European Commission

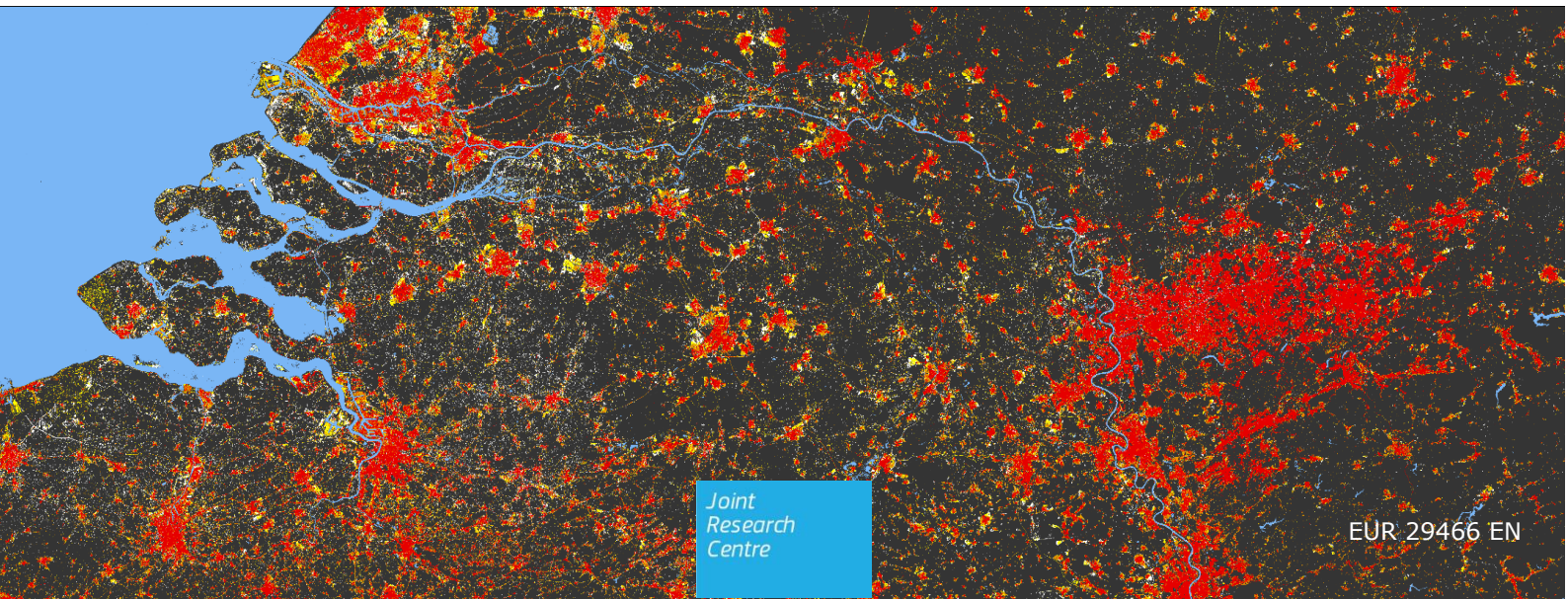
JRC TECHNICAL REPORTS

Community pre-Release of GHS Data Package (GHS CR2018) in support to the GEO Human Planet Initiative

Version 1.0

Florczyk A.J., Ehrlich D., Corban C., Freire S., Kemper T., Melchiorri M., Pesaresi M., Politis P., Schiavina M., Zanchetta L.

2018



Joint Research Centre

EUR 29466 EN

Contact information
Name: Thomas Kemper
Address: Via Fermi, 2749 21027 ISPRA (VA) - Italy - TP 267
European Commission - DG Joint Research Centre
Space, Security and Migration Directorate
Disaster Risk Management Unit E.1
Email: thomas.kemper@ec.europa.eu
Tel.: +39 0332 78 5576

GHSL project: JRC-GHSL@ec.europa.eu
GHSL Data: JRC-GHSL-DATA@ec.europa.eu

JRC Science Hub
https://ec.europa.eu/jrc

JRC111451

EUR 29466 EN

PDF	ISBN 978-92-79-97547-9	ISSN 1831-9424	doi:10.2760/777868
Print	ISBN 978-92-79-97546-2	ISSN 1018-5593	doi:10.2760/421521

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Florczyk A.J., Ehrlich D., Corban C., Freire S., Kemper T., Melchiorri M., Pesaresi M., Politis P., Schiavina M., Zanchetta L., *Community pre-Release of GHS Data Package (GHS CR2018) in support to the GEO Human Planet Initiative*, EUR 29466 EN, Publications Office of the European Union, Luxembourg, 2018,
ISBN 978-92-79-97547-9, doi:10.2760/777868,

Contents

Authors

Aneta J. Florczyk[a], Daniele Ehrlich, Christina Corbane[a], Sergio Freire[a], Thomas Kemper[a], Michele Melchiorri[b], Martino Pesaresi[a], Panagiotis Politis[c], Marcello Schiavina[a], Luigi Zanchetta[a]

[a] European Commission-Joint Research Centre, Disaster Risk Management Unit
[b] Piksel S.r.l
[c] Arhs Developments S.A.

Abstract

The GEO Human Planet Initiative is committed to developing a new generation of measurements and information products that provide new scientific evidence and a comprehensive understanding of the human presence on the planet and that can support global policy processes with agreed, actionable and goal-driven metrics. The Human Planet Initiative relies on a core set of partners committed in coordinating the production of the global settlement spatial baseline data, and an enlarged community of partners developing experimental activities on using the new baseline data for derived post-2015 indicators. One of the core partners is the European Commission, Directorate General Joint Research Centre, Global Human Settlement Layer project. The Global Human Settlement Layer project produces global spatial information, evidence-based analytics, and knowledge describing the human presence in the planet.

This document describes the Community pre-Release of the GHSL Data Package 2018 - GHS CR2018 - created for the members of the GEO Human Planet Initiative. The data in this data package have the purpose to gather feedbacks from the GEO Human Planet Initiative community, in preparation for the public release of these data.

Disclaimer: the data included in the Community Pre-Release of the GHSL Data Package 2018 should be considered preliminary and not yet validated. It may differ substantially with the data that will be included in the public final Release 2018. The JRC data are provided "as is" and "as available" in conformity with the JRC Data Policy[1] and the Commission Decision on reuse of Commission documents (2011/833/EU). Although the JRC guarantees its best effort in assuring quality when publishing these data, it provides them without any warranty of warranty of any kind, either express or implied, including, but not limited to, any implied warranty against infringement of third parties' property rights, or merchantability, integration, satisfactory quality and fitness for a particular purpose. The JRC has no obligation to provide technical support or remedies for the data. The JRC does not represent or warrant that the data will be error free or uninterrupted, or that all non-conformities can or will be corrected, or that any data are accurate or complete, or that they are of a satisfactory technical or scientific quality. The JRC or as the case may be the European Commission shall not be held liable for any direct or indirect, incidental, consequential or other damages, including but not limited to the loss of data, loss of profits, or any other financial loss arising from the use of the JRC data, or inability to use them, even if the JRC is notified of the possibility of such damages.

[1]JRC Data Policy https://doi.org/10.2788/607378

1 Introduction

1.1 Overview

The Global Human Settlement Layer (GHSL) project produces global spatial information, evidence-based analytics, and knowledge describing the human presence in the planet. The GHSL relies on the design and implementation of new spatial data mining technologies that allow automatic processing, data analytics and knowledge extraction from large amounts of heterogeneous data including global, fine-scale satellite image data streams, census data, and crowd sources or volunteered geographic information sources.

This document accompanies the community pre-Release of the GHSL data package 2018 (GHS CR2018), made available to the members of the GEO Human Planet Initiative (HPI).

1.2 Rationale

The data openness and free access lays in the core of principles of GHSL and HPI, in order to favour an open forum for discussion for academy, international stakeholders, governmental bodies and private firms.

Before releasing the new GHSL data package as open and free data, the members of the HPI community are asked for feedback on these data.

1.3 History and Versioning

In 2016 the first GHSL data package was released (GHS P2016), as open and free data. It consists in several multi-temporal and multi-resolution products, including built-up area grids (GHS-BUILT), population grids (GHS-POP), settlement model (GHS-SMOD) and selected quality grids (data mask and confidence grids for GHS-BUILT).

The GHS-BUILT product is the result of a large scale experiment aimed at extracting information on built-up from Landsat in 2014/2015 (Pesaresi *et al.,* 2016a), producing the first multi-temporal explicit description of the evolution of built-up presence in past 40 years. The main product is the GHS_BUILT_LDSMT_GLOBE_R2015B[2] (Pesaresi *et al.,* 2015), and two quality grids accompany it: (1) a built-up confidence layer (GHS_BUILTt_LDSMTCNFD_GLOBE_R2015B[3]) and (2) data mask layer (GHS_BUILTt_LDSMTDM_GLOBE_R2015B[4]).

The population grids (GHS_POP_GPW41MT_GLOBE_R2016A5) were produced in collaboration with Columbia University, Center for International Earth Science Information Network (CIESIN) in 2015, and the GHS-SMOD grids (GHS_SMOD_POP_GLOBE_R2016A6) present an assessment of the REGIO degree of urbanization model using as input the population grid cells.

The datasets from the GHS P2016 are available at GHSL collection in JRC Open Data Repository[7].

In 2017, a revised image processing workflow was implemented in the JRC Earth Observation Data and Processing Platform (JEODPP), and applied the Landsat multi-temporal imagery collection. As a result, an updated version of the multi-temporal built-up and population grids has been produced, GHS_BUILT_LDSMT_GLOBE_R2018A and GHS_POP_GPW41MT_GLOBE_R2018A respectively. These datasets are distributed in the community pre-Release of the GHSL data package 2018 (GHS CR2018). The final public

[2] http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmt_globe_r2015b
[3] http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmtcnfd_globe_r2015b
[4] http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmtdm_globe_r2015b
[5] http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a
[6] http://data.europa.eu/89h/jrc-ghsl-ghs_smod_pop_globe_r2016a
[7] http://data.jrc.ec.europa.eu/collection/ghsl

release of the GHSL data package 2018 (GHS P2018) is planned towards the end of 2018.

1.4 Main Characteristics

There is a clear interdependency between the GHSL products starting from the training set. Therefore, this data package pre-release includes also a product mapping the built-up area presence at 20-m spatial resolution, derived from Sentinel-1 imagery (GHS_BUILT_S12016NODSM_GLOBE_R2018A). The Sentinel-1 built-up area grid was used as training set for the production of the new Landsat GHS built-up (GHS_BUILT_LDSMT_GLOBE_R2018A).

In order to facilitate the data analytics, as it was in the GHS P2016, the release includes a set of multi-resolution datasets produced by aggregation of the main products. Additionally, the density grids are produced in an equal-area projection as grids of 250m and 1-km spatial resolution. For example, the multi-temporal population grids were produced in grids of 250-m spatial resolution, later aggregated to $1km^2$.

The main difference between the products in GHS P2016 and the current products (GHS CR2018) are:

- Improved workflow for built-up area extraction from satellite image, for example, refined learning datasets (e.g., GHS_BUILT_S12016NODSM_GLOBE_R2018A), production at 30m spatial resolution;
- Improved approach for production of population grids;
- Technical specification of the grids (i.e., the grid origin);
- Encoding of NoData values (e.g., projection domain, *NoData* within the data domain).

The subsections of the Section 2 introduce briefly each product (including more details on differences with the corresponding past version). Dedicated reports are under preparation.

1.5 Terms of Use

 The pre-release data have the purpose to allow the GEO Human Planet community to perform early testing before the public release. Redistribution of these data outside the GEO HPI community should be done only with explicit agreement from the GHSL team. The pre-release data are not validated and only partially documented. In order to reduce the risk of misuse, it is strongly recommended to involve the GHSL data producer team in case these data will be included in experimentations contributing to future scientific publications. For any inquiry related to the use of these data please contact the GHSL data producer team at the electronic mail address:

JRC-GHSL@ec.europa.eu

2 Products

2.1 GHS Built-up grid derived from Sentinel-1 (2016), R2018A [GHS_BUILT_S12016NODSM_GLOBE_R2018A]

The Sentinel-1 product is a layer grid that contains a built-up area classification derived from Sentinel-1 backscatter images. This product increases the spatial coverage of the dataset produced in 2016, referred to as GHS_BUILT_S12016NODSM_GLOBE_R2016A.

The information extraction of Sentinel-1A data at global scale is described in Corbane *et al.* (2017). The main workflow builds on a new artificial intelligence approach for the satellite data classification process named "Symbolic Machine Learning" (SML)(Pesaresi at al., 2017). The SML classifier automatically generates inferential rules linking the image data to available high-abstraction semantic layers used as training sets.

The SML workflow was adapted to exploit the key features of the Sentinel-1 Ground Range Detected (GRD) data which are: i) the spatial resolution of 20m with a pixel spacing of 10m and ii) the availability of dual polarisation acquisitions (VV and VH) widely used for monitoring urban areas since different polarizations have different sensitivities and different backscattering coefficients for the same target.

The learning data at the global level consisted of the union of the built-up obtained from the GHSL-Landsat for 2014 and the Global Land Cover map at 30 meter resolution (GLC30). The latter has been also derived from Landsat imagery through operational visual analysis techniques (Chen *et al.*, 2015).
The massive processing of more than 7000 Sentinel-1 scenes was enabled by JEODPP platform developed in the framework of the JRC Big Data Pilot Project. The platform is set-up to answer the emerging needs of the JRC Knowledge Production units following the new challenges posed by Earth Observation entering the big data era.

Figure 1. Mosaic of the S1 scenes processed within the SML for extracting built-up areas

2.1.1 Input Data

The input imagery collection consists of Sentinel-1A (S1A) and Sentinel-1B (S1B) images:

- 5026 S1A images from December 2015 to October 2016;
- 1695 S1A and 329 S1B images from November 2016 to December 2017.

2.1.2 Technical Details

Author: C. Corbane, P. Politis, V. Syrris; Joint Research Centre (JRC) European Commission

Product name: GHS_BUILT_ S12016NODSM _GLOBE_R2018A

Spatial extent: Global

Temporal extent: 2016

Projection: Spherical Mercator (EPSG:3857)

Resolutions available: 20-m

*Encoding**: Built-up area classification map (integer) [0,1];

Data organisation (*): VRT file (with TIFF tiles); pyramids. **ArcGIS users of the 30-m product: *ESRI.vrt. file**

The grid is provided as a VRT file (with GeoTIFF tiles), and with pyramids. Table 1 below outlines the technical characteristics of the datasets pre-Released in this data package.

Table 1. Technical details of the datasets in GHS_BUILT_ S12016NODSM _GLOBE_R2018A

GHS_BUILT_S12016NODSM_GLOBE_R2018A			***34 GB***
ID	**Description**	**Resolution (projection)**	**Size (ZIP)**
GHS_BUILT_S12016NODSM _GLOBE_R2018A _3857_20_V_1_0	Classification map depicting built-up presence. 0 = no built-up or no data 1 = built-up are **ArcGIS users: *ESRI.vrt.file**	20-m (Pseudo Mercator)	34 GB

2.2 GHS built-up grid, derived from Landsat, multi-temporal (1975-1990-2000-2014), R2018A [GHS_BUILT_LDSMT_GLOBE_R2018A]

The Landsat product contains a set of multi-temporal and multi-resolution grids. The main product is the multi-temporal classification layer on built-up presence derived from the Global Land Survey (GLS) Landsat[8] image collections (GLS1975, GLS1990, GLS2000, and ad-hoc Landsat 8 collection 2013/2014). A detailed report is under preparation, and here we briefly introduce the product.

2.2.1 Improvements comparing to the previous version

The satellite-derived information extraction tasks included in the GHSL production workflow used to produce GHS_BUILT_LDSMT_GLOBE_R2015B and GHS_BUILT_LDSMT_GLOBE_R2018A builds on the Symbolic Machine learning (SML) method that was designed for remote sensing big data analytics (Pesaresi *et al.*, 2016b). For the purpose of the GHS_BUILT_LDSMT_GLOBE_R2018A, a revised image processing workflow was implemented in the JRC Earth Observation Data and Processing Platform (JEODPP).

Comparing to the previous version (R2015B), this dataset includes a number of improvements, namely:

[8] http://glcf.umd.edu/data/gls/

- Improved spatial coverage (additional Landsat 8 scenes)
- Improved spatial resolution (30-m)
- Improved methods (e.g., improved learning data set), which resulted in:
    - Reduction in omission error (i.e. more built-up areas were detected)
    - Reduction in commission error (i.e. less detection of false built-up areas)

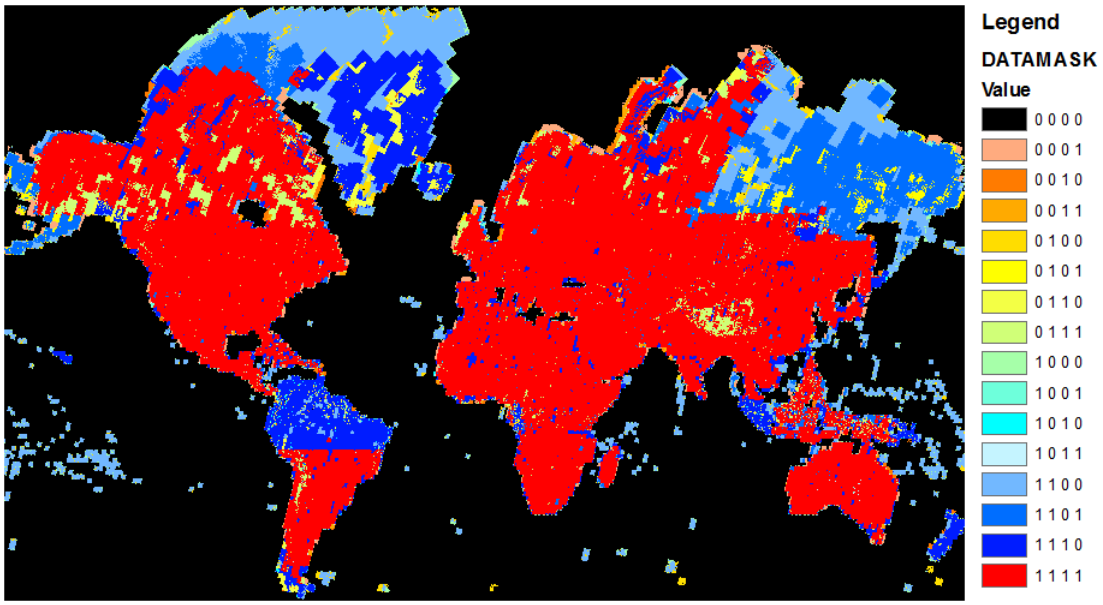
However, the main improvement is the usage of the GHSL Sentinel-1 data set (GHS_BUILT_S12016NODSM_GLOBE_R2018A) as a learning dataset.

2.2.2 Input Data

In total, 32,808 images organized in four Landsat data collections centred at 1975, 1990, 2000 and 2014 were processed with the SML classifier as follows:

- 7,588 scenes acquired by the Multispectral Scanner (collection 1975);
- 7,375 scenes acquired by the Landsat 4-5 Thematic Mapper (TM) (collection 1990);
- 8,756 scenes acquired by the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) (collection 2000) and;
- 9,089 scenes acquired by Landsat 8 (collection 2014).

Figure 2. Overview of Landsat data availability by collection; the binary codes shown in the legend correspond to the four data collections in the following order: 2014-2000-1990-1975



2.2.3 Technical Details

Author: C. Corbane, A.J. Florczyk, M. Pesaresi, P. Politis, V. Syrris; Joint Research Centre (JRC) European Commission

Product name: GHS_BUILT_LDSMT_GLOBE_R2018A

Spatial extent: Global

Temporal extent: 1975-1990-2000-2014

*Projection**: Spherical Mercator (EPSG:3857), World Mollweide (EPSG:54009)

*Resolutions available**: 30-m, 250-m, 1-km

*Encoding**: Multi-temporal built-up area classification map (integer): [1,6], NoData: 0; Built-up density grid (float32): [0-100], NoData [-201, -200]

Data organisation (*): VRT file (with GeoTIFF tiles) or GeoTIFF files; pyramids **ArcGIS users of the 30-m product: *ESRI.vrt.file**

(*) product dependent, see Table 2.

Table 2. Technical details of the datasets in GHS_BUILT_ LDSMT_GLOBE_R2018A

GHS_BUILT_LDSMT_GLOBE_R2018A			***8.3 GB***
ID	**Description**	**Resolution (projection)**	**Size (ZIP)**
GHS_BUILT_LDSMT _GLOBE_R2018A _3857_30_V_1_0	Multi-temporal classification of built-up presence. 0 = no data 1 = water surface 2 = land no built-up in any epoch 3 = built-up from 2000 to 2014 epochs 4 = built-up from 1990 to 2000 epochs 5 = built-up from 1975 to 1990 epochs 6 = built-up up to 1975 epoch **ArcGIS users: *ESRI.vrt.file**	30-m (Pseudo Mercatore)	6.0 GB
GHS_BUILT_LDS2014 _GLOBE_R2018A _54009_250_V_1_0	Built-up area density for epoch 2014, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – no data	250-m (World Mollweide)	674 MB
GHS_BUILT_LDS2000 _GLOBE_R2018A _54009_250_V_1_0	Built-up area density for epoch 2000, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	250-m (World Mollweide)	64 MB
GHS_BUILT_LDS1990 _GLOBE_R2018A _54009_250_V_1_0	Built-up area density for epoch 19990, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	250-m (World Mollweide)	575 MB
GHS_BUILT_LDS1975 _GLOBE_R2018A _54009_250_V_1_0	Built-up area density for epoch 1975, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	250-m (World Mollweide)	56 MB
GHS_BUILT_LDS2014 _GLOBE_R2018A _54009_1K_V_1_0	Built-up area density for epoch 2014, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	1-km (World Mollweide)	491 MB
GHS_BUILT_LDS2000 _GLOBE_R2018A _54009_1K_V_1_0	Built-up area density for epoch 2000, aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	1-km (World Mollweide)	48 MB
GHS_BUILT_LDS1990 _GLOBE_R2018A	Built-up area density for epoch 1990. Aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100	1-km (World Mollweide)	399 MB

_54009_1K_V_1_0	NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain		
GHS_BUILT_LDS1975 _GLOBE_R2018A _54009_1K_V_1_0	Built-up area density for epoch 1975. Aggregated from 30m. Values are expressed as decimals (Float) from 0 to 100 NoData [-201, -200]: -200 – out of projection domain -201 – NoData within the projection domain	1-km (World Mollweide)	40 MB

2.3 GHS population grid, derived from GPW4.1, multi-temporal (1975-1990-2000-2014), R2018A [GHS_POP_GPW41MT_GLOBE_R2018A]

This spatial raster dataset depicts the distribution and density of population, expressed as the number of people per cell. Residential population estimates for target years 1975, 1990, 2000 and 2015 provided by CIESIN Gridded Population of the World, version 4.10 (GPWv4.10) were disaggregated from census or administrative units to grid cells, informed by the distribution and density of built-up as mapped in the Global Human Settlement Layer (GHSL) global layer per corresponding epoch (for disaggregation method see Freire *et al.*, 2016).

2.3.1 Improvements comparing to the previous version

The new version of the GHSL population distribution grids aimed at incorporating improvements originating from input datasets, namely population estimates and built-up presence. While the disaggregation relied essentially on the same clear and simple approach, there were significant differences to the input data that had a positive effect on the final quality and accuracy of population grids. Here, we describe the main differences between the current pre-Release (GHS_POP_GPW41MT_GLOBE_R2018A) and the previous one (GHS_POP_GPW41_GLOBE_R2015A).

For the new GHS-POP (GHS_POP_GPW41MT_GLOBE_R2018A), the new Landsat based GHS-BU (GHS_BUILT_LDSMT_GLOBE_R2018A) was used as target for disaggregation of population estimates. Cells declared as "NoData" in built-up layers were treated as zero for population disaggregation.

The base source of population estimates (both counts and geometries) for the four epochs was the Gridded Population of the World, version 4.10 (GPWv4.10), from CIESIN/SEDAC. Respect to the previous release of GHSL grids (R2015A), this pre-Release used GPW source data that incorporated boundary or population updates for 67 countries.

Due to the previous GHSL population grids being produced in last quarter of 2015, before the final GPWv4 data set was fully assembled, more changes were included in population sources in the current pre-Release than those incorporated in the GPW data between GPWv4 and the current GPWv4.10. For detailed information on what has changed in GPWv4.10, refer to:

http://beta.sedac.ciesin.columbia.edu/data/collection/gpw-v4/whatsnewrev10

2.3.1.1 Harmonisation of Coastlines

Seashore and waterfront can be especially intense and dynamic zones, contributing to making census or administrative geometries outdated and inaccurate. Inconsistencies between census data and GHSL along coastlines (including inland water bodies) were detected and reconciled accordingly. The high-resolution GHSL layer on built-up areas for 2014 (from R2015B) was used to detect significant human presence (i.e. BU) beyond censuses' coastlines and these lines were reconciled accordingly. This harmonization was carried out in the following countries:

Albania	Finland	Romania
Austria	France	Russia
Azerbaijan	Guinea-Bissau	Singapore
Bulgaria	Iceland	Sweden
Bahrain	Japan	Tunisia
Switzerland	Republic of Korea	Ukraine
Germany	Malaysia	USA
Denmark	Netherlands	Venezuela
United Arab Emirates	Norway	Viet Nam

2.3.1.2 Revision of Unpopulated Areas

Units deemed as "uninhabited" in the census data were critically assessed for presence of residential population, based on ancillary data and high-resolution imagery. Inconsistencies between census data and contradicting evidence were detected and reconciled accordingly. An automated method was devised to split and merge these polygons, based on geographical proximity, with those ones adjacent and containing population. This procedure was implemented while minimizing changes to source geometry, preserving the regional distribution of population, and the overall counts. This procedure was carried out in the following countries:

Afghanistan	Georgia	Nepal
Armenia	Guyana	Rwanda
Democratic Republic of the Congo	Iraq	Thailand
Colombia	Lebanon	Ukraine
Cyprus	Mali	
Egypt	Malawi	

2.3.2 Input Data

The new product GHS_BUILT_LDSMT_GLOBE_R2018A was used as target for disaggregation of population estimates. The base source of population estimates for the four epochs was the Gridded Population of the World, version 4.10 (GPWv4.10), from CIESIN/SEDAC, with some modifications as described above.

2.3.3 Technical Details

Author: *S. Freire, M. Schiavina,* Joint Research Centre (JRC) European Commission; Columbia University, Center for International Earth Science Information Network - CIESIN.

Product name: GHS_POP_GPW41MT_GLOBE_R2018A

Spatial extent: Global

Temporal extent: 1975-1990-2000-2015

Projection: World Mollweide (EPSG: 54009)

Resolutions available: 250-m, 1-km

Encoding: Population data float32 [0, ∞); No Data: -200 outside World Mollweide projection domain, -201 outside Population source area (census)

Data organisation: The grids are provided as GeoTIFF files, together with pyramids.

Table 3 outlines the technical characteristics of the datasets pre-Released in this data package.

Table 3. Technical details of the datasets in GHS_POP_GPW41MT_GLOBE_R2018A

GHS_POP_GPW41MT_GLOBE_R2018A			***2.3 GB***
ID	**Description**	**Resolution (Projection)**	**Size (ZIP)**
GHS_POP_GPW41E2015_ GLOBE_R2018A _54009_250_V_1_0	Population density for epoch 2015 Values are expressed as decimals (Float) from 0 to 442590.9375 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	250-m (World Mollweide)	510 MB
GHS_POP_GPW41E2000_ GLOBE_R2018A _54009_250_V_1_0	Population density for epoch 2000 Values are expressed as decimals (Float) from 0 to 303161.25 N oData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	250-m (World Mollweide)	462 MB
GHS_POP_GPW41E1990_ GLOBE_R2018A _54009_250_V_1_0	Population density for epoch 1990 Values are expressed as decimals (Float) from 0 to 237913.46875 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	250-m (World Mollweide)	431 MB
GHS_POP_GPW41E1975 _GLOBE_R2018A _54009_250_V_1_0	Population density for epoch 1975 Values are expressed as decimals (Float) from 0 to 899329.375 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	250-m (World Mollweide)	401 MB
GHS_POP_GPW41E2015 _GLOBE_R2018A _54009_1K_V_1_0	Population density for epoch 2015 Values are expressed as decimals (Float) from 0 to 442590.9375 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	1-km (World Mollweide)	136 MB
GHS_POP_GPW41E2000 _GLOBE_R2018A _54009_1K_V_1_0	Population density for epoch 2000 Values are expressed as decimals (Float) from 0 to 338233.03125 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	1-km (World Mollweide)	132 MB
GHS_POP_GPW41E1990 _GLOBE_R2018A _54009_1K_V_1_0	Population density for epoch 1990 Values are expressed as decimals (Float) from 0 to 1013921.625 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	1-km (World Mollweide)	131 MB
GHS_POP_GPW41E1975 _GLOBE_R2018A _54009_1K_V_1_0	Population density for epoch 1975 Values are expressed as decimals (Float) from 0 to 3017848.75 NoData [-201, -200]: -200 – out of projection domain -201 – no population data (from census)	1-km (World Mollweide)	132 MB

2.4 Country Urbanisation and Agglomeration Dynamics Database, for epochs (1975-1990-2000-2015), R2017A [GHS_STAT_CNTRDUMT_GLOBE_R2017A]

The country urbanisation and agglomeration dynamics database describes the amount of population and built-up surface per each class of "Degree of Urbanization" and per

country, in the four standard GHSL epochs 2015, 2000, 1990, and 1975. The data were created by the application of the "Degree of Urbanization" model (featuring (1) cities vs. (2) towns and suburbs vs. (3) rural areas) to the GHSL baseline data (GHS_SMOD_POP_GLOBE_R2016A).

Note: During 2018, this database will be consolidated based on the new GHSL baseline data release 2018 and taking into account ongoing discussions in the working group for a global definition of cities and settlements.

2.4.1 Input Data

The input data are the multi-temporal grids of the GHSL Data Package 2016 (GHS P2016): GHS-BUILT, GHS-POP and GHS-SMOD. The source of the GIS country layer is the database of Global Administrative Areas (GADM)[9] version 2.8 (November 2015).

2.4.2 Technical Details

Author: A.J. Florczyk, M. Melchiorri, M. Pesaresi; Joint Research Centre (JRC) European Commission.

Product name: GHS_STAT_UCDB2015_GLOBE_R2017A

Spatial extent: Global

Temporal extent: 1975-1990-2000-2015

File Name: GHS_STAT_CNTRDUMT_GLOBE_R2017A_V1_0.xlsx

File Format: XLS

This dataset is formatted as an XLS file, containing:

i. description sheet – general description of the data
ii. data sheet – data table with multi-temporal attributes

2.5 Urban Centres Database, multi-temporal attributes, R2017A [GHS_STAT_UCDB2015_GLOBE_R2017A]

The Urban Centres Database describes more than 10.000 urban centres identified by the application of the "Degree of Urbanization" model to the GHSL baseline data publicly released in 2016 (GHS P2016).
Each city is characterised by a set of consistent attributes reporting about resident population, land cover, physical land characteristics (slope, elevation), night light emissions, air quality and presence of vegetation. Most of the indicators have a temporal depth until 1990s. Some of them go back to 1975. The urban centres database provides new open data supporting the monitoring of Sustainable Development Goals, the New Urban Agenda and the Sendai Framework for Disaster Risk Reduction. The pre-release Urban Centres Database is the same data showcased by the GHSL web site at the 9th session of the UN World Urban Forum in February 2018.

Note: During 2018, the City Centres Database will be consolidated based on the new GHSL baseline data release 2018 and taking into account ongoing discussions in the working group for a global definition of cities and settlements.

2.5.1 Input Data

The input data are the multi-temporal grids of the GHSL Data Package 2016 (GHS P2016): GHS-BUILT, GHS-POP and GHS-SMOD. The source of the GIS country layer is the database of Global Administrative Areas (GADM)[10] version 2.8 (November 2015).

[9] http://gadm.org/
[10] http://gadm.org/

2.5.2 Technical Details

Author: A.J. Florczyk, M. Melchiorri, M. Pesaresi, S. Freire, C. Corban, T. Kemper, M. Schiavina; Joint Research Centre (JRC) European Commission.
Contributors: Digital Observatory for Protected Areas (DOPA), Joint Research Centre (JRC) European Commission; Organisation for Economic Co-operation and Development (OECD); German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Germany; Columbia University, Center for International Earth Science Information Network - CIESIN.

Product name: GHS_STAT_UCDB2015_GLOBE_R2017A

Spatial extent: Global

Temporal extent (*): 1975-1990-2000-2015

File Name: GHS_STAT_UCDB2015_GLOBE_R2017A_V1_0.xlsx

File Format: XLS

(*) It depends on attribute: see Metadata Sheet of the main file.

This dataset is formatted as an XLS file, containing:

i. description sheet – general description of the data
ii. data sheet – data table with multi-temporal attributes
iii. metadata – description of the attributes

References

Chen J., Chen J., Liao A., Cao X., Chen L., Chen X., He C., Han G., Peng S., Lu M., Zhang W., Tong X., Mills J., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, pp. 7-27.

Corbane C., Pesaresi M., Politis P., Syrris V., Florczyk J. A., Soille P., Maffenini L., Burger A., Vasilev V., Rodriguez D., Sabo F., Dijkstra L. and Kemper T. , 2017, "Big Earth Data Analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping", Big Earth Data, Volume 1, issue 1, doi: 10.1080/20964471.2017.1397899.

Esch T., Marconcini M., Felbier A., Roth A., Heldens W., Huber M., Schwinger M., Taubenböck H., Müller A., Dech S., 2013. Urban Footprint Processor – Fully Automated Processing Chain Generating Settlement Masks from Global Data of the TanDEM-X Mission. IEEE Geoscience and Remote Sensing Letters, 10(6), pp. 1617-1621.

Freire S, MacManus K., Pesaresi M, Doxsey-Whitfield E., Mills J., 2016. Development of new open and free multi-temporal global population grids at 250 m resolution. Proceedings of the 19th AGILE Conference on Geographic Information Science. Helsinki, Finland, June 14-17, 2016.

Pesaresi, M., Ehrlich, D., Florczyk, A.J., Freire, S., Julea, A., Kemper, T., Soille, P. and V. Syrris (2015). GHS built-up grid, derived from Landsat, multi-temporal (1975, 1990, 2000, 2014). European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmt_globe_r2015b

Pesaresi, M., D. Ehrlich, S. Ferri, A. Florczyk, Manuel Carneiro Freire Sergio, S. Halkia, Andreea Julea, T. Kemper, Pierre Soille, and Vasileios Syrris (2016a). *Operating Procedure for the Production of the Global Human Settlement Layer from Landsat Data of the Epochs 1975, 1990, 2000, and 2014*. Publications Office of the European Union. http://publications.jrc.ec.europa.eu/repository/handle/111111111/40182

Pesaresi, M., Syrris, V. and A. Julea (2016b). "A New Method for Earth Observation Data Analytics Based on Symbolic Machine Learning." *Remote Sensing, 8*(5), 399; doi:10.3390/rs8050399

Syrris, V., Corbane, C., & Soille, P. (2017). A global mosaic from copernicus Sentinel-1 data. In *Proc. of the 2016 Conference on Big Data from Space (BiDS'17)*. Vol. in press. Toulouse: European Union.

Other useful Resources

GHSL project

http://ghsl.jrc.ec.europa.eu

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

On the phone or by email

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

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: http://europa.eu/contact

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

KJ-NA-29466-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/777868
ISBN 978-92-79-97547-9