Updated UMZs and corresponding methodological documentation.

Method based on CLC v16

Task: 261-1#2.3_1.
Further development of IUME and urban assessment

Prepared by:

Alejandro Simón, María Leal (UMA)

Date:

05.12.2011

Project Manager:

Marie Cugny-Seguin
supported by Óscar Gómez

European Environment Agency

UNIVERSIDAD DE MÁLAGA

Universidad de Málaga

ETCSIA

PTA - Technological Park of Andalusia

C/ Marie Curie, 22 (Edificio Habitec)

Campanillas

29590 - Malaga

Spain

Telephone: +34 952 02 05 48

Fax: +34 952 02 05 59

Contact: etc-sia@uma.es

UAB
Universitat Autònoma de Barcelona

ALTERRA
WAGENINGEN UR

fömi

GeoVille
Group

gisat

umweltbundesamt

con terra

ifgi
Institut für Geoinformatik
Universität Münster

IGN
France
INTERNACIONAL

ISPRA

NATIONAL ENVIRONMENTAL
RESEARCH INSTITUTE
AARHUS UNIVERSITY

Société de Calcul Mathématique, S. A.
Algorithmes et Optimisation

JUNTA DE ANDALUCIA
CONSEJERIA DE MEDIO AMBIENTE
REGIONAL MINISTRY OF ENVIRONMENT

UWE
BRISTOL

TABLE OF CONTENTS

LIST OF FIGURES

1 OVERVIEW

This working document describes the methodology to create the Urban Morphological Zones (UMZ). A UMZ describes the urban tissue of an area and is based on Land Cover classification, in this case, Corine Land Cover version 16 (delivered by EEA in April 2012). This document is an updated of version of Simon and Milego, 2010 (Old paper can be downloaded here: http://www.eea.europa.eu/data-and-maps/data/urban-morphological-zones-2006-umz2006-f3v0/)

To obtain the final UMZ vector layer, one model has been developed in ArcGIS 10 to create it automatically. The model is named "Building_UMZ" (. Before running the model, it is needed to define the output folder (and to create it if it doesn't exist). In the following paragraphs the main intermediate outputs are explained but not all of them as the model has many steps.

UMZ definition
An Urban Morphological Zone can be defined as "A set of urban areas laying less than 200m apart". Those urban areas are defined from land cover classes contributing to the urban tissue and function.
The Corine Land Cover classes (Figure 1) used to build the Urban Morphological Zone dataset are the following ones:

- **Core Classes**
    - o 111 – Continuous urban fabric
    - o 112 – Discontinuous urban fabric
    - o 121 – Industrial or commercial units
    - o 141 – Green urban areas
- **Enlarged core classes**, CLC Classes 123 (Port areas), 124 (Airports) and 142 (Sport and leisure facilities), are also considered if they are neighbours to the core classes or to one of them touching the core classes.
- **Road and rail** networks, CLC class 122 when neighbours to the enlarged core classes, roads and rails will be included (cut by 300m buffer) as part of the UMZ due to their artificial nature and *connectivity* properties.
- **Water courses**, CLC class 511, when neighbours to the enlarged core classes and acting as a link between them, they are selected by this fact.
- **Forests & scrub**, CLC classes 311,312,313,322,323,324, when they are completely within the core classes.

111	211	311	411
112	212	312	412
121	213	313	421
122	221	321	422
123	222	322	423
124	223	323	511
131	231	324	512
132	241	331	521
133	242	332	522
141	243	333	523
142		334	
		335	
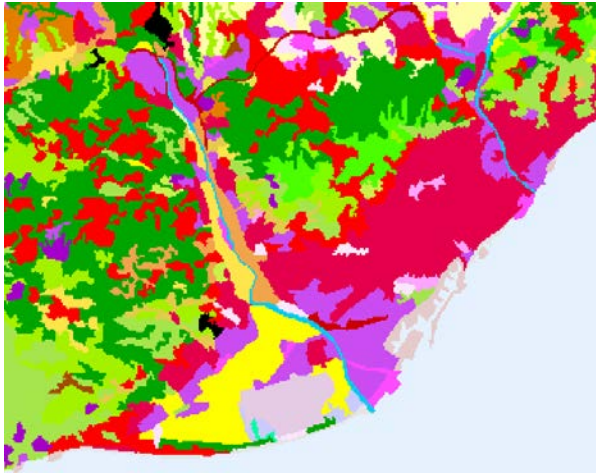
Figure 1. Land Cover (Barcelona Area)   *Figure 2. Legend*

2 PROCEDURAL STEPS

1.1 PREPROCESSING

Due to the vast dimensions of the data for the whole Europe, the described procedure should be applied in zones. Different areas are delineated in Europe and the CLC is firstly selected by those areas through an extract by mask.

1.2 CLC RECLASSIFICATION

The starting point is the reclassification of Corine Land Cover raster data (*Input data: CLC in raster format*)

CLC is reclassified according to the definition described above to obtain the following set of land cover classes:

- *Core classes* (Figure 3)
- *Ports, Airports and sport facilities* (Figure 4)
- *Transportation (*Figure 5)
- *Water bodies* (Figure 6)
- *Forests & scrub* (Figure 7)

Each set of land cover classes in raster format is vectorised in order to obtain the corresponding vector layer and stored in a file geodatabase.

- *Core classes raster.* Directly vectorised. **Output: Core_Classes**
- *Ports, Airports and sport facilities* → Enlarge Core Classes

  Starting from the Core classes described above, the first step consists on adding those ports, airports and sport facilities which are neighbours to the core classes. This step is done iteratively until all neighbouring polygons are selected1[1]

  Output: Enlarged_final
- *Transportation.* Directly vectorised. **Output: Roads**
- *Water bodies.* Directly vectorised **Output: Water**
- Forests & scrub. Directly vectorised. **Output: Forestscrub.shp**

---

[1] Due to the expertise applied, 30 iterative steps are enough to include all the enlarge objects to the core classes. Test done with CLC 90, CLC 00, and CLC 06 showed that 25 steps were enough. However 5 buffer steps were added.
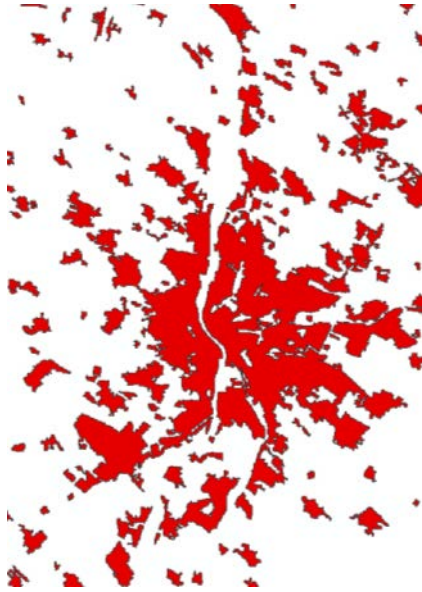
Figure 3. Core classes
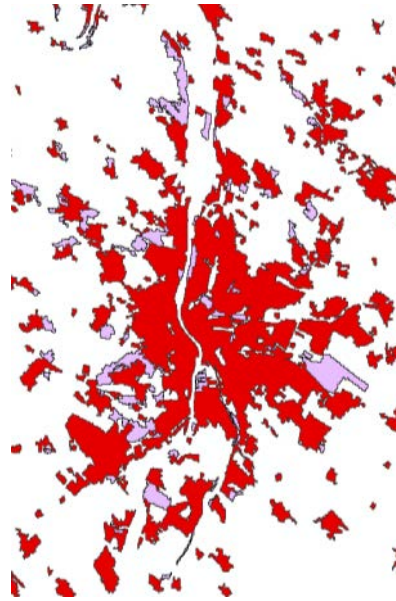
Figure 4. Enlarge Core classes
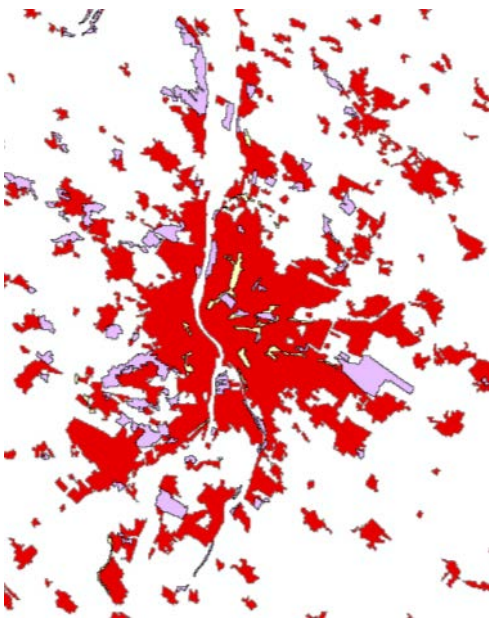
Figure 5. Enlarge Core Classes + Road classes
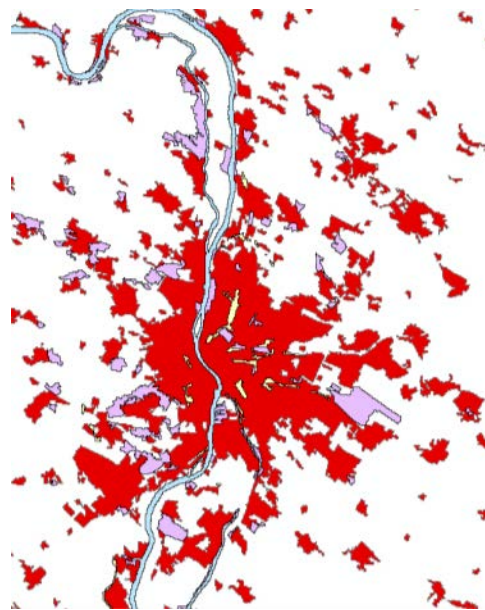
Figure 6. Enlarge Core Classes + Road+ water classes
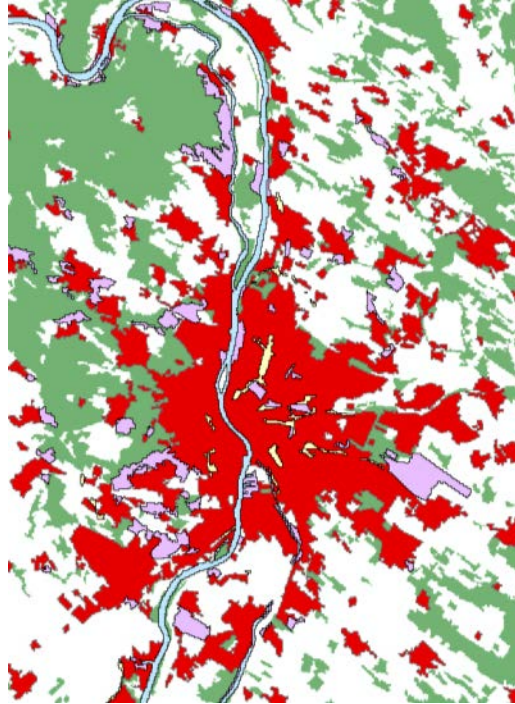
Figure 7. Enlarge Core Classes + Road+ water classes+ Forest and scrub classes

1.3 FINDING GAPS IN THE CORE CLASSES

For the final delineation of the UMZ, it is needed to identify the gaps inside the UMZs. These can be forest areas and therefore they should be included (Figure 8)

The methodology to find the gaps is very easy. Only 4 steps:

- Core classes shapefile (Core_Classes.shp) is converted from polygon to lines → *Core_Classes_Lines.shp*

- Core_Classes_Lines.shp is converted again to a polygon feature → *core_classes_POLYGONtotal*

- Finally, the core_classes_POLYGONtotal is erased with the Core_Classes geometry → *Gaps_Core_Classes*

Enlarge classes are not taken in account to allocate forest gaps inside them, to limit the forest delineation and consider only those core urban areas related ones.

1.4 BUILD-UP UMZ

This last step is a list of processes to include all the features that will participate in a UMZ

1.4.1 Addition of Forest to the Core Classes

All the forest geometries which are inside the UMZ will be added. Therefore all the forest contained by the "Gaps_core_classes" will belong to the UMZ (Figure 8). ***Output: "Forest_selected.shp"***

Figure 8. Forest classes in the Gaps of the Core classes

1.4.2 Addition of water bodies to the enlarge core classes

Water bodies (lakes, rivers, etc.) are natural items that can either serve as a link between different urban areas or as isolation, thus they must be carefully analysed.
The main problem found during the analyses appears when considering UMZ separated by rivers bigger than 200m (e.g. Budapest or Bonn areas). In those cases it is necessary to decide if the different UMZs are linked or not by the river and whether to include the river in the UMZ or not.

The elementary process to incorporate water bodies to the UMZs selected the features in touch with the enlarged core classes and clipped them by a 300m buffer around the enlarged core classes, to finally merge all. The clip prevents having all UMZ merged together via linear features such as motorways or wide rivers. After several trials in different European cities with predominant linear features (e.g. big rivers), 300m has been considered as the optimal distance for the aggregate polygon and clip functions in the process of adding water bodies to the UMZs. Some examples can be seen in section 8.3).

Despite this circumstance, when considering rivers as linear units and using the previously defined approach to include them in the UMZs, some not linked urban zones were gathered by the river. To solve this problem, the latest approach applied initially identifies the junction elements between the urban areas that correspond to water bodies and uses them to select the rivers´ segments that should be included in the UMZ. Hence, junction elements are defined by aggregating the enlarged core class´ polygons that are within a 300m distance and identifying the connexion components[2]. Those corresponding to water bodies are selected and used to identify the water features that need to be included in the UMZ by a buffer and clipping procedure, analogous to the elementary process above described.
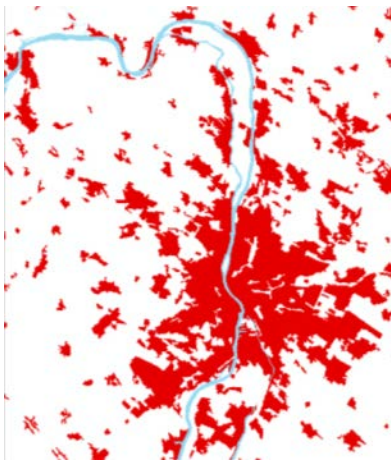
Figure 9. Enlarge Core Classes + Water classes
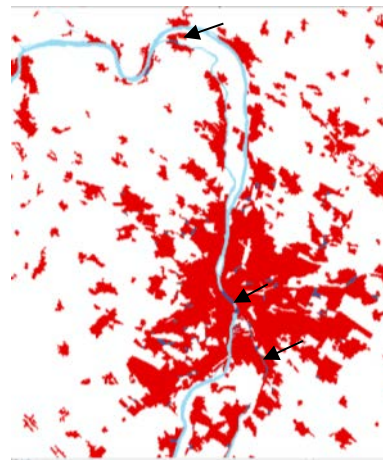
Figure 10. Identification of junction elements, corresponding with water areas
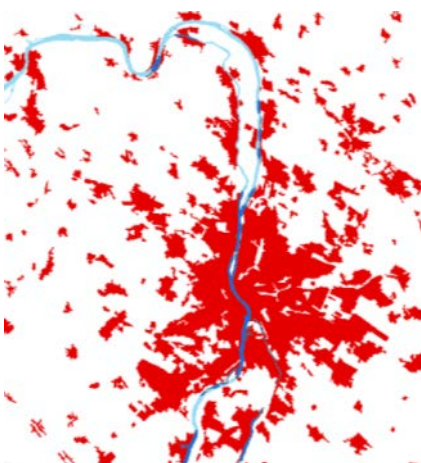
Figure 11. Water Selected cut by 300m buffer following urban link properties.
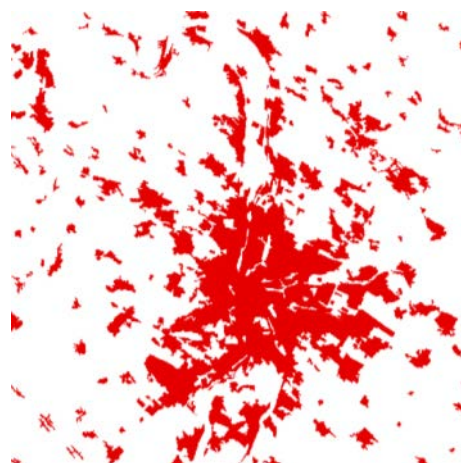
Figure 12. Enlarge Core Classes + additional of water bodies features.

---

[2] See Annex III for detailed description.

1.4.3   Addition of roads and railways to the enlarge core classes

Following the previously described method for including water bodies in the UMZ delineation process, those roads and railway polygons neighbours to the enlarged core classes are selected, then clipped by a 300m buffer around the enlarged core classes, and finally merged to them. The clip prevents having all UMZ merged together via linear features such as motorways[3] (Figure 9 - Figure 12) **Output: "Enlarge_CoreClasses_plus_lines.shp"**
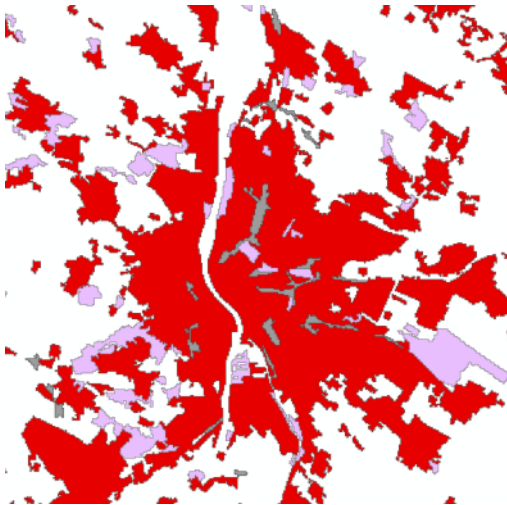
Figure 13. Enlarge Core Classes + Road classes
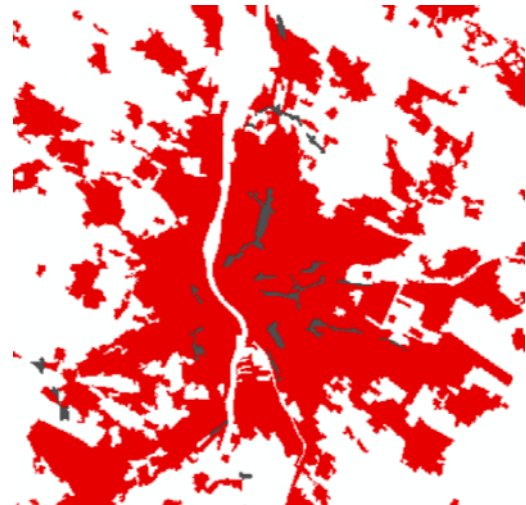
Figure 14. Roads selected (Selection by neighbourhood Linear – Enlarge Core Classes)

---
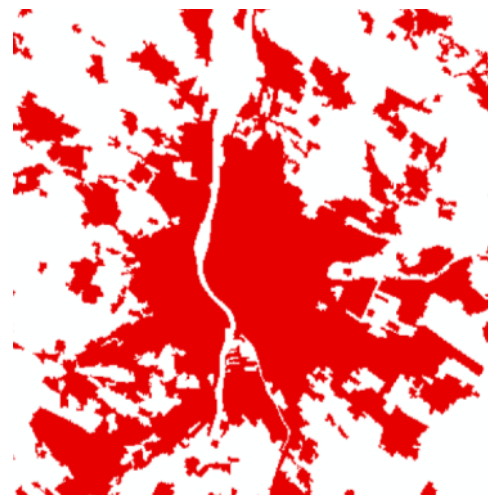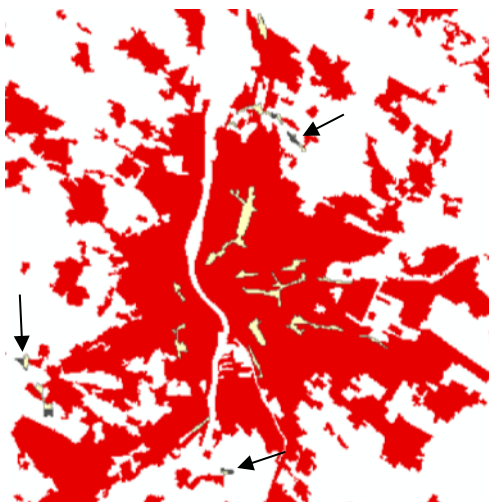
[3] Buffering process of the linear future. Due to complex nature of the UMZ, this step took too much time and sometimes it failed. Now, instead of a 300m buffer above the vector geometry, a 300m expand (3 pixels of 100 m resolution) from the raster version was done.

1.4.4 Final Built-up UMZ

The outputs from the steps 3.3.1, 3.3.2 and 3.3.3 are merged in one feature class, *"PreUMZ.shp"*. With this step a pre-UMZ layer is obtained, which is rasterised (100m resolution) to undertake the so-called UMZ built-up, explained in the next paragraph and Figure 17- Figure 19

Once all the land cover classes to become part of the UMZ have been merged, it is time to apply the definition that tells us a UMZ is a set of urban areas laying less than 200m apart. The best way to put in practice this definition is through an expansion and shrinking process of the rasterised pre-UMZ by 1 pixel (i.e. 100m). Doing so, in the expansion process some areas will be in touch and will remain in touch after the shrinking process as well, while areas not touching after the expansion process will come back to their original "place" after the shrinking process. In the end, urban areas laying less than 200 m apart will be connected without any modification of their original shapes (except for the connection zones).
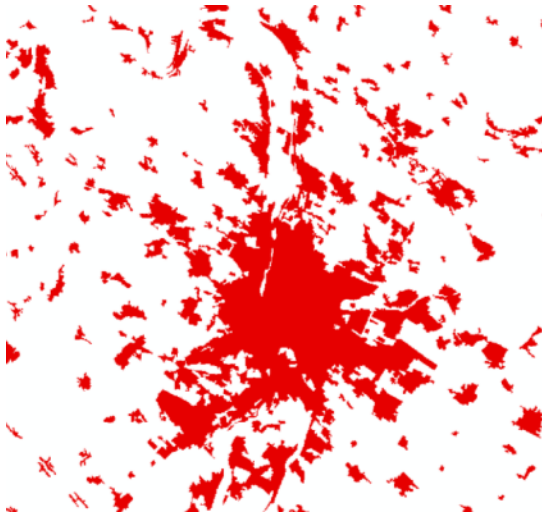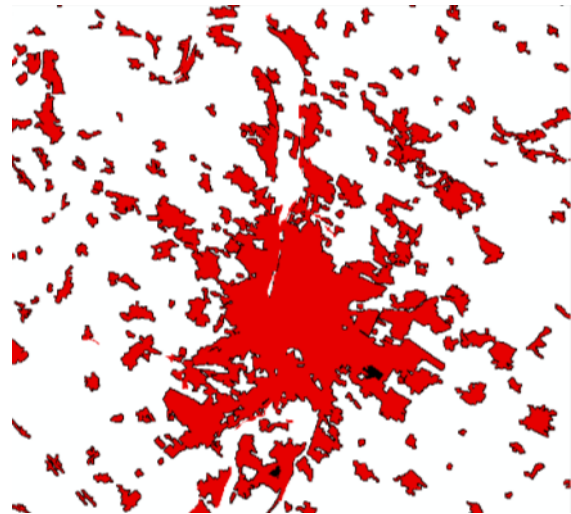
Figure 17. Pre-UMZ   *Figure 18. Expanded UMZ*
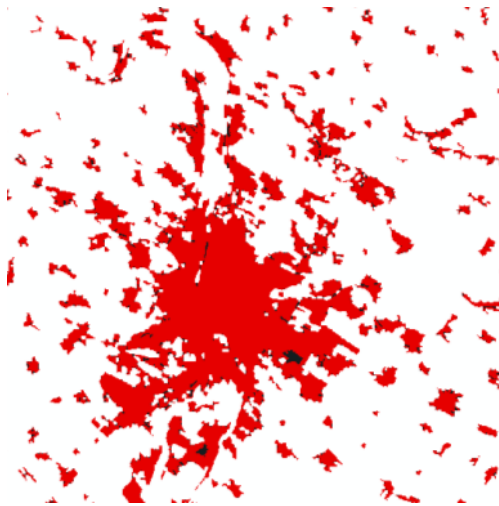
Figure 19. Shrinked UMZ

After this process, the final partial rasters (*shrinkedUMZ)* of the different zones are merged in a mosaic (*Merge to new mosaic*) and vectorised (*Raster to polygon).* Two new fields are calculated

- "IDXX" = [FID] + 1
- "Area_ha" = Area in hectares

All the polygons below 25 ha (CLC limit) are removed, obtaining the UMZ dataset (Figure 20)

FID	Shape	Area_ha	IDXX
0	Polygon	32	1
1	Polygon	70	2
2	Polygon	99	3
3	Polygon	46	4
4	Polygon	87	5
5	Polygon	56	6
6	Polygon	205	7
7	Polygon	62	8
8	Polygon	135	9

Figure 20. Attributes of the UMZ

1.4.5   Cleaning UMZ_final.shp

Finally optional steps can be made manually:

- The field IDXX can change the name to another more comprehensive. For example: UMZ00_ID

- The name "UMZ_Final.shp" can be changed to a more comprehensive one. For example. UMZ00.shp

- The shapefile can be exported to a geodatabase.

3 OTHER UMZ GEOMETRIES

3.1 CHANGES 1990 – 2000

UMZ Changes have been calculated as well, by unioning both UMZ90 and UMZ00 and looking for those polygons existing only in one year (i.e. 1990 or 2000). Most changes are Positive changes, understood as areas of urban sprawl (i.e. new UMZ areas between 1990 and 2000), while negative changes describe the reduction of a certain UMZ between 1990 and 2000 (warning: some negative changes might be due to different interpretations between 1990-2000). Only countries with CLC in both years have changes in UMZ

Steps to follow:
1. Create a file geodatabase to keep the geometries
2. Union UMZ 1990 – 2000. *Output:UMZ90_00_Full*
3. Select ("FID_UMZ1990" = -1 OR "FID_UMZ2000" = -1)
4. Export the selection. *Output:UMZ90_00_FULL_select.shp*
5. Selection of the real changes. Erasing
   a. Select the countries with **NO** CLC in both years (90-00 and 00-06) from a NUTS0 Layer (as accurate as possible)→ selected_countries.shp
   b. Clip *UMZ90_00_FULL_select.shp* + selected_countries.shp → Output: *Changes_90_00.shp*
   c. Manually work to remove the changes belonging to the selected countries. Small feature classes are left in these countries due to not perfect overlay among layers (NUTS0 and CLC) mainly in the coastal areas
6. Add and Calculate the Field "Area_ha"
7. Add a new field name "Change" type = Text with length = 1 to Change_*90_00.shp*
8. Delete unnecessary fields
9. Select "UMZ90_ID" = 0 and Calculate Value for "Change" = + . This means Urban Sprawl
10. Select "UMZ00_ID" = - 1 and Calculate Value for "Change" = - . This means Reduction of Urban Areas
11. Select "UMZ00_ID" <> - 1 and "UMZ90_ID" <> 0 and Calculate Value for "Change" = 0. This means no changes of Urban Areas

3.2 Changes 2000 – 2006

Same methodology than in the step 4.1
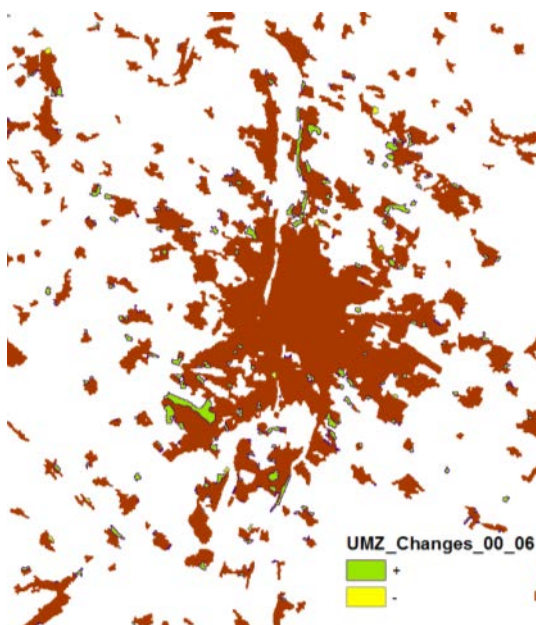
Figure 21. UMZ changes 2000- 2006 (Budapest area)

3.3 UMZs Core areas

This section shows the methodology followed to identify the main core areas inside the large UMZ combining the UMZ, 1km Reference Grid and the CORILIS smoothing. Large UMZ are defined by a population threshold of 50000 habitants.

The complete procedure in included in the Annex I, Section 0

3.3.1 Selection of Inputs
- EU GRID Reference 1km^2
- CORILIS Level1, 1km
- UMZ Population >= 50000 hab.

3.3.2 Description of the Methodology
Main steps are:
- to split each UMZ in 1km^2 cells, what allows to identify them with the EU reference Grid 1km^2. Each cell will contain the value of the CORILIS smoothing providing information of the intensity of the use of the land

- To use the same resolution in the reference grid as the Corilis, what avoids calculating zonal statistics for each UMZ cell, since the base is the same, but taking the Corilis values directly by spatial intersection.

- Final Result Output number 1 – Core Areas reference by 1km grid cell

The final results will be vector layer for each year spitted by 1km^2 cells according the 1km reference grid. This does not mean that all the cells will be 1km^2. In fact, edges cells will have smaller size (*Figure 23*).

OBJECTID *	Shape *	Area_ha	UMZ00_ID	GRIDCODE_1	Pop00	Interval	UrDom	Shape_Length	Shape_Area
1	Polygon	1514	120177	26711940	109820	>=25 and <50	27	2800	310000
2	Polygon	1514	120177	26721940	109820	>=25 and <50	26	400	10000
3	Polygon	1514	120177	26681939	109820	<25	23	2200	160000
4	Polygon	1514	120177	26691939	109820	>=25 and <50	30	3400	340000
5	Polygon	1514	120177	26701939	109820	>=25 and <50	33	3800	840000

Figure 22. Output format of the vector file, UMZUrban_dom_final
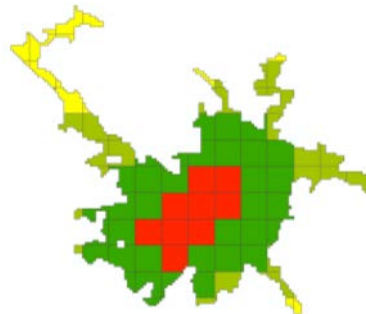
Figure 23. Example of a medium-large city. Zaragoza (Spain)

450.000 inhabitants (UMZ 2000)

- Final Result Output number 2 – Core Areas Iso-Polygons

All the cells of a UMZ inside the range 75-100% of urban dominance are aggregated in order to create new layers. Thus the Urban cores will be obtained by dissolving the selected features according to the UMZ they belong to.

OBJECTID *	Shape *	UMZ00_ID	Shape_Length	Shape_Area
1	Polygon	318	23000	20140000
2	Polygon	593	11600	4770000
3	Polygon	1401	10000	4000000
4	Polygon	2101	8000	1350000
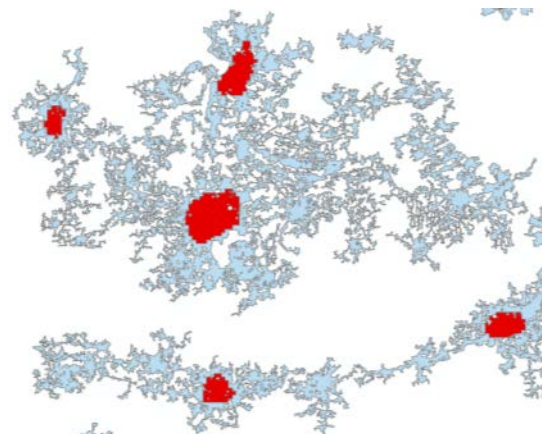
Figure 24. Core areas (red) after dissolve in Belgium area

4 RELATIONSHIPS BETWEEN UMZ90, 00, 06

The methodology is described in the annex II, Section 7. These intermediate related tables allows linking UMZ 1990, 2000 and 2006 through the IDs

5 SELECTION OF LARGE URBAN AREAS

The criterion is based on the work developed under the project ESPON Database for the UMZ[4]. UMZ 2000 version 13 are described by a name and some basic indicators.

Input data required:
- UMZ ESPON Database
- UMZ version 16
- Population Grid 2001 version 5.0

UMZ_v16 geometries are spatially different to version 13 due to the optimization of the process held in 2012. Therefore, it is not straight forward to do the link between both datasets. Each year has a different procedure starting with year 2000 as base reference.

5.1 UMZ 2000

The selection is based in spatial overlaying and expert knowledge.
- Spatial Join based on identical UMZs. One to one relationship between UMZ ESPON and UMZ v16. The selection is almost automatic.
- The non-identical implies three possible relationship:
  - Many (UMZ00v16) to one (UMZ00v13) relationship – The most common
  - One (UMZ00 v16) to many (UMZ00v13) relationship
  - Many (UMZ00 v16) to many (UMZ00v13) relationship
- The following selection was based on the expert knowledge, deciding which name is the correct one
  - Firstly, only those UMZs with population >= than 50000 hab were selected
  - Some countries like Norway, Switzerland and the Balkans were not in the population Grid 2011
    - These were added manually based on the statistics of the municipality
- A selection of 951 Large Urban areas was decided

5.2 UMZ 1990

The selection is based in the relationship between the IDs UMZ1990 and UMZ2000
The three possible relationships:
  - Many (UMZ90) to one (UMZ00) relationship – The most common

---

[4] Data is available at: http://database.espon.eu/espondb/urban

     o One (UMZ90) to many (UMZ00) relationship
     o Many (UMZ90) to many (UMZ00) relationship

The final selection about which UMZ 90 should be kept was based on:
- Population 2001 for UMZ 00 >= 45000 (This means a maximum variation of 10% of the previous threshold)
- For countries (Balkans), not included in the population grid 2001, the selection was based in a double criteria
  1. the expert knowledge
  2. the total area of the geometry. When relationships were many to one, it was selected the biggest UMZ. In all the previous cases, one UMZ was clearly bigger than the rest
- A selection of 732 Large Urban areas was decided

5.3 UMZ 2006

The selection is based in the relationship between the IDs UMZ2000 and UMZ2006

The three possible relationships:

     o Many (UMZ06) to one (UMZ00) relationship
     o One (UMZ06) to many (UMZ00) relationship – The most common Many
     o Many (UMZ06) to many (UMZ00) relationship

The final selection about which UMZ 06 should be kept was based on:
- Spatial Join based on identical UMZs. One to one relationship between UMZ 00 and UMZ 06. The selection is almost automatic.
- The non-identical is based in the relationship between the IDs UMZ2000 and UMZ2006
  - For countries (Balkans, Switzerland and Norway), not included in the population grid 2001, the selection was based in a double criteria
    - the expert knowledge
    - the total area of the geometry. When relationships were many to one, it was selected the biggest UMZ. In all the previous cases, one UMZ was clearly bigger than the rest
  - For the rest of countries,
    - a dissolve process was done to avoid the UMZ06 with IDs replicated.
    - Calculation of the population based on the grid 2001
    - Selection of the UMZ >= 50000 hab.
- A selection of 942 Large Urban areas was decided

6 ATTRIBUTES

6.1 POPULATION FIGURES

Intersecting the UMZ polygons with JRC's 2001 Population density grid (version 5.0 made by Javier Gallego) we can obtain a population figure within every single UMZ (Figure 25). In the case no population figures are available; the value -9999 has been recorded instead.

The tool used is Zonal Statistics of the UMZ vector layer above the Population GRID v5.0 to calculate the basic statistics per UMZ (Fields created: count, area, min, max, range, sum, mean, and std fields, Majority, minority, median, and variety (This 4 only if the input contain integer data (Raster or Vector))

Steps to follow:
- Zonal Statistics between the UMZ 2000 vector layer and the Population Grid
  - Due to the large number of feature class (more than 100000) this process is very slow or it could not work properly. Therefore, a new script (UMZ_Stats) in python has been created to facilitate the work. It has been included in the ArcToolBox "Urban_UMZ" to run directly from ArcGIS
- Identified the UMZ90_ID in the output table as Value_
- Sum field keeps the population for each UMZ. Notice that the real value is the number divided by 100. For further information see the Downscaling methodology for the population GRID:
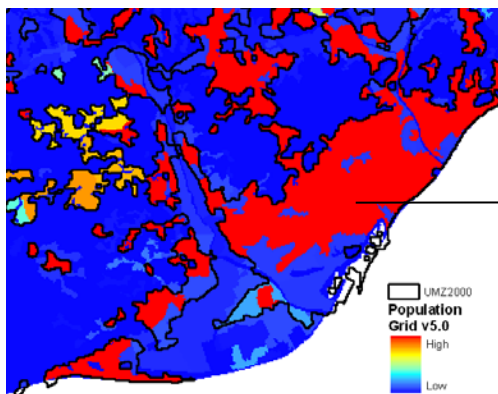
Figure 25. Population distribution in the UMZ   Figure 26. Population as an attribute
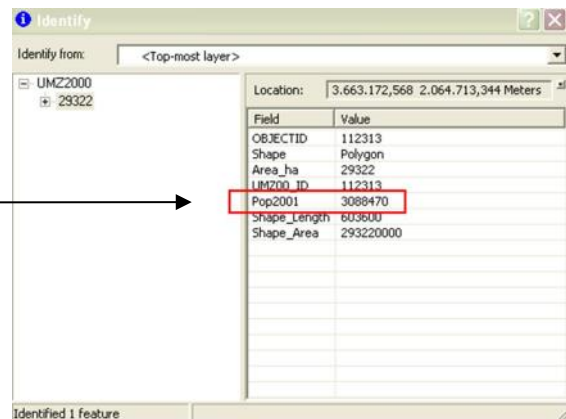
6.2 COMMUNE CODES

Using the EuroBoundaryMap version 6.0 dataset or any other source (Commune boundaries), commune attributes can be added to each UMZ. A relationship class between both datasets has been created in the way that once you identify a UMZ you get the list and attributes of all communes or parts of commune being part of it:

Steps to follow (Figure 27):

1. Make sure that the UMZs feature classes are in a geodatabase
2. Intersect the UMZ with the commune boundaries: "UMZ2000_vs_Communes"
3. Remove unnecessary fields
4. Export the table of this new feature class. Name: "Communes_UMZ2000"
5. Create the relationship in ArcCatalog
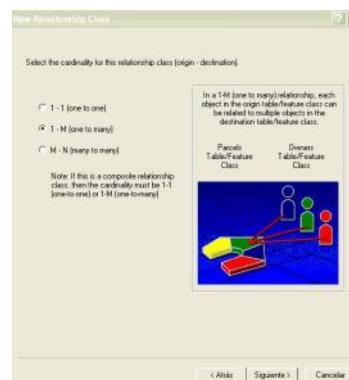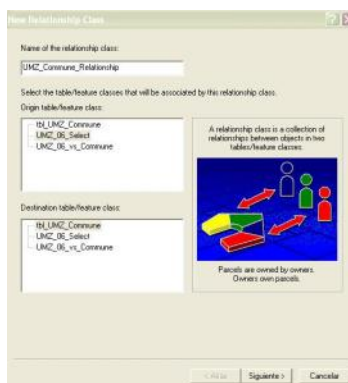
Figure 27. Procedure to create the relationship UMZ – Commune 2006

Therefore, each UMZ is related with its commune, the relationship is 1 to many (Figure 28 , Figure 29)

Figure 28. UMZ 2006 Barcelona and the Communes 2006
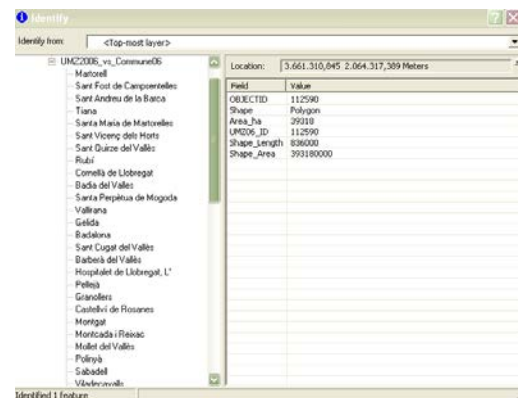
Figure 29. Attributes of the UMZ 2006

This step is redone with each UMZ feature class (1990, 2000, 2006)

6.3 AVAILABLE ATTRIBUTES

The attributes available for each UMZ dataset follow this structure (example for UMZ2000):

All_UMZ_Level1

UMZOO: All UMZs processed for CLC 2000

Feature class - polygon

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Area_ha: Area in hectares

UMZOO_centroid: centroids of all UMZs processed for CLC 2000

Feature class - points

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Area_ha: Area in hectares

UMZOO_ Pop2000: All UMZs processed for CLC 2000, with the Population.

Feature class - polygon

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Area_ha: Area in hectares
- Pop_2000: population based on year 2001

UMZOO_Urban_Dominance

Feature class - polygon

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- ID_1kmGRID: 1km grid identifier
- UrDom: Percentage of urban dominance
- Interval: Each 1km grid cell belongs to an interval of urban dominance (25, 50, 75, 100)
- Area_ha: Area in hectares

UMZOO_Urban_Dominance75_gridID

Feature class – polygon – Only when Interval >= 75 and by divided by 1km grid cells

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- ID_1kmGRID: 1km grid identifier
- UrDom: Percentage of urban dominance
- Interval: Each 1km grid cell belongs to an interval of urban dominance (25, 50, 75, 100)
- Area_ha: Area in hectares

UMZOO_Urban_Dominance75_UMZID

Feature class – polygon – Only when Interval >= 75 and by divided by 1km grid cells

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Area_ha: Area in hectares

Large_UMZ_Level2

UMZ00_Large: Large UMZs processed for CLC 2000

Feature class - polygon

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Name: It is based on ESPON DB
- Population2001: population based on year 2001
- Area_ha: Area in hectares

UMZ00_Large_Urban_Dominance
Feature class – polygon. Only Large Urban Areas

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- ID_1kmGRID: 1km grid identifier
- UrDom: Percentage of urban dominance
- Interval: Each 1km grid cell belongs to an interval of urban dominance (25, 50, 75, 100)
- Area_ha: Area in hectares

UMZ00_ Large_Urban_Dominance75_gridID

Feature class – polygon – Only Large Urban Areas and when Interval >= 75 and by divided by 1km grid cells

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- ID_1kmGRID: 1km grid identifier
- UrDom: Percentage of urban dominance
- Interval: Each 1km grid cell belongs to an interval of urban dominance (25, 50, 75, 100)
- Area_ha: Area in hectares

UMZ00_ Large_Urban_Dominance75_UMZID

Feature class – polygon – Only Large Urban Areas and when Interval >= 75 and by divided by 1km grid cells

Attributes

- UMZ00_ID: Unique identifier of each UMZ 2000
- Area_ha: Area in hectares

UMZ CHANGES 1990-2000

Feature class - polygon

Attributes

- UMZ90_ID: Unique identifier of each UMZ 1990
- UMZ00_ID: Unique identifier of each UMZ 2000
- CHANGES: + (Urban sprawl between 1990 and 2000) ,– (UMZ reduction), 0 (no changes, original UMZ 2000)
- Area_ha: Area in hectares

UMZ CHANGES 2000-2006

Feature class - polygon

Attributes

- UMZ06_ID: Unique identifier of each UMZ 2006
- UMZ00_ID: Unique identifier of each UMZ 2000
- CHANGES: + (Urban sprawl between 2000 and 2006) ,– (UMZ reduction), 0 (no changes, original UMZ 2000)
- Area_ha: Area in hectares

7 ANNEX I. NEW APPROACHES TO DEFINE UMZS CORE AREAS

7.1 INTRODUCTION

The goal of this paper is to show the methodology followed to identify the main core areas inside the large UMZ combining the UMZ, 1km Reference Grid and the CORILIS smoothing. Large UMZ are defined by a population threshold of 50000 habitants.

The purpose of CORILIS is to calculate "intensities" or "potentials" of a given theme in each point of a territory. CORILIS results into probability surfaces (varying from 0 to 100) for the presence of a certain Corine land cover (CLC) class within a defined smoothing radius.

The outcome will be a sublevel 2 of UMZ dependant of the UMZ itself and the EU 1Km Reference Grid. The core areas of the UMZ depending on the level of artificial predominance according to some predefined intervals created to facilitate the comprehension of the sublevel2 of UMZ.

7.2 SELECTION OF INPUTS

- **EU GRID Reference 1km²** recommended by EEA[5] and elaborated by ETC/SIA – UAB.

- **CORILIS Level1, 1km** CORILIS level 1 is a customized aggregation level for land accounting purposes. Aggregation of all artificial areas has been done by simple addition, according to the below mapping table. The resolution used is 1km. Versions for years 1990, 2000 and 2006 are used according to each case´s needs.

- **UMZ Population >= 50000 hab.** The UMZ feature classes are crosses with the Population Grid 2000 (EEA & Gallego, 2010) in order to summarise the total population of each UMZ. Population grid is only available for 2000, so the UMZ2000 with population >= 50000 hab (**UMZ00_50k**) is taken as a reference to calculate the main core areas for the years 1990 and 2006, based on the spatial concordance with the UMZ 2000

7.3 DESCRIPTION OF THE METHODOLOGY

The purpose of this section is not to describe in detail all the GIS steps done but to give a general overview of the process.

The goals were two:

- to split each UMZ in 1km² cells, what allows to identify them with the EU reference Grid 1km². Each cell will contain the value of the CORILIS smoothing providing information of the intensity of the use of the land

---

[5] http://www.eea.europa.eu/data-and-maps/data/eea-reference-grids

- To use the same resolution in the reference grid as the Corilis, what avoids calculating zonal statistics for each UMZ cell, since the base is the same, but taking the Corilis values directly by spatial intersection.

7.3.1 **Focal statistics**

Due to Corilis resolution, some areas of the UMZ do not match any of its values. To solve this fact a focal statistics is done. It calculates for each input cell location a statistic of the values within a specified neighborhood around it. Thus when no matching values exist, the average of the neighbors will be assigned.

7.3.2 **Reference grid**

The reference Grid must be vectorized and its extent adjusted, to perfectly match the Corilis values.

7.3.3 **Intersection of inputs**

Once every UMZ cell matches a CORILIS value, an intersection operation will be performed among CORILIS; UMZ and the reference Grid.
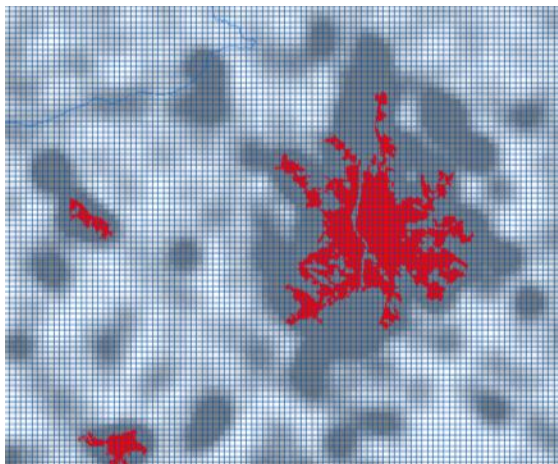

Figure 30. Examle of the Interset inputs, (UMZ, Corilis and Grid 1km)

7.3.4 **Fine tunning**

The feature derived from the intersec operation must be organized, new fields properly renamed and unnecessary fields deleted. The Corilis values are identified as Urban dominance (UrDom).
The output layer will contain, at least, the next attributes:

FID	Shape	Area_ha	UMZ00_ID	Pop00	GRIDCODE_1	UrDom
0	Polygon	1514	120177	109820	26711940	27
1	Polygon	1514	120177	109820	26721940	26
2	Polygon	1514	120177	109820	26681939	23
3	Polygon	1514	120177	109820	26691939	30
4	Polygon	1514	120177	109820	26701939	33
5	Polygon	1514	120177	109820	26711939	35
6	Polygon	1514	120177	109820	26721939	33

Figure 31. Output format of the vector file, UMZ2000 versus Grid1km

Thus every cell can be identified in a unique way with the grid gridcode and the UMZ ID

7.3.5 Making pre-defined interval

The user will be able to create their own intervals though easy GIS operations or simply by visualisation techniques. However, a first rank of intervals is created after analysing the results by urban experts in order to facilitate next analysis and visualization.
Several intervals are created according with Urban dominance values:

Value of CORILIS Level 1	GIS Syntaxes
< 25	"UrDom" < 25
>= 25 AND < 50	"UrDom" >= 25 AND "UrDom" < 50
>= 50 AND < 75	"UrDom" >= 50 AND "UrDom" < 75
>= 75"	""UrDom" >= 75

7.4 FINAL RESULT (OUTPUT)

7.4.1 Create Core Areas 1km2 grid reference -Polygons

Following the explanation of section 3.3.2, the final results will be vector layer for each year spitted by 1km^2 cells according the 1km reference grid. This does not mean that all the cells will be 1km^2. In fact, edges cells will have smaller size (*Figure 23*).

OBJECTID *	Shape *	Area_ha	UMZ00_ID	GRIDCODE_1	Pop00	Interval	UrDom	Shape_Length	Shape_Area
1	Polygon	1514	120177	26711940	109820	>=25 and <50	27	2800	310000
2	Polygon	1514	120177	26721940	109820	>=25 and <50	26	400	10000
3	Polygon	1514	120177	26681939	109820	<25	23	2200	160000
4	Polygon	1514	120177	26691939	109820	>=25 and <50	30	3400	340000
5	Polygon	1514	120177	26701939	109820	>=25 and <50	33	3800	840000

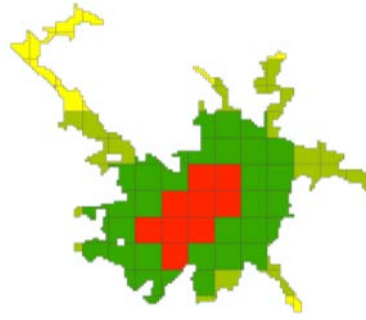*Figure 32. Output format of the vector file, UMZUrban_dom_final*

Figure 33. Example of a medium-large city. Zaragoza (Spain), 450.000 inhabitants (UMZ 2000)

In the previous picture, it is easy to identify only one core area for each UMZ. However, the purpose of this exercise was to focus in very large UMZ, for example Belgium case. In Figure 34 is really easy to identify the core areas:
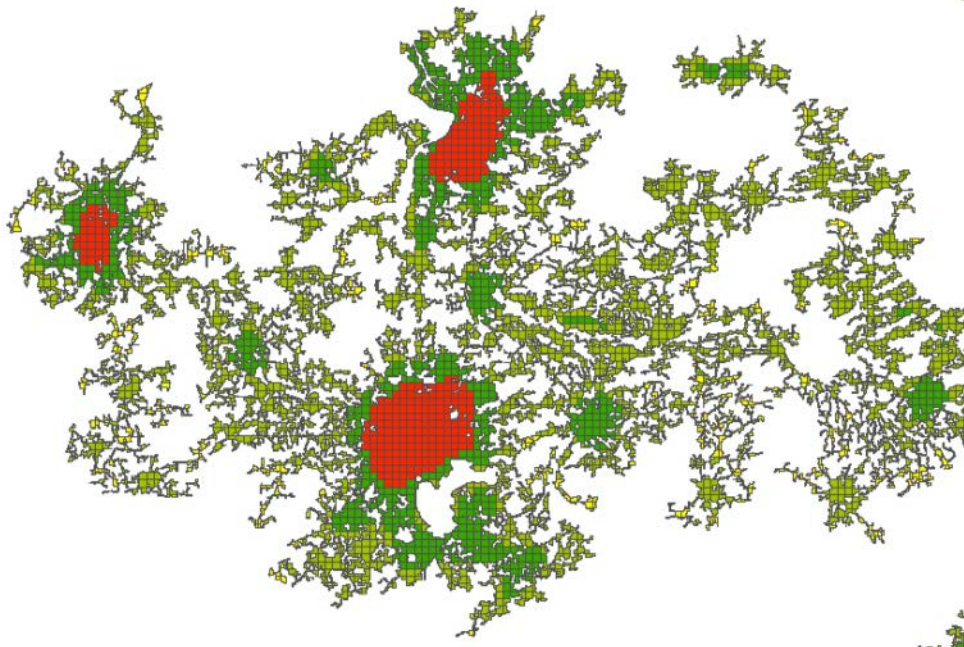

Figure 34. Example of a very large UMZ. Brussels-Gent-Antwerpen, (Belgium)
3.769.016 inhabitants in 2000

7.4.2 **Create Core Areas Iso-Polygons**

All the cells of a UMZ inside the range 75-100% of urban dominance are aggregated in order to create new layers. Thus the Urban cores will be obtained by dissolving the selected features according to the UMZ they belong to.

OBJECTID *	Shape *	UMZ00_ID	Shape_Length	Shape_Area
1	Polygon	318	23000	20140000
2	Polygon	593	11600	4770000
3	Polygon	1401	10000	4000000
4	Polygon	2101	8000	1350000

Figure 35. Core areas (red) after dissolve in Belgium area
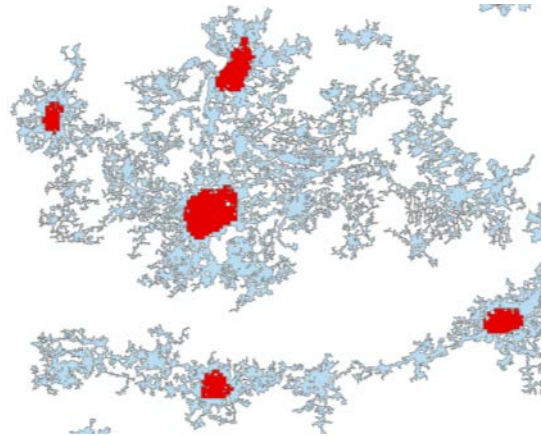
7.4.3 **Identify secondary cores**

This fact can be done playing around with the intervals or using a more research approach. Population can be used or analysing the STD of the cells per object/country or any other border.

8 ANNEX III : HOW TO DEFINE THE JUNCTION ELEMENTS LINKED TO WATER

8.1 INTRODUCTION

This document describes how to obtain the areas linked by rivers, by identifying the junction elements between them corresponding to water bodies and use them to select the rivers that should be included in the UMZ.

The output feature classes will be:

- Water_ECC_Clip.shp,
- Junction_elements300_water.shp

After several trials in different European cities with predominant linear features (e.g. big rivers), 300m has been considered as the optimal distance for the aggregate polygon and clip functions in the process of adding water bodies to the UMZs. Some examples can be seen below. (see section 8.3)

8.2 METHODOLOGY

8.2.1 Input: Feature Class

Shapefiles and Geodatabase are accepted to work with, although the work was done with shapefile due to its faster processing.

- **Enlarged_core_classes.shp**
- **Water.shp**

8.2.2 ArcGIS

- Aggregate polygons of *Enlarged_core_classes.shp* with 300m distance, choosing to keep orthogonal shape avoid strange shapes.

  This way, areas within 300m will be gathered in Enlarged_core_classes_agg300.shp, to be used afterwards.
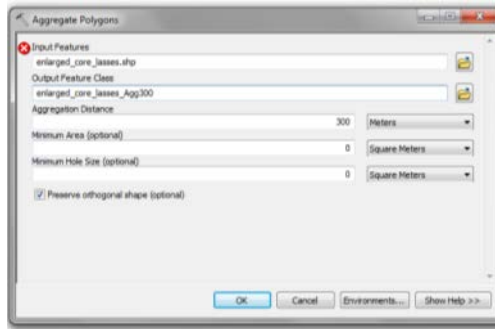
Figure 36. Aggregate polygons tool

- **Erase**: To obtain the **junction_elements300.shp** erase **Enlarged_core_classes.shp** from **Enlarged_core_classes_agg300.shp**.

Figure 37. Erase tool

Figure 38. Output example after aggregate and erase tools

- **Select by location: junction_elements300.shp** *can help to determine the areas linked by water bodies, hence to include them in the UMZ when necessary and avoid other junction elements different than water its necessary to select*

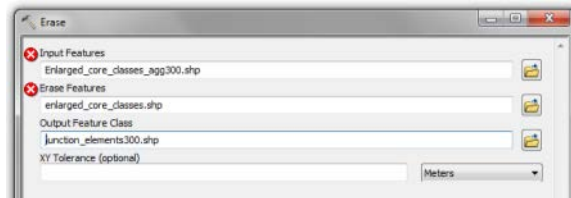*those elements **touching water bodies**, what reflects the case of rivers working as a link. This selection will help to determine the urban areas where the river is a link, and only select those ones, avoiding merging urban areas not related, and separated by rivers. **junction_elements_water.shp.***

Figure 39. Output example after select by location number 1

- **Select by location:** to determine the urban areas where the river is a link ***Enlarged_core_classes.shp*** that touch ***junction_elements_water.shp*** have to be selected→ ***Enlarged_core_classes_water_linked.shp***

Figure 40. Output example after select by location number 2

- 300m **Buffer**[6] to ***Enlarged_core_classes_water_linked.shp*** → ***Enlarged_core_classes_water_linked_Buf300.shp*** will provide the clip shape to determine the river area connected to the urban zones**.**

---

[6] Some strange results, try with raster format, expand.

- **Select by location: water bodies in contact with the** areas linked by water (***Enlarged_core_classes_water_linked.shp)*** will be selected, ***water_selected.shp***.

- **Clip:** following a similar procedure as the one applied to the road features; they will be clipped with ***Enlarged_core_classes_water_linked_Buf300.shp.*** This way only the rivers related to those areas as a linking element will be included, not just by simple touching. (although some exceptions still remain). Output ***Water_ECC_Clip.shp***

8.3 MAIN DIFFERENCES BETWEEN USING 200M INSTEAD OF 300M FOR WATER CLASSES AND THE NEW WATER CLASSIFICATION
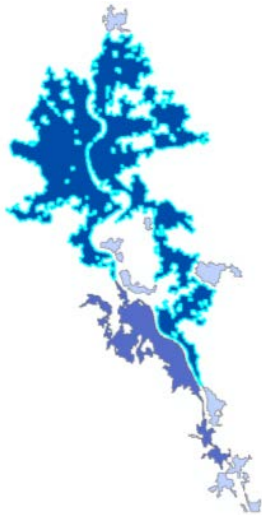
- Köln – Bonn area

Figure 41. UMZ delineation using 200m distance   *Figure 42. UMZ delineation using 300m distance*
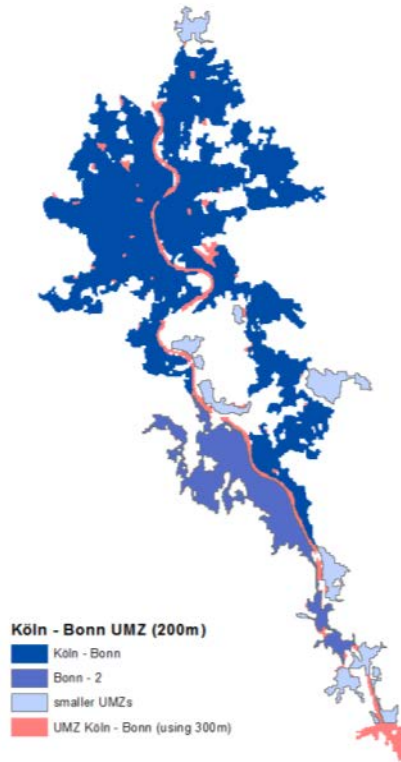
Figure 43. Comparison of UMZ delineated with 200m or 300 m approach.

Using 200m distance in the water bodies addition phase of the UMZs delineation process splits the UMZ Köln-Bonn in two (Köln-Bonn and Bonn-2) due to river discontinuity, although the right side of Bonn is still connected to Köln through other terrains non related to water bodies. Thus the urban area of Bonn is divided in two. Moreover, for the 200m approach, the river, working as a link in the UMZ, is not completely included in it. The 300m method gathers the UMZs in one, including the river within them, which in most of the cases works as a link. This result is considered more appropriate from a conceptual point of view.

As mentioned, in previous versions some problems were found when delineating UMZs at river´s edges, especially with the small ones. In some cases the river was mistakenly included between them, in consequence some not linked urban zones were gathered by the river. With this new approach some of these problems were solved by including the river only in the cases when it acts as a link between UMZs. As explained this decision was based on their proximity. Below some examples to illustrate these improvements:

- Köln – Bonn

Both UMZs are connected, in version 16 the river is not the link, but the terrains around the area.
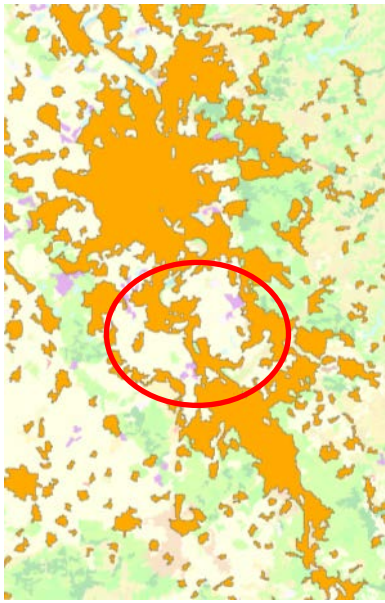
UMZ version 15
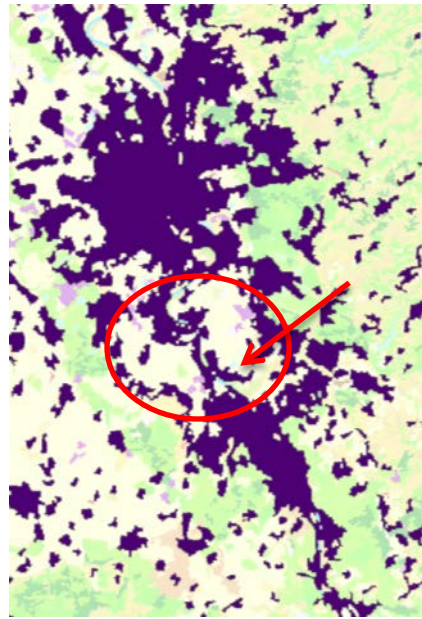
UMZ version 16

UMZ version 15 and 16

Figure 44. Köln – Bonn UMZ versions comparison

- Budapest area (Hungary)

UMZ version 15   UMZ version 16

UMZ version 15 and 16

Figure 45. Budapest UMZ versions comparison

- Hamburg outskirts (Germany)

UMZ version 15   UMZ version 16

UMZ version 15 and 16

Figure 46. Hamburg UMZ versions comparison

- Rouen (France)
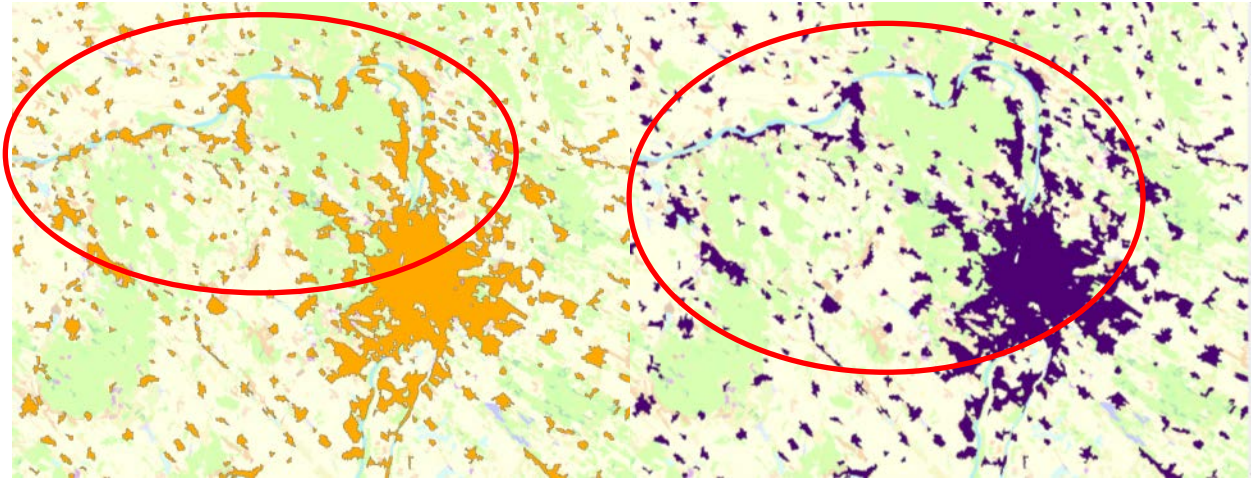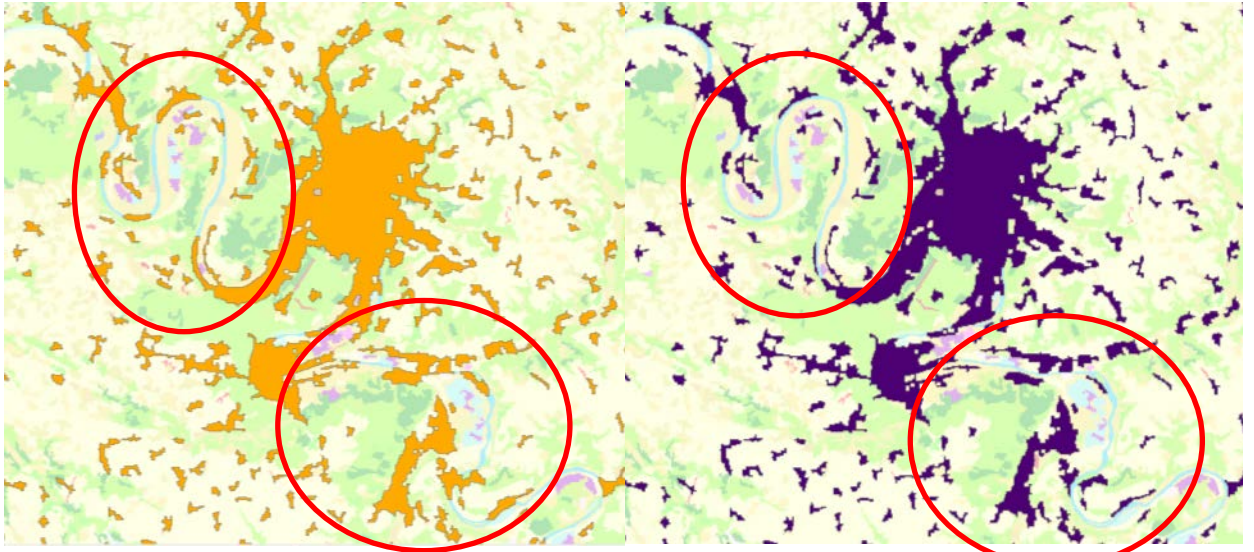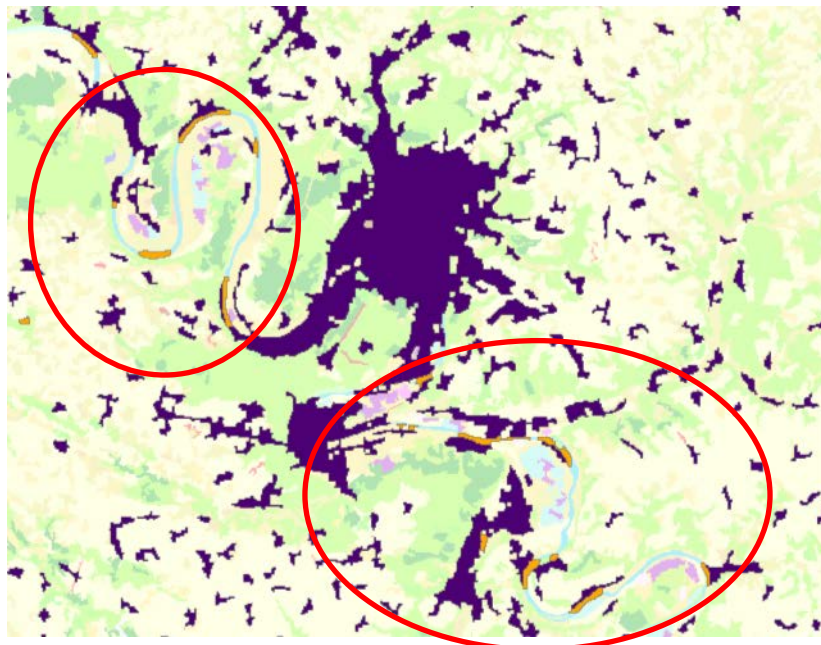

UMZ version 15   UMZ version 16

UMZ version 15 and 16
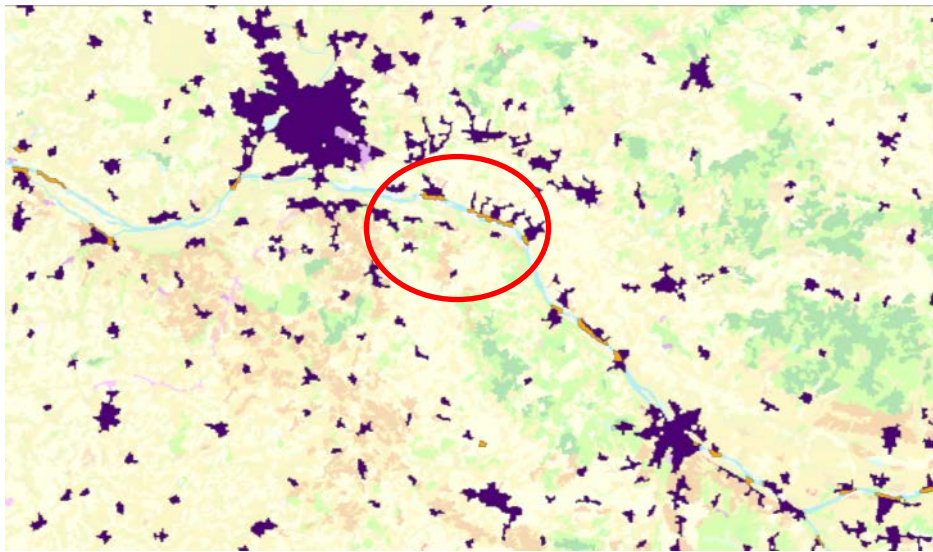
Figure 47. Rouen UMZ versions comparison

- Angers (France)

UMZ version 15   UMZ version 16

UMZ version 15 and 16

Figure 48. Angers UMZ versions comparison

9 ANNEX II. HOW TO CREATE THE RELATED TABLES BETWEEN UMZ 1990, 2000, 2006

9.1 INTRODUCTION

This document describes how to create the related table between the following feature classes: UMZ90, UMZ00, UMZ06. The output tables will be:

- UMZ1990_UMZ2000_relationship
- UMZ2000_UMZ2006_relationship

9.2 METHODOLOGY

9.2.1 Input: Feature Class

Shapefiles and Geodatabase are accepted to work with.

- UMZ1990. Last version with 98959 features
- UMZ2000. Last version with 112168 features
- UMZ2006. Last version with 81603 features

Shapefile

Create a specific folder to keep the UMZ feature classes.

Geodatabase

Create a personal database. It can be done in the ArcCatalog. It has to include all the input feature classes.

Name: *"UMZs_join.mdb"*

9.2.2 ArcGIS

- Load UMZ1990, UMZ2000, UMZ2006
- Create a new feature class: Union UMZ1990 – UMZ2000
  Name: "*UMZ90_00Union_total*"
- Create a new feature class: Union UMZ2000 – UMZ2006
  Name: "*UMZ00_06Union_total*"
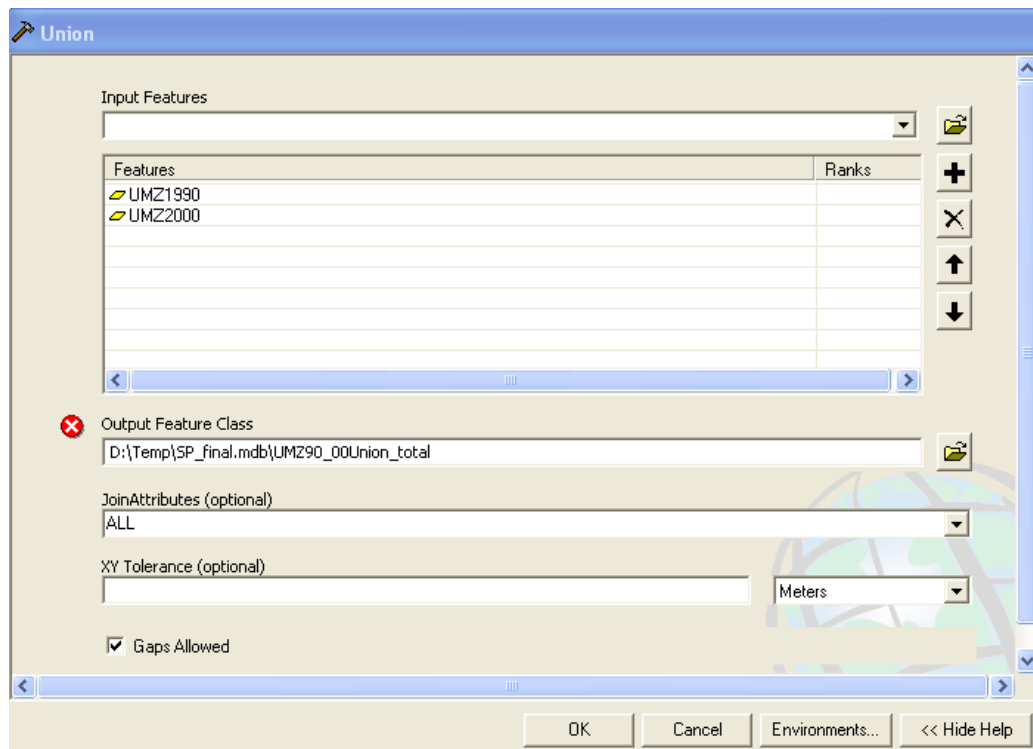
Figure 49. Union tool for ID relationship calculation

9.3 MS ACCESS

9.3.1.1 Inputs

- *UMZs_Relations_BASE.mdb.* Last location in:

  O:\2009_subvention\2_7_2_Urban\Tasks\UMZ\UMZ_table_relation

  This database will be used as template to create the final tables.

- *UMZ90_00Union_total* feature class

- *UMZ00_06Union_total.*

Notice two issues:

✓ This database and the respective queries have been developed to create the relationship table between UMZ1990 – UMZ2000 and UMZ2000 – UMZ2006. To create new tables these queries have to re-created or updated.

✓ Changing the names of the tables or the queries, makes the DB not to work

9.3.1.2 Methodology

- Open the database: "*UMZs_Relations_BASE.mdb*"
- Save as "UMZs_Relations_1990_2000_2006.mdb"
- Import the next tables:

- o UMZ90_00Union_total

  Name: *"tbl_UMZ90_00Union_total"*.

- o UMZ00_06Union_total

  Name: *"tbl_UMZ00_06Union_total"*

- o UMZ1990

  Name: *"tbl_UMZ1990"*. Required fields OBJECTID, and UMZ90_ID (at least).

- o UMZ2000

  Name: *"tbl_UMZ2000"*. Required fields OBJECTID, and UMZ00_ID (at least).

- o UMZ2006

  Name: *"tbl_UMZ2006"*. Required fields OBJECTID, and UMZ06_ID (at least).

- o Delete unnecessary fields to save space.

  These tables come from shapefiles or geotadabases. ArcCatalog can be used to export, or access, or any other tool. The field names should be respected to those mentioned above

- Run the queries according to the respective order. The description of how to use each of them is the picture below or in the description inside the *"UMZs_Relations_BASE.mdb"*

Two are the tables that related the IDs of the UMZ 1990, 2000, 2006.

- Relation IDs UMZ1990 - UMZ2000
- Relation IDs UMZ2000 - UMZ2006

Example of the methodology with UMZ1990 – UMZ2000.

9.3.1.3 Steps

GIS Software

In a GIS software, union UMZ1990 and UMZ2000 --> UMZ90_00Union_Total. This geometry contains all the attributes from both layers.

DB Software

In a DB software serveral queries are developed to obtain the final result.

1. Query to create a table named "tbl_UMZ90_00Union_total"" with the records with ID90 and ID00 > 0 from "UMZ90_00Union_Total" table --> "qry_01_Union_with_ID90_ID00"

2. Query to create an identical table (only the structure of the table) to "tbl_UMZunion_with_ID90_ID00" with the name "UMZ1990_UMZ2000_Relation_all". It will be used to add new records which don´t fit the sentence: ID90 > 0 AND ID00 > 0 --> "qry_02_create_final_UMZ1990_UMZ2000_Relation_all"

3. Query to identify the records (feature classes) from UMZ1990 which are not in tbl_UMZunion_with_ID90_ID00. These records only exist in UMZ1990 and not in UMZ2000 --> "qry_UMZ1990_not_coincident_tbl_UMZunion_with_ID90_ID00"

4. Query to identify the records (feature classes) from UMZ2000 which are not in tbl_UMZunion_with_ID90_ID00. These records only exist in UMZ2000 and not in UMZ1990 --> "qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID90_ID00"

5. Add the selected records in "qry_UMZ1990_not_coincident_tbl_UMZunion_with_ID90_ID00" to the table "tbl_UMZ1990_UMZ2000_Relation_all" --> "qry_03_add_missing_UMZ1990_to_UMZ1990_UMZ2000_relation_all"

6. Add the selected records in "qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID90_ID00" to the table "tbl_UMZ1990_UMZ2000_Relation_all" --> "qry_04_add_missing_UMZ2000_to_UMZ1990_UMZ2000_relation_all"

7. Query to create an identical table to "tbl_UMZ1990_UMZ2000_Relation_all" named "Relation_IDs_UMZ1990_UMZ2000" which will store the final relationship table UMZ1990 - UMZ2000 --> "qry_05_create_final_UMZ1990_UMZ2000_relationship_IDs"

8. Query to update "tbl_UMZ1990_UMZ2000_relationship_IDs". Records with value 0 for ID90 update to -9999 --> "qry_06_update_final_UMZ1990_UMZ2000_relationship_ID90"

9. Query to update tbl_UMZ1990_UMZ2000_relationship_IDs. Records with value 0 for ID00 update to -9999 --> "qry_07_update_final_UMZ1990_UMZ2000_relationship_ID00"

qry_01_Union_with_ID90_ID00	Create a table with the records with ID90 and ID00 > 0 in "tbl_UMZ90_00Union_total"
qry_02_create_final_UMZ1990_UMZ2000_Relation_all	Create an identical table to "tbl_UMZunion_with_ID90_ID00". It will be used to add new records which don´t fit the sentence: ID90 > 0 AND ID00 > 0
qry_03_add_missing_UMZ1990_to_UMZ1990_UMZ2000_relation_all	Add the selected records in "qry_UMZ1990_not_coincident_tbl_UMZunion_with_ID90_ID00" to the table "tbl_UMZ1990_UMZ2000_Relation_all"
qry_04_add_missing_UMZ2000_to_UMZ1990_UMZ2000_relation_all	Add the selected records in "qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID90_ID00" to the table "tbl_UMZ1990_UMZ2000_Relation_all"
qry_05_create_final_UMZ1990_UMZ2000_relationship_IDs	To create an identical table to "tbl_UMZ1990_UMZ2000_Relation_all" named "tbl_UMZ1990_UMZ2000_relationship_IDs" which will store the final relationship table UMZ1990 - UMZ2000
qry_06_update_final_UMZ1990_UMZ2000_relationship_ID90	To update tbl_UMZ1990_UMZ2000_relationship_IDs. Records with value 0 for ID90 update to -9999
qry_07_update_final_UMZ1990_UMZ2000_relationship_ID00	To update tbl_UMZ1990_UMZ2000_relationship_IDs. Records with value 0 for ID00 update to -9999
qry_08_Check_all_UMZ1990_NOTin_UMZ1990_UMZ2000_relationship_IDs	To check if any UMZ1990 is not in tbl_UMZ1990_UMZ2000_relationship_IDs
qry_09_Check_all_UMZ2000_NOTin_UMZ1990_UMZ2000_relationship_IDs	To check if any UMZ2000 is not in tbl_UMZ1990_UMZ2000_relationship_IDs
qry_11_Union_with_ID00_ID06	Create a table with the records with ID00 and ID06 > 0 in "tbl_UMZ00_06Union_total"
qry_12_create_final_UMZ2000_UMZ2006_Relation_all	Create an identical table to "tbl_UMZunion_with_ID00_ID06". It will be used to add new records which don´t fit the sentence: ID00 > 0 AND ID06 > 0
qry_13_add_missing_UMZ2000_to_UMZ2000_UMZ2006_relation_all	Add the selected records in "qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID00_ID06" to the table "tbl_UMZ2000_UMZ2006_Relation_all"
qry_14_add_missing_UMZ2006_to_UMZ2000_UMZ2006_relation_all	Add the selected records in "qry_UMZ2006_not_coincident_tbl_UMZunion_with_ID00_ID06" to the table "tbl_UMZ2000_UMZ2006_Relation_all"
qry_15_create_final_UMZ2000_UMZ2006_relationship_IDs	To create an identical table to "tbl_UMZ2000_UMZ2006_Relation_all" named "tbl_UMZ2000_UMZ2006_relationship_IDs" which will store the final relationship table UMZ2000 - UMZ2006
qry_16_update_final_UMZ1990_UMZ2000_relationship_ID00	To update tbl_UMZ2000_UMZ2006_relationship_IDs. Records with value 0 for ID00 update to -9999
qry_17_update_final_UMZ1990_UMZ2000_relationship_ID06	To update tbl_UMZ2000_UMZ2006_relationship_IDs. Records with value 0 for ID06 update to -9999
qry_UMZ1990_not_coincident_tbl_UMZunion_with_ID90_ID00	To identify the records (feature classes) from UMZ1990 which are not in tbl_UMZunion_with_ID90_ID00. These records only exist in UMZ1990 and not in UMZ2000
qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID00_ID06	To identify the records (feature classes) from UMZ2000 which are not in tbl_UMZunion_with_ID00_ID06. These records only exist in UMZ2000 and not in UMZ2006
qry_UMZ2000_not_coincident_tbl_UMZunion_with_ID90_ID00	To identify the records (feature classes) from UMZ2000 which are not in tbl_UMZunion_with_ID90_ID00. These records only exist in UMZ2000 and not in UMZ1990
qry_UMZ2006_not_coincident_tbl_UMZunion_with_ID00_ID06	To identify the records (feature classes) from UMZ2006 which are not in tbl_UMZunion_with_ID00_ID06. These records only exist in UMZ2006 and not in UMZ2000

Figure 50. Queries developed to calculate the ID relationship

9.3.1.4   Final Step.
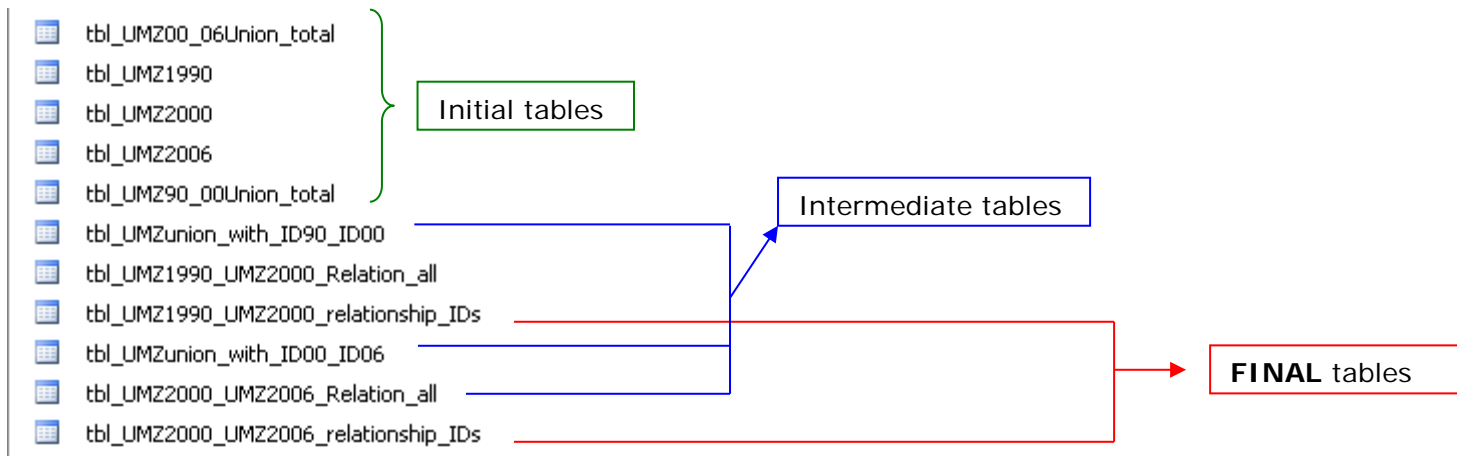
The final tables in our database will be:

Figure 51. query schema to calculate the ID relationship

This database can be quite big in size, 1GB. To reduce the size, just the important tables can be extracted to other database, excel sheets, dbf, etc. or delete unnecessary fields and tables

- tbl_UMZ1990_UMZ2000_relationship_IDs - UMZ1990_UMZ2000_relationship
- tbl_UMZ2000_UMZ2006_relationship_IDs - UMZ2000_UMZ2006_relationship

For example, a new database named *"UMZs_Relations_90_00_06_Final.mdb"* can be created and the previous tables imported.

10 ANNEX IIII. MODEL BUILDER UMZ

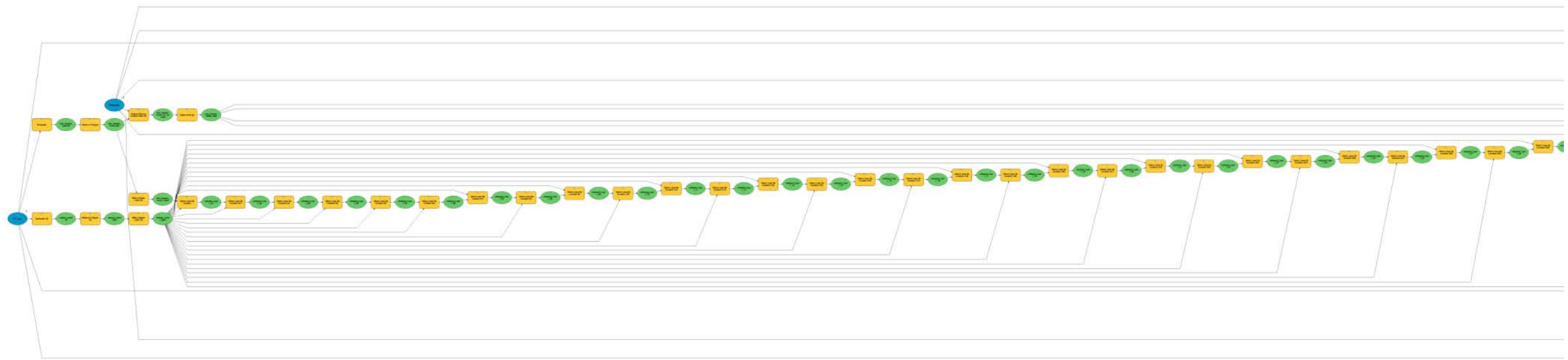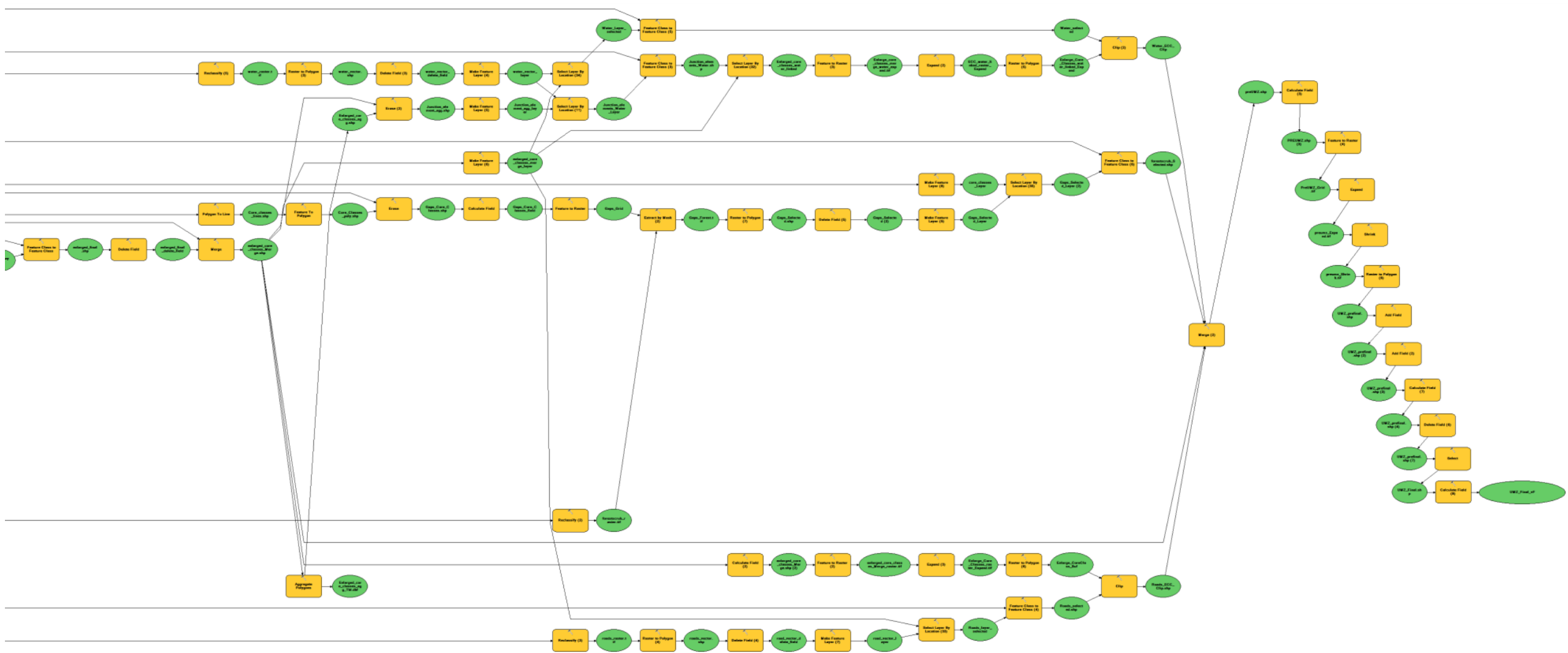*Figure 52. UMZ Model. First steps including iterative process, Select by location*

Figure 53. UMZ Model. Last steps the aggregation of extra classes