estimates for the reference years of 2010 and 2015. The use of the JRC formula allowed to deal with those situations with zero growth and, thus, provided consistent results for the different municipalities.

The result for Portugal's mainland for the period 2010-2015 was -10% and only 15 municipalities scored positive LUE values, i.e., an increment of population faster than the increase of artificial land [Figure 12]. The indicator on artificial land *per capita* was calculated using data from COS 2015 and population data was also derived from the annual resident population estimates for the reference year of 2015. The results show that 60 municipalities, mainly located in the metropolitan areas of Lisboa and Porto, recorded less area of artificial land *per capita* than the one registered for Mainland Portugal [Figure 13].

The dissemination of this new set of Land Use and Land Cover Statistics comprised several challenges on complying with the standard statistical methodological document made available describing all the methodological procedures, concepts and classifications associated with a statistical operation, as this constitutes the first statistical operation disseminated by Statistics Portugal based on a geospatial data source and on its integration with statistical data.

Figure 12 |

LUE by municipality 2010-2015

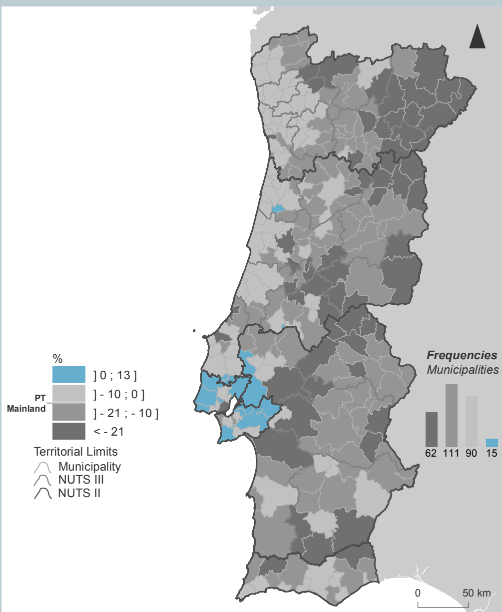
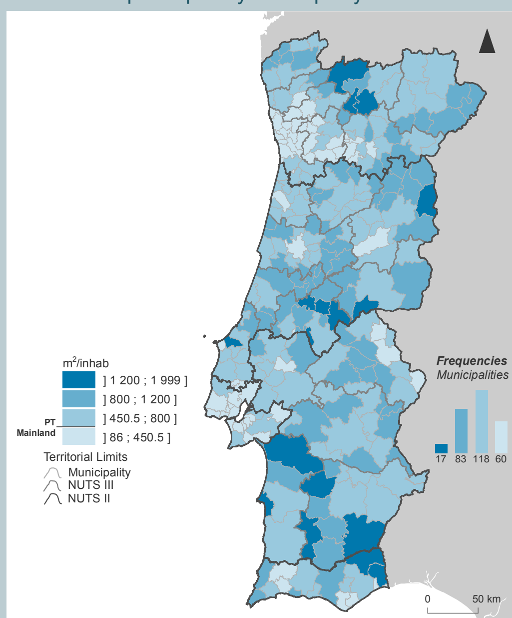


Figure 13 |

Artificial land per capita by municipality 2015



Source: Statistics Portugal/Directorate-General for Territorial Development, Land use and Land Cover Statistics.

Addressing challenging indicators based on land use and cadastral data – the case of indicator 11.7.1 *Average share of built-up area of cities that is open space for public use*

Make cities and human settlements inclusive, safe, resilient and sustainable

Target 11.7 I By 2030, provide universal access to safe, inclusive and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities persons

The monitoring framework for the 2030 Agenda on Sustainable Development includes a number of more challenging indicators due to lack of data availability and existing established methodology. This is the case of the selected SDG former tier III, now tier II indicator 11.7.1 *Average share of the built-up area of cities that is open space for public use*, for which land use and cadastral data obtained using different geospatial based products can play a significant contribution for its operationalization.

Analysis at the Global level

At the global level, UN-Habitat and other partners including private and regional commissions are leading the efforts of building national capacity to monitor and report this indicator. Conceptually, the indicator requires the estimation of the area of public space based on:

i) spatial analysis to delimit the built-up area of the city (see also point on indicator 11.3.1 on this section);

ii) estimation of the total open public space;

iii) estimation of the total area allocated to streets.

This indicator has been categorised under tier III, meaning that internationally established methodology or standards are not yet available for the indicator, but methodology/standards are being developed and tested. A request to upgrade the indicator to tier II has been made by UN-Habitat and at the 8th IAEG-SDG meeting (5-8 November 2018) the indicator was reviewed and upgraded to tier II.

At the global level, the harmonization of the following data sources is proposed to ensure more consistent reporting on this indicator:

- For estimating the total surface of built-up area, it is proposed the use of existing layers of satellite imagery ranging from open sources, such as Google Earth and US Geological Survey/NASA Imagery Landsat, to more sophisticated and higher resolution land cover datasets. Images are to be analysed for the latest year.
- For the Inventory of open public space, information can be obtained from legal documents outlining publicly owned land and well-defined land use plans. In some cases, where this information is lacking, incomplete or outdated, open sources, informants in the city and community-based maps, which are increasingly recognised as a valid source of information, can be a viable alternative.
- The share of land in public open spaces cannot be obtained directly from the use of high-resolution satellite imagery, because it is not possible to determine the ownership or use of open spaces by remote sensing. However, fieldwork to validate and verify the open spaces derived from satellite imagery helps to map out land that is for public and non-public use.

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

- At the conceptual level:

There is a need for clear definition on the underlying concepts of this indicator. Regarding the concept of built-up area, it is not clear what categories should be included in the definition of a built-up area, but it should be in line with indicator 11.3.1 *Ratio of land consumption rate to population growth rate* as it aims to capture the same entity.

Concerning the definition of urban areas/cities, the territorial classifications and/or methodology for its delimitation is also not clearly stated. At the European level there is the Degree of urbanization (DEGURBA) and countries may also have national classifications. Therefore, common and harmonised territorial typologies should be used to capture the urban dimension and, in the case of EU, the TERCET regulation defines territorial typologies to be used and published by the Commission (Eurostat) [see Box 2].

The concept of public open space is very complex and difficult to measure, especially at the global level. It can be challenging, and in some cases impossible, to classify the types of public open space without conducting field inventories. In this context, other proxies might be considered that still grasp the idea of quality of life in cities and that are easier to measure, such as open green space. Additionally, the global metadata included the estimation of area allocated to streets, in order to better capture the use of public spaces for leisure activities, but the indicator might benefit from excluding the area allocated to streets. In this vein, a calculation with and without this dimension could be carried out in order to evaluate differences in terms of results.

Accessibility is typically measured using rules of spatial proximity between objects, such as between peoples' permanent place of residence and public parks. In the metadata description, accessibility is measured at each identified individual public open space, such as a square. If no restrictions for the public to access the square is found, it is considered as accessible open space. Regarding this approach, the scope of the indicator is ambiguous. There are two different types of objects to measure (according to the title of the indicator), the open space, on the one hand, and the people that have access to the open space, on the other hand. Another point of view could be to measure accessibility to this type of space based on a similar methodology as proposed for indicator 11.2.1 *Proportion of population that has convenient access to public transport*, which would entail changing the perspective on the indicator.

- At the data source and geospatial processing levels:

The data sources listed in the metadata are relevant, however further clarifications may be needed. Specifically, for estimating the total surface of built-up area, use of satellite imagery is recommended. However, it can be discussed whether use of raw earth observation data should be the first hand choice or if global earth observation data derived products could provide a more efficient option. Such data could be the Global Human Settlement Layer (GHSL), which is freely available and has global coverage. Another option could be the grid cluster approach jointly developed by EU, OECD and partners to define urban areas following ESS DEGURBA typology [see Box 6]. This approach also builds on the GHSL, but with additional modelling of population density.

For the inventory of open public space, the metadata description relies heavily on use of semi-structured, non-geospatial data such as information obtained from legal documents outlining publicly owned land or well-defined land use plans and fieldwork to validate and verify the open spaces derived from satellite imagery to map out land that is for public and non-public use. In a European context, it can be assumed the willingness from governments to spend resources on additional field data collection will be low, considering the large investments that have been done in geospatial information during the recent years. Investments in geospatial information are expected to pay-off in terms of reduction of the need to use of semi-structured, non-geospatial data and expensive fieldwork.

The metadata description should benefit from better recognising the use of alternative geospatial data sources to describe the accessibility of open space. Such data sources can be cadastral information (where ownership is defined), as in the case of Spain [see Box 8].

Box 8 |

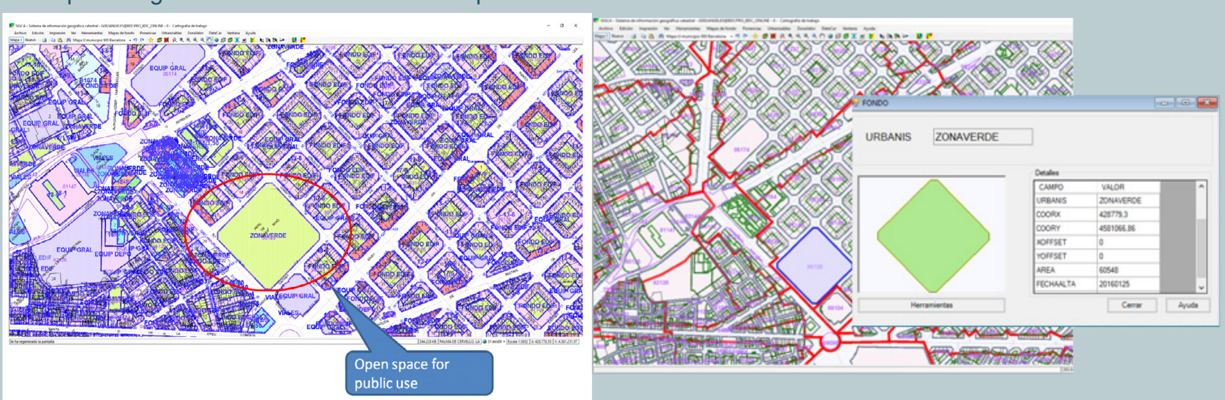
Identifying open public space with cadastral information: the case of Spain

In the case of Spain, the cadastre includes in its database the urban qualification of each real estate, including green areas. The Spanish cadastre provides information on the status, area (m^2) and coordinates [Figure 14], and information can be analysed at different territorial levels (e.g. by municipality, building blocks).

In contrast to fieldwork, the use of such data will result in less precise classification of types of public open space, but on the other hand it will have a better coverage and possibly also a more consistent and objective classification.

Figure 14 |

Example of a green area as included in the Spanish Cadastre



Source: Dirección General del Catastro (http://www.catastro.meh.es).

Comparability is best guaranteed with the use of available global/EU data sources, as well as territorial typologies, but at the national level data sources might have more detailed information and be able to better address the conceptual definitions of the indicator, including the proposed segmentations. The indicator might be calculated based on different sources at different levels of analysis, but to guarantee comparability across countries it is relevant to use harmonized definitions, typologies and data sources.

Analysis at the European level

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. Worth to mention, however, is that the EU collects data possibly useful for calculation of this indicator. Urban atlas is one of the relevant data sources compiled by the European Environment Agency (EEA) in a harmonised way across Member-States. The European Settlement Map is another new data source, produced by the Joint Research Centre (JRC) that provides an opportunity for calculations based on very detailed and harmonized urban data. However, presently none of these data sources has a complete coverage of all urban areas. Additionally, a further European approach based on geospatial data has been developed by DG REGIO in the working paper *A walk to the park? Assessing access to green areas in Europe's cities*. The study relies on the concept Urban Audit grid-based concept of urban centre, but also produces results for the city and greater city level as defined according to administrative boundaries. The definition of green areas relies on the one provided by the Copernicus Urban Atlas and population is also derived from the Copernicus Urban Atlas [Box 9].

Analysis at the National level

At the national level, from the national practices collected (Ireland, Sweden and Switzerland), it was possible to identify that this indicator has not been calculated, disseminated or reported by countries, which is explained by the fact that the indicator was until November 2018 in tier III mode. Countries have identified national data sources to address this indicator. In the case of Ireland, data related to built-up areas and public space can be sourced from the PRIME2 database maintained by Ordnance Survey Ireland (Irish NMCA), which corresponds to a central database of spatial information [Box 10].

Box 9 I

Assessing access to green areas in Europe's cities (DG REGIO)

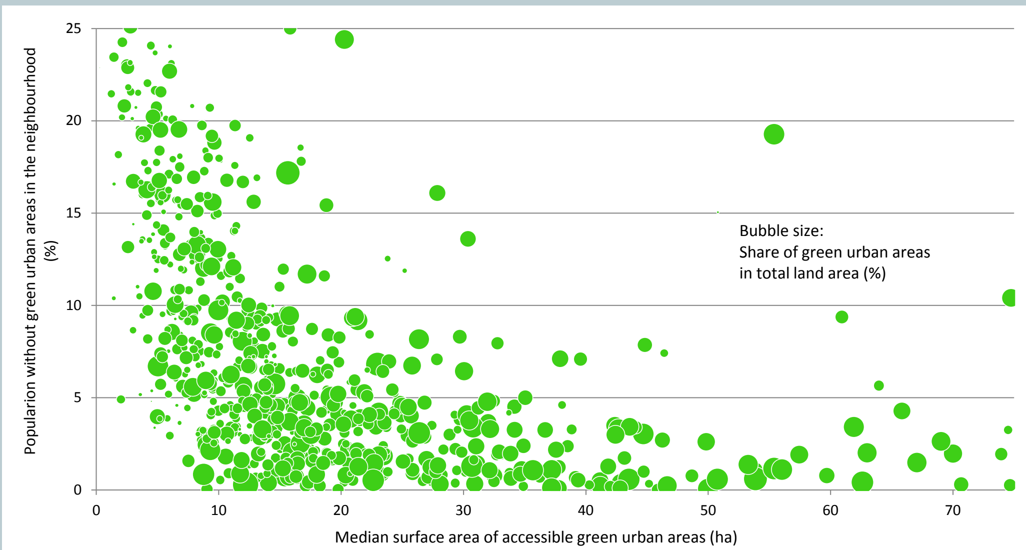
The DG REGIO study on access to green areas is based on Copernicus Urban Atlas data. The definition used in Copernicus Urban Atlas refers to "public green areas for predominantly recreational use such as gardens, zoos, parks, castle parks; suburban natural areas that have become and are managed as urban parks".

Nevertheless, because at the fringe of cities, the distinction between "green urban areas" and forests is not easily made, the DG REGIO study also included the Urban Atlas class "forests" in the analysis with a minimum mapping unit of 0.25 ha.

To measure proximity to urban areas, the study determined an area of easy walking distance (defined as 10 minutes along the street network) around each Urban Atlas polygon. For each polygon an estimation of the total resident population was available, making it possible to calculate the population-weighted median surface of green areas by urban centre or by city/greater city that can be reached within 10 minutes walking (the median rather than the average is used as the study argues that the latter tends to be influenced by outliers in the distribution of green areas). The study also calculates the distribution of the urban population compared to the level of access to green areas and the share of green areas in total land area [Figure 15].

Figure 15 I

Proximity of green areas, population without green areas nearby and share of green areas in the total land area



Source: Poelman, H. (2018). *A walk to the park? Assessing access to green areas in Europe's cities.* Regional Working Paper, DG REGIO.

Box 10 |

Measuring built-up areas that are open space for public use based on the Irish PRIME2 database

For the national context of Ireland, the identification of built-up areas that are an open space for public use can be derived from the PRIME2 geospatial database maintained by the Irish NMCA - Ordnance Survey Ireland (OSi). 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 and lane ways. Figure 16 shows the objects classified as 'artificial' and 'way' in the centre of Dublin.

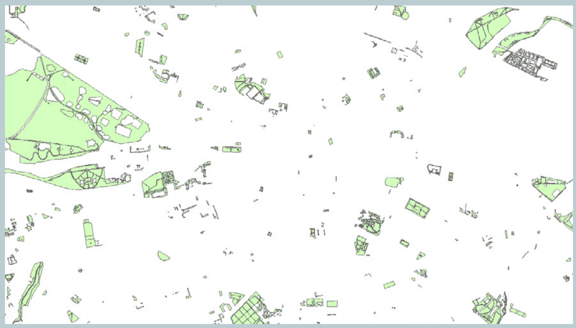
PRIME2 contains over 50 million attributed objects in point (building), line (road) and polygon (parcel) format, and includes form and function object classification which describes the physical form (e.g. building) and its use (e.g. residential, hospital, church etc.). There are over 1 000 different function types recorded in PRIME2. The following categories can be selected for open public space: green space, public park, beach, and cemetery. Figure 17 shows the objects classified as open public space in the centre of Dublin.

Figure 16 |

Built-up areas in Dublin city centre

Figure 17 |

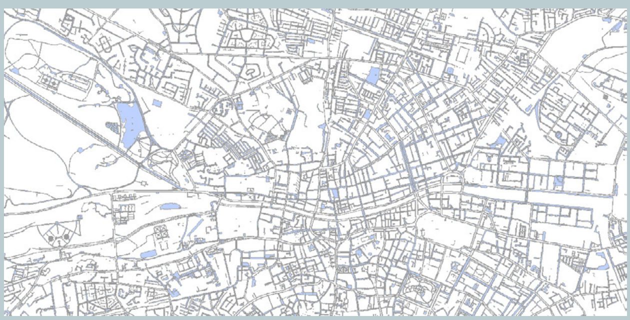
Objects classified as open space in Dublin city centre



The following categories of objects can be selected in the PRIME2 that resemble land allocated to streets that are considered as public space: car park, roundabout, sidewalk, traffic island, pier and street. Figure 18 shows the objects classified as land allocated to streets in the centre of Dublin.

Figure 18 |

Objects classified as land allocated to streets in Dublin city centre



Source: PRIME2 database (OSi – Ordnance Survey Ireland).

In the context of Switzerland and Sweden complementary indicators within their national monitoring framework have been proposed. In the case of Switzerland, the national complementary indicator is defined as *Urban recreational areas* and the Swiss Federal Office for Spatial Development holds thematic responsibility while the Swiss Federal Statistical Office is responsible for the data sources used. The indicator relies on national land use statistics which are established by visual interpretation of high-resolution aerial photography and distinguishes 46 categories of land use and 27 categories of land cover. The percentage of the surface occupied by the land use classes referring to *Recreational areas and cemeteries* and to *Surroundings of residential buildings* is used to calculate the indicator for the urban municipalities. Urban municipalities are defined according to the Swiss Federal Statistical Office urban-rural typology. In the case of Sweden, the national complementary indicator is based on access to urban green areas. Access is defined as spatial proximity to public green areas and calculation of this indicator is conducted as part of the existing official land use statistics [Box 11].

Besides Sweden, the GEOSTAT 3 testing regarding the principles of GSGF on this indicator carried out by Estonia and Norway have also taken into consideration the proxy concept of green areas to calculate this indicator and have considered the perspective of population access to this type of areas.

Both countries provided an exercise using national data sources to capture green areas – namely the Estonian Topographic Database (maintained by the Estonian NMCA and available as open data) and the Land use land cover dataset compiled by Statistics Norway based on data from the National geospatial infrastructure (Digital Norway cooperation coordinated by the NMCA) and data made available through agreements with the Norwegian NMCA – and land use and land cover data from the Urban Atlas – a Pan-European product developed in the frame of the EU Copernicus programme with the support of the European Space Agency and the European Environment Agency.

Population data is available at point level for both national cases and urban agglomerations were defined using national classifications of urban areas and using the European grid-based typology to define urban centres and urban clusters [see Box 6]. The testing evaluation of both Estonia and Norway highlighted the importance of having geocoded data at unit level in order to get reliable results at a detailed territorial level.

Box 11 |

Sweden national complementary indicator on Access to urban public green areas

For the context of the national SDG indicator framework and monitoring, Statistics Sweden proposed a complementary national indicator measuring people's access to public green areas in urban areas.

This indicator is based on existing official statistics. To calculate these statistics, a method has been developed, based on a combined use of Earth Observation (EO) data and geocoded register data and other geospatial information (geocoded population data, cadastral parcels, buildings, road networks etc).

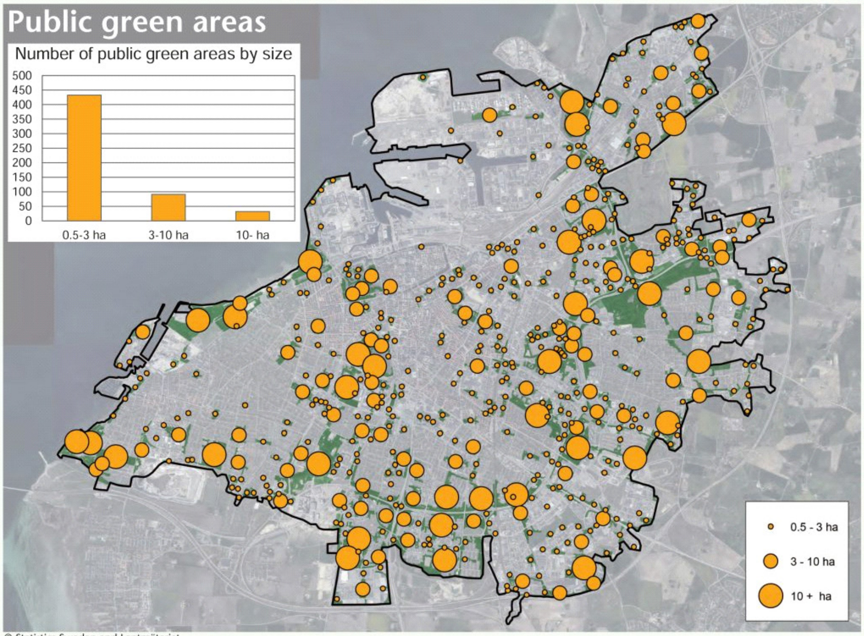
The data compiled for the green area statistics can also be used as a foundation to estimate the global indicator on the totality of urban public space. Some of the categories of public open space are already calculated [Figure 19] and some can be retrieved on the basis of the data compiled.

In brief, green spaces are mapped using EO data. Public access is determined using cadastral parcels in combination with ownership data. Large enough adjacent green spaces (>= 0.5 ha) accessible for the public are qualified for "green areas".

Proximity to population is calculated using Euclidian distance between population data linked to physical address location and green areas. The share of urban population with access to green areas within 200 meters from the place of permanent residence is calculated.

Figure 19 |

Public green areas in Malmö by size



© Statistics Sweden and Lantmäteriet

Footnote: A public green area is defined as an area of contiguous green space of at least 0.5 hectares which is available for the public.

Increasing the scope of indicators disaggregation with earth observation data – the case of indicator 15.1.1 *Forest area as a proportion of total land area*

15 LIFE ON LAND

Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss

Target 15.1 | By 2020, ensure the conservation, restoration and sustainable use of terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains and drylands, in line with obligations under international agreements

Earth observation data is a relevant source of information to monitor progress towards the SDGs, as this type of data can provide an objective and consistent view of the earth for different periods in time, at different scales and ensuring a coherent basis for comparability between different countries. The selected SDG tier I indicator 15.1.1 *Forest area as a proportion of total land area* is a very good example to showcase the potential of earth observation data to increase the scope of territorial disaggregation of statistical indicators.

Analysis at the Global level

At the global level, data for this indicator will be provided by FAO (Food and Agriculture Organization). Conceptually, the indicator measures the relative presence of forest area in a country and is based on two components:

i) Forest area to be computed according to FAO definition, i.e., “land spanning more than 0.5 ha with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use”;

ii) Total land area corresponds to the total surface area of a country excluding inland waters such as rivers and lakes.

The indicator is proposed to be provided at national level, with no further disaggregation and monitoring is set to be repeated at regular intervals of five years, allowing for three reporting points until the year 2030.

This indicator is categorized under tier I, meaning the indicator is conceptually clear, has an internationally established methodology and standards are available. FAO has been collecting and analysing data on forest area, as part of the Global Forest Resources Assessment (FRA) since late 40’s and the collection frequency has been every five years since 2000. The FRA is based on two primary sources of data: country reports prepared by national correspondents and remote sensing analysis that is conducted by FAO together with national focal points and regional partners. FRA collects country data following a standard format. It includes the original data and reference sources and descriptions of how these have been used to estimate the forest area for different points in time. Detailed methodology and guidance on how to prepare the country reports and to convert national data according to national categories and definitions to FAO’s global categories and definitions can be found in the Guide for country reporting for FRA (2015). Data is available for all 234 countries and the last FRA 2015 includes around 120 variables covering the following periods: 1990, 2000, 2005, 2010 and 2015.

The analysis of the WG members regarding the metadata on this indicator has pointed:

- At the conceptual level:

Taking into account the definitions provided by FAO, it is important to consider that not all countries report data to FAO every five years; therefore a five year release of the indicator is not possible everywhere, and the collection strategy, valid at country level, does not always allow a proper disaggregation of the indicator over smaller spatial units. In addition, the extension of the total land, not provided by FAO, must have a unique reliable data source for each country, in particular for the exclusion of water bodies, which must also be clearly defined.

- At the data source and geospatial processing levels:

Geospatial layers, based on remote sensing techniques, could be used for this indicator, allowing a more detailed result and an improvement in the computation frequency. Geospatial layers that can be used for this purpose are different at global and at European level and, in general, European geospatial data layers are much more detailed and reliable. In this context, there are several initiatives that map, on a regular basis, the world land cover. The most interesting initiative is the Land Cover generated by European Space Agency (ESA) in support to the Climate Change Initiative [Box 12]. It is based on automatic workflows for the generation of land cover with a resolution of 250 meters (down to 30 meters and less in the very near future), and a land cover map on a worldwide basis is generated annually from which geospatial layers related to forestry areas can be extracted. Nevertheless, it is important to take into consideration that when adopting land cover maps to extract forest areas the legend may not perfectly fit with the FAO forestry definition.

Box 12 |

Calculation of the indicator based on ESA Land cover data: the case of Italy

A simulation of the indicator computation has been carried out by e-GEOS using ESA Land Cover and is based on the following workflow:

- 2012 ESA Land Cover data and Italy's NUTS 1, NUTS 2, NUTS 3 and commune administrative borders provided by ISTAT (Italian NSI)
- extraction of the Italy area subset of the ESA Land Cover
- computation of the forest area over Italy by selecting the following classes:

  ESA Land Cover classes: 40 (Mosaic natural vegetation (tree, shrub, herbaceous cover)/cropland), 50 (Tree cover, broadleaved, evergreen, closed to open), 60 (Tree cover, broadleaved, deciduous, closed to open), 70 (Tree cover, needle-leaved, evergreen, closed to open), 80 (Tree cover, needle-leaved, deciduous, closed to open), 90 (Tree cover, mixed leaf type), 100 (Mosaic tree and shrub/herbaceous cover), 110 (Mosaic herbaceous cover/tree and shrub)
- computation of the internal water area to be subtracted by considering class 210
- computation of the total reference unit area from administrative borders
- calculation of the indicator by applying the formula: 100 x (forest area)/(administrative unit area - internal water area)

At national level, the indicator for Italy computed for 2012 based on ESA Land Cover is 33.4%, a value that should be compared with the one reported in FRA 2010 (36.7%) and FRA 2015 (37.7%). The following figures show the results obtained for the case of Italy considering a disaggregation at NUTS 3 [Figure 20] and commune level [Figure 21].

Figure 20 |

Forest area as a proportion of total land area based on 2012 ESA Land Cover data, NUTS 3 level

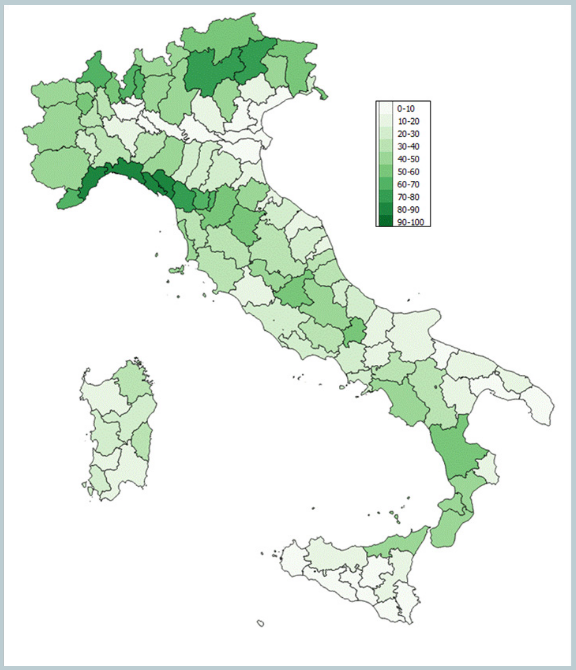
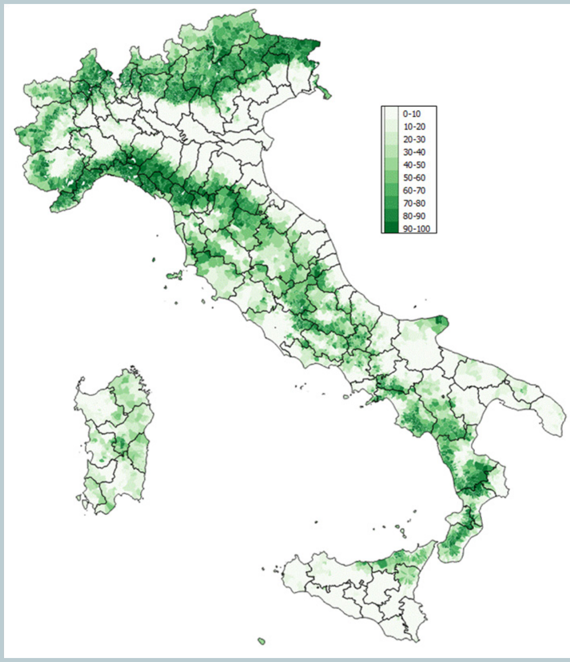


Figure 21 |

Forest area as a proportion of total land area based on 2012 ESA Land Cover data, Commune level



Source: ESA Land Cover.

Analysis at the European level

At the EU level, the EU SDG indicator *Share of forest area,* defined by Eurostat, has a direct correspondence with the one defined at the global level. Forests in the EU are covered under the EU Forest Strategy, which stresses the importance and multi socioeconomic and environmental benefits of sustainable forest management. The indicator measures the proportion of forest ecosystems in comparison to the total land area. The data used for this indicator derives from the Land Use and Cover Area Survey (LUCAS) and has been mapped according to FAO definitions, distinguishing between the categories 'forests' and 'other wooded land'. LUCAS survey is based on in situ data, meaning that observations are registered on the ground by field surveyors. A mixed panel approach is used, so some points are visited in subsequent years. In the field, the surveyor classifies the land cover and the visible land use according to the harmonized LUCAS land cover and land use classifications. The indicator is delivered every three years, with data at NUTS 2 level comparable for all EU Member-States.

In the framework of Copernicus Pan-European High Resolution Layers, geospatial data layers are available at the EU level, allowing the whole computation of the indicator [see Box 1 and Box 13], namely:

- the geospatial data layer Copernicus Forest High Resolution Layer (HRL)
- the geospatial data layer Copernicus Water Bodies HRL

These datasets, available for all EU Member-States, can provide very detailed statistics on forest cover, allowing for a disaggregation below NUTS 3 level. The Copernicus HRL permanent water bodies' geospatial layer provides a detailed measure of inland waters (to be subtracted to total land area for indicator computation).

Copernicus Forest HRL is updated with a three years frequency (2012-2015-2018). A Commission task team has been launched to understand if Copernicus information could complement statistical data with the goal to improve coverage, timeliness and resolution.

The Copernicus HRL Forest is based on two products:

- The tree cover density geospatial layer maps the level of tree cover density in a range from 0-100% and minimum mapping width of 20 m.
- The forest type geospatial layer allows to get as close as possible to the FAO forest definition, with a Minimum Mapping Unit (MMU) of 0.5 ha, as well as a 10% tree cover density threshold applied. For the final 100 m product trees under agricultural use and urban context from the support layer are removed, in line with the FAO forest definition.

Box 13 I

Calculation of the indicator based on Copernicus High Resolution Layers (HRL): the case of Italy

A simulation of the indicator computation has been carried out by e-GEOS using Copernicus HRL and based on the following workflow:

- 2012 Copernicus HRL Forest Type geospatial layer and Copernicus HRL Water Bodies geospatial layer
- administrative borders geospatial layer provided by ISTAT
- computation of the total country area by administrative borders geospatial layer
- computation of the total land area by subtracting water bodies geospatial layer from the total country area
- computation of indicator based on the ratio HRL Forest Type / Total land area

At national level, the indicator for Italy computed for 2012 based on Copernicus HRL was 36.9%, also a close values to those reported in FRA2010 (36.7%) and FRA2015 (37.7%).The following figures show the results obtained by using only Copernicus geospatial layers for the case of Italy considering a disaggregation at NUTS 3 [Figure 22] and commune level [Figure 23].

Figure 22 I

Forest area as a proportion of total land area based on 2012 Copernicus HRL, NUTS 3 level

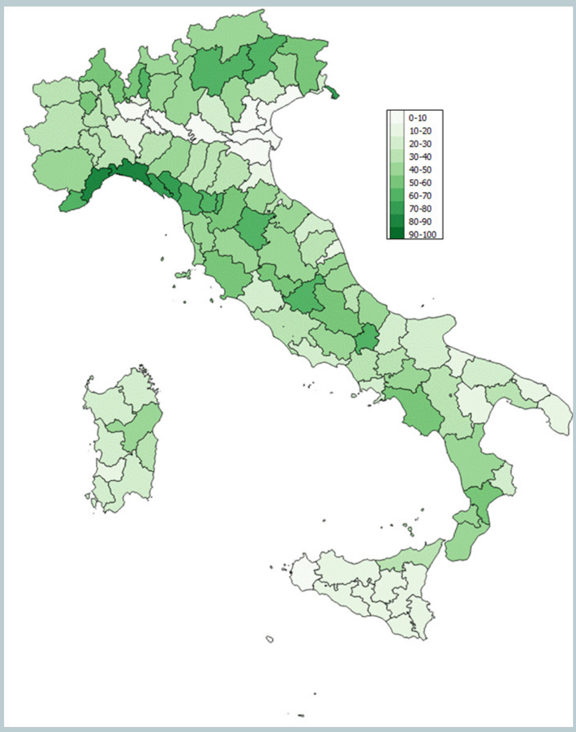
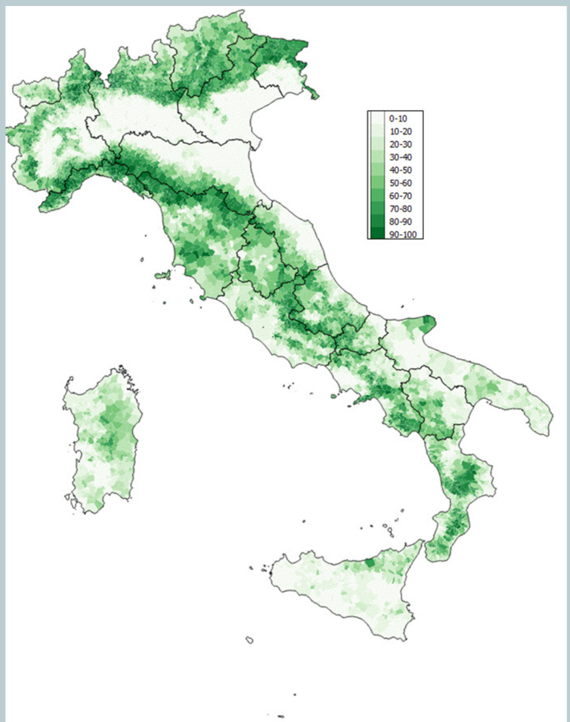


Figure 23 I

Forest area as a proportion of total land area based on 2012 Copernicus HRL, Commune level



Source: Copernicus High Resolution Layers.

Additionally, at the EU level, the CORINE Land Cover (CLC) data could also be used to calculate this indicator, since it provides a strategy for the identification of inland waters, and in particular of rivers, that are considered or not in the global measure according to their size. CORINE Land Cover data are ready to be used, but spatial resolution is relatively coarse (25 ha). Nevertheless, the next generation of CLC (CLC+ for reference year 2018) is expected to provide features at a 0.5 ha spatial resolution.

Both experiences described in Box 12 and Box 13 show how the adoption of geospatial information allows a proper disaggregation on a geographical basis of the indicator, providing a more powerful representation of its spatial variability, for its better understanding and management for a single country. It must also be considered that the usage of geospatial layers generated with a stable methodology enable an easy comparison of data across different countries at high spatial detail.

Analysis at the National level

At the national level, from the cases analysed (Austria, Finland, France, Germany, Italy and Spain), it is possible to identify that this indicator can be calculated based on national data sources. Countries have identified different agencies with responsibility for relevant data for the indicator calculation, namely ministries, NMCA and NSI. Cases have stressed the relevance of geospatial layers and remote sensing techniques in order to improve national data sources disaggregation over small reference units. Finland, France, Italy and Spain have identified National Forest Inventories (NFI) as the main data sources to derive the indicator and this is also the case for Sweden. Regarding the cases of Austria and Germany, forest areas are derived from cadastre information.

The Austrian cadastre is updated every year and is based on a thematic description, so no geometric representation is available. In the case of Germany, the cadastral data comes from the land surveying authorities based on analysis of orthophotos and in situ measurements. The data is transferred from the cadastral institutions to the Federal Statistical Office and each parcel object within the cadastral data contains a land use type, such as “forest”. The indicator relies on the sum of all cadastral parcels with the use type “forest”. The data is updated either on a cyclic basis or occasion-based, namely when a parcel is divided or other cadastral surveying is accomplished.

Forest inventories collect data by applying statistical procedures without a geometric representation. Remote sensing based techniques, when not used for the delimitation of forest area, are used for the collection of additional information over sampled area. For example, in the case of Finland, in addition to sample plot field measurements, the NFI multi-source inventory method employs remote sensed data and other digital data sources such as land-use maps and elevation models, and with the aid of satellite images, the forest characteristics can be estimated for areas lying between the relatively sparse networks of NFI sample plots. In the case of Spain, besides the NFI, an Information System on Land Cover and Land Use (SIOSE) is also available. The SIOSE is part of the National Monitoring Plan (PNOT), managed and coordinated by the NMCA (National Geographic Institute of Spain), and integrates different data of regional and national administrations. It includes 85 classes and a resolution of 2 ha for forest and natural areas. SIOSE is produced in conformity with INSPIRE implementing rules on interoperability of spatial datasets and services and four versions have been made available (2005, 2009, 2011 and 2014).

4 OUTCOMES AND FINDINGS

resulting from the SDG indicators analysis

The following outcomes and findings can be summarised regarding the analysis of the four selected SDG indicators on a global, European and national perspective on the contribution of geospatial data analysis and its integration with statistical data.

SDG indicator 11.2.1 ***Proportion of population that has convenient access to public transport***

- Regarding the concept of urban area/city the underlying definitions of Global Human Settlement Layer (GHSL) could be used at the global level. At the European level, a common definition of urban areas should be considered based on available common territorial typologies, such as the Degree of urbanization (DEGURBA). The global metadata would benefit from a clear definition of stops, especially the methodological approach to deal with large stops with multiple entrances, and of terms such as ‘comfortable environment’ and ‘frequent service’.

- At the data source level, public transport data on stops can entail different levels of detail on available routes, number of services and timetable information whether the indicator is being computed at the global, European and national level. A more general approach on the selection of public transport stops can be followed for global and European comparability and at the national level a more refined and differentiating approach can be applied when detailed data on public transport is available. Point-based population data provides a more flexible and easier way to calculate the indicator as no disaggregation procedure is needed.

- The calculation of network distances, such as the shortest path, can provide more accurate results than using the Euclidian distances, but it requires high quality data on the road network, including pedestrian walks to account for walking access. The proposed global computation takes into account the places of residence as reference points, but other reference points could be considered, such as schools, workplaces, markets and the indicator would then measure the proportion of schools, workplaces or markets with convenient access to public transport.

- The following main issues and challenges were identified regarding the use of geospatial data for this indicator:

  i) availability of comparable point-based data on public transport data, especially including information on routes, frequencies and timetables;

  ii) availability of road network data for computation of distances, including for walking distances;

  iii) complying with the proposed global disaggregation by disability.

SDG indicator 11.3.1 *Ratio of land consumption rate to population growth rate*

- At the conceptual level, the global metadata should be more precise on the definition and use of urban area/city and built-up area as they refer to two different concepts. In this context, urban area/city operationalization would benefit from an international statistical definition to deal with issues of comparability. The EU definition as presented in the territorial typologies regulation (TERCET) can be taken as a reference for a worldwide concept following the discussions in the UN Statistical Division for a global definition of cities as proposed by the European Commission, OECD and partners. On the other hand, the concept of built-up area should be used as a metric to capture artificial land and the expansion of land consumption over time and it is, therefore, a clear distinct concept from the one relating to urban area/city.

- In terms of data sources, the GHSL should be considered as a ready to use product and/or the European Space Agency (ESA) Land Cover products but special attention should also be given to stimulate European remote sensing derived products initiatives worldwide, such as Copernicus Imperviousness High Resolution Layer (HRL) and CORINE Land Cover. National data sources, as exemplified with the Land Use and Land Cover Map for the case of Portugal, can provide more detailed and high quality geospatial data for greater territorial disaggregation of the indicator.

- Regarding the indicator calculation the algorithm on Land Use Efficiency (LUE) as proposed by the Joint Research Centre (JRC) should be considered as it deals with issues of zero population growth and periodicity of the available information by recommending a normalization of the results for a 10 year reference. In addition JRC developed a toolbox in open code format which provides a good way to increase indicator harmonization and comparability.

- The following main issues and challenges were identified regarding the use of geospatial data for this indicator:

  i) accessibility, periodicity, timeliness and methodological stability of the data sources in order to measure progress over time;

  ii) the identification of precise land cover components or categories to derive built-up areas based on different geospatial data products, including at the national level;

  iii) the need to take into account the spatial resolution of input data and its impact on the quality of the territorial disaggregation of statistical outputs.

SDG indicator 11.7.1 *Average share of the built-up area of cities that is open space for public use*

- At the conceptual level, the two different concepts of urban area/city and built-up area should be defined in line with other indicators of Goal 11. In this vein, the concept of built-up area should be measured as artificial land and available common and harmonised territorial typologies (e.g. EU typologies) should be used to capture the urban dimension.

- The global metadata should benefit from a more detailed description on the different dimensions that are supposed to be captured by the concept of open space for public use, especially, in order to achieve comparability across countries at the global level. In this context, other proxies could be considered, such as open green space as shown in the national case for Sweden. Another point of view could be to measure accessibility to this type of spaces in a similar approach as the one defined at indicator 11.2.1 regarding accessibility to public transport.

- In terms of data sources the GHSL provides global coverage to measure built-up areas and other remote sensing derived products, such as Copernicus Imperviousness HRL and CORINE Land Cover provide ready to use comparable data at the European level. National data sources might have more detailed information and are able to better address the conceptual definitions of the indicator, namely the differentiation between public and private open space. Specifically, the case Spain showed the potential of having national cadastral data to derive the data needed for the calculation of the global indicator.

- Cadastral data can provide better data coverage for a more detailed territorial disaggregation and a more consistent and stable classification for measurements over time. The global metadata should, in this vein, include a reference on the substitutability of field data or non-geospatial inventory data for geospatial information as a valuable approach to increase comparability across countries.

- The following main issues and challenges were identified regarding the use of geospatial data for this indicator:

  i) the definition and availability of data sources in order to measure the concept of public open space ensuring international data comparability;

  ii) the availability of data sources with ownership (public vs. private), usually not available in land cover and land use products;

  iii) the combination of different geospatial data sources in order to grasp the necessary definitions to calculate the indicator: urban delimitation, built-up areas and open space for public use.

SDG indicator 15.1.1 *Forest area as proportion of total land area*

- At the conceptual level, it is important to have common technical specifications for targets to be managed by the indicator, in particular forest and inland waters. In addition, a shared forest definition according to standards of the Food and Agriculture Organization (FAO) that could be captured by geospatial layers would provide a way to increase data comparability at the global level.

- In terms of data sources, global (e.g. ESA Land cover) and European (e.g. Copernicus HRL, CORINE Land Cover) geospatial layers, based on remote sensing techniques, could be used for this indicator allowing a more detailed territorial segmentation of the indicator as shown by the different results obtained for the case of Italy. At the European level, geospatial data layers are generated on a regular basis and the situation is the same for satellite data, since EU Sentinel satellites are planned to remain in the orbit for several years.

- At algorithm level, vector or raster format geospatial layers should be used in order to allow proper data aggregation/disaggregation. Different levels of technical algorithm specifications can be considered for a single country depending on its objectives, but common, harmonized specifications and its availability in open formats provide a way to increase data comparability across countries.

- The following main issues and challenges were identified regarding the use of geospatial data for this indicator:

  i) the stability of the geospatial data sources and of its methodology regarding processing workflows with a high degree of automation;

  ii) clear definition of who is responsible for the generation of the output product quality at the global level (at the EU level the production centres validate the products generated);

  iii) capacity building in using earth observation data and derived products to produce statistical indicators.

5 RECOMMENDATIONS

towards a more effective geospatial data integration to address SDG statistical indicators

The outcomes and findings of the analysis carried out on the territorial dimension of SDG indicators have allowed to agree on the following set of recommendations to enhance the contribution of geospatial data analysis and its integration with statistical data to address the SDG indicators which can be directly linked to Global Statistical Geospatial Framework principles:

1. Harmonize relevant geospatial data themes in Europe, such as Buildings, Addresses, Land Use and Land Cover data as well as Cadastral data, according to the UN-GGIM: Europe Core Data recommendations.

2. Implement Cadastral and Land Cover data as key national authoritative data for the operationalization of SDG indicators measurement and encourage European institutions to financially support the Member-States on the implementation of this type of data and its regular update.

3. Use geospatial layers generated from Earth Observation data with a stable and validated methodology at global (e.g. Global Human Settlement Layer) and European level (e.g.,Copernicus High Resolution Layers, CORINE) to enable data comparability across countries.

4. Create capacity building initiatives for National Statistical Institutes to take full advantage of Earth Observation based data to produce new statistical indicators and to increase territorial disaggregation of traditional indicators already reported by NSIs.

5. Define and implement National Spatial Data Infrastructures having in mind the requirements for statistical production to meet the needs of the Sustainable Development Goals monitoring framework and to improve the modernisation of official statistics.

6. Implement consistent and stable sub-national spatial units based on different geographical levels of detail (including grid systems), and its correspondent models of codification, to produce and disseminate coherent and comparable statistical data and indicators over time.

7. Develop and use population grids and other grid-based statistics as a way to increase statistical and geospatial data integration, including geospatial data processing analysis to calculate relevant indicators for the Sustainable Development Goals monitoring framework.

8. Adopt harmonised and comparable concepts, definitions and classifications and build consensus among Geospatial Agencies and National Statistical Institutes in common thematic and technical domains within statistical and geospatial communities.

9. Ensure availability and accessibility of processing workflows, including open formats of programming codes, allowing the automatic or semi-automatic extraction of information from satellite images, the development of algorithms for indicator calculation and territorial classifications (e.g. ESS Degree of urbanization) and of its associated metadata, as a way to improve reporting harmonization and comparability of data.

10. Develop initiatives that promote availability, accessibility and usability of geospatial data by making use of standard metadata and quality reference frameworks aligned with the requisites of the Generic Statistical Business Process Model and Metadata Reporting Standards for statistical production.

11. Increase the collaboration with researchers and data providers to take full advantage of the available data and processing infrastructures and also for tuning operational workflows and regular computation of SDG indicators.

12. Increase cooperation between National Statistical Institutes and Geospatial Agencies for the calculation of SDG indicators, to better address the territorial dimension of SDG indicators and to promote the relevance of geographical data in institutional national forums for SDG reporting.

In general, it is fundamental to increase collaboration between National Statistical Institutes and Geospatial Agencies (comprising the National Mapping and Cadastral Agencies) within countries and across European and Global institutions.

This could be a way to improve processes and methodologies, to harmonize concepts, definitions and procedures, to develop relevant new statistical indicators and to assure consistent points of view in international forums. Action plans of cooperation in these domains can be defined and implemented with the aim of modernizing the statistical production chain and of promoting better data integration.

Annex I

Template for indicator analysis

1. CURRENT REPORTING SITUATION
Responsibility: (Identify the agency responsible for the indicator and the situation regarding the ESS and NSS projects (including dissemination) and /or INSPIRE conformance)
Indicator disaggregation: (List the indicator disaggregation by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts to support the monitoring of the implementation of the SDGs)
Frequency of dissemination: (Describe the time interval at which information is disseminated over a given time period)
Timeliness: (Length of time between data availability and the event or phenomenon they describe. Describe the average production time for each release of data)
Data sources: (List the data sources and themes or variables in use, including conditions of access, timeliness and frequency of dissemination, situation regarding the ESS and NSS projects (including dissemination) and /or INSPIRE conformance)
Geospatial data analysis and integration: (Describe spatial analysis methods, procedures and computations, including regarding data integration)
Data quality requirements: (List in general terms the requirements for the sources and themes in use with relevant parameters: Resolution, completeness, logical consistency, positional accuracy, temporal accuracy etc. List if certain international standards are being followed, including classifications/nomenclatures. Data quality should allow computing results to the needed level of resolution and disaggregation). Please take into account the EURO-SDMX Metadata Structure (ESMS) 2.0
Current use of geospatial data for the indicator: (Describe the current use of geospatial data, as suggested by the existing metadata – the "as-is" situation)
2. SUGGESTED METHODOLOGY
GAP analysis: (Describe what changes in use of **applied methods** are needed to go from the suggested/current procedure for monitoring the indicator, to a future procedure which better fulfils the reporting requirements - going from the "as-is" situation in the present metadata proposal to a "to-be" situation)
3. SUGGESTED GEOSPATIAL DATA INTEGRATION
GAP analysis: (Describe what changes in use of **data** needed to go from the suggested/current procedure for monitoring the indicator, to a future procedure which better fulfils the reporting requirements - going from the "as-is" situation in the present metadata proposal to a "to-be" situation)
List required geospatial data: (Develop a list from the GAP analysis, which lists the geospatial data sources and themes which are required to support the to-be situation, including INSPIRE conformance)
Data quality requirements: (List in general terms the requirements for the suggested sources and themes with relevant parameters: Resolution, completeness, logical consistency, positional accuracy, temporal accuracy etc. List if certain international standards should be followed including classifications/nomenclatures. Data quality should allow computing results to the needed level of resolution and disaggregation). Please take into account the EURO-SDMX Metadata Structure (ESMS) 2.0
Data availability: (List the data availability for the suggested sources and themes or variables: 1) Geographically: national/ regional/global (as well as comparability across countries), 2) Source: Accessible through services or download, 3) Commercial/ legally: license conditions - are data free or are there restriction on use; 4) Timeliness; 5) Frequency of dissemination)
Data collection: (Describe how the geospatial data for the indicator can be collected/made available, and issues to overcome – are there many sources to collect from, do they need to be integrated and normalized etc.)
Geospatial data analysis and integration: (Describe which analysis, procedures and computations are needed to provide the results needed to support the reporting requirements - "to-be" situation)

Annex II

Table of contributions for each selected SDG indicator

SDG Indicator	Contributors	Documents of analysis
11.2.1	Austria (NSI) France (NMCA) Ireland (NSI) Sweden (NSI) Switzerland (NSI)	✓ Detailed indicator analysis ✓ Brief discussion paper
11.3.1	Finland (NMCA) Ireland (NSI) Italy (e-GEOS) Portugal (NSI and NMCA)	✓ Detailed indicator analysis ✓ Brief discussion paper
11.7.1	Ireland (NSI) Sweden (NSI and NMCA) Switzerland (NSI)	✓ Detailed indicator analysis ✓ Brief discussion paper
15.1.1	Austria (NMCA) Finland (NMCA) France (NMCA) Germany (NMCA) Italy (e-GEOS) Spain (NMCA)	✓ Detailed indicator analysis ✓ Brief discussion paper

Annex III
List of contributors

	Name	Member-State	Institution
1	Francisco Vala	Portugal	Statistics Portugal (INE)
2	Cátia Nunes	Portugal	Statistics Portugal (INE)
3	Pier-Giorgio Zaccheddu	Germany	Federal Agency for Cartography and Geodesy (BKG)
4	Sabine Afflerbach-Thom	Germany	Federal Agency for Cartography and Geodesy (BKG)
5	Jeanette Kretz	Germany	Federal Agency for Cartography and Geodesy (BKG)
6	Jerker Moström	Sweden	Statistics Sweden
7	Fabio Volpe	Italy	Geo Content Innovation at e-GEOS
8	Ingrid Kaminger	Austria	Statistics Austria
9	Lars Hägg	Sweden	Lantmäteriet (Swedish mapping, cadastral and land registration authority)
10	Amalia Velasco Martín-Varés	Spain	Spanish Directorate General for Cadastre
11	Paloma Abad Power	Spain	Instituto Geográfico Nacional
12	Rainer Humbel	Switzerland	Swiss Federal Statistical Office (FSO)
13	Esa Tiainen	Finland	National Land Survey of Finland
14	Igor Kuzma	Slovenia	Statistics Slovenia
15	Dominique Laurent	France	Institut Geographique National (IGN)
16	Dermont Corcoran	Ireland	Statistics Ireland
17	Ekkehard Petri	EU	Eurostat
18	Vlado Cetl	EU	European Commission (Joint Research Centre - JRC)
19	Jose Miguel Rubio-Iglesias	EU	European Commission (EEA)
20	David Nuno	Portugal	Directorate-General for Territory (DGT)
21	Mário Lucas	Portugal	Statistics Portugal (INE)
22	Bart Schoenmakers	Portugal	Statistics Portugal (INE)

Annex IV
Acronyms

Acronym	Full name
ABS	Australian Bureau of Statistics
API	Application Programming Interface
BKG	Federal Agency for Cartography and Geodesy
CLC	CORINE Land Cover
CLMS	Copernicus Land Monitoring Service
CORINE	Coordination of Information on the Environment
COS	Land Use and Land Cover Map
DEGURBA	Degree of Urbanization
DG REGIO	Directorate-General for Regional Development
DIAS	Copernicus Data and Information Access Service
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecasts
ECOSOC	United Nations Economic and Social Council
EEA	European Environmental Agency
EFTA	European Free Trade Association
EG-ISGI	Expert Group on Integration of Statistical and Geospatial Information
EO	Earth Observation
ESA	European Space Agency
ESMS	Euro-SDMX Metadata Structure
ESQRS	ESS Standard for Quality Reports Structure
ESRI	Environmental Systems Research Institute
ESS	European Statistical System
ESTAT	Eurostat - Statistical Office of the European Union
EU	European Union
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
EU-SILC	European Union Survey on Income and Living Conditions
FAO	Food and Agriculture Organization
FRA	Global Forest Research Assessment
FUA	Functional Urban Areas
GEO	Group on Earth Observations
GEOSTAT	Project of Eurostat in cooperation with the European Forum for Geography and Statistics
GHSL	Global Human Settlement Layer
GROW	Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs
GSBPM	Generic Statistical Business Process Model

Annex IV
Acronyms

Acronym	Full name
GSGF	Global Statistical Geospatial Framework
GTFS	Google General Transit Feed Specification
HRL	High Resolution Layer
IAEG-SDG	Inter-Agency and Expert Group on SDG Indicators
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
JRC	Joint Research Centre
LAU	Local Administrative Unit
LUCAS	Land Use and Land Cover Survey
LUE	Land Use Efficiency
MMU	Minimum Mapping Unit
NASA	National Aeronautics and Space Administration
NFI	National Forest Inventories
NMCA	National Mapping and Cadastral Agency
NSI	National Statistical Institute
NSS	National Statistical System
NUTS	ESS Classification of Territorial Units for Statistics
OECD	Organization for Economic Co-operation and Development
ONS	Office for National Statistics - United Kingdom
OSi	Ordnance Survey Ireland
PNOT	National Monitoring Plan
QGIS	Quantum Geographic Information System
SDG	Sustainable Development Goal
SDMX	Statistical Data and Metadata eXchange
SIMS	Single Integrated Metadata Structure
SIOSE	Information System on Land Cover and Land Use
TERCET	ESS Territorial Typologies
UN	United Nations
UN DESA	United Nations Department of Economic and Social Affairs
UNECE	United Nations Economic Commission for Europe
UN-GGIM	United Nations - Global Geospatial Information Management
UN-Habitat	United Nations Human Settlements Programme
WG	Working Group
WG GI	IAEG-SDG Working Group on Geospatial Information

UN-GGIM EUROPE

UNITED NATIONS
COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

Edited by Statistics Portugal