



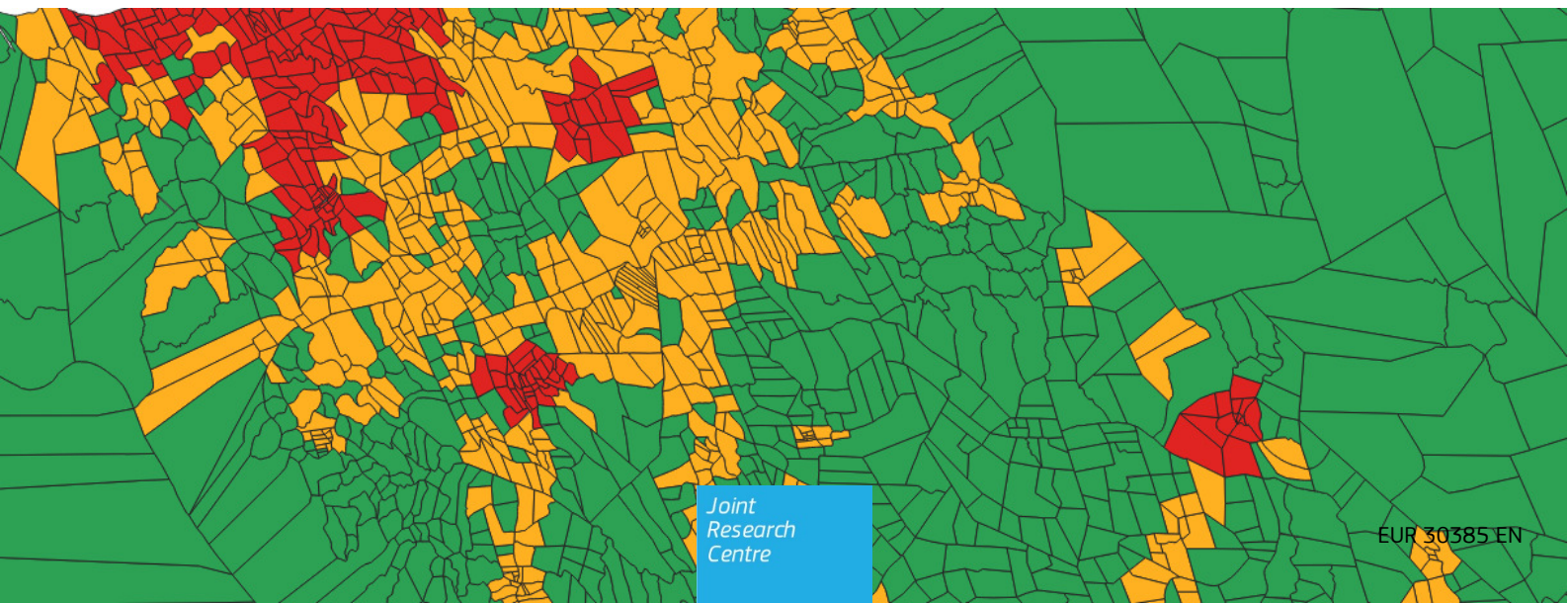
JRC SCIENTIFIC INFORMATION SYSTEMS AND DATABASES REPORT

GHS-DU-TUC User Guide

*Degree of Urbanisation
Territorial Units Classifier
User Guide
Version 1*

Maffenini, L.; Schiavina, M.; Melchiorri, M.;
Pesaresi, M.; Kemper T.

2020



This publication is a Scientific Information Systems and Databases report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information

Name: Thomas Kemper
Address: Via Fermi, 2749 21027 ISPRA (VA) - Italy - TP 267
European Commission, Joint Research Centre (JRC), Ispra
Email: Thomas.kemper@ec.europa.eu
Tel.: +39 0332 785576

EU Science Hub

<https://ec.europa.eu/jrc>

JRC121486

EUR 30385 EN

PDF	ISBN 978-92-76-22421-1	ISSN 1831-9424	doi:10.2760/633624
Print	ISBN 978-92-76-22422-8	ISSN 1018-5593	doi:10.2760/638429

Luxembourg: Publications Office of the European Union, 2020

© European Union, 2020



The reuse policy of the European Commission is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

All content © European Union, 2020

How to cite this report: Maffeni, L., Schiavina, M., Melchiorri, M., Pesaresi, M. and Kemper, T., GHS-DU-TUC User Guide, EUR 30385 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-22421-1, doi:10.2760/633624, JRC121486

Contents

Abstract.....	2
1 Introduction.....	3
1.1 Degree of Urbanisation.....	4
1.1.1 Step 1: grid classification Level 1.....	4
1.1.2 Step 1: grid classification Level 2.....	4
1.1.3 Step 2: classifying local units.....	5
1.1.3.1 Territorial units classification Level 1.....	5
1.1.3.2 Territorial units classification Level 2.....	6
1.1.3.3 Operational classifier workflow.....	7
1.2 A consistent nomenclature for the Degree of Urbanisation.....	7
2 Standalone tool.....	9
2.1 Installation.....	9
2.1.1 System requirements.....	9
2.1.2 Installation procedure.....	9
2.1.2.1 Install MATLAB Runtime.....	9
2.1.2.2 Install GHS-DU-TUC.....	9
2.1.2.3 Remove GHS-DU-TUC.....	10
2.2 The graphical user interface.....	11
2.2.1 Menu bar.....	12
2.2.1.1 File menu.....	12
2.2.1.2 Input.....	12
2.2.1.3 Output.....	12
2.2.1.4 Options.....	13
2.2.1.5 Help.....	13
2.2.2 Buttons.....	14
2.2.3 Editable text fields.....	14
2.2.4 Guided run.....	15
3 ArcGIS toolbox.....	20
3.1 Installation.....	20
3.1.1 System requirements.....	20
3.1.2 Installation procedure.....	20
3.1.2.1 Install MATLAB Runtime.....	20
3.1.2.2 Install GHS-DU-TUC ArcGIS toolbox.....	20
3.1.2.3 Remove GHS-DU-TUC ArcGIS toolbox.....	21
3.2 The toolbox user interface.....	21
3.2.1 Run the toolbox.....	23
4 Input and output.....	25

4.1	Input.....	25
4.1.1	Territorial Units shapefile	25
4.1.2	Population grid	25
4.1.3	Degree of Urbanisation grid.....	26
4.2	Output.....	27
4.2.1	Classified shapefile	27
4.2.2	Statistics table.....	27
5	Conclusion.....	29
	References.....	30
	List of boxes.....	31
	List of figures	32
	List of tables	33
	Annexes	34
	Annex 1. Error messages.....	34

Authors

Maffenini, Luca^b; Schiavina, Marcello^a; Melchiorri, Michele^c; Pesaresi, Martino^a; Kemper Thomas^a

^a European Commission-Joint Research Centre, Ispra, Italy

^b UniSystems Luxembourg SàRL

^c Engineering S.p.a

The following institutions provided useful reviews to improve the usability of the tool: Statistics South Africa (Policy Research & Analysis), EUROSTAT Unit E-4 (Regional statistics and geographical information); University of Southampton (WorldPOP)

Abstract

The Degree of Urbanisation Territorial Units Classifier (GHS-DU-TUC) Tool (– version 1) is an information system developed in the framework of the Global Human Settlement Layer (GHSL) to produce the classification of territorial units based on the Degree of Urbanisation and extract related statistics. The tool classifies territorial units by Degree of Urbanisation at Level 1 (3 classes) and Level 2 (7 classes) based on population majority by settlement classes derived from the “*Degree of Urbanisation*” method and ported to the GHSL environment through the *GHSL Settlement Model (GHSL SMOD)*.

The GHS-DU-TUC v1 is designed as an operational tool to perform the second step required to apply the Degree of Urbanisation released as standalone tool and as ArcGIS Toolbox. Once the first step produces the settlement classification grid (i.e. with the GHS-DUG Tool), the user runs the GHS-DU-TUC that requires this settlement classification grid, the population grid used to produce the settlement classification grid (i.e. produced with the GHS-POP2G Tool) and a geometry of territorial units to be classified by Degree of Urbanisation. This tool is conceptualised to be deployed after the application of the GHSL tools GHS-POP2G and GHS-DUG but it accepts in input population grids produced by means of any other procedure respecting the described constrains.

This document contains the description of the GHS-DU-TUC Tool use, the rationale for the second step to apply the Degree of Urbanisation (the classification of territorial units) and the comprehensive description of the outputs. The tool is a capacity enhancement asset in the framework of the multi-stakeholder effort for the uptake of the Degree of Urbanisation, the people-based harmonised definition of cities and settlements recommended by the 51st Session of the United Nations Statistical Commission as the method to delineate cities and rural areas for international statistical comparison.

The GHS-DU-TUC, as all GHSL Tools, is issued with an end-user licence agreement, included in the download package.

1 Introduction

The Global Monitoring Framework of the 2030 Agenda for Sustainable Development includes several indicators that require disaggregation in urban and rural classes, and several others that are sensitive to how an urban area is delineated. The broad array of different criteria applied in national definitions of rural-urban areas poses serious challenges to cross-country comparisons (ILO 2018). Both the Action Framework of the Implementation of the New Urban Agenda (UN-Habitat 2017) and the Global Strategy to improve Agricultural and Rural Statistics (GSARS 2018) highlight the need for a harmonised method to facilitate international comparisons and to improve the quality of rural and urban statistics in support of national policies and investment decisions.

Under these circumstances the European Union, The Food and Agriculture Organization of the United Nations (FAO), the International Labour Office (ILO), the Organisation for Economic Co-operation and Development (OECD), UN-Habitat and the World Bank have joined forces to develop a new method to delineate cities, urban and rural areas in a harmonised way. This work was launched at the Habitat III conference in 2016 with the explicit aim to organise global consultations and present the new method to the UN Statistical Commission for endorsement. The 51st Session of the United Nations Statistical Commission recommended this people-based harmonised definition of cities and settlements called “*Degree of Urbanisation*” as the method to delineate cities and rural areas for international statistical comparison.

The goal is to facilitate international statistical comparisons of the performance of cities, urban and rural areas across a selection of global indicators. This method is meant to complement and not replace existing national definitions. National definitions typically rely on a much wider set of indicators and can be adjusted to take into account specific national characteristics. This richness makes national definitions more suitable for national policies, but less suitable for international or global comparisons.

The “*Degree of Urbanisation*” (DEGURBA) is a statistical classification method originally introduced by the European Commission, Directorate of Regional and Urban Policy and successively refined by the joint work of the EU, the Food and Agriculture Organization of the United Nations (FAO), the International Labour Office (ILO), the Organisation for Economic Co-operation and Development (OECD), UN-Habitat and the World Bank. The DEGURBA aims to identify the spatial extents of “Urban Centres”, “Urban Clusters” and “Rural Areas” based on resident population density grids at a scale (spatial detail or resolution) of 1 square kilometre (Dijkstra and Poelman, 2014). Following the methodology indicated by the statistical office of the European Union (EUROSTAT), the population grids classified by the DEGURBA method can be used for statistical characterisation of local administrative units¹.

The *GHSL Settlement Model* (GHSL SMOD) is the porting of the DEGURBA in the Global Human Settlement Layer (GHSL) framework developed by the European Commission, Joint Research Centre². The GHSL SMOD supports the international multi-stakeholder discussion on the DEGURBA operationalization parameters and on the DEGURBA derived metrics and indicators using the GHSL baseline information as common global data frame (European Commission, Joint Research Centre, 2018) (Melchiorri et al., 2018) (Corbane et al., 2018) (Melchiorri et al., 2019) (Schiavina et al., 2019).

The Global Human Settlement “*Degree of Urbanisation Territorial Units Classifier*” GHS-DU-TUC 1 described in this report is released as a standalone free software tool and as ArcGIS Toolbox that operationalise the EUROSTAT methodology to classify local typologies by Degree of Urbanisation. The classification of territorial units based on population majority rules and the Degree of Urbanisation Settlement grid. The tool is integrated in the family of GHSL tools that allow the user to apply the Degree of Urbanisation to an area of interest providing as input only the geocoded population data (and possible covariates), and the geometry of the territorial units. The synergy among the GHS-POP2G tool (that outputs the population grid), and the GHS-DUG tool (that taking the population grid as input outputs the settlement classification grid) makes it possible with a streamlined workflow that is computationally sustainable to apply the Degree of Urbanisation.

1 <https://ec.europa.eu/eurostat/web/degree-of-urbanisation/background>

2 <https://ghsl.jrc.ec.europa.eu/>

1.1 Degree of Urbanisation

The Degree of Urbanisation classifies the entire territory of a country along the urban-rural continuum. It combines population size and population density thresholds to capture the full settlement hierarchy. **It is applied in a two-step process:** First, 1 km² grid cells are classified based on population density, contiguity and population size (such step can be performed using the GHS-DUG tool available at <https://ghsl.jrc.ec.europa.eu/tools.php>). Subsequently, local units are classified based on the type of grid cells their population resides in. This method works best with small administrative or statistical units, such as municipalities or census enumeration areas. Additional information on cluster types and local typologies is contained in the EUROSTAT Methodological Manual on Territorial Typologies (European Commission and Statistical Office of the European Union, 2019).

The Degree of Urbanisation has two hierarchical levels. The Degree of Urbanisation level 1 uses three classes, and the Level 2 uses 7 classes.

1.1.1 Step 1: grid classification Level 1

The Degree of Urbanisation Level 1, as operationalised in the GHS-DUG tool, version 3.2, identifies three types of grid cells (Figure 1):

- An urban centre consists of contiguous grid cells (4-connectivity cluster) with a density of at least 1,500 inhabitants per km² of permanent land, and has at least 50,000 inhabitants in the cluster with smoothed boundaries (3-by-3 majority filtering) and <15 km² holes filled³.
- An urban cluster (or moderate density clusters) consists of contiguous grid cells (8-connectivity cluster) with a density of at least 300 inhabitants per km² of permanent land and has a population of at least 5,000 in the cluster. (The urban centres are subsets of the corresponding urban clusters).
- Rural grid cells (mostly low density cells) are cells that do not belong to an urban cluster. Most of these will have a density below 300 inhabitants per km². Some rural cells will have a higher density, but they are not part of cluster with a large enough population size to be classified as an urban cluster.

1.1.2 Step 1: grid classification Level 2

The Degree of Urbanisation level 2 is a sub-classification level 1 (Figure 3). It was created to identify medium and small settlements, i.e. towns and villages.

An urban centre is identified in the identical manner as in the Degree of Urbanisation level 1:

- An urban centre consists of contiguous grid cells (4-connectivity cluster) with a density of at least 1,500 inhabitants per km² of permanent land, and has at least 50,000 inhabitants in the cluster with smoothed boundaries (3-by-3 majority filtering) and <15 km² holes filled³.

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

- A dense urban cluster consists of contiguous cells (4-connectivity cluster) with a density of at least 1,500 inhabitants per km² of permanent land, with a population of at least 5,000 and less than 50,000 in the cluster.
- A semi-dense urban cluster consist of contiguous grid cells (8-connectivity cluster) with a density of at least 300 inhabitants per km² of permanent land and has a population of at least 5,000 (i.e. an urban cluster) and this cluster is neither contiguous with nor within 2 km of a dense urban cluster or an urban centre⁴.
- Suburban or peri-urban cells are the remaining urban cluster cells i.e. those not part of dense or semi-dense cluster. These cells are part of an urban cluster that is contiguous or within 2 km of a dense urban cluster or an urban centre.

³ In a few countries with relatively low-density urban development and a strong separation of land use functions, the Degree of Urbanisation generates multiple urban centres for a single city. Creating urban centres using both cells with a density of at least 1,500 inhabitants and cells that are at least 50% built-up on permanent land resolves this issue. Such highly built-up cells typically contain office parks, shopping malls, factories and transport infrastructure.

⁴ Measured as outside a buffer of three grid cells of 1 km² around dense urban clusters and urban centres.

Rural grid cells can be categorised into three types.

- A rural cluster consists of contiguous cells (8-connectivity cluster) with a density of at least 300 inhabitants per km² of permanent land and a population between 500 and 5,000 in the cluster.
- Low density rural grid cells are rural grid cells with a density of at least 50 inhabitants per km² of permanent land and are not part of a rural cluster.
- Very low density rural grid cells are cells with a density of less than 50 inhabitants per km² of permanent land.

1.1.3 Step 2: classifying local units

The Degree of Urbanisation classifies local units based on population majority applied to the grid level classification. Each local unit belongs exclusively to one class at Level 1 and one at Level 2 (hierarchy based). Local units can be administrative units - such as municipalities - or statistical units - such as census enumeration areas.

1.1.3.1 Territorial units classification Level 1

Once all grid cells have been classified as urban centres, urban clusters and rural grid cells with the GHS-DUG Tool, the next step concerns overlaying these results onto local units, as follows (Figure 2):

- **Cities (or densely populated areas):** local units that have at least 50% of their population in urban centres, **code 3**.
- **Towns and semi-dense areas (or intermediate density areas):** local units that have less than 50% of their population in urban centres and less than 50% of their population in rural grid cells, **code 2**.
- **Rural areas (or thinly populated areas):** local units that have at least 50% of their population in rural grid cells, **code 1**.

Urban areas consist of cities plus towns and semi-dense areas.

Figure 1 Urban centre, urban cluster and rural grid cells around Cape Town, South Africa

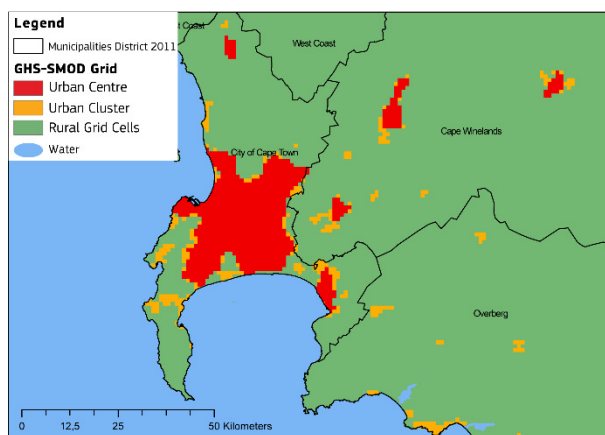
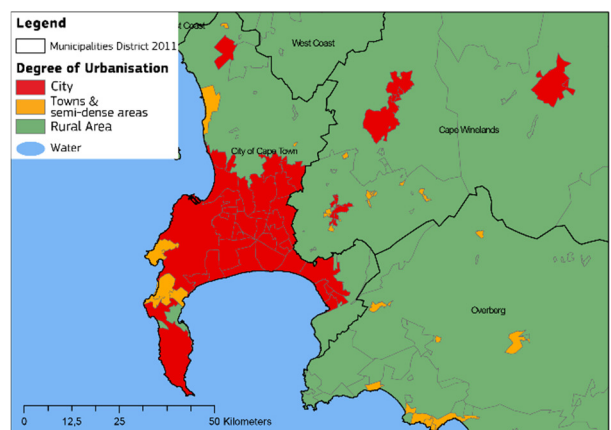


Figure 2 City, towns & semi-dense areas and rural areas around Cape Town, South Africa (classification of Main Places units, note that Cape Peninsula is part of Cape Town Main Place)



1.1.3.2 Territorial units classification Level 2

Local units are classified as cities in the identical manner as in Degree of Urbanisation level 1 (Figure 4):

- A city consists of one or more local units that have at least 50% of their population in an urban centre, **code 30**.

Local units classified as “towns and semi-dense areas” can be divided into three classes:

- Dense Towns have a larger share of their population in dense urban clusters than in semi-dense urban clusters (i.e. it is dense) and a larger share in dense plus semi-dense urban clusters than in suburban or peri-urban cells (i.e. it is a town), **code 23**.
- Semi-dense Towns have a larger population share in semi-dense urban clusters than in dense urban clusters (i.e. it is semi-dense) and a larger share in dense plus semi-dense urban clusters than in suburban or peri-urban cells (i.e. it is a town), **code 22**.
- Suburban or peri-urban areas have a larger population share in suburban or peri-urban cells than in dense plus semi-dense urban clusters, **code 21**.

Dense and semi-dense towns can be combined into towns. This reduces the number of classes and maybe useful especially if the population share in semi-dense towns is low.

Local units classified as “rural areas” can be divided into three classes:

- Villages have the largest share of their rural grid cell population living in a rural cluster, **code 13**.
- Dispersed rural areas have the largest share of their rural grid cell population living in low density rural grid cells, **code 12**.
- Mostly uninhabited areas have the largest share of their rural grid cell population living in very low density rural grid cells, **code 11**.

Figure 3 Degree of urbanisation level 2 grid classification around Toulouse, France

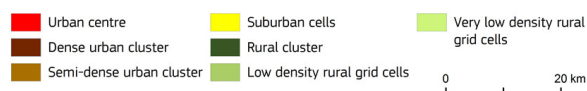
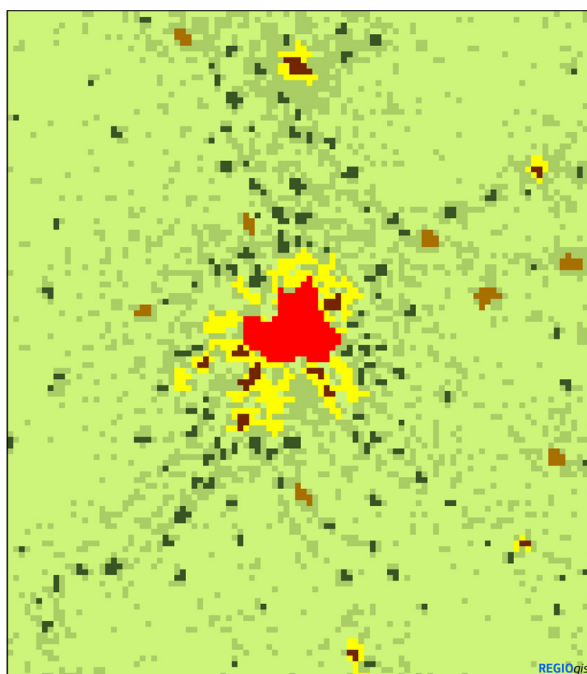
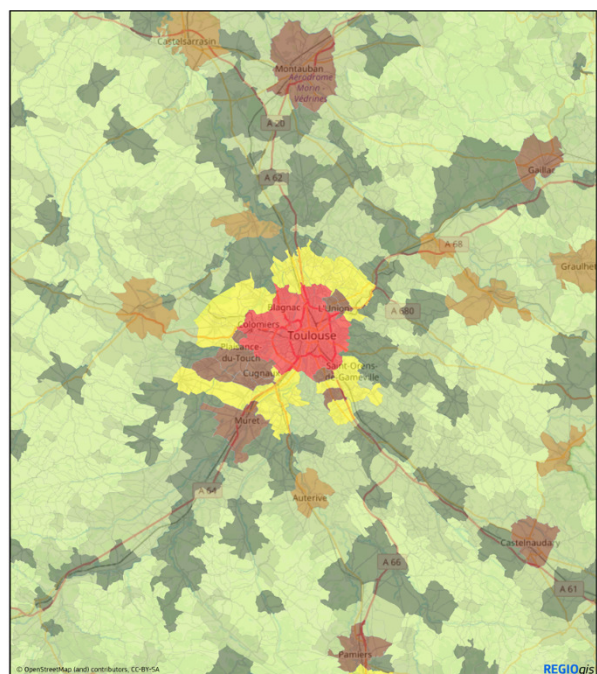


Figure 4 Degree of urbanisation level 2 local unit classification around Toulouse, France



1.1.3.3 Operational classifier workflow

The classifier combines three inputs of different nature:

- two raster layers (potentially at different resolutions)
 - a settlement classification grid at 1 km;
 - a population grid at the same or higher resolution than the settlement classification grid;
- one vector layer of territorial units.

The working projection selected is the population grid projection to avoid warping of population grid which usually introduce undesirable distortions (i.e. total population volume variations) or requires huge computational resources (e.g. volume preserving warping).

The procedure begins with the rasterization of the vector layer. As suggested, the unit classification works better with small units, therefore the rasterization is performed by targeting a resolution of 50 m, snapped to the population grid, to reduce the number of units that won't have a representation in the raster layer. The settlement classification grid (generated at 1 km resolution, by definition) is oversampled and, if necessary, warped to the population grid projection using nearest neighbour algorithm to align with the 50-m territorial units' raster. The warping procedure is usually not required as the population layer should be the same (or at the native resolution of the layer) used to produce the settlement classification grid (e.g using the GHS-DUG tool). Using different population grid could produce meaningless results. The population grid (if not provided at 50m) is oversampled using nearest neighbour algorithm then adjusted by dividing all cell with the oversampling ratio (e.g. from 250 m to 50 m the ratio is 25, as 25 grid cells of 50 m compose each 250 m cell) assuming a uniform distribution of population within each cell.

Once the pre-processing of the layers is completed, the algorithm computes for each unit the share of population in each class of the settlement classification grid (both at Level 1 and 2), through zonal statistics, and assigns the class of the unit accordingly (i.e. following the classification rules described in sections 1.1.3.1 and 1.1.3.2).

Even if the working resolution is set at 50 m, it could happen that some small polygons could not be rasterised due to geospatial constrains. In such cases the algorithm evaluates the class of these polygons by running separately the zonal statistics (i.e. one polygon at a time) and performing the rasterization procedure with "all touching cells" option. To avoid population double counting no population is assigned to such polygons.

1.2 A consistent nomenclature for the Degree of Urbanisation

Two sets of terms have been developed to describe each of the classes of the Degree of Urbanisation. The first set uses simple and short terms such as city, town, suburb and village. The second set uses a more neutral and technical language. The second set can be helpful to avoid overlap with the terms used in the national definition.

Table 1 Short and technical terms for Degree of Urbanisation level 1 and 2 for the local unit classification

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

Table 2 Short and technical terms for Degree of Urbanisation level 1 and 2 for grid cell classification

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

2 Standalone tool

The standalone version of the tool allows the user to perform the classification using a simple graphical user interface.

2.1 Installation

2.1.1 System requirements

The Global Human Settlement “*Degree of Urbanisation Territorial Units Classifier*” (GHS-DU-TUC) is graphical tool developed in MATLAB and created for Windows, and its system requirements are listed below (Table 3)

Table 3 System requirements

Requirements	Operating system	Processor	RAM	Disk space
Minimum	Windows 7	Any Intel or AMD x86-64 processor	6 GB	150 Mb (+ 700 Mb for Matlab Runtime)
Recommended	Windows 10	Any Intel or AMD x86-64 processor	16 GB	150 Mb (+ 700 Mb for Matlab Runtime)

GHS-DU-TUC operates on vector files and raster grids: the larger the grid the higher the required memory. Few examples of RAM requirements in the most computational expensive scenario are below in Table 4. The input raster size (in number of 1 km grid cells) is intended per each of the raster layer(s).

Table 4 Requirements by input raster size

Input raster size	RAM needed	Disk space for outputs
1000x1000 pixels	6 GB	< 10 MB
5000x5000 pixels	12 GB	< 50 MB

2.1.2 Installation procedure

2.1.2.1 Install MATLAB Runtime

GHS-DU-TUC is developed in MATLAB and the MATLAB Runtime is required to run the tool.

The specific version to be installed is **R2018b (9.5)**, Windows 64-bit version.

Check for the official MATLAB website to download and install:

<https://it.mathworks.com/products/compiler/matlab-runtime.html>

2.1.2.2 Install GHS-DU-TUC

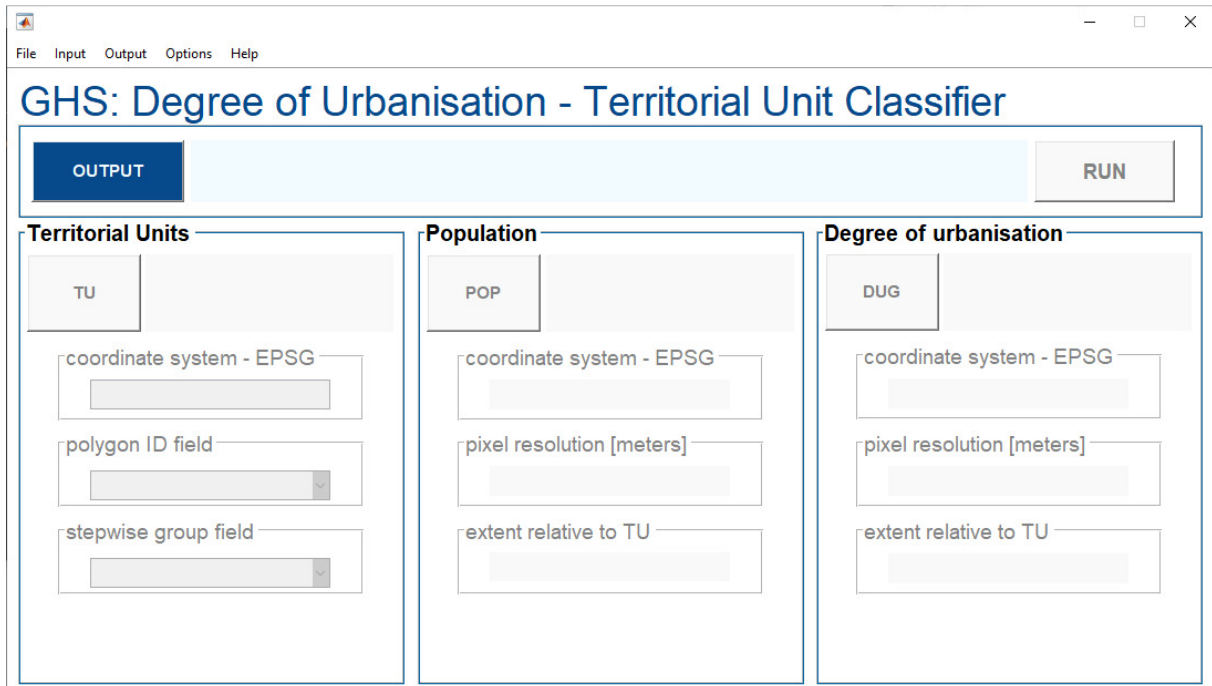
Once the MATLAB Runtime is installed, if you have not done so already, download the GHS-DU-TUC tool from the GHSL website (Tools section):

<https://ghsl.jrc.ec.europa.eu/tools.php>

Unzip the GHS-DU-TUC zip file and run the executable directly by double click the *exe* file.

The first time GHS-DU-TUC is launched on a new computer it performs several checks and initialization steps, therefore it can take up to 2 minutes to start.

Figure 5 First startup GUI



2.1.2.3 Remove GHS-DU-TUC

To uninstall GHS-DU-TUC simply delete the entire folder containing the *.exe* file.

MATLAB Runtime can also be removed if not necessary for other applications.

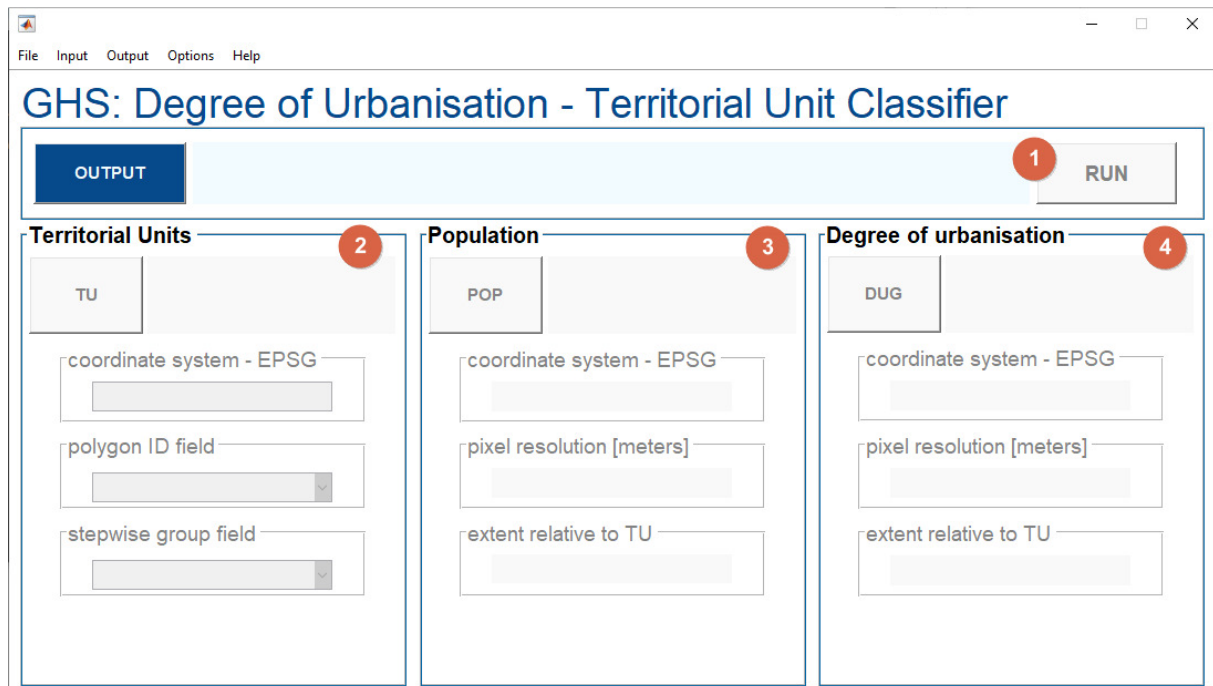
2.2 The graphical user interface

The GHS-DU-TUC graphical user interface is created in MATLAB and is distributed as a standalone application.

The main window is divided in three areas:

1. Top panel, which contains the buttons to select the output folder and to start the process;
2. Left panel, which allow to select the territorial units input layer to be classified;
3. Middle panel, where the population raster grid is selected;
4. Right panel, where the Degree of urbanisation raster grid is selected

Figure 6 GHS-DU-TUC main windows and areas



After the start, the user has to first choose the output folder, as it is used to store results and temporary files as well. Then all other inputs are activate and can be selected.

2.2.1 Menu bar

The menu bar is organized as in five sections:

1. File
2. Input
3. Output
4. Options
5. Help

2.2.1.1 File menu

Table 5 File menu bar

Item	Action	Shortcut
Run	Starts the territorial units classification process	CTRL + R
Clear all	Clear all inputs	
Quit	Exit GHS-DU-TUC and close the main window	CTRL + Q

2.2.1.2 Input

Table 6 Input menu bar

Item	Action	Shortcut
Territorial Units	Select and load the Territorial Units to classify	CTRL + T
Population Grid	Select and load the Population raster	CTRL + P
Degree of Urbanisation Grid	Select and load the Degree of Urbanisation raster	CTRL + D

2.2.1.3 Output

Table 7 Output menu bar

Item	Action	Shortcut
Output	Set the output folder	CTRL + O

2.2.1.4 Options

Table 8 Options menu bar

Item	Action	Shortcut
Open output folder when process completes	The output folder is automatically opened in windows explorer when the process is completed (default: activated)	
Multicore	Use all available machine core (default: single core)	

2.2.1.5 Help

Table 9 Help menu bar

Item	Action	Shortcut
Online help	Open a web page to the DUG documentation website using the default browser	CTRL + H
About GHS-DU-TUC	Open a new window containing GHS-DU-TUC version number, release date and contact information	

2.2.2 Buttons

All panels contain buttons some of them provide the same functionality of the single menu entries and are an alternative way to set input data and start the classification process.

Table 10 Buttons

Item	Action	Shortcut
OUTPUT	Set the output folder	CTRL + O
RUN	Starts the GHSL SMOD computation	CTRL + R
TU	Select and load the Territorial Units to classify	CTRL + T
polygon ID field	Select the shapefile field to be used as index in the objects rasterization.	
Stepwise group field	Select the shapefile field to be used to split the rasterization process in smaller steps (useful with low memory machine). . When not used all layer is processed at once. By default is the same value used for polygon ID field	
POP	Select and load the Population raster	CTRL + P
DUG	Select and load the Degree of Urbanisation raster	CTRL + D

2.2.3 Editable text fields

The Territorial Units panel contain one editable text field that show the EPSG code of the shapefile loaded by the user. If the EPSG code cannot be automatically read by the tool the user can write its values in this field.

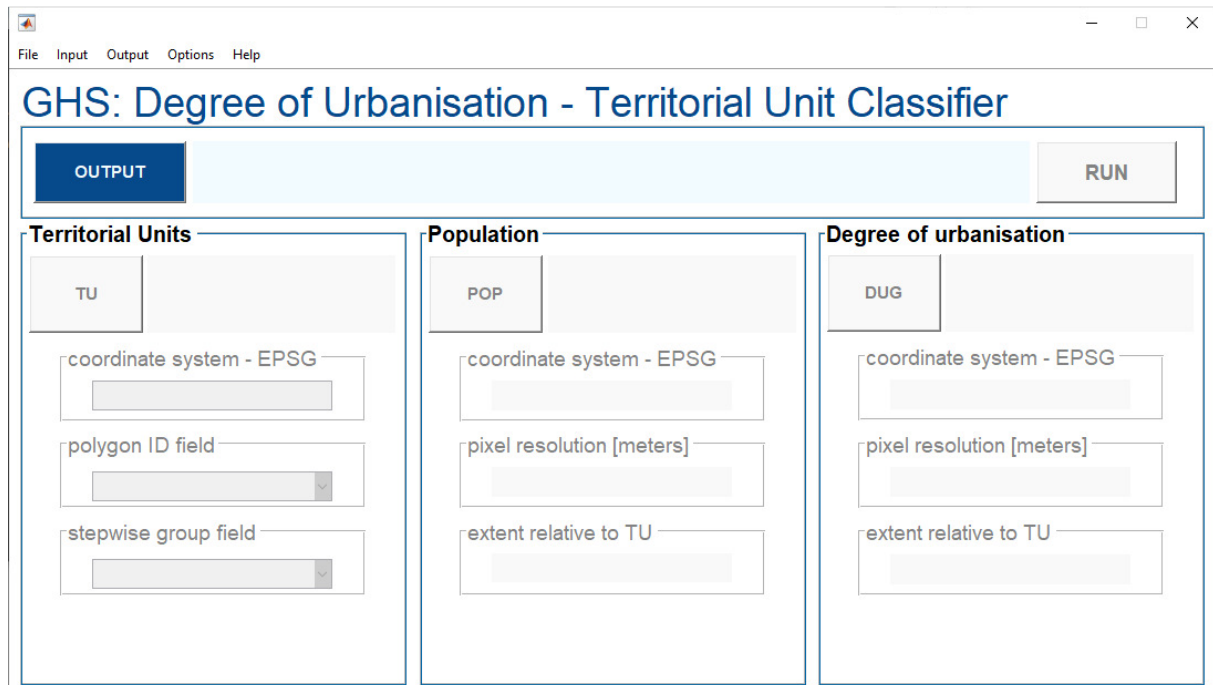
Table 11 Editable text fields

Item	Action
coordinate system - EPSG	Display the territorial units' coordinate system: it can be manually changed if the value is incorrect or missing.

2.2.4 Guided run

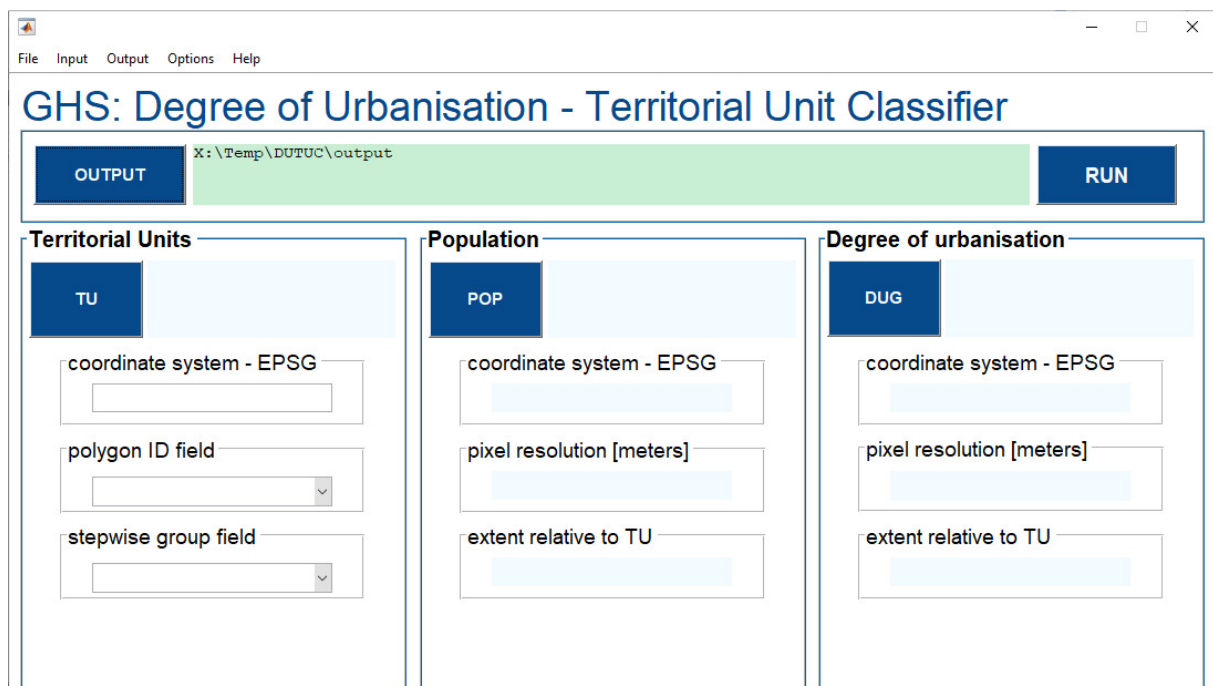
After starting GHS-DU-TUC, the only active command is the output button (Figure 7): this is to set the output folder (path) first in order to create the log file and check disk write permission for results. Below in the shaded italic text lines the actions and commands the user should prompt to run the GHS-DU-TUC Tool.

Figure 7 GHS-DU-TUC startup



The user clicks on OUTPUT and selects the output folder. (Figure 8)

Figure 8 Output folder selected



Once output is set, all other input panels become active.

The user chooses the Territorial Units shapefile. (Figure 9)

Figure 9 Territorial Units selected

The screenshot shows the 'GHS: Degree of Urbanisation - Territorial Unit Classifier' software interface. The window title is 'GHS: Degree of Urbanisation - Territorial Unit Classifier'. At the top, there is a menu bar with 'File', 'Input', 'Output', 'Options', and 'Help'. Below the menu bar, the title 'GHS: Degree of Urbanisation - Territorial Unit Classifier' is displayed. A green bar contains the text 'X:\Temp\DUTUC\output' and a blue 'RUN' button. Below this, there are three main panels: 'Territorial Units', 'Population', and 'Degree of urbanisation'. The 'Territorial Units' panel has a blue 'TU' button and a green bar with the file path 'X:\Temp\DUTUC\TERR_Italy.shp'. It contains three input fields: 'coordinate system - EPSG' with the value '3035', 'polygon ID field' with a dropdown menu showing 'OBJECTID', and 'stepwise group field' with a dropdown menu showing 'OBJECTID'. The 'Population' panel has a blue 'POP' button and contains three input fields: 'coordinate system - EPSG', 'pixel resolution [meters]', and 'extent relative to TU'. The 'Degree of urbanisation' panel has a blue 'DUG' button and contains three input fields: 'coordinate system - EPSG', 'pixel resolution [meters]', and 'extent relative to TU'.

The coordinate system box reports the EPSG value, the user can modify it if the value is missing or wrong. The user can select the polygon ID field to be used as index in the object rasterization and the stepwise group field. If the latter is set as "NOT USED" the entire shapefile is processed at once: this is faster but requires more resources, especially more RAM.

The user chooses the Population grid (Figure 10).

Figure 10 Population grid selected

The screenshot shows the 'GHS: Degree of Urbanisation - Territorial Unit Classifier' software interface. The window title is 'GHS: Degree of Urbanisation - Territorial Unit Classifier'. The menu bar includes 'File', 'Input', 'Output', 'Options', and 'Help'. The main interface is divided into three columns: 'Territorial Units', 'Population', and 'Degree of urbanisation'. At the top, there is an 'OUTPUT' field containing 'X:\Temp\DUTUC\output' and a 'RUN' button. The 'Territorial Units' column has a 'TU' field with 'X:\Temp\DUTUC\TERR_Italy.shp', a 'coordinate system - EPSG' field with '3035', a 'polygon ID field' dropdown with 'OBJECTID', and a 'stepwise group field' dropdown with 'OBJECTID'. The 'Population' column has a 'POP' field with 'X:\Temp\DUTUC\ENACT_POP_D012011_EU28_R2019A_3035_1R_V1.0.tif', a 'coordinate system - EPSG' field with '3035', a 'pixel resolution [meters]' field with '1000', and an 'extent relative to TU' field with 'OK'. The 'Degree of urbanisation' column has a 'DUG' field, a 'coordinate system - EPSG' field, a 'pixel resolution [meters]' field, and an 'extent relative to TU' field.

The coordinate system box reports the EPSG value, the pixel resolution box the raster cell size in meters.

The extent relative to TU box gives an indication about the raster geographic extent. If the Population raster bounding box is smaller than the Territorial Units' one, or simply it does not cover all polygons, then the message is "TOO SMALL / NO OVERLAP": in this case the process cannot start as some polygons would not be classified. Otherwise the message is "OK" and the user can proceed.

The user chooses the Degree of Urbanisation grid (Figure 11)

Figure 11 Degree of urbanisation grid selected

The coordinate system box reports the EPSG value, the pixel resolution box the raster cell size in meters.

The extent relative to TU box gives an indication about the raster geographic extent. If the Degree of Urbanisation raster bounding box is smaller than the Territorial Units' one, or simply it does not cover all polygons, then the message is "TOO SMALL / NO OVERLAP": in this case the process cannot start as some polygons would not be classified. Otherwise the message is "OK" and the user can proceed.

The user hit run (Figure 12)

All commands are disabled when running, the RUN button becomes orange with label RUNNING and the mouse pointer transform into wheel (Figure 12). Once the run completed everything is reactivated again.

Figure 12 GHS-DU-TUC running

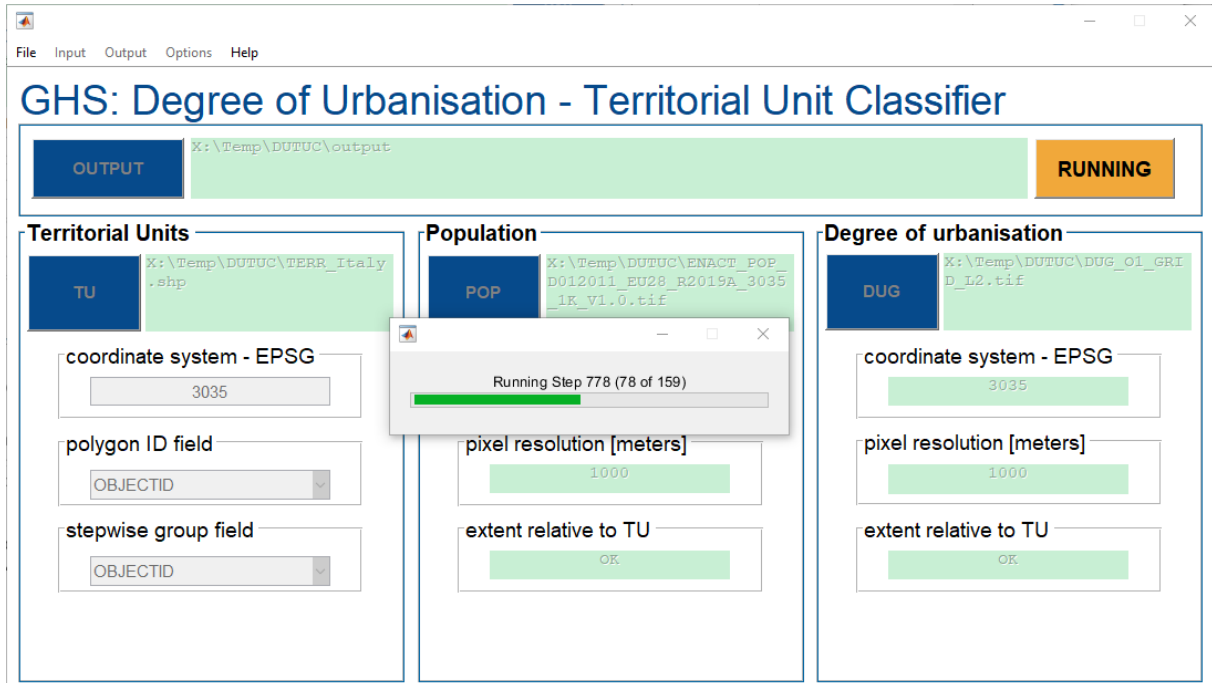
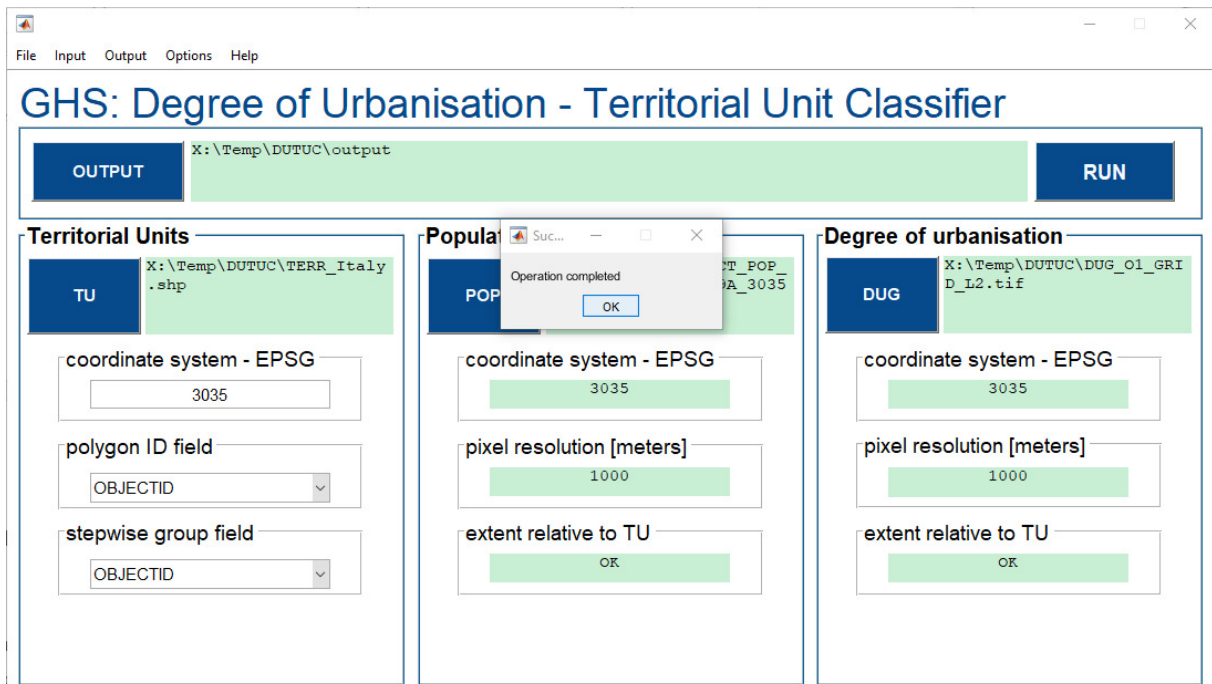


Figure 13 Run successful



3 ArcGIS toolbox

The ArcGIS toolbox version of the tool allows the user to perform the classification directly from the ArcGIS software (Esri). The main algorithm is still run through the same MATLAB code used for the standalone version to ensure equivalent results from both tools, therefore the MATLAB Runtime is required. The ArcGIS toolbox version has been developed in ArcGIS 10.6 environment and tested for back compatibility in ArcGIS 10.3.

3.1 Installation

3.1.1 System requirements

The requirements are the same as listed in 2.1.1.

3.1.2 Installation procedure

3.1.2.1 Install MATLAB Runtime

GHS-DU-TUC is developed in MATLAB and the MATLAB Runtime is required to run the tool.

The specific version to be installed is **R2018b (9.5)**, Windows 64-bit version.

Check for the official MATLAB website to download and install:

<https://it.mathworks.com/products/compiler/matlab-runtime.html>

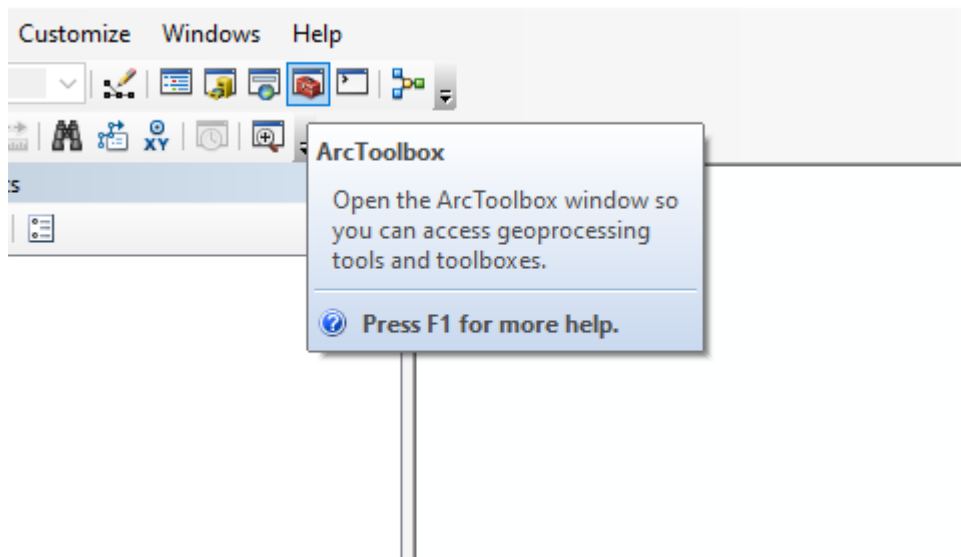
3.1.2.2 Install GHS-DU-TUC ArcGIS toolbox

Once the MATLAB Runtime is installed, if you have not done so already, download the GHS ArcGIS toolbox from the GHSL website (Tools section):

<https://ghsl.jrc.ec.europa.eu/tools.php>

Unzip the file, then open ArcGIS and from the Tools panel open the ArcToolbox window:

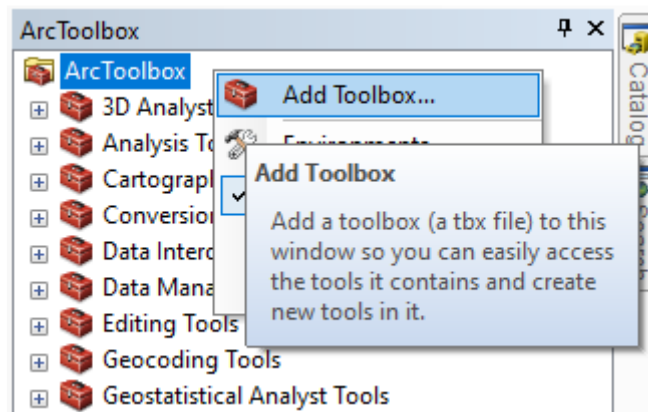
Figure 14 Show ArcToolbox window



From the ArcToolbox window right click, press Add Toolbox and select the *GHS-DU-TUC.pyt* file.

The toolbox will be listed under GHS Tools section.

Figure 15 Add toolbox to ArcGIS



3.1.2.3 Remove GHS-DU-TUC ArcGIS toolbox

To uninstall GHS ArcGIS toolbox simply remove the toolbox from ArcGIS and delete the entire folder containing the toolbox files.

MATLAB Runtime can also be removed if not necessary for other applications.

3.2 The toolbox user interface

The toolbox user interface includes the same input and parameters than the standalone version. The main windows cannot be split in sections, therefore all fields are together but disposed in a logical order. The user should start to fill them from top to bottom, but that's not mandatory.

For a detailed description of input data and their constraints check section 4.1.

Figure 16 ArcGIS toolbox window

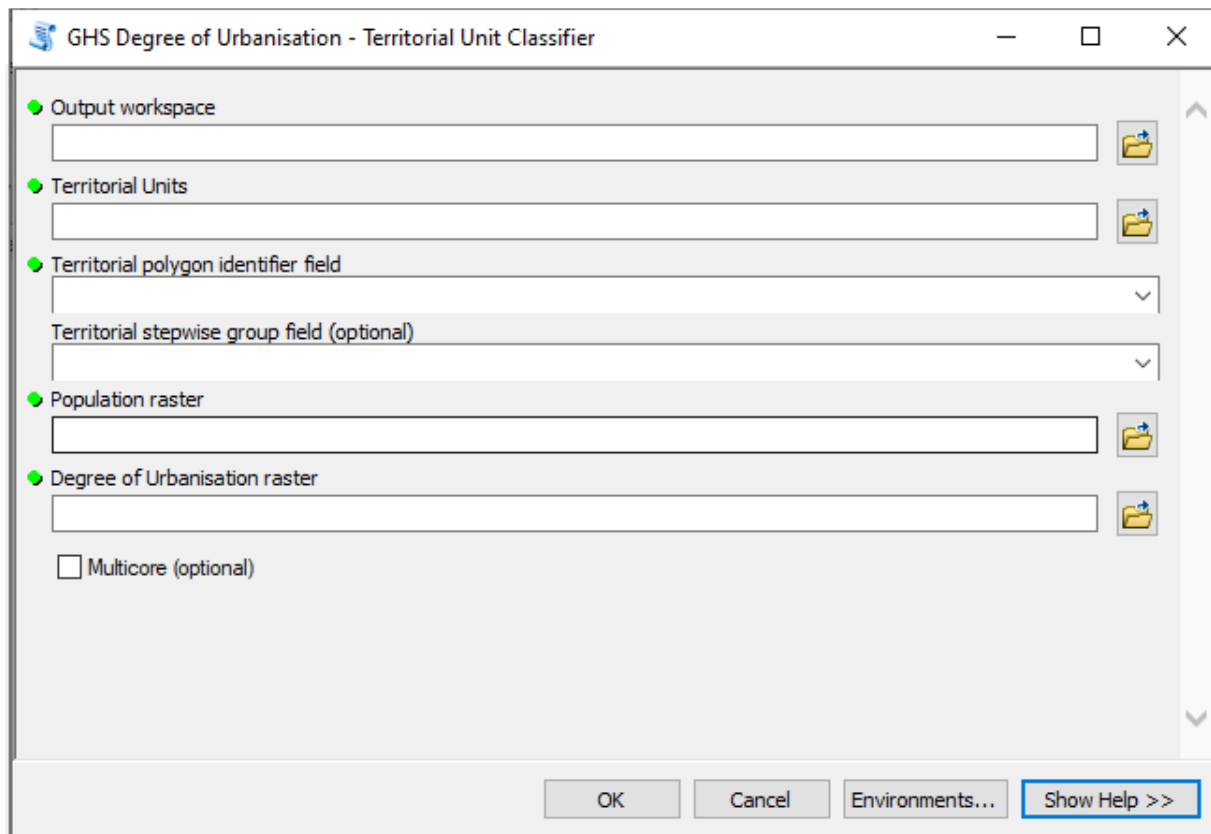


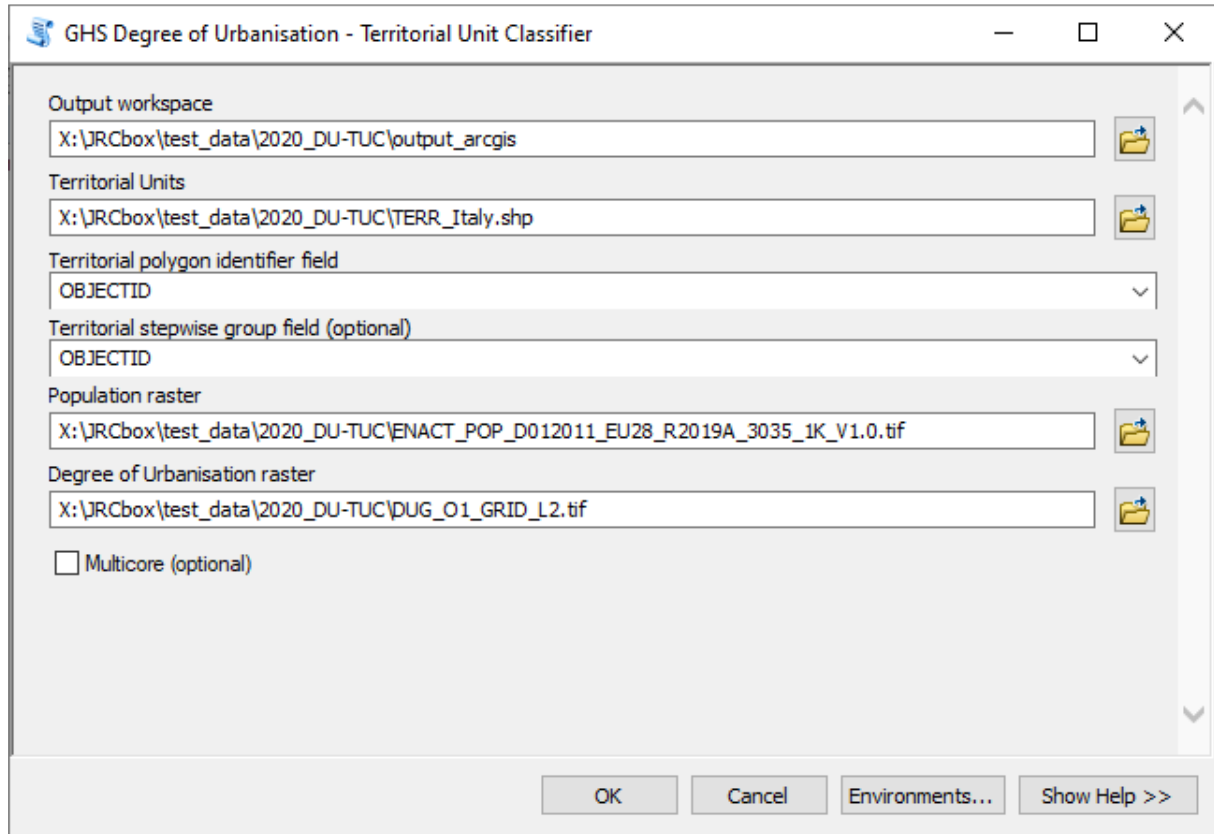
Table 12 Toolbox items description

Item	Description
Output workspace	Set the output folder. This is the output path where results and intermediate files are saved. The user needs to have write permission to this folder.
Territorial Units	Select and load the Territorial Units to classify
Territorial polygon identifier field	Select the shapefile field to be used as index in the objects rasterization.
Territorial stepwise group field	Select the shapefile field to be used to split the rasterization process in smaller steps (useful with low memory machine). . When not used all layer is processed at once. By default is the same value used for polygon ID field
Population raster	Select and load the Population raster
Degree of Urbanisation raster	Select and load the Degree of Urbanisation raster
Multicore	Use all available machine cores. By default is deactivated (single core).

3.2.1 Run the toolbox

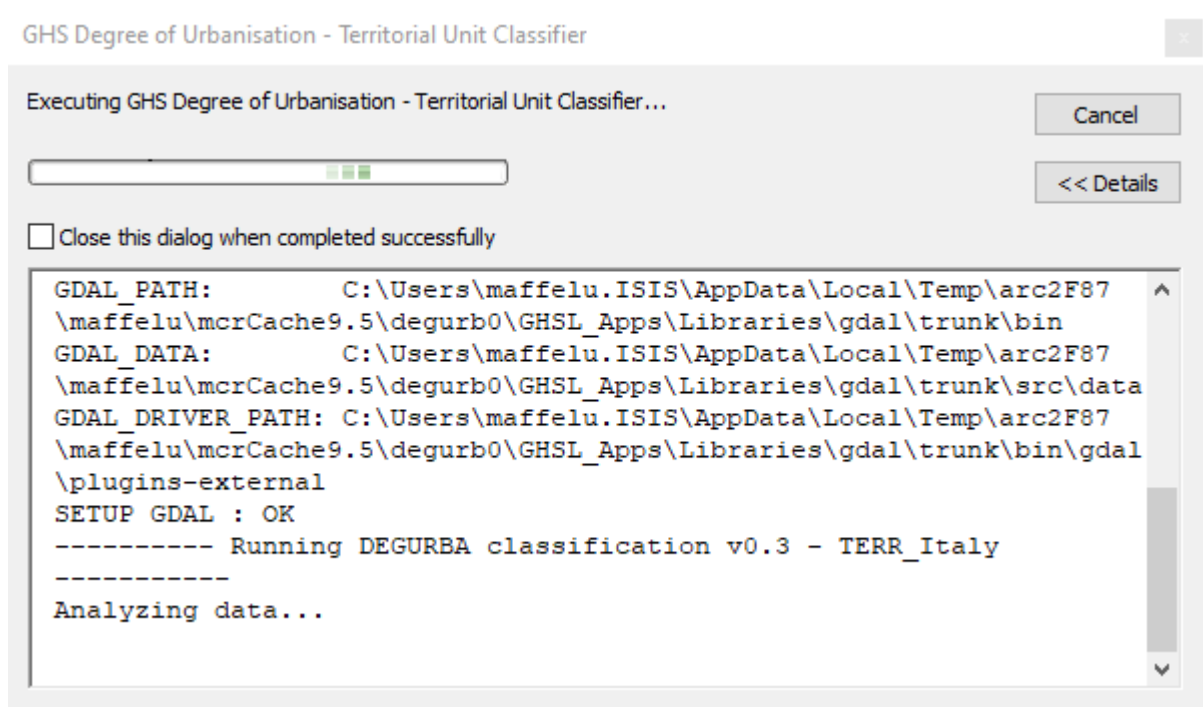
Before running the toolbox all inputs and parameters must be filled. The toolbox performs internal checks to ensure the values are correct. In case of doubt, the help section contains a short description about each field.

Figure 17 Toolbox setup with all inputs and parameters



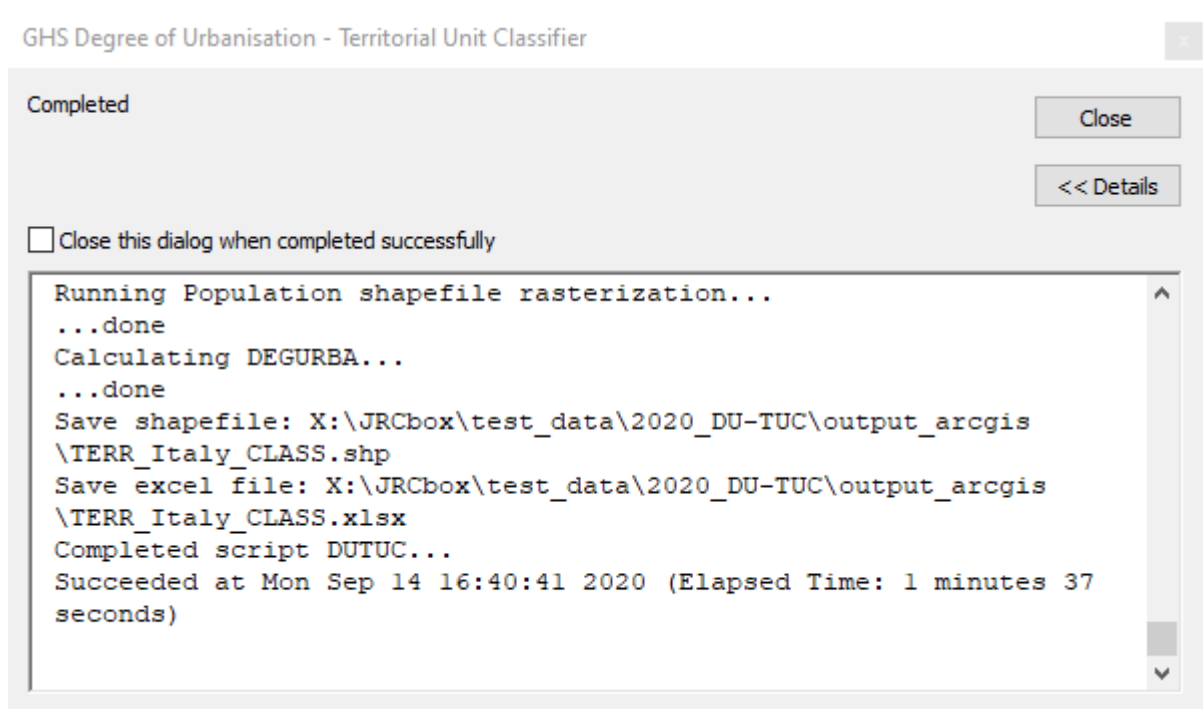
If no error messages appears while select the input data and the parameters the user can hit OK and start the computation. The user can follow the run progress by checking the computation logs.

Figure 18 Toolbox running



Once the process is completed the windows can be closed and the results can be loaded in ArcGIS or any other GIS to be inspected.

Figure 19 Toolbox run successfully completed



4 Input and output

4.1 Input

4.1.1 Territorial Units shapefile

The territorial units shapefile is the main input: it represents the polygon that will be classified. It can be in any coordinate system supported by GDAL.

Box 1 Supported projections

Check the projection supported by DUG through GDAL at:

<https://proj.org/operations/projections/index.html>

Table 13 Regions vector file requirements

File format	Data type	Map projection	Geometry
Shapefile	Any numeric (int or float)	Any supported by GDAL, it will be re-project to population's projection	Polygons

To be able to classify all polygons the layer extent must be fully contained inside both population and degree of urbanisation grid.

4.1.2 Population grid

The population grid has to be a GeoTIFF file, with integer or float data type, in World Mollweide (EPSG:54009) or LAEA Europe (EPSG:3035) map projection. The pixel resolution (i.e. cell size) has to be a multiple of 50 meters (e.g. 100, 250 or 1000, Table 14). Each pixel shall contain the population count expressed as absolute number (e.g. "3125.9" means 3125.9 people).

The population input for the GHS-DU-TUC must be the same population grid used as input to generate the settlement classification grid (i.e. 1-km-resolution population grid input to the GHS-DUG Tool). **If available, higher resolution versions (e.g. 250 m, 100 m or 50 m) of the same population grid are highly recommended (e.g. population grid produced at 250 m and aggregated to 1 km to run the GHS-DUG tool).**

If the input population grid does not comply with these requirements the process cannot start and the tool will highlight the issues with error messages. In such case the user needs to re-project and/or resample the grid in order to comply with these requirements or provide another grid. The re-projection of geospatial layers requires specific technical knowledge. No responsibility is taken for workflows performed by users.

The best way to be sure that the grid comply with all requirements is to generate it using the **GHS-POP2G tool** (Maffenini et al., 2020a).

Be careful: re-projecting population grid is a sensitive task, as the resampling method can affect population distribution and total count.

Table 14 Population grid requirements

File format	Data type	Map projection	Pixel size
GeoTIFF	Any numeric (int or float)	<ul style="list-style-type: none"> — World Mollweide (EPSG:54009) — ETRS89 / LAEA Europe (EPSG:3035) 	multiple of 50 meters

4.1.3 Degree of Urbanisation grid

The degree of urbanisation grid is the last input and the characteristics of the layer are displayed in Table 15. The grid must contain the pixel classification according to the Degree of Urbanisation Level 2 as explained in section 1.1.2 to obtain the territorial units classification at Level 2. To obtain the classification at Level 1 the settlement grid can be provided as input either at Level 1 or at Level 2. The coordinate systems has to be in World Mollweide (EPSG:54009) or LAEA Europe (EPSG:3035) projection and it has to be the same as the one for the population grid. The pixel resolution must be square and of 1 km x 1 km.

If the input population grid does not comply with these requirements the process cannot start and the tool will highlight the issues with error messages. In such case the user needs to re-project and/or resample the grid in order to comply with these requirements or provide another grid.

The best way to be sure that the grid comply with all requirements is to generate it using the **GHS-DUG tool** (Maffenini et al., 2020b)

Table 15 Degree of urbanisation grid requirements

File format	Data type	Map projection	Pixel size
GeoTIFF	Any numeric (int or float)	<ul style="list-style-type: none"> — World Mollweide (EPSG:54009) — ETRS89 / LAEA Europe (EPSG:3035) 	1 km x 1 km

4.2 Output

The GHS-DU-TUC take a territorial units shapefile in input and export the same shapefile as outputs, where all polygons are classified according to their Degree of Urbanisation (see section 1.1.3).

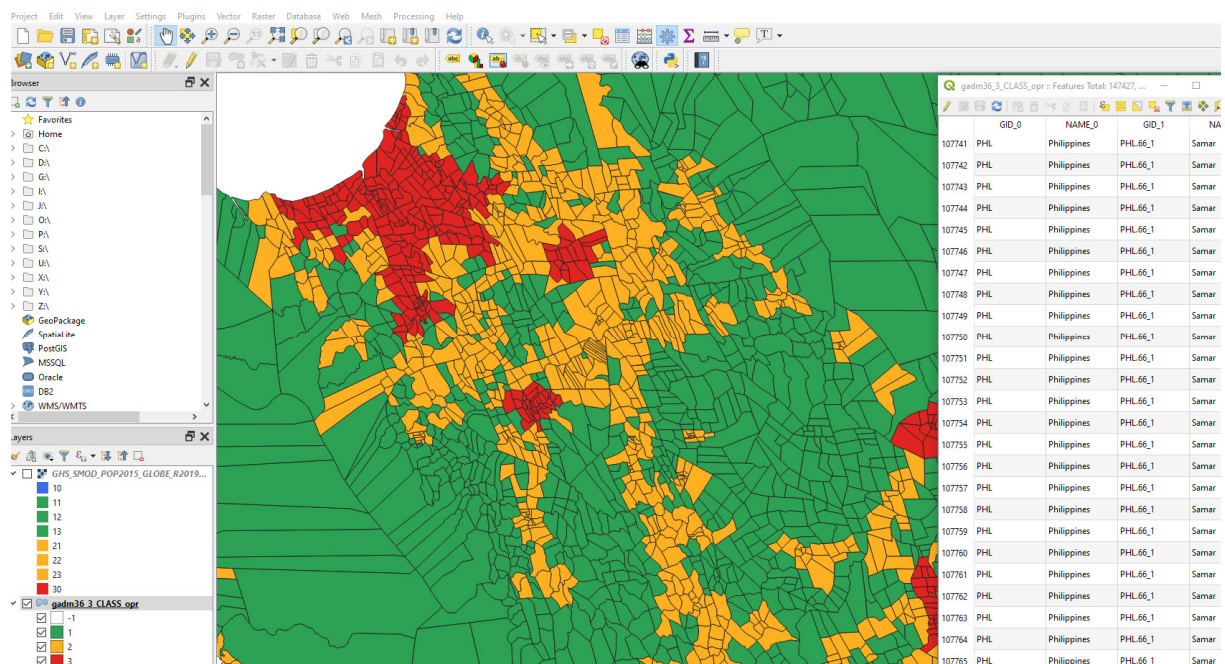
It also compute several statistics about share of population for each polygon that are stored in the output shapefile as well as in an excel table.

4.2.1 Classified shapefile

This is the input Territorial Units shapefile where each polygon has been classified. The shapefile attributes table contains the classification code and several statistics, including area and population share for each class.

The same statistics information, as well as the classification code, are reported also in the excel table. The list of such values and their names is reported in 4.2.2.

Figure 20 QGIS Session showing the results of a GADM 3.6 Level 3 geometry classified with the GHS-DU-TUC and coloured by Degree of Urbanisation Level 1 in the area of Dagupan (Philippines)



4.2.2 Statistics table

This table is saved as an excel file, which is usually easier to handle for statistics analysis, but the exact same information is also contained in the output shapefile.

Table 16 describes all statistics and information extracted for each polygon.

Table 16 List of all statistics and information of both the excel and shapefile output

Column / field name	Value per polygon
Tot_Pop	Total population within the polygon
UCentre_Pop	Population count in Urban Centre
UCluster_Pop	Population count in Urban Cluster
Rural_Pop	Population count in Rural
UCentre_share	Population share in Urban Centre
Urban_share	Population share in Urban Cluster (population share in the urban domain)
Rural_share	Population share in Rural (population share in the rural domain)
DEGURBA_L1	Polygon classification value according to Degurba Level 1
DUC_Pop	Population count in Dense Urban Cluster
SDUC_Pop	Population count in Semi Dense Urban Cluster
SUrb_Pop	Population count in Suburban or Peri-Urban Grid Cell
RC_Pop	Population count in Rural Cluster Grid Cell
LDR_Pop	Population count in Low Density Rural Grid Cell
VLDR_Pop	Population count in Very Low Density Rural Grid Cell
DUC_share	Population share in Dense Urban Cluster
SDUC_share	Population share in Semi Dense Urban Cluster
SUrb_share	Population share in Suburban or Peri-Urban Grid Cell
RC_share	Population share in Rural Cluster Grid Cell
LDR_share	Population share in Low Density Rural Grid Cell
VLDR_share	Population share in Very Low Density Rural Grid Cell
DEGURBA_L2	Polygon classification value according to Degurba Level 2

5 Conclusion

The GHS-DU-TUC Tool 1 adds to the family of GHSL open and free tools to enact the “open input, open method, open output” paradigm of the Global Human Settlement Layer framework. With the GHS-DU-TUC Tool presented in this report the user can produce an operational classification of territorial units at Degree of Urbanisation Level 1 and Level 2 and obtain additional statistics. This process comply with the second step for the application of the Degree of Urbanisation in an area of interest. This capability supports the endeavours of the European Commission and the partner Organisations for the uptake of the Degree of Urbanisation, the people-based harmonised definition of cities and settlements recommended by the 51st Session of the United Nations Statistical Commission as the method to delineate cities and rural areas for international statistical comparison. The tool specifically supports the capacity enhancement activities currently carried out jointly by the European Commission Directorate General for Regional and Urban Policy, the Directorate General Joint Research Centre and partner organisation to apply the Degree of Urbanisation in support of global monitoring of SDGs and the New Urban Agenda urban targets as recommended by the 51st Session of the United Nations Statistical Commission.

References

- Corbane, C., Pesaresi, M., Politis, P., Florczyk, A.J., Melchiorri, M., Freire, S., Schiavina, M., Ehrlich, D., Naumann, G., Kemper, T., 2018. The grey-green divide: multi-temporal analysis of greenness across 10,000 urban centres derived from the Global Human Settlement Layer (GHSL). *International Journal of Digital Earth* 1–18. <https://doi.org/10.1080/17538947.2018.1530311>
- Dijkstra, L., Poelman, H., 2014. A harmonised definition of cities and rural areas: the new degree of urbanisation (Working Papers), Regional Working Paper 2014.
- European Commission , Joint Research Centre, 2018. Atlas of the Human Planet 2018, a World of Cities. Publications Office of the European Union.
- European Commission, Statistical Office of the European Union, 2019. Methodological manual on territorial typologies: 2018 edition.
- Maffenini, L., Schiavina, M., Freire, S., Melchiorri, M., Pesaresi, M., Kemper, T., 2020a. GHS-POP2G User Guide. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2760/500887>
- Maffenini, L., Schiavina, M., Melchiorri, M., Pesaresi, M., Kemper, T., 2020b. GHS-DUG User Guide. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2760/916217>
- Melchiorri, M., Florczyk, A., Freire, S., Schiavina, M., Pesaresi, M., Kemper, T., 2018. Unveiling 25 Years of Planetary Urbanization with Remote Sensing: Perspectives from the Global Human Settlement Layer. *Remote Sensing* 10, 768. <https://doi.org/10.3390/rs10050768>
- Melchiorri, M., Pesaresi, M., Florczyk, A.J., Corbane, C., Kemper, T., 2019. Principles and Applications of the Global Human Settlement Layer as Baseline for the Land Use Efficiency Indicator—SDG 11.3.1. *ISPRS International Journal of Geo-Information* 8, 96. <https://doi.org/10.3390/ijgi8020096>
- Schiavina, M., Melchiorri, M., Corbane, C., Florczyk, A.J., Freire, S., Pesaresi, M., Kemper, T., 2019. Multi-Scale Estimation of Land Use Efficiency (SDG 11.3. 1) across 25 Years Using Global Open and Free Data. *Sustainability* 11, 5674.

List of boxes

Box 1 Supported projections..... 25

List of figures

Figure 1 Urban centre, urban cluster and rural grid cells around Cape Town, South Africa5

Figure 2 City, towns & semi-dense areas and rural areas around Cape Town, South Africa (classification of Main Places units, note that Cape Peninsula is part of Cape Town Main Place).....5

Figure 3 Degree of urbanisation level 2 grid classification around Toulouse, France6

Figure 4 Degree of urbanisation level 2 local unit classification around Toulouse, France6

Figure 5 First startup 10

Figure 6 GHS-DU-TUC main windows and areas 11

Figure 7 GHS-DU-TUC startup..... 15

Figure 8 Output folder selected..... 15

Figure 9 Territorial Units selected 16

Figure 10 Population grid selected..... 17

Figure 11 Degree of urbanisation grid selected..... 18

Figure 12 GHS-DU-TUC running 19

Figure 13 Run successful..... 19

Figure 14 Show ArcToolbox window 20

Figure 15 Add toolbox to ArcGIS..... 21

Figure 16 ArcGIS toolbox window..... 21

Figure 17 Toolbox setup with all inputs and parameters..... 23

Figure 18 Toolbox running 24

Figure 19 Toolbox run successfully completed..... 24

Figure 20 QGIS Session showing the results of a GADM 3.6 Level 3 geometry classified with the GHS-DU-TUC and coloured by Degree of Urbanisation Level 1 in the area of Dagupan (Philippines) 27

List of tables

Table 1 Short and technical terms for Degree of Urbanisation level 1 and 2 for the local unit classification.....7

Table 2 Short and technical terms for Degree of Urbanisation level 1 and 2 for grid cell classification.....8

Table 3 System requirements 9

Table 4 Requirements by input raster size..... 9

Table 5 File menu bar 12

Table 6 Input menu bar 12

Table 7 Output menu bar 12

Table 8 Options menu bar 13

Table 9 Help menu bar 13

Table 10 Buttons 14

Table 11 Editable text fields 14

Table 12 Toolbox items description 22

Table 13 Regions vector file requirements 25

Table 14 Population grid requirements 26

Table 15 Degree of urbanisation grid requirements 26

Table 16 List of all statistics and information of both the excel and shapefile output..... 28

Annexes

Annex 1. Error messages

Error	Description	Solution
Cannot create log file! Check you have write permission	The tool has no write access to the selected folder and cannot create log file	Check and change folder permission or chose another folder with write permission
Shapefile geometry is not polygon	The provided shapefile does not contain polygons but other geometries (e.g. points or lines)	Use a shapefile which contains only polygon geometries
Shapefile has no attribute	The geometry contained in the shapefile has no attribute	Use a shapefile with at least one attributes for each geometry
Cannot read data projection! Please manually enter the EPSG code	The tool cannot detect the Spatial Reference of the provided layer	Manually enter the layer EPSG code in the dedicate box
Invalid data projection! Please insert a valid EPSG code	The EPSG code provided is invalid (e.g. 9, 123 or abcd)	Provide a valid EPSG code (e.g. 54009 or 3035)
Raster is not fully contained inside the Territorial Unit layer!	The provided raster is too small or does not completely overlap to the Territorial Unit layer	Choose a raster with an extension that completely overlap or is bigger than the Territorial Unit's one
Wrong data projection! Valid values are only 54009 and 3035'	The layer has spatial reference that is different from the two allowed	Provide the layer in the required spatial reference or re-project the data
Different data projections! Population and degurba must be in the same projection	Population and degurba rasters have different spatial reference	Provide both rasters in the same spatial reference
Wrong pixel resolution! Pixel resolution must be multiple of 50 meters	The rasters has a pixel resolution that is not a multiple of 50 meters (e.g. 10, 30 or 140)	Provide a raster that has pixel resolution multiple of 50 meters (e.g. 50, 100 or 250)
Wrong pixel resolution! Pixel resolution must be 1 Km (1000 meters)	The rasters has a pixel resolution that is different than 1000 meters	Provide a raster with 1000 meters pixel resolution
Check missing data or error in Territorial Units panel!	Generic error that shows Territorial Units panel has some incorrect inputs or is missing data from the user	Check the specific error messages and take action accordingly
Check missing data or error in Population panel!	Generic error that shows Population panel has some incorrect inputs or is missing data from the user	Check the specific error messages and take action accordingly
Check missing data or error in Degurba panel!	Generic error that shows Degurba panel has some incorrect inputs or is missing data from the user	Check the specific error messages and take action accordingly

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: https://europa.eu/european-union/contact_en

On the phone or by email

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

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

FINDING INFORMATION ABOUT THE EU

Online

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

EU publications

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

The European Commission's science and knowledge service

Joint Research Centre

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



EU Science, Research and Innovation



EU Science Hub



Publications Office
of the European Union

doi:10.2760/633624

ISBN 978-92-76-22421-1