TECHNICAL REPORT

ESPON

THE FUNCTIONAL URBAN AREAS DATABASE

CONTENT

- As a joint venture between 3 challenges of the Espon DB project (Urban data, Local data and Time series) and starting from the results of the previous Espon program we provide an update of the database of the Functional Urban Areas (FUAs) and Morphological Urban Areas (MUAs), as well as their inter-relations. Not only is it enhanced, it is also fundamentally enriched by the quality of the data provided, as the Functional Urban Areas (FUAs) are now delineated for most of the European countries of the Espon space at the LAU2 level.
- The FUAs are defined as labor basins of the MUAs which are themselves defined as densely populated areas, all this independently from any national, administrative or political definitions, but based instead on pure statistics.
- The main quality and advantage of these FUAs are their simple and universal definition throughout Europe, making them comparable in all the countries where they were delineated.
- Moreover we have also produced a list of indicators for these FUAs.

ESPON 2013 DATABASE

MARCH 2011

EUROPEAN UNION
Part-financed by the European Regional Development Fund
INVESTING IN YOUR FUTURE

18 pages

LIST OF AUTHORS

Didier Peeters, Free University of Brussels, IGEAT

Contact

dpeeter1@ulb.ac.be
tel. +32 2 650 50 77

TABLE OF CONTENTS

1 General Methodology

Theoretically each MUA has a FUA, each FUA has a MUA near its center. But reality can be less straightforward and we find a good number of slightly more or much more complex cases.

First a MUA can be a secondary MUA in another's FUA. This happens when the extent of a FUA takes in a small MUA or when the population of the MUA has a commuting rate higher than the chosen threshold to another MUA. Most cases are obvious but a few are not, since the commuting flows are measured at the level of the LAU2 and a MUA can have several of these: one can therefore see that one part of a MUA is commuting toward another one while the rest of it is not. Those cases were solved by first checking if the secondary MUA was landlocked into the rest of the «main» FUA which would lead to consider the secondary MUA as actually part of the main FUA, second considering the literature and the commonly accepted facts. We don't see any major rule that could lead us to elaborate a methodology to classify these rare cases.

Second a FUA can have two «twin» MUAs (or even more) when the commuting flows from one to another are crossing each other (they are usually touching each other) and the LAU2s around them send commuters on a relatively high and equivalent level to the both of them. An example of this is Douai and Lens, or Béthune and Bruay-la-Buissière near Lille in France, or Locarno-Bellinzona in Switzerland.

We must here mention an unfortunate limitation of the database: except for the French residents there are no transnational commuters data provided, which is a non-sense for a European perspective. We have trans-border FUAs (see project Espon Metroborder) around the French frontier but only for the French residents working in another country, not for the opposite although we know that there are commuters from Germany, Belgium or Italy working in French FUAs (Strasbourg, Nice, ...). In the Espon Metroborder project this weakness was bypassed by assuming that the FUAs on both sides of the border are forming one entity, considering the results of Espon 1.4.3 based themselves on the literature.

Besides all the data limitations mentioned above we have encountered here and there difficulties in some areas like Scotland or former East-Germany, due to too big incoherencies between the different database, and this because of radical reshaping of the LAU2s during the last ten years.

The poly-FUAs : Espon 1.4.3 provided also a higher urbanization level called the "Polycentric Metropolitan areas" (poly-FUAs), which where made of groups of FUAs of the same neighborhood. These were based on a distance criteria :

" In some cases, we have to consider the situation where different metropolises, with the centre of their cores distant from less than 60 km, are contiguous, or are only separated one from the other by other cities, with their own labour pool, or yet are bordered by other large, medium or small cities, distant from less than 30 km, also with their own individualised manpower basin. In these cases, we have identified conurbations of POLYCENTRIC METROPOLITAN AREAS (poly-FUAs). We have also considered as forming a POLYCENTIRC METROPOLITAN AREA two large cities distant one from the other less than 30 km and reaching together the level of 500,000 inhabitants. For the rest, we don't have considered as being a polycentric metropolitan

area two or more large, medium or small cities with contiguous manpower basins, even if they reach together the threshold of 500,000 inhabitants.
So to form a poly-FUA structure we must have either :

- *metropolises (> 500 000 inh.) with their centres less than 60 km apart, and labour basins touching each other*
- *2 large cities (> 250 000 inh.) with their centres less than 30 km apart, and labour basins touching each other*
- *1 metropolis and 1 large or medium city (> 100 000 inh.) with their centres less than 30 km apart, and labour basins touching each other*
- *metropolises with their centres less than 60 km apart, labour basins separated only by the labour basin of a smaller FUA touching the both of them »*

These poly-FUAs where not taken into account into this project since the delimitation of the FUA has no impact on them but they remain useful and necessary

1.1 Structure of the FUAs

The FUAs may be elements of the higher structures named poly-FUA (see above General Methodology) but this "super-structure" is also related to their internal structure which reflects their functional nature.

The following diagrams summarize for instance four different situations in a high-density area, implying quite different realities as regards functions, economy, management of mobility and territorial planning, but which could be confused if the analysis did not sufficiently explicit the definitions used. Even if these four patterns are purely theoretical, they are respectively globally based on the situation of an old coal basin for the first one (type1), the Ile-de-France Region for the second (type 2, with new cities functionally not much independent from Paris), the Belgian central metropolitan area (type 3) and the big London metropolitan area (type 4), where secondary centers of the external fringe of the FUA have more decisional autonomy and are moreover doubled by a belt of important or specialized cities (cf. Cambridge, Oxford) inside the FUA.

These should be understood as milestones of possible urbanization situations or evolution paths, rather than an exhaustive typology. There can be numerous intermediary states between any of these types and it would be vain to classify all the FUAs.

The following diagrams summarize for instance four different situations in a high-density area, implying quite different realities as regards functions, economy, management of mobility and territorial planning, but which could be confused if the analysis did not sufficiently explicit the definitions used. Even if these four patterns are purely theoretical, they are respectively globally based on the situation of an old coal basin for the first one (type1), the Ile-de-France Region for the second (type 2, with new cities functionally not much independent from Paris), the Belgian central metropolitan area (type 3) and the big London metropolitan area (type 4), where secondary centers of the external fringe of the FUA have more decisional autonomy and are moreover doubled by a belt of important or specialized cities (cf. Cambridge, Oxford) inside the FUA.

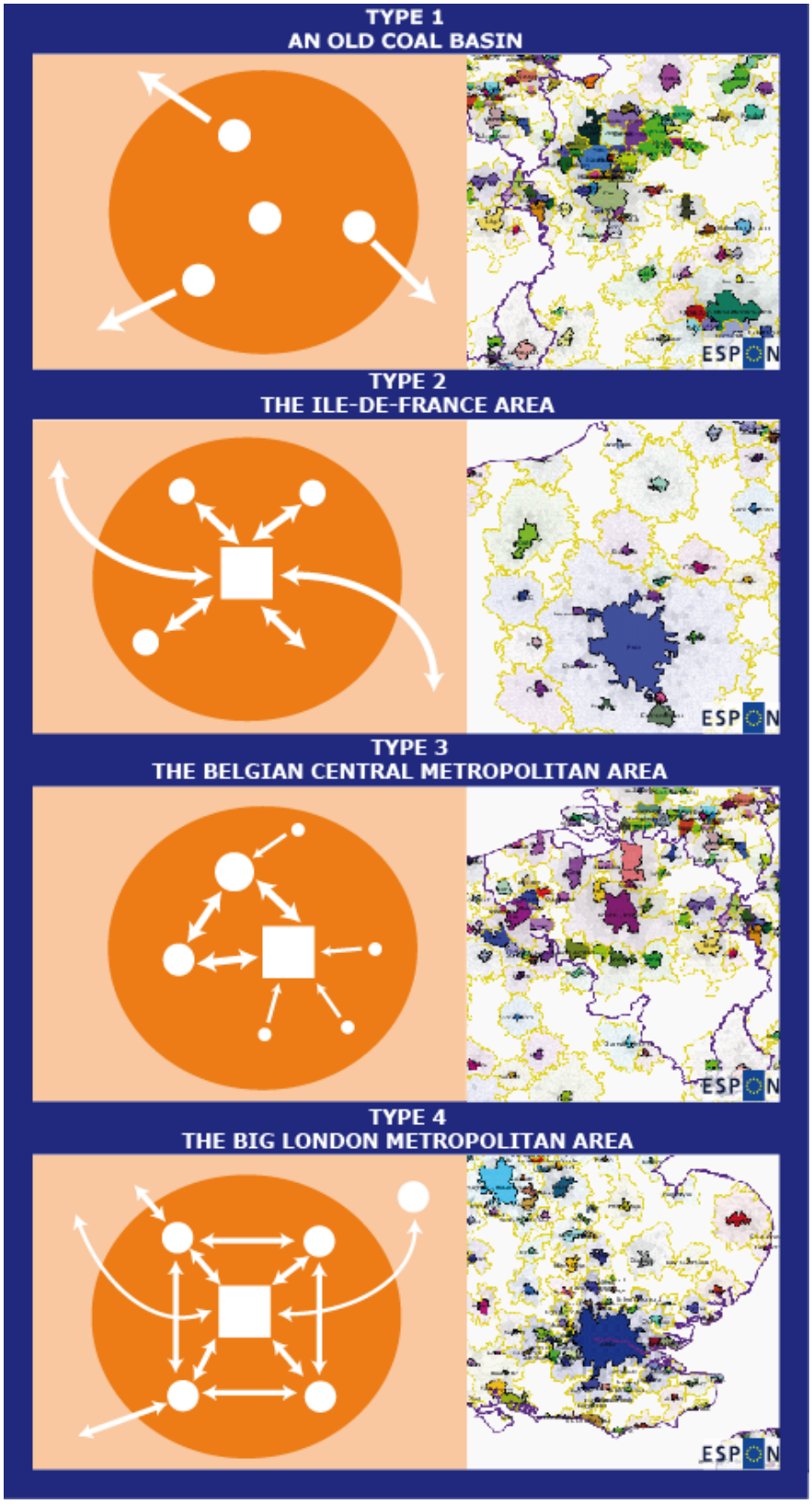

Figure 2 - Four different FUA types illustrated by the Ostrava region (1), the Ile-de-France Region (2), the Belgian central metropolitan region (3) and the London region (4).

Available databases

Different databases had to be joined together to achieve this work:

- Eurogeographics (GISCO) 2008 geometries of the LAU2s with identification code "SHN"
- MUAs' compositions in NUTS-5s with identification code from "CMRGCD97"
- SIRE database from Eurostat, tables emp_place_tot_mat_01 and c_emp_place_tot_mat_01 for the commuting numbers between LAU2s, emp_place_tot_01 and c_emp_place_tot_01 for the non-commuters numbers (working in area), LAU2s with identification code "CODCOM"

The database providing the commuting numbers between LAU2s is named SIRE and is provided by Eurostat. Most of its content comes initially from the 2001 censuses.

Data are provided only at the LAU1 level for Bulgaria and Slovenia, the data seem to be insufficient for Hungary. 2001 is quite old now but we know by experience that the FUAs are relatively stable, they are built on a quite low commuting rate and made simpler and coherent by filling the holes in them (LAU2s that don't reach the commuting threshold but landlocked in the FUA) and by eliminating the LAU2s remotely detached from the rest of the FUA. So chances are that any changes in the commuting rates would merely affect the general shape of the FUA, moreover if any they would probably affect small LAU2s at the fringes that don't have too much influence. The objective is not to compute the populations of the FUAs with accuracy close to one inhabitant but to have a good magnitude and to establish functional relations. But still changes can happen and it would be a good thing to have the opportunity to make this work on more recent data.

Entity	*Year*	*Source*
LAU2s spatial definition	2008	Eurogeographics
MUAs population	2001	Espon 1.4.3
FUAs population	2001 and 2006*	LAU2s population, Eurostat
Commuting numbers	2001*	SIRE database, Eurostat

Table 1 - Sources of the different data used in the elaboration of the FUAs and the MUAs

1.2 Linking the different databases

The first step of the work was to link all these tables with their different identification codes. We got the help from Eurostat which provided us with a table of correspondence between CODCOM and 3 other codes (COMM_ID, LAU2_code and nsi_code), as well as the NUTS-3 code for 2004 and 2006, which was easier to match

with the other available codes although not straightforward. A resulting table was produced, certainly not perfect, especially because of the difficulty to deal with complex cases were the matching was a n-to-n type. This is of course a time series issue, since we are dealing with data from 1997, 1999 and 2006 and it is not always possible to determine what became what. At first we made a spatial join between the old CMRGCD97 geometries and the 2008 geometries but the matching was not very good, due to some distortion in the projection. By checking the names and the code similarities we kept what appeared to be good and worked the imperfectly matched codes differently. For this we carefully used the names of the entities in the different sources and the code similarities, starting with the steadier methods and using progressively less steady solutions. We always used the code similarities and checked with the names. Progressively we eliminated most of the difficulties.

The objective of this first step was to transpose the LAU2s components of the old MUAs from the 1997 system into the new 2008 one, and also to match the CODCOM from the SIRE database with the EUROGEOGRAPHICS 2008 boundaries. So we stopped when we got a result satisfying for this objective but it should be improved by splitting this single big table into 3 or 4 more simple tables, but we didn't have enough time to produce something more rigorous.

1.3 The transposition of the MUAs into the new system

Espon 1.4.3 had produced the delineation of about 2000 MUAs (the 300 smaller not published) with a total of about 11000 NUTS-5 in them. All these NUTS-5 were transposed into the EUROGEOGRAPHICS 2008 LAU2s delimitations as explained above. Only a fistful of new MUAs was added during the FUA production but no further research was made on the identification of the MUAs. This might be done by inserting the UMZ database (produced in this same Espon Database project) into this urban areas delimitation tool.

1.4 The correspondence between the SIRE codes and the eurogeographics codes

CODCOM (SIRE) and SHN (EUROGEOGRAPHICS) share no similarities but the correspondence file provided by EUROSTAT was used and we have completed some relations due to differences between 2006 (SHN in the EUROSTAT file) and 2008 (EUROGEOGRAPHICS), by the same method based on the code similarities and the names correspondences.

2 The data processing

All the following steps are realized in SQL (MySQL or PotgreSQL) with postgis 1.5 for the spatial steps.

2.1 The selection of the LAU2s in relation to the MUAs

The initial work was done with the SIRE matrix. The table emp_place_tot_mat_01 (and c_ emp_place_tot_mat_01) is a simple matrix of about 2 millions lines containing the residence LAU2 (CODCOM), the working place LAU2 (DESCODCOM) and the number of commuters. The table emp_place_tot_mat_01 (and c_ emp_place_tot_mat_01) was completed with the active population working in their residence LAU2, in order to take into account all the working places. The MUAs were added in another column (this is not an orthodox way to proceed but it made things easier) in relation to the work place LAU2.

The process then consisted in computing the total number of commuters going in a MUA, for each CODCOM and each MUA. In the same time a commuting rate is computed by comparing this number to the «economically active population» from the tables emp_place_tot_01 (EU15) and c_ emp_place_tot_01 (former candidate countries). Only the rates of 10 % or higher were kept.

Then specific cases are processed because of the incoherencies between the different databases that were not solved by the correspondence table mentioned above, especially the countries were the commuting data are provided at the LAU1 level or cases of big cities considered as a single entity in one database and a set of multiple LAU2s in the other (Budapest, Bratislava, Paris), or transborder data provided by other sources (Luxembourg, Belgian border), or Hungary were there is no active population data provided.

We then obtain a list of LAU2s with their respective MUA work places and their commuting rates, all this forming kind of «proto-FUAs» since each MUA is considered as the seed of a potential FUA and since the LAU2s can be part of several «proto-FUAs». This is a raw material to build the real FUAs, with their complex elements.

2.2 The spatial operations

The geographical objects for a map representation are produced.

The MUAs are simply aggregated from the LAU2s in the Eurogeographic 2008 layer. This step is easy because the MUAs were previously delineated as explained above.

The FUAs are much more complicated. The objective is to obtain coherent areas, without holes in them and no isolated parts. Here are the different steps.

- Selection of the main destination of the commuters for each LAU2, among all the MUAs toward which the commuting rate is higher than 10 % . Actually the different destinations are ranked and the first one is kept. In case of equality (same number of commuters toward 2 MUAs) the bigger MUA is selected as main destination.
- The MUAs are grouped into FUAs, according to the main destination of the commuters of each MUA and/or according to the literature as explained in the main report and above.
- The exterior rings of the FUAs are created and the bigger part is kept. Several parts can be kept where the separation is due to geographical particularities (islands, both sides of a fjord,...).
- The LAU2s spatially enclosed in the exterior rings are selected and attributed to their FUA, so that the holes are filled.

3 Known issues

Besides the unavoidable incoherencies between the different databases already mentioned above or in the main report there are specific problems we could not bypass :

- The original definition of the FUA is based on a proportion of the 'occupied' active population but all we have is the 'economically' active population, which includes the unemployed population. So this distorts the commuting rate by overestimating it by a maximum factor of 1.1, considering a 10 % unemployment.
- The active population in Germany is wrong in at least half of the municipalities! This problem is somehow minimized because, first, when a LAU2 has a commuting flow toward more than one MUA (for example in Rhein-Ruhr region) what matters is the highest number of commuters whatever the active population is, considering that this happens in areas with a strong peri-urbanization and without free space between the FUAs. Second, the process of building the FUAs eliminates the LAU2s that are not among the others for a same FUA and homogenizes the area by making a ring around the LAU2s of the FUA. Third, the LAU2s where the commuters' number is higher than the active population are ignored. Fourth, the overestimation of the FUAs (there can't be underestimation) happens by including probably small municipalities at the fringes of the areas. But still, errors remain.
- Same remark for Slovakia, but apparently with a smaller proportion of erroneous data.
- In former East-Germany, especially in the Saxe-Anhalt, there were many municipalities merging after the reunification. This leads to many incoherencies between SIRE and EUROGEOGRAPHICS, i.e. we have commuting numbers for LAU2s that don't exist anymore and we didn't spend time on trying to redistribute the commuters in the new LAU2s. This might perhaps be done but not in this project.
- In the area between Glasgow and Edinburg we have the same problem than in Saxe-Anhalt (see above).
- In Portugal the commuting numbers in the core city areas are provided at the LAU1 level but this is not a problem because in every case we have a ring of municipalities (LAU2s) around this LAU1 that are included in the FUA. So any central municipality that would be out of the MUA but into that LAU1 area would be included in the FUA anyway.
- Technically speaking, instead of making a ring surrounding the FUAs with the Postgis ST_ExteriorRing function we could explore the possibility to use the ST_convexHull function that minimizes the quirks in the shape of some FUAs, but it could lead also to exaggerate their size.

Whenever new data would become available, like especially the active population in Germany, we would rebuild the FUAs immediately.

4 The indicators

In collaboration with the project Espon METROBORDER we have produced a list of indicators for all the FUAs. Since the FUAs are defined at the LAU2 level any information available at that same level can be used to characterize the FUAs, otherwise it is sometimes possible to use the NUTS3 values, as explained below.

Indicator	*Year**	*Source*	*Unit*
Population	2001	LAU2, EUROSTAT	inhab.
	2006	LAU2, EUROSTAT	inhab.
Population variation	2001 - 2006	LAU2, EUROSTAT	%
FUAs Areas	2008	EUROGEOGRAPHICS	km^2
MUAs Areas	2008	EUROGEOGRAPHICS	km^2
Compactness	synthetic indicator		%
GDP	2006	NUTS-3, EUROSTAT	Euro
GDP/inhab.***	synthetic indicator		Euro/inhab.
Economical structure	2006	NUTS-3, EUROSTAT (6 big NACE sectors)	%
Unemployment**	2006	NUTS-3, EUROSTAT	%

Table 2 - list of the indicators available in the Espon database

* : For some indicators some data are coming from different years. See the missing data issues below.

** : Unemployment values should be used «with caution» !

*** : see the map below

4.1 Methodology

For the population and area indicators we have simply computed the FUA values from the LAU2 values. The compactness is the % of the population of the FUA actually living in the main MUA.

The economical indicators are computed by using the NUTS-3 values on which we apply a population ratio between the NUTS-3 and the intersection of the FUA and the NUTS-3. This is possible because we now have the LAU2 composition of the FUAs which make the link between those and the FUAs possible. The methodology is

therefore useful to assess the urban areas throughout Europe (except in the countries where no commuting data are available – see above).

By way of illustration let's look at the Basel case. Basel is a complex case, it is a transborder metropolitan urban region with 5 MUAs and involving 7 NUTS-3.

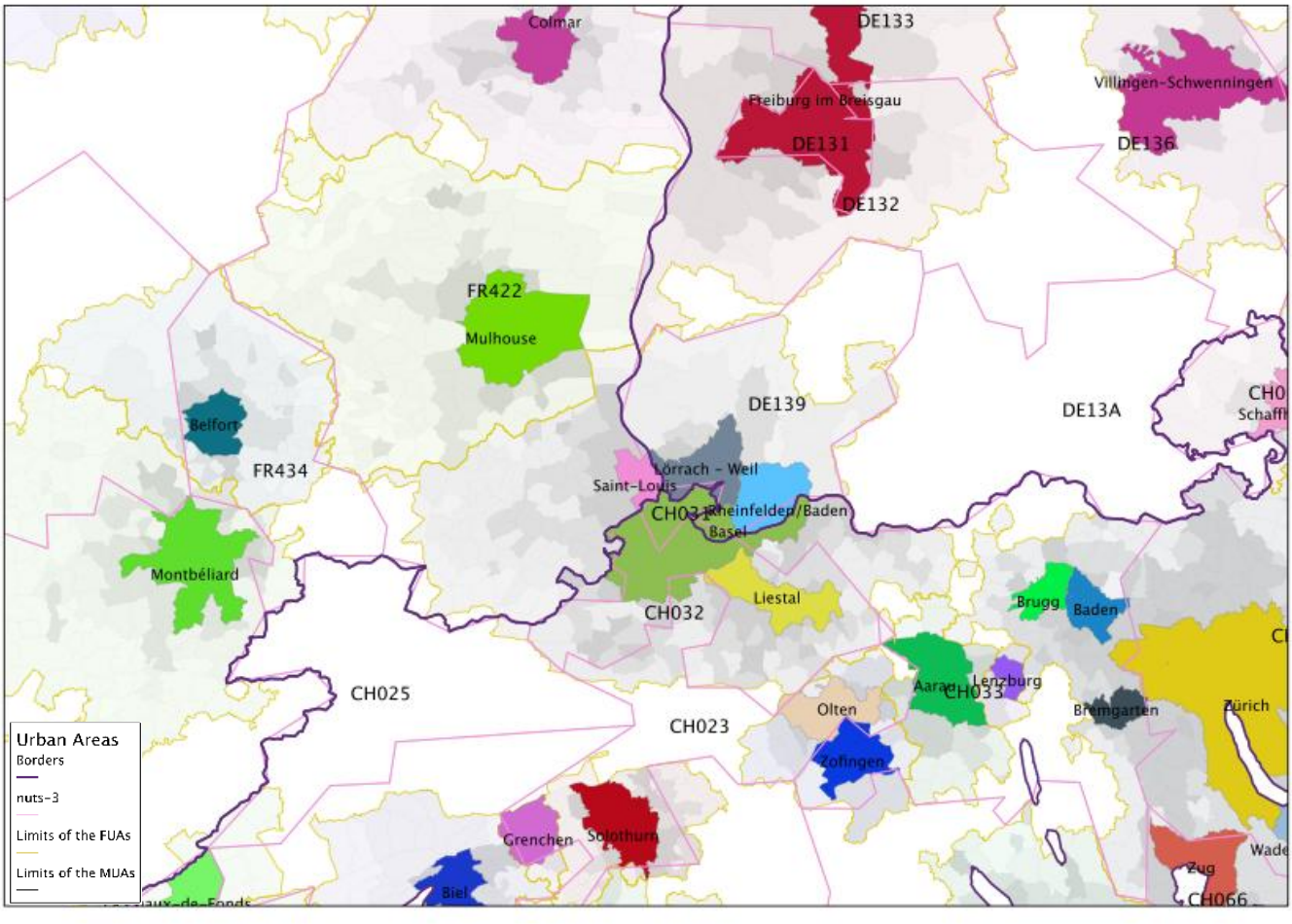

Figure 1 - The FUA of Basel in its region

The MUAs are shown in plain colors, the FUAs are in light colors delimited by a yellow line, the NUTS-3 are represented by the pink lines, and finally the gray shade shows the intensity of the commuting (the more grayish the higher the number of commuters is). The national borders are violet.

nuts3_06	pop_nuts3	pop_fua	pop_fua in nuts3	coef
CH023	248	827,3	31,7	13
CH025	69	827,3	0,1	0
CH031	185	827,3	188,1	100
CH032	266	827,3	258,7	100
CH033	572	827,3	52,6	10
DE139	221	827,3	215,6	100
FR422	738	827,3	91,4	13

Table 3 - correspondence between the nuts-3 and the FUA of Basel.

We see here (table 3) that the FUA of Basel extends over 7 nuts-3 in 3 countries, and we see in the table that this corresponds to different population values according to the nuts-3. We have computed for each of them a coefficient ('coef') in % giving what part of the nuts-3 indicator (for instance the GDP) we take from every nuts-3, the total giving the indicator value (for instance the GDP) of the FUA. In the particular case of the GDP or added values we assume that the productivity is equal everywhere inside each NUTS-3, which is not a bold assumption. For every other indicator we use the same methodology as long as we can make the same kind of uniformity hypothesis.

4.2 Displaying FUAs database on maps, two examples

The FUAs database allows displaying innovative results for the all ESPON Area. For instance, the maps of GDP per capita 2006 (figure 3) and evolution of population 2001-2006 (figure 4) show strong contrasts between FUAs, and give an additional picture of the situation as compared to classical maps produced at NUTS3 level. However, it is important to have a look to keep in mind that some of the values are estimated (cf section 4.3 below) and such results must be interpreted carefully.

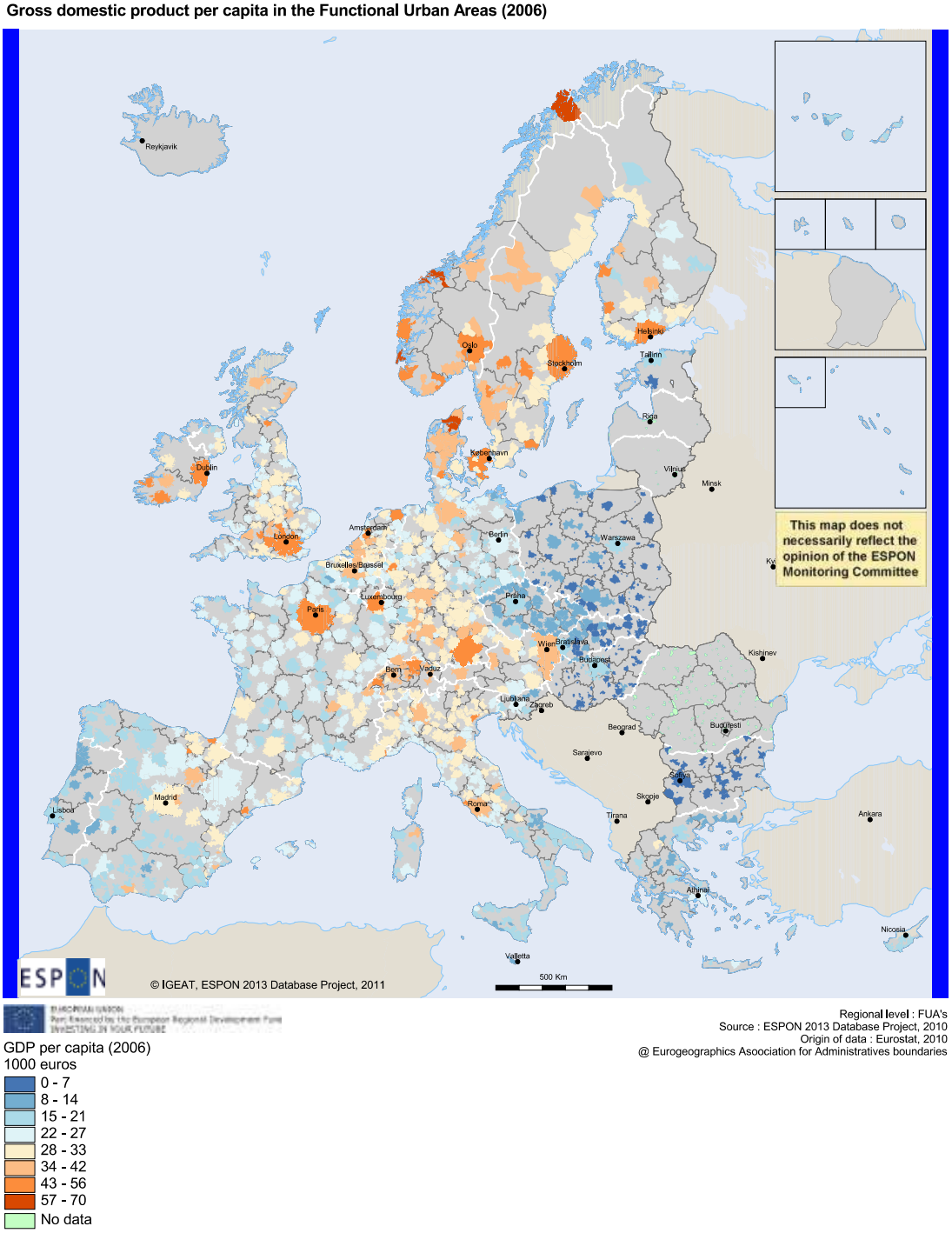

Figure 3 – Gross Domestic Product per capita 2006 in the FUA delineation

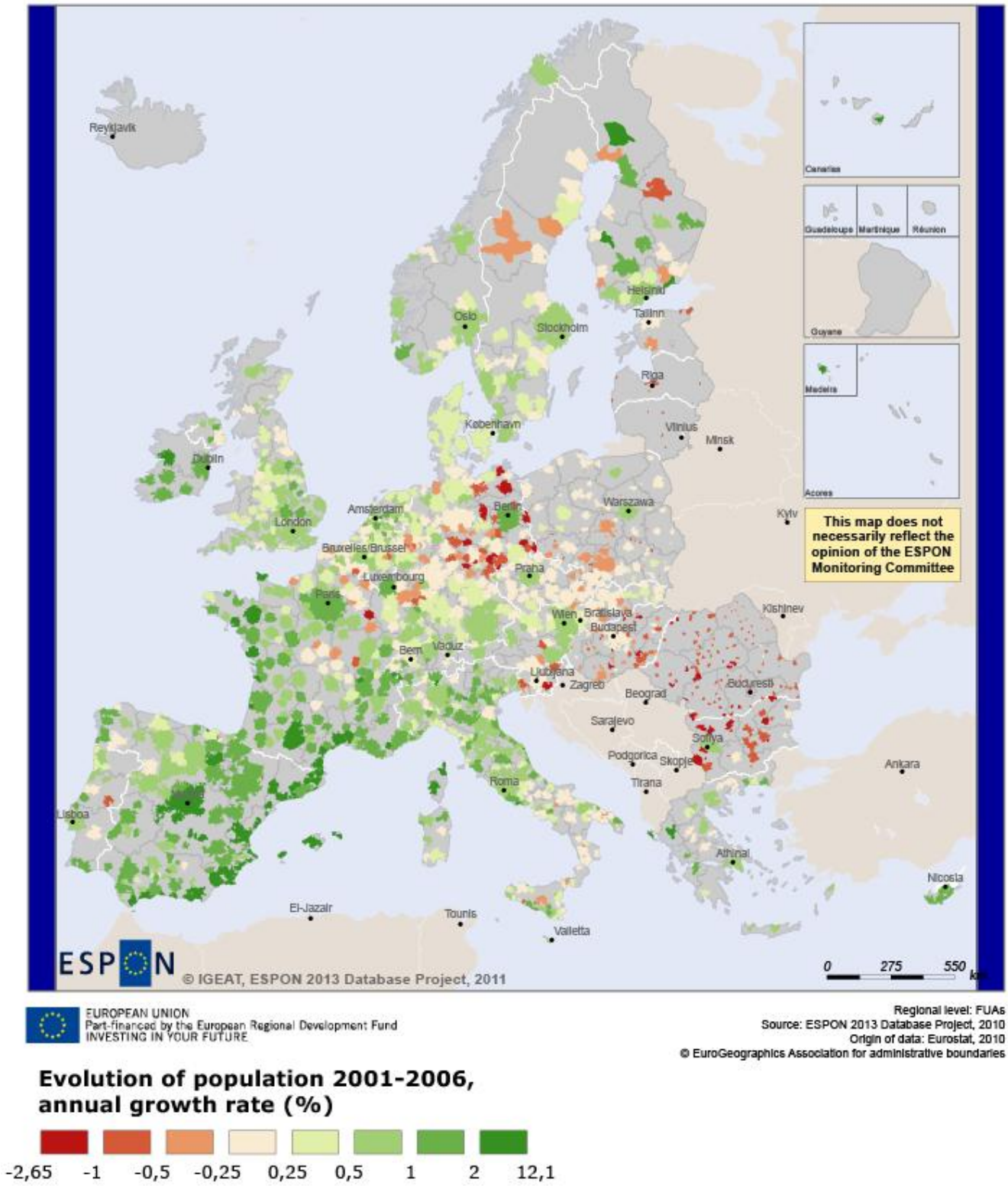

Figure 4 – Evolution of population 2001-2006 in the FUA delineation

4.3 Missing data issues

Poland: So far there are no commuting data in Poland. We received from Przemyslaw Sleszynski (Polish Academy of Sciences) a set of functional areas based on the last census data (2002) and other data from 2004 like the socio-cultural profiles of the population. We have included these data 'as is' in our database.

Romania, Latvia and Lithuania: No commuting data are provided for Romania, Latvia and Lithuania and we couldn't find any substitute, so we used the population data from Espon 1.4.3..

LAU2 Population in 2006: We don't have population numbers at the LAU2 level for Bulgaria, Cyprus, Denmark, United Kingdom, Lithuania, Latvia, Portugal and Romania. So we used the NUTS-3 replacement methodology described above to obtain the population of the FUA in 2006.

GDP in Switzerland and Norway: The GDP for Swiss and Norwegian FUAs are from 2005, due to missing data in Eurostat for the NUTS-3.

Unemployment: At the NUTS-3 level the unemployment values are not available in 2006 for Denmark, Sweden, Ireland, Switzerland, Sachsen-Anhalt and two Italian provinces (Sassari and Cagliari). See the synthesis table below for the replacement years.

Indicators	Countries	Substituting Years
Commuting data	Poland Romania Latvia Lithuania	-
2006 LAU2 population	Bulgaria Cyprus Denmark Lithuania Latvia Portugal Romania United Kingdom	NUTS-3 replacement
GDP	Switzerland Norway	2005
Unemployment	Denmark Sweden Sachsen-Anhalt	2007
	Ireland	2004
	Switzerland	2005
	Sassari province (ITG25) Cagliari province (ITG27)	2008
	Northern Ireland	2003

Table 4 - synthesis of the missing data issues

4.4 Synthesis of the problems encountered to the attention of the Espon MC national representatives

For every country except France it should be very appreciated to have the commuters working in a foreign country, at the LAU2 level, so that we could delineate transnational Urban Areas.

It would be also interesting to enrich to local data by providing population figures for more years (2010 ?) than 2006 in order to be able to make time series, and of course any local (LAU2 level) indicators would be useful, such as economical, employment and social indicators.

Country	Issue	Description	Request
Bulgaria	Commuting data at LAU1 level only		Commuting and active population data at LAU2 level
	Missing population numbers for 2006		Population numbers for 2006
Cyprus	Missing population numbers for 2006		Population numbers for 2006
Germany	Active population at LAU2 level in SIRE are wrong	At least 50 % of the LAU2s have active population (much) lower than the numbers of commuters, which is impossible.	Correct active population at LAU2 level.
	Many LAU2s reshaping	The SIRE LAU2s are quite different from the GISCO LAU2s because of municipality merging in former east Germany regions. This leads to incoherences between the databases	Commuting statistics at LAU2 level according to the current LAU2s delimitations
	The origin of the commuting data.	We heard that the german commuting data are provided by the insurance companies, which probably biases the figure.	Commuting data at LAU2 level based on better sources.
Denmark	Many LAU2s reshaping	The SIRE LAU2s are quite different from the GISCO LAU2s because of municipality merging and Nuts reshaping. This leads to incoherences between the databases.	Commuting statistics and active population numbers at LAU2 level according to the current LAU2s delimitations
	Missing population numbers for 2006		Population numbers for 2006
Estonia	Apparently insufficient commuting data	The FUAs of Tartu and Tõrva seem to be placed away from these cities, which probably comes from inacuurate commuting data. We don't know the origin of this problem.	Better commuting and active population data at LAU2 level
Hungary	Apparently insufficient commuting data	Globally the Hungarian FUAs don't appear to be well delineated, probably because of inaccurate commuting and/or active population data.	Better commuting and active population data at LAU2 level
Lithuania	No commuting data.		Commuting and active population data at LAU2 level
	Missing population numbers for 2006		Population numbers for 2006

Latvia	No commuting data.		Commuting and active population data at LAU2 level
	Missing population numbers for 2006		Population numbers for 2006
Netherlands	Many LAU2s reshaping	The SIRE LAU2s are quite different from the GISCO LAU2s because of municipality merging and Nuts reshaping. This leads to incoherences between the databases.	Commuting statistics and active population numbers at LAU2 level according to the current LAU2s delimitations
Poland	No commuting data.		Commuting and active population data at LAU2 level
Portugal	Commuting data at LAU1 level only in the urban areas.		Commuting and active population data at LAU2 level everywhere
	Missing population numbers for 2006		Population numbers for 2006
Romania	No commuting data.		Commuting and active population data at LAU2 level
Slovenia	Commuting data at LAU1 level only		Commuting and active population data at LAU2 level
Slovakia	Many active population at LAU2 level in SIRE are wrong	Some LAU2s have active population (much) lower than the numbers of commuters, which is impossible.	Correct active population at LAU2 level.
United Kingdom	Many LAU2s reshaping	The UK LAU2s change very frequently but the commuting data are old. This leads to incoherences between the databases, especially in Scotland.	Commuting statistics and active population numbers at LAU2 level according to the current LAU2s delimitations
	Missing population numbers for 2006		Population numbers for 2006

Table 5 - synthesis of the problems encountered