



JRC SCIENTIFIC INFORMATION SYSTEMS AND DATABASES

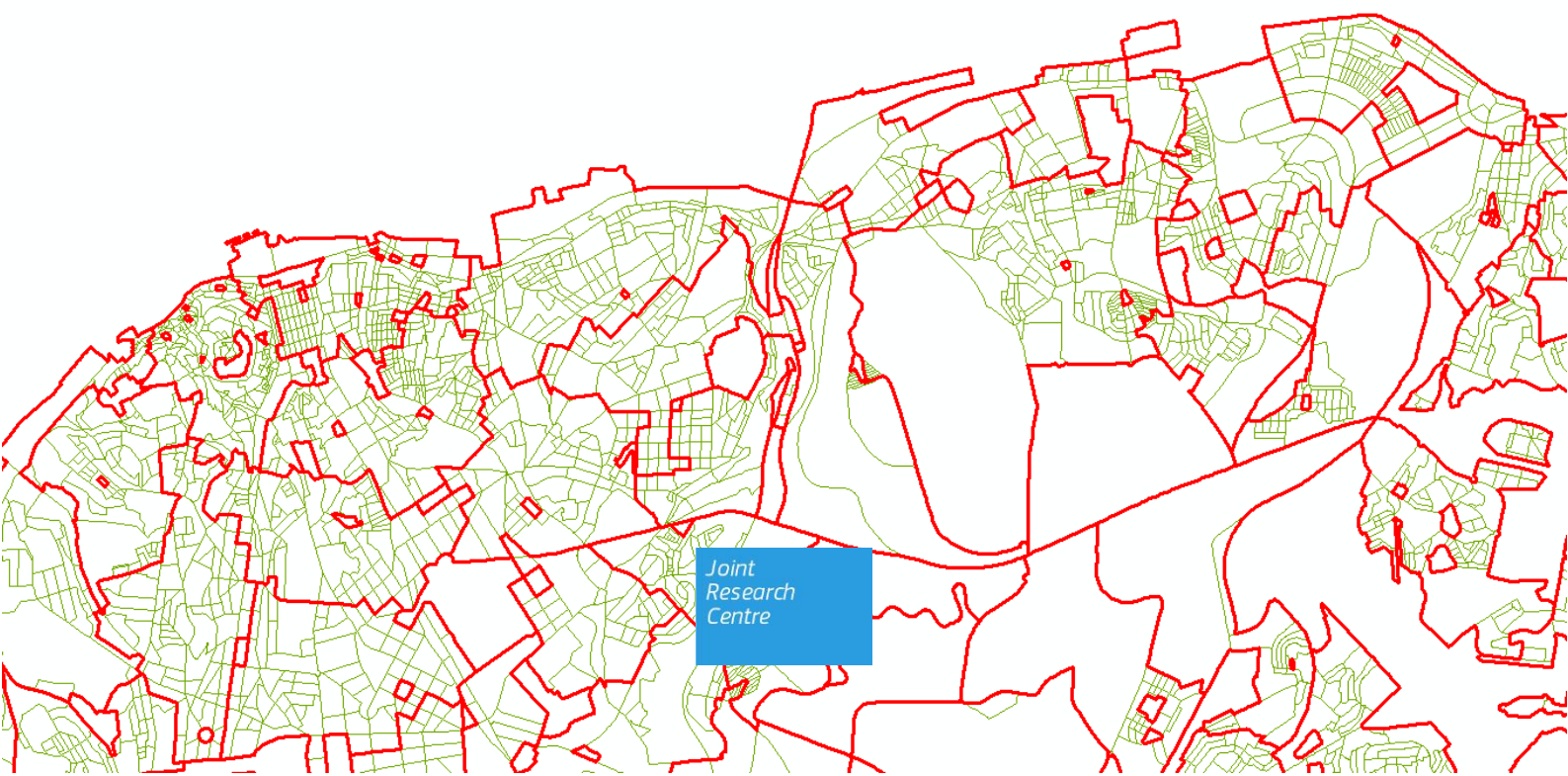
SmartDissolve User Guide

Documentation
Version 1.2

Marcello Schiavina

Sérgio Freire

2017



Joint
Research
Centre

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 policy-making 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 which might be made of this publication.

Technical Support

For technical support send an email to ghsl-tools@jrc.ec.europa.eu

JRC Science Hub

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

JRC106896

PDF ISBN 978-92-79-69304-5 doi:10.2760/176749

Luxembourg: Publications Office of the European Union, 2017

© European Union, 2017

Reproduction is authorised provided the source is acknowledged.

How to cite: Marcello Schiavina, Sérgio Freire; *SmartDissolve User Guide*; doi:10.2760/176749

All images © European Union 2017

Table of contents

Abstract	1
1 Introduction	2
2 Installing SmartDissolve	3
2.1 System Requirements.....	3
2.1.1 SmartDissolve command line tool	3
2.1.2 SmartDissolve toolbox for ArcGIS	3
2.1.3 SmartDissolve function for MATLAB.....	3
2.2 Installing MATLAB Runtime.....	3
2.3 Installing SmartDissolve	4
2.3.1 Installing SmartDissolve command line tool	4
2.3.2 Installing SmartDissolve toolbox for ArcGIS	4
2.3.3 Installing SmartDissolve function for MATLAB.....	4
2.4 Uninstalling SmartDissolve	5
3 Using the SmartDissolve.....	6
3.1 Using SmartDissolve command line tool or function for MATLAB.....	6
3.1.1 Syntax for SmartDissolve command line tool	6
3.1.2 Syntax for SmartDissolve function for MATLAB	6
3.1.3 List of inputs and examples.....	7
3.1.3.1 Example as command line tool	8
3.1.3.2 Example as MATLAB function.....	8
3.2 Using SmartDissolve toolbox for ArcGIS.....	9
References	11

Abstract

SmartDissolve is a polygon aggregation tool developed in the frame of the Global Human Settlement Layer (GHSL) project and it is being used to dissolve polygons setting an areal threshold.

SmartDissolve is a tool that handles minimum mapping unit, resolution mismatch between layers, or spatial uncertainty problems in GISc. This tool automatically dissolves polygons below a threshold area, updating fields' values. The toolbox allows to select the ordering of polygon analysis (i.e. from the smallest to the largest area, vice versa, or order of IDs), different dissolve rules (i.e. with smallest, largest, or maximum-border-share adjacent polygon, minimum total perimeter or maximum compactness) and different field updating operations (i.e. sum, mean or text concatenation). The software is available as toolbox for ArcGIS 10.X, a standalone command line tool or a MATLAB function.

More details about the algorithm, the method and its actual applications can be found in the paper in the Reference section.

This guide provides instructions about installing and using the SmartDissolve toolbox for ArcGIS and the SmartDissolve command line tool on a Windows computer and the SmartDissolve MATLAB function in a MATLAB environment.

1 Introduction

The 'dissolve' spatial operation is one of the more common and useful procedures applied in Geographic Information Systems (GIS) (Davis, 2001). In available dissolve tools, the aggregation is based on polygons sharing the same category or code (i.e. boundaries are removed between adjacent polygons that have the same value for a specified attribute).

However, many spatial analyses require more flexible and sophisticated aggregation by dissolution of adjacent polygons storing quantitative data, such as reaching defined target areas (surface) while correctly handling multi-part features and accurately summing count variables.

SmarDissolve is a tool developed to meet such requirements: it handles minimum mapping unit, resolution mismatch between layers or spatial uncertainty problems in GISc. This tool automatically dissolves polygons below a given threshold area with an adjacent polygon (according to a user defined rule), updating their fields' values. The toolbox allows to select the ordering of polygon analysis (i.e. from the smallest to the biggest, vice versa or ordered by IDs), different dissolution rules (i.e. with smallest, largest, maximum-border-share adjacent polygon, minimum total perimeter or maximum compactness) and different functions for updating attributes (e.g. sum, mean, max, min or text concatenation). The algorithm correctly processes multi-part polygons but not polygons with intersections. For more details about the motivation, the potential uses and the algorithm refer to the paper in the Reference section.

The tool is developed using MATLAB and it is available as Windows executable, as a toolbox for ArcGIS 10.X using a Python script and as MATLAB function.

2 Installing SmartDissolve

2.1 System Requirements

To use SmartDissolve your system must meet the following minimum requirements:

Operating system	Processor	Disk space	RAM
Windows 7 or Windows 10 (any platform for the MATLAB function)	Any Intel or AMD x86-64 processor	30 Mb (+ 700 Mb for Matlab Runtime)	2 GB

2.1.1 SmartDissolve command line tool

To have SmartDissolve command line tool running on your computer you need to:

1. Install MATLAB Runtime
2. Install SmartDissolve

2.1.2 SmartDissolve toolbox for ArcGIS

To have SmartDissolve toolbox for ArcGIS on your computer you need to:

1. Install ArcGIS 10.X (not included in the package)
2. Install Matlab Runtime
3. Install SmartDissolve

2.1.3 SmartDissolve function for MATLAB

To have SmartDissolve function for Matlab on your computer you need to:

1. Install MATLAB 2015b or later (not included in the package). The function has been tested with 2015b, 2016a and 2016b versions
2. Install SmartDissolve

2.2 Installing MATLAB Runtime

Verify the MATLAB Runtime is installed and ensure you have installed version **9.1 (R2016b)**.

If not already installed, download the Windows 64-bit version of the MATLAB Runtime for R2016b from the MathWorks Web site:

<http://www.mathworks.com/products/compiler/mcr/index.html>

For more information about the MATLAB Runtime and the MATLAB Runtime installer, see Package and Distribute in the MATLAB Compiler documentation in the MathWorks Documentation Center.

2.3 Installing SmartDissolve

SmartDissolve has no Windows installer, you just need to download the package from the GHSL project webpage (<http://ghsl.jrc.ec.europa.eu>) and extract the archive.

2.3.1 Installing SmartDissolve command line tool

SmartDissolve command line tool is a portable tool and does not need any particular installation process.

2.3.2 Installing SmartDissolve toolbox for ArcGIS

SmartDissolve toolbox can be imported into the ArcToolbox by using the "Add Toolbox..." command within ArcToolbox as shown in Figure 1. Once the dialog box is open you will be asked to insert the path of the "SmartDissolve.pyt" file.

2.3.3 Installing SmartDissolve function for MATLAB

SmartDissolve function for MATLAB can be used by setting the MATLAB working directory to the location of the "smartdissolve.m" file.

In order to use the function in any working directory, you can add the path of "smartdissolve.m" to the MATLAB search path list, using the following command (replace "C:\SmartDissolvePath\" with the actual location of the "smartdissolve.m" file on your machine):

```
>> addpath('C:\SmartDissolvePath\')
```

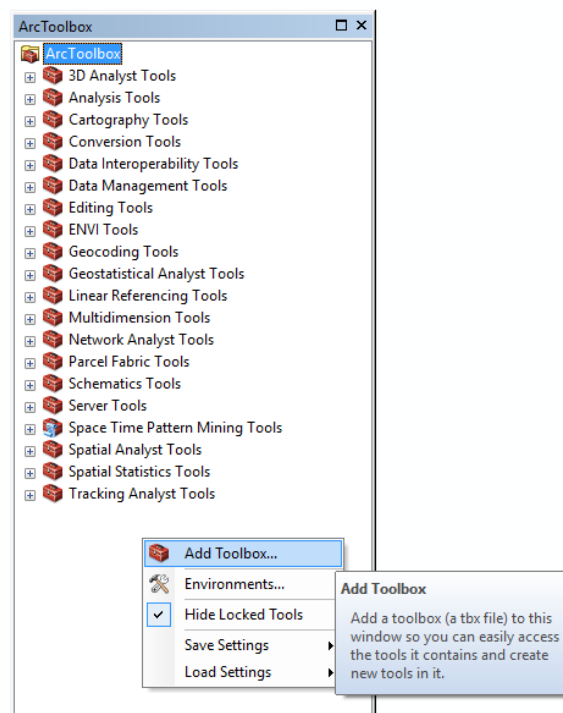


Figure 1 - ArcToolbox import procedure

2.4 Uninstalling SmartDissolve

To remove SmartDissolve just delete all SmartDissolve files, remove the SmartDissolve toolbox for ArcGIS from ArcToolbox and uninstall the MATLAB Runtime.

3 Using the SmartDissolve

3.1 Using SmartDissolve command line tool or function for MATLAB

3.1.1 Syntax for SmartDissolve command line tool

The syntax from command line could be one of the following:

1. C:\SmartDissolvePath\> smartdissolve.exe original_shape area_ref diss_rule
2. C:\SmartDissolvePath\> smartdissolve.exe original_shape area_ref diss_rule fields rules
3. C:\SmartDissolvePath\> smartdissolve.exe original_shape area_ref diss_rule options optionvalues
4. C:\SmartDissolvePath\> smartdissolve.exe original_shape area_ref diss_rule fields rules options optionvalues

The minimum syntax (1) defines the path of the shapefile of polygons to work on, *original_shape*, the threshold area to determine polygons to be dissolved, *area_ref*, and the rule to select the adjacent polygon to be dissolved with, *diss_rule*. The syntax (2) defines also the fields to be updated, *fields* (for example: "{ 'field1', 'field2' }"), and the rule of update for each field, *rules* (for example: "{ 'sum', 'max' }", see section [List of inputs and examples](#) for more details). In (3) and (4) a set of options, described in section [List of inputs and examples](#), can be defined.

3.1.2 Syntax for SmartDissolve function for MATLAB

The syntax within MATLAB could be one of the following:

1. >> dissolved_shape = smartdissolve(original_shape, area_ref, diss_rule);
2. >> dissolved_shape = smartdissolve(original_shape, area_ref, diss_rule, fields, rules, ...)
3. >> dissolved_shape = smartdissolve(original_shape, area_ref, diss_rule, options, optionvalues, ...)

The three different syntaxes return respectively:

1. an N-by-1 geographic data structure (a mapstruct) containing polygons from '*original_shape*' shapefile dissolved with adjacent polygons selected according to '*diss_rule*' when original polygon area is smaller than a threshold area ('*area_ref*') expressed in square meters. No attribute is saved in the newly created mapstruct;
2. an N-by-1 geographic data structure as (1), updating all fields listed in '*fields*', an M-by-1 cell array, according to each rule listed in '*rules*', an M-by-1 cell array, see [List of inputs and examples](#) for a list of possible updating rules;
3. an N-by-1 geographic data structure as (1), with specific options for the algorithm (polygon order, subset, multiprocessing, saving options, ...). See option~optionvalues pair arguments in section [List of inputs and examples](#) for more details. The geographic data structure, '*dissolved_shape*', is a mapstruct or a geostruct according to '*original_shape*'. Intersection among polygons will throw an error unless '*Interactive*' option is set to 'true' to open an interactive session to handle them.

3.1.3 List of inputs and examples

The SmartDissolve command line tool and MATLAB function "smartdissolve.m" have the same functionalities and options.

Here is the list of possible inputs:

- **original_shape:** is the path of the original shapefile of polygons. The algorithm does not automatically handle polygons with intersections.
- **area_ref:** is the threshold area, in square meters, below which polygons are selected to be dissolved.
- **diss_rule:** defines the aggregation rule (i.e. given a polygon *i* below the threshold area, the rule to select among all its adjacent polygons the one to be dissolved with):
 - *MinArea* selects the adjacent polygon with the smallest area;
 - *MaxArea* selects the adjacent polygon with the largest area;
 - *MaxBorder* selects the adjacent polygon that shares the longest border with polygon *i*;
 - *MinPerimeter* selects the polygon that results in the minimum total perimeter (i.e. the perimeter of the dissolved polygon);
 - *MaxCompactness* selects the polygon that results in the highest isoperimetric quotient (i.e. the ratio of the total area over the area of the circle having the same perimeter).
- **fields:** is a cell array with all polygon attribute field's names to be updated.
 - Example: {'field1', 'field2'}
- **rules:** is a cell array with the updating rule for each attribute field in *fields*. They can be:
 - *sum* to sum values
 - *mean* to calculate the average
 - *min* to calculate the minimum value
 - *max* to calculate the maximum value
 - *stdev* to calculate the standard deviation
 - *concatenate* to concatenate values
 - Example: {'sum', 'concatenate'}
- **options~optionvalues** pair arguments:
 - **OrderRule:** is a string that specifies the ordering rule of polygons:
 - *ID*, default;
 - *MinMax* from smallest to largest;
 - *MaxMin* from largest to smallest.
 - **Tolerance:** is a *double* value to set the tolerance for vertices comparison. In planar coordinates it is expressed in meters (1e-2 -> 1 cm). In Geographic Coordinate System is expressed in degrees (1e-7 -> 1 cm at the equator).
 - Default: 1e-2 for planar; 1e-7 for GCS.
 - **Multiprocessing:** is the number of cores the algorithm will use (in MATLAB function, if a parallel session is already open it will use it).

- Default: 1.
- **SaveShape:** is a logical that determines if the result is saved into a shapefile, if set to true, or not, false. The dissolved shape will be saved in the same folder as '*original_shape*'_dissolved.shp (command line tool will not produce any results if this option is set to false).
 - Default: true.
- **PartialLoad:** is an N by 3 cell matrix of attribute names, relationship and specific value to limit the dissolving procedure to a subset of polygons. For *string*, possible relationships are '==' or '~='. For scalars possible relationships are '==', '~=', '>', '<', '>=' and '<='. All conditions work with AND logic.
 - Example: {'Country', '==', 'Italy'; 'Region', '~=', 'Lombardy'; 'Population', '>=', '100'; 'Population', '<', '1000'}. To select polygons in Italy not in Lombardy with population between 100 and 1000 inhabitants.
- **Interactive:** Logical scalar that specifies opening an interactive session whenever an intersection between polygons is found, plotting the problem on a graph and letting the user fix it manually. Interactive section is meant to fix small vertex problems (aligning two polygons and removing slivers), or to skip the problematic polygon.
 - Default: false.

3.1.3.1 Example as command line tool

Here is an example of the SmartDissolve command line tool:

Polygons below an area of 10 km² in the shapefile "*Shape.shp*", where attribute "ISO" is equal to "ITA", will be dissolved using the *MaxCompactness* rule, starting from the smallest to the largest; fields "Pop" and "LAU2" will be updated by summing the first and concatenating the second, during polygon aggregation; the process will use 4 cores:

```
C:\SmartDissolvePath> smartdissolve.exe Shape.shp 1e7 MaxCompactness {'Pop',
'LAU2'} {'sum', 'concatenate'} OrderRule MinMax
Multiprocessing 4 PartialLoad {'ISO', '==', 'ITA'}
```

3.1.3.2 Example as MATLAB function

Here is an example of the SmartDissolve function for MATLAB:

Polygons below an area of 10 km² in the shapefile "*Shape.shp*", where attribute "ISO" is equal to "ITA", will be dissolved using the '*MaxCompactness*' rule, starting from the smallest to the largest; fields "Pop" and "LAU2" will be updated by summing the first and concatenating the second, during polygon aggregation; the process will use 4 cores and won't save the results in a shapefile, returning the structure "*Shape_dissolved*" into Matlab workspace; interactive session is turned on to handle possible polygon intersection:

```
>> Shape_dissolved = smartdissolve('Shape.shp', 1e7, 'MaxCompactness', {'Pop',
'LAU2'}, {'sum', 'concatenate'}, 'OrderRule', 'MinMax', 'Multiprocessing', 4,
'SaveShape', false, 'Interactive', true, 'PartialLoad', {'ISO', '==', 'ITA'});
```

3.2 Using SmartDissolve toolbox for ArcGIS

Running the SmartDissolve toolbox for ArcGIS will open a dialogue box as in Figure 2. The resulting feature layer will be added to the Table Of Contents and saved as a shapefile, if the input is a shapefile, or within the same gdb of the original vector layer, with the original layer name plus “_dissolved”.

Here it is the list of inputs:

- **Input Features:** is the feature layer to be dissolved. The feature layer can be a shapefile (shp) or a vector layer in GeoDatabase (gdb). If the input layer is a vector within gdb, a shapefile is exported in a temporary folder "gdb_name" gdb_temp/"vector_filename".shp. A shapefile will be temporarily created as result of computation in that folder and then exported in the gdb. All temporary files will be removed at the end of computation.

Field	Statistic Type
code	Concatenate
P_night	Sum

Field	Relationship	Value
P_WS	>	0

Figure 2 - ArcGIS dialog box for SmartDissolve toolbox. The example shows the SmartDissolve of the shapefile 'LAU2', dragged from the Table of Content. Polygons below 10 km² with field 'P_WS' value greater than 0 will be dissolved using the 'MinMax' polygon ordering and the 'MinArea' dissolve rule, updating the fields 'code' and 'P_night' by, respectively, concatenation and sum. The algorithm will run on 4 cpus.

- **Threshold area:** is the threshold area, along with its unit of measure from the list (the list allows to select "Unknown", the value will be interpreted as square meters), below which polygons are selected to be dissolved
- **Order rule:** is the ordering rule to cycle all polygons of the input shapefile checking which polygons are below the threshold area value. The areal analysis cycle can be set:
 - following IDs order (*ID*), default;
 - from smallest to largest polygon (*MinMax*);
 - from largest to smallest polygon (*MaxMin*).
- **Dissolve rule:** is the rule to select the adjacent polygon to be dissolved with. Five different aggregation rules are available:
 - with adjacent polygon that has the smallest area (*MinArea*), default;
 - with adjacent polygon that has the largest area (*MaxArea*);
 - with adjacent polygon that shares the longest border (*MaxBorder*);
 - with adjacent polygon that results in the minimum total perimeter (*MinPerimeter*);
 - with adjacent poly that results in the highest isoperimetric quotient, i.e. the ratio of the total area over the area of the circle having the same perimeter (*MaxCompactness*).
- **Update fields (optional):** selection from the list of the shapefile's attribute fields those to be updated, along with the update rule (sum, mean, minimum, maximum, or text concatenation), the algorithm updates the field values of the resulting dissolved polygon by applying the calculation to the original polygons' field values. If none is selected the output will not have any attribute.
- **Multiprocessing (number of CPUS):** enables multiprocessing by setting this value greater than 1 (the algorithm automatically checks if the value exceeds the maximum possible CPUS).
- **Subset (optional):** selection from the list of the shapefile's attribute fields along with a list of possible *Relationship* with a *Value*, to select a subset of polygons to work on. For string, possible *Relationships* are '=' or '~='. For scalars possible *Relationships* are '==', '~=', '>', '<', '>=' and '<='. All conditions work with AND logic.

References

SmartDissolve is based on the studies performed at the Joint Research Center (JRC) in Ispra, Italy.

For more details about the implemented algorithm please refer to:

Development of a new and open approach to dissolve polygons storing count data based on areal threshold

Schiavina, M; Freire, S

Proceedings of the 20th AGILE conference, Wageningen, The Netherlands 9-12 May 2017

ISBN 978-90-816960-7-4

<http://publications.jrc.ec.europa.eu/repository/handle/JRC105799>

GHSL Project

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

External components and libraries:

ArcGIS 10.X

<https://www.arcgis.com/features/index.html>

MATLAB Runtime

<http://it.mathworks.com/products/compiler/mcr>

Europe Direct is a service to help you find answers to your questions about the European Union
Free phone number (*): 00 800 6 7 8 9 10 11
(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server <http://europa.eu>

How to obtain EU publications

Our publications are available from EU Bookshop (<http://bookshop.europa.eu>),
where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.
You can obtain their contact details by sending a fax to (352) 29 29-42758.

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

