



Belgeo

Revue belge de géographie

4 | 2017

La Belgique : quelques nouveaux regards sur un pays en forte mutation – Belgium: some new looks at a highly evolving country

Migration and commuting interactions fields: a new geography with community detection algorithm?

Isabelle Thomas, Arnaud Adam and Ann Verhetsel



Electronic version

URL: <http://journals.openedition.org/belgeo/20507>

DOI: 10.4000/belgeo.20507

ISSN: 2294-9135

Publisher:

National Committee of Geography of Belgium, Société Royale Belge de Géographie

Brought to you by Université catholique de Louvain

UCL

Université
catholique
de Louvain

Electronic reference

Isabelle Thomas, Arnaud Adam and Ann Verhetsel, « Migration and commuting interactions fields: a new geography with community detection algorithm? », *Belgeo* [Online], 4 | 2017, Online since 30 September 2017, connection on 25 February 2019. URL : <http://journals.openedition.org/belgeo/20507> ; DOI : 10.4000/belgeo.20507

This text was automatically generated on 25 February 2019.



Belgeo est mis à disposition selon les termes de la licence Creative Commons Attribution 4.0 International.

Migration and commuting interactions fields: a new geography with community detection algorithm?

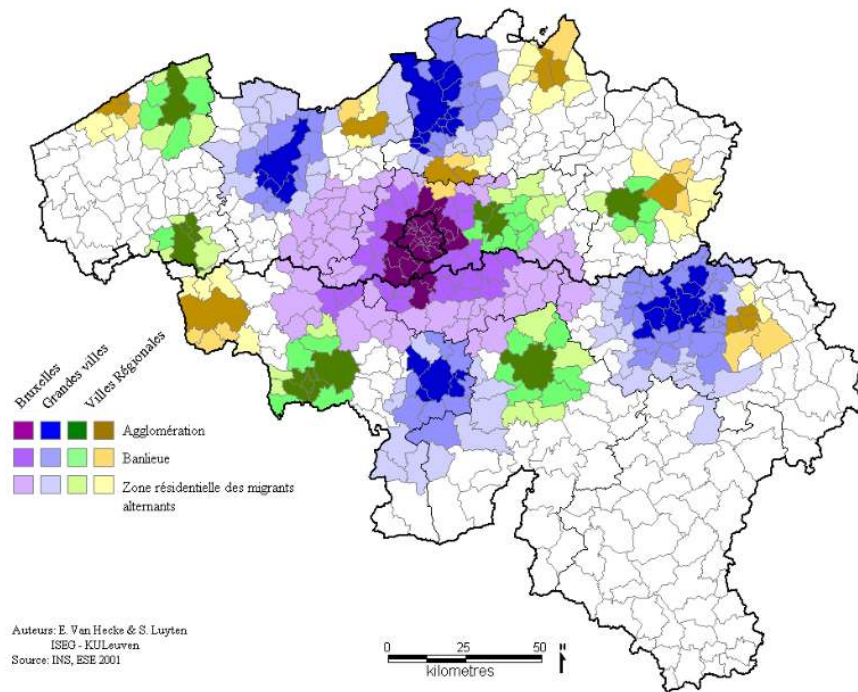
Isabelle Thomas, Arnaud Adam and Ann Verhetsel

Introduction

- 1 In Belgium, a lot of mainstream social and economic geography is anchored into the context of the so-called *city regions* (“stadsgewesten”/“aires fonctionnelles urbaines”) introduced by Van der Haegen and Pattyn (1979, 1980). These authors observed that some basic activities take place outside the traditional city, and hence people operate in a much wider spatial urban environment. As a consequence, these authors proposed to deny the use of administrative borders when studying a city and identified the real “cities in new forms”. Their approach has led to a delineation of *functional urban regions* around major Belgian cities, internally organised from centre to periphery into city centre, central city, agglomeration, urban fringe, and “commuter area” still connected to the city centre. Key variables used for delineation were characteristics of the places such as population and job densities, population evolution, dwelling morphology and percentage of commuters going to the CBD. This methodology was repeated every 10 years since, after each new housing and population census (Figure 1) (Leemans *et al.*, 1989; Van der Haegen *et al.*, 1996; Luyten and Van Hecke, 2007). The success of the concept of “city regions” lies not only in the repeated application of a clear methodology since its creation, but also in the fact that it was used and became a standard concept in many academic and administrative geographical analyses in Belgium.
- 2 Nowadays, the way of conducting the Census in Belgium has totally changed and the data needed for repeating the abovementioned urban delineation are not easily available anymore. We have cried and could still complain about this (Thomas *et al.*, 2009), but as in

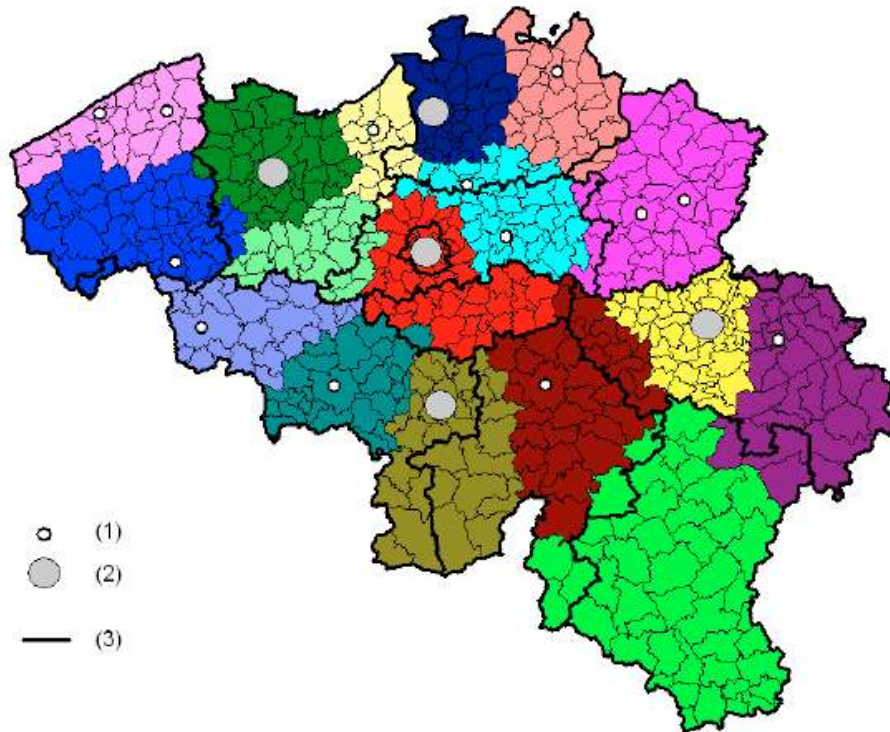
every crisis new opportunities appear. In recent years, *newly collected (big) data* allow us to measure and reveal new aspects of the urban complexity and especially the pulses of the territory in time and space and the interrelationships between people and places. The development of ICT technologies, of increasing data storage and computer capacities indeed enable *to collect very large volumes* of data and hence better follow people and firms in their everyday activities. Hence, we know where people are (GPS in smartphones) and with whom they are in contact (social media, phone data, etc.). The ICT data constitute a new opportunity for characterizing nodes (X_i), and the interrelationships between them (w_{ij}). In this sense, the partition of the Belgian telephone territory has become a well-known example based on the frequency of communication between municipalities (Blondel *et al.*, 2010), undoubtedly reflecting a geography of “phone basins”, closely related to the daily urban system. Comparing the phone map with the “city regions” reveals challenging similarities, but also differences and complementarities (Figure 2).

- 3 If “big data” has the advantage of opening new research opportunities, they are often unfortunately not in the public domain (privacy rules) and need a very clear understanding of what they are really measuring (see e.g. Kitchin, 2013). Big data and communication technologies have also led to the development of new methods in Data Sciences for *optimally* detecting groups of items that interact (see e.g. Blondel *et al.*, 2008); they enable to avoid a priori defined thresholds often used traditionally by spatial analysts and they have the advantage of letting the data speak for themselves (data driven models). These new methods offer interesting opportunities to study interactions fields in and around cities and hence better understand the economics of agglomeration (see e.g. Fujita and Thisse, 2013, chapter 6) and verify theory driven models. It is now time for geographers to catch up this new opportunity and link these new results to spatial knowledge and theory critically.

Figure 1. Urban Regions.

Source: Van Hecke *et al.*, 2007

Figure 2. Phone basins based on cell phones billing addresses.



(1) Regional city, (2) major city, (3) provincial border

Source: Blondel *et al.*, 2010

- 4 The objective of this paper is hence to construct a socio-economic geography of Belgium using very traditional relational data (migrations and commuting) and new methods developed in Data Sciences (Louvain Method) for partitioning Belgium into interactions fields. Results are discussed in terms of methodological choices, geography but also in terms of political issues.

Data and methods

Data: migration and commuting

- 5 We here limit ourselves to *census* data, and leave the debate about ICT/Big data aside: ICT data have high potentials, but also strong limitations especially when dealing with geolocations (see e.g. Kitchin (2013) and also discussion in companion papers such as that about phone calls in Belgium (Blondel *et al.*, 2010; Adam *et al.*, 2017a), logistic fluxes (Beckers *et al.*, 2017), Twitter data (Decuyper *et al.*, 2017) or rail journeys (Jones *et al.*, 2017)).
- 6 Census data have attracted many criticisms in hitherto time, but we have to keep in mind that they also have many advantages (see e.g. Thomas *et al.*, 2009) among which being comparable in time and space, and being very well documented (see e.g. the Statbel website). This is especially true up to 2001, the date of the last decennial census and monographs in Belgium. In 2011, it was indeed replaced by an “administrative census” (named *Census11*), based essentially on existing data collected by different Belgian and

Regional administrations. The transition from survey to administrative census data questions the temporal comparability for many variables and raises major criticisms in the data and their definition. Specifically, the place of work suffers from multiple (new) source errors, biases and estimations. There is a clear need for a detailed critical approach which goes beyond the objective of this paper.

- 7 Two *Census11* data are here used: the number of *commuters* and the number of changes of residence (here called *migrations*) between pairs of nodes. Migration data are based on the National Register that centralises information collected by the municipalities, and commuting data rely on the coupling the National Register and the Social Security Register (BCE). Commuting or migrating with another country is unfortunately too inaccurately spatially coded to be considered here; our analyses are hence limited to movements *within Belgium*.
- 8 Both migration and commuting movements are officially available at the *scale* of the municipalities. Commuting is also available at the level of the statistical sectors (neighbourhoods); these latter are also used here to show how far the level of aggregation influences the final results. Hence, we end up with three data sets: migrations and commuting between the 589 municipalities, and commuting between the 19,782 statistical sectors.

Method: community detection

- 9 It is more and more common nowadays to use methods based on *modularity* maximisation for detecting *communities* of nodes that tightly interact in large data sets (Newman, 2004). It is important to insist on the fact that these methods do not require any location information as input; they only require the importance of the links between the nodes (in other words: the weights associated to the edges). If nodes are geocoded, it is possible to map a posteriori the results and hence obtain a spatial partition of the studied area. This kind of method is quite useful for studying large origin-destination matrices (networks), for obtaining groups of places that are connected without predefining a priori central places and/or thresholds.
- 10 Methods based on the modularity search for a compromise between the minimisation of the connections between nodes classified in different communities (cut) and the maximisation of the diversity (the number of communities) (Blondel *et al.*, 2008, p. 2). Let us here remind that the number of communities is endogenous: it is internally detected by the modularity without any a priori defined threshold.
- 11 Among the many existing methods, the *Louvain Method* is here used (Blondel *et al.*, 2008); it is nowadays considered as a standard heuristic despite its limitations (Fortunato, 2009). If the method is indeed efficient and fast, it suffers from the *resolution limit* (Traag *et al.*, 2013): results might be generated by the method itself and not by the structure of the data and hence small communities might not show up. We here deal with this methodological problem by using the solution proposed by Delvenne *et al.*, 2013): a parameter (ρ) is implemented in the objective function. Small values of ρ lead to the detection of large numbers of (small) communities, and the opposite for large values of ρ . Hence, by controlling the value of ρ , we better control the number of communities. The choice of ρ hence plays a capital role in the veracity of the final partition.

- 12 A second problem of the Louvain method is its sensitivity to the *order in which the nodes* enter the analysis (Blondel *et al.*, 2008). In order to avoid this problem, each map presented in this paper is the synthesis of 1,000 independent runs of the method. In each run, the order of the nodes is randomly changed. It is then easily possible to compute for each node the number of times the node is allocated to a given community and hence to create a stability measure. Unstable zones are hatched on the choromatic maps presented in this paper, or lead to a choropleth map. For further methodological issues, we refer to companion papers (see e.g. Adam *et al.*, 2017a and Decuyper *et al.*, 2017).
- 13 The following section summarises the most relevant results obtained while analysing commuting and migration in Belgium; numerous analyses are not reported here but are available upon request. We here want to illustrate and remind the advantages of the method and the robustness of the results, the importance of the data definition on the results, that of the control of the value of ρ and also the role of the scale of analysis in the final results.

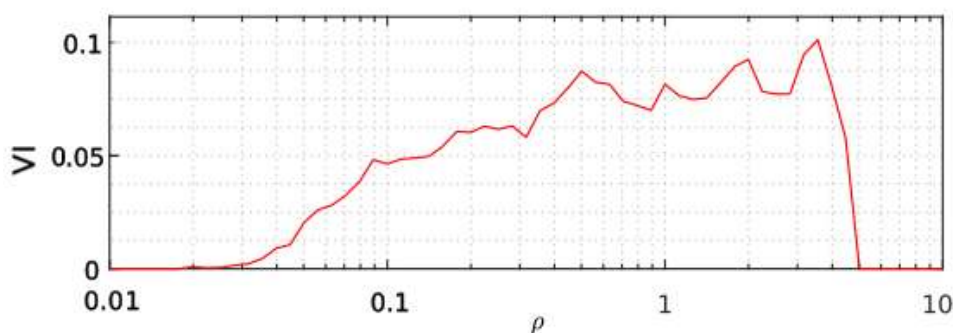
Results

Commuting basins

Optimal number of communities

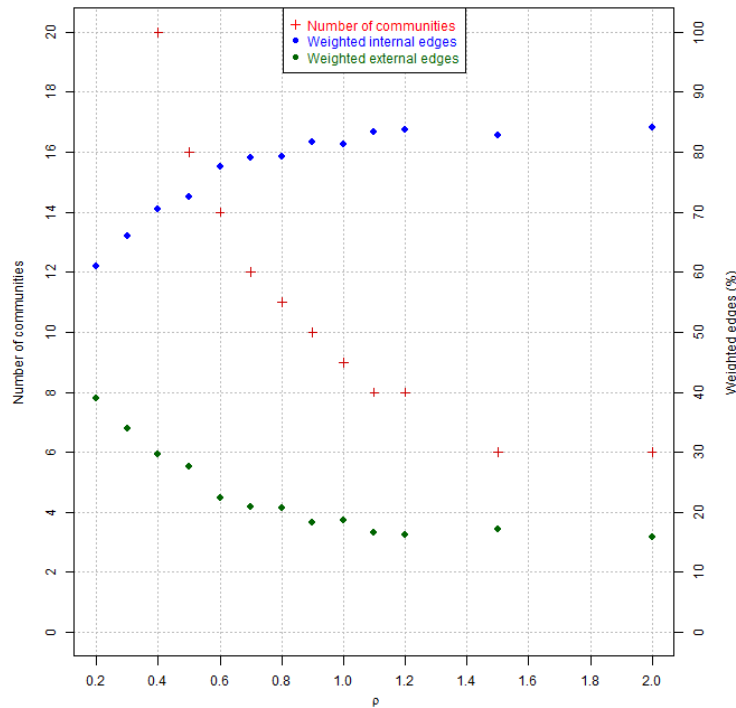
- 14 The problem is first to find the optimal ρ value, if any. Figure 3 illustrates this search following the Delvenne *et al.* (2013) method. As already mentioned, 1000 runs are applied for each value of ρ . The similarity between each set of 1000 runs (partitions) is computed by means of a Variation of Information (noted VI) and iteratively computed for each value of ρ . The optimal partition corresponds to the minimal value of VI.
- 15 $\rho = 1$ corresponds to the standard version of the Louvain Method; it is a default value for the method, corresponding to a situation where the optimization of the modularity is not modified. When ρ is smaller than 0.02, VI values are always equal to 0 because modularity is always equal to 0 and the Louvain method can't detect communities in the network; in this case, the number of communities is equal to the number of nodes (589). When ρ is very large (larger than 5), the Louvain method has merged all the nodes in one single community. It is hence impossible to observe a positive value of VI. Figure 3 clearly shows that there are 2 main and clear local minimal values of VI: $\rho = 0.3$ and $\rho = 0.9$.

Figure 3. Variation of Information (VI) versus ρ values for commuting movements between municipalities.



- 16 We also know that the best partition has more (weighted) edges within the communities (internal edges) than between communities (external edges) (Newman and Girvan 2004). Figure 4 shows the proportion of internal and external edges, for increasing values of ρ and hence for decreasing values of the number of communities. The proportion of internal edges is higher when $\rho = 0.9$ than when $\rho = 0.3$. Based on the VI method (also called Stability Method) and on the proportion of internal edges, we hence consider that the best community structure (optimal) is found when $\rho = 0.9$.

Figure 4. Proportion of internal and external edges and number of communities for different values of ρ for the 2011 commuting movements (at municipality level).

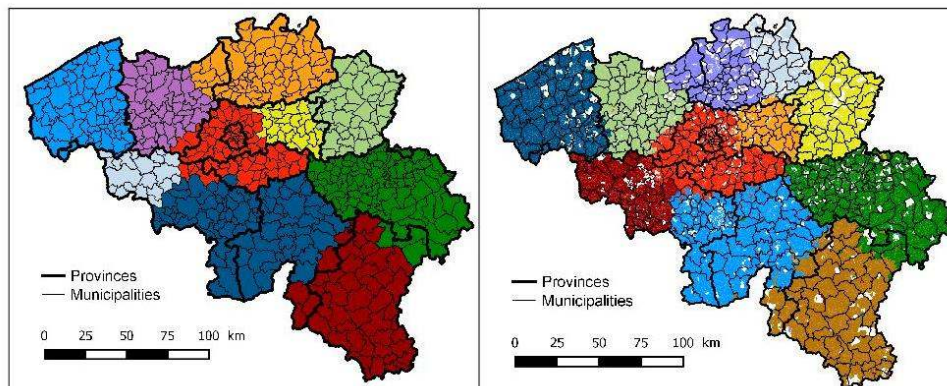


Spatial partition

- 17 With the Louvain Method and the commuting data, Belgium is optimally partitioned in about ten “job basins” which is much less than the 47 defined by de Wasseige *et al.* (2000) but more than the five large cities often defined as job centers (see e.g. Van Hamme *et al.*, 2011 and Figure 1). Contrarily to other publications using thresholds (Marissal *et al.*, 2006) or other methodologies (see e.g. Charlier *et al.*, 2016 and Adam *et al.*, 2017b), the method allocates all municipalities to a community.
- 18 Results of the Louvain partition are mapped in Figure 5, on the left at the level of municipalities and on the right hand side at the level of the statistical sectors. Comparison of the two maps is straightforward: almost the same number of communities (respectively 10 when nodes are municipalities, and 11 for statistical sectors) is observed and the two spatial partitions clearly look alike. Data aggregation little impacts the partitioning of Belgium in terms of commuting movements. Note that uncolored statistical sectors (Figure 5 – below) are due to the absence (or too small numbers).

- 19 Surprisingly enough, the limits of the communities correspond quite well to provincial borders: commuters tend to stay in the same province at some exceptions mainly located in the vicinity of the two largest cities (Brussels and Antwerp). Communities around these two big cities are indeed large, and they step over regional/provincial borders: Brussels extends to Halle-Vilvoorde and encompasses the entire Brabant Wallon stepping over the linguistic border, but it does not include the area of Leuven which is a job basin on its own. The community centered on Antwerp includes Waasland (across the provincial border) while the Eastern part of the province of Antwerp is a job basin on its own (Antwerp Kempen) when we work at the scale of the statistical sectors. Also note that in Wallonia, the provinces of Liège and Luxembourg correspond to clearly defined communities, while the limits of the provinces of Hainaut and Namur do not correspond to those of the communities: Charleroi and Namur form one job basin, while Mons and Tournai form another one.
- 20 Let us remind here that location is not an input of the method; the intensity of the commuting fluxes is the only input. Hence, commuting interactions in Belgium lead to a very clear structure dominated by the contiguity of the basins (small distances are more important than large ones), and an imprint of the provinces; moreover, the urban hierarchy (Brussels and Antwerp) and transportation infrastructure (Mons-Tournai, Namur-Charleroi) also intervene.

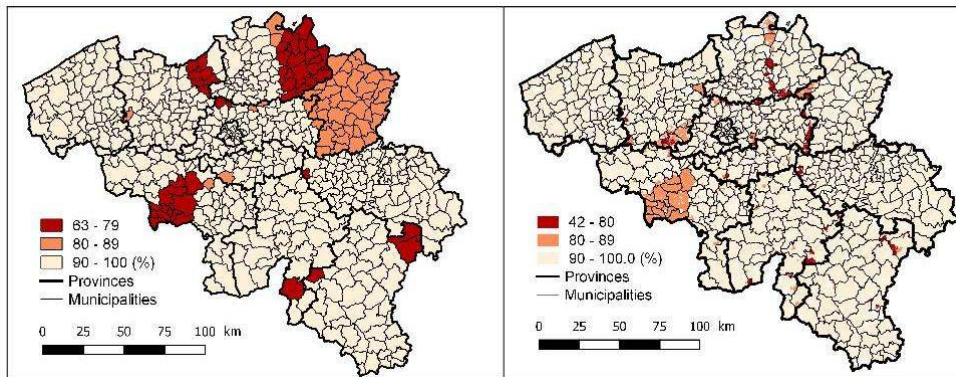
Figure 5. Optimal partition in communities when nodes are municipalities (*above* – $\rho = 0.9$) and statistical sectors (*below* – $\rho = 1.0$). 2011 commuting data.



Sensitivity analyses

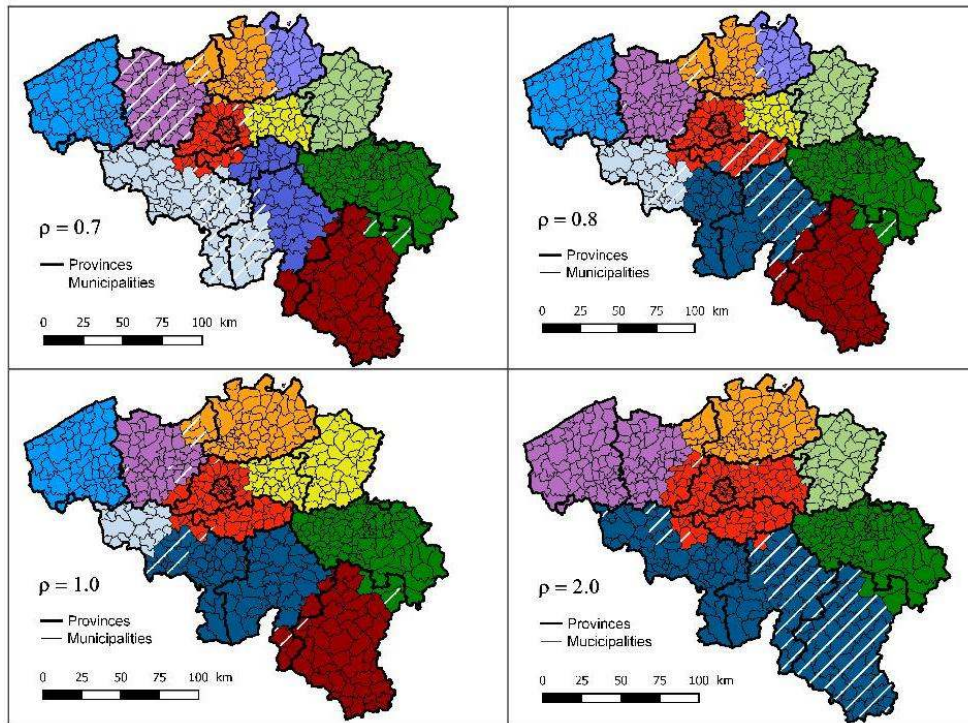
- 21 First and as already stated, 1000 runs of the method have been performed for each value of ρ . Each node can hence be characterized by the percentage of runs for which it is associated to the same community, hence measuring the stability of the partition. Choropleth maps of the stability of node in the partition can hence be drawn (Figure 6). Note that it is hard to get the same discretisation of the values observed at the municipality level (*above*) and for statistical sectors (*below*) because of the large differences in the number of nodes and in the statistical distribution of the variable. By visually comparing the two maps in Figure 6 we see that the coloured surface is much smaller and lighter for statistical sectors and it mainly corresponds to statistical sectors located at the limits of the communities.

Figure 6. Stability of the partitions reported in Figure 5: % of times a node belongs to the same community in the 1000 runs. 2011 commuting data.



- 22 Second, the role of ρ in the final partitioning of Belgium is illustrated by 4 maps in Figure 7. When $\rho = 1$ (default option), 9 communities are delineated. The job basin of Brussels sprawls out of the Brussels Capital Region into the arrondissement of Halle-Vilvoorde and encompasses Brabant Wallon and even extends to Flanders (West) and Hainaut (South-West). The borders of the communities of Luxembourg, Liège, Antwerp and the two Flanders correspond at some exceptions (such as Waasland where recently major expansions of the Antwerp port are located) to the provincial limits, while Leuven and Limburg merge and so do also the job basins of Charleroi and Namur. For these two latter cases, rail/road infrastructure might strongly explain the merge. When ρ is smaller (0.7), 11 communities are delineated; increasing ρ of 0.1 unit ($\rho = 0.8$) doesn't change the number of communities (11) but the delineation is slightly different (Hainaut shrinks, Namur and Brussels enlarge). In both cases ($\rho = 0.7$ and 0.8) there are more communities in the Northern part of the country than in Wallonia in accordance with the structure of the urban network (more job centres in the North, closer to each other; for historical and geographical reasons, denser city network in the North). When choosing a much larger value of ρ ($\rho = 2$), 6 communities appear: Brussels, West and Oost Vlaanderen (together), Antwerp (including Waasland), Limburg, Liège and Hainaut-Namur. Hatched municipalities correspond to less robust allocations: the classification of these nodes is less certain because they share links with several attracting Belgian job centres (let us here remind that no foreign job centre was considered).

Figure 7. Commuting communities for different values of ρ hatched (municipalities are those where the robustness is < 90%).

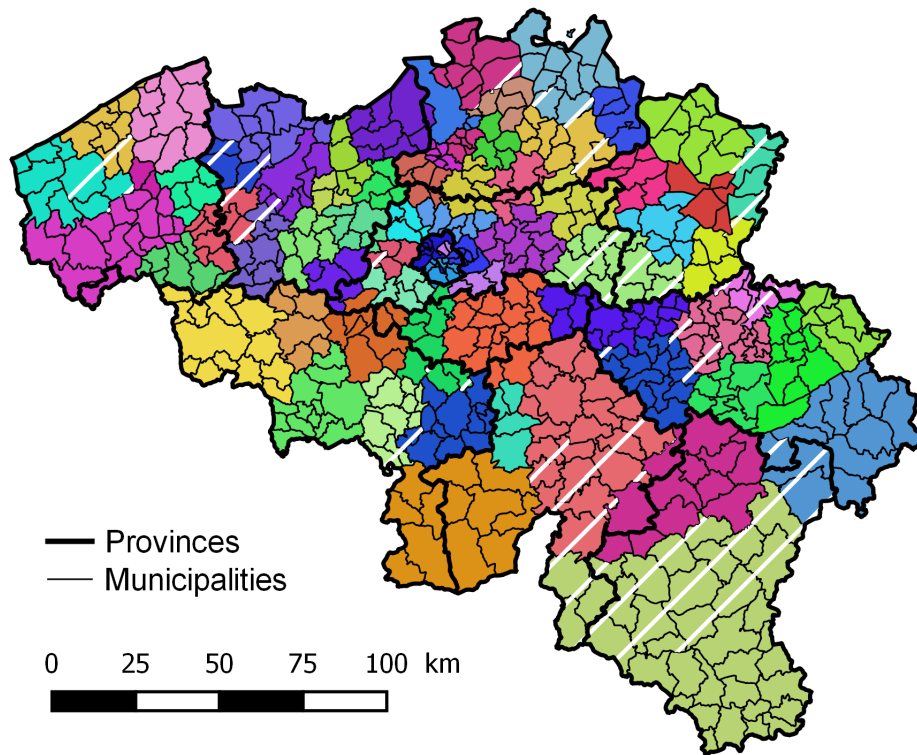


- 23 As concluding comments, we insist on three points about commuting basins: (1) as expected, the larger the value of ρ , the smaller the number of communities. Hence, the search of an optimal value is clearly required in order to avoid misinterpretation. (2) All communities are made up of contiguous municipalities meaning that links between contiguous municipalities are always stronger than between distant places. Space matters. *Gravity type* relationships extraordinarily still matter despite transportation and ICT infrastructures: workers live and work in the same area. (3) The limits of the communities fit quite well those of the provinces as well as that of the linguistic border, at the exception of the southern part of the community of Brussels that sprawls into Brabant Wallon and steps over the linguistic border.

Migratory fields

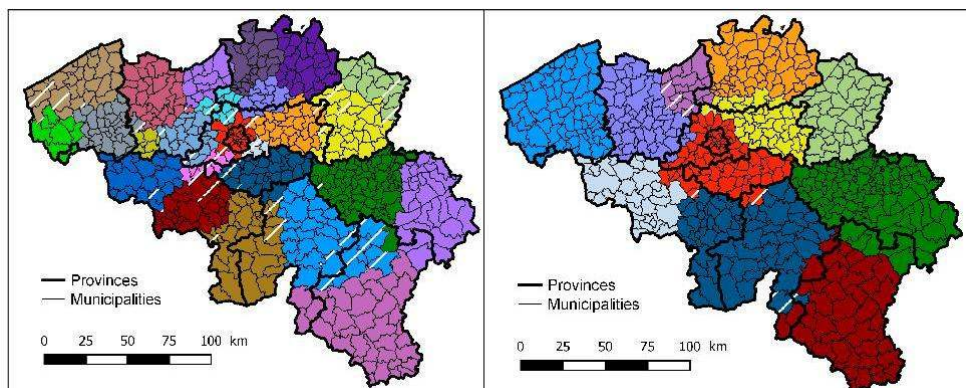
- 24 Let us now take another O-D matrix and apply the Louvain Method on the number of migrations between Belgian municipalities. When considering the standard Louvain Method ($\rho = 1$), 71 migrations fields are delineated (Figure 8): this means that most people decide to change their place of residence for another place in the close vicinity and that “migration basins” are very small compared to commuting basins. Commuting and migration correspond to two different spatial choices processes. Another difference with commuting maps is that the linguistic (regional) border is much more impermeable. This does not mean that nobody moves from Wallonia to Flanders or inversely, but those moves are anecdotic compared to other fluxes.

Figure 8. Migration communities when $\rho = 1.0$ ρ (municipalities are hatched when the stability is $< 90\%$).



- 25 Based on the Stability Method, the optimal value of ρ is this time equal to 3, which corresponds to 24 “migration fields”; this is still much more than the optimal 11 “commuting basins”. Finally, we force the Louvain Method to run with a ρ value equal to 8 in order to get 11 “migration fields” (Figure 9). We end up with a partition resembling that with $\rho = 1$ computed for commuting at the exceptions of Waasland and Leuven that form specific migration fields and also some municipalities at the border of the delineations that are differently allocated. Hence, surprisingly enough, the Louvain Method run on migration data with $\rho = 3$ corresponds to an *optimal* partition, but with a relatively large percentage of instability (hatched surfaces), while the solution with $\rho = 8$ is not considered as optimal but spatially more robust (less hatched surfaces). This robustness has to be related with the number of communities detected: there are more communities when $\rho = 3$ than when $\rho = 8$.

Figure 9. Migration communities for $\rho = 3$ and $\rho = 8$ (municipalities are hatched when the robustness is $< 90\%$).



Comparison

- 26 Differences between migration and commuting “fields” are here quantified as well as their resemblance with provincial delineations. Table 1 refers to the pairwise comparisons of the chorochromatic maps by means of the Normalized Mutual Information (NMI) index which quantifies the amount of information that one partition shares with the other. The NMI index varies between 0 and 1, 1 standing for perfectly identical maps. NMI values confirm the visual observations: they are high between commuting and migration fields, the most important differences being observed for ρ values equal to 1.0 (NMI = 0.640). The maximum value of the NMI is reached for commuting and migrations fields observed with a ρ value of 0.8 for commuting and 8.0 for migration: in this case 88.8 % of the partitioning is similar (NMI = 0.888). The last column of Table 1 gives the NMIs between the provincial delineation and the computed migration/commuting fields. When the number of communities is comparable to the number of provinces (11), NMIs are generally larger than 80% meaning that maps largely look alike, confirming our visual analysis.

Table 1. NMI values between pairs of partitions (municipality level).

	Commuting $\rho = 1.0$, 9c	Migration $\rho = 1.0$, 71c	Migration* $\rho = 3.0$, 28c	Migration $\rho = 8.0$, 11c	Provinces
Commuting* $\rho = 0.9$, 10c	0.896	0.688	0.766	0.872	0.861
Commuting $\rho = 1.0$, 9c	1	0.640	0.711	0.819	0.848
Migration $\rho = 1.0$, 71c		1	0.823	0.715	0.660
Migration* $\rho = 3.0$, 28c			1	0.792	0.731
Migration $\rho = 8.0$, 11c				1	0.834

* OPTIMAL SOLUTION IN FIGURE 4; C STANDS FOR THE NUMBER OF COMMUNITIES

- 27 Comparison is pursued in Table 2. We see that for migration data, the number of communities is much larger (*diversity*) and hence the size of the communities is smaller. In other words, and geographically speaking, distances are smaller when selecting a new place of residence: when households move they want to keep their relatives, their social relationships, neighbours, sports or collective activities and hence do not move far away. When selecting a job place (commuting), workers are ready to travel daily a longer distance; note that this trend is supported in Belgium by several policy tools such a reimbursement of commuting travel, company cars, etc.
- 28 Note also that the second line of Table 2 refers to the stability of the solution on the 1 000 runs is another type of information, different from modularity. For instance, migration data analysed with a $\rho = 3.0$ corresponds to an optimal solution, but with a relatively large percentage of instability, while solution with $\rho = 8.0$ is not optimal but more robust. This once again illustrates that one map (one partition) should always be interpreted carefully, taking into consideration the methodological choices made the researcher. Mathematical optimality and robustness (stability) do not always converge. Data may indeed speak by themselves, but what they reveal is significantly framed by the researcher himself (Miller, 2017).

Table 2. Comparing solutions for partitions based on municipalities.

	Commuting* ($\rho = 0.9$)	Commuting ($\rho = 1.0$)	Migrations ($\rho = 1.0$)	Migrations* ($\rho = 3.0$)	Migrations ($\rho = 8.0$)
# Communities	10	9	71	28	11
% of hatched municipalities	17	5,85	16,87	15,15	2,58

* optimal solution

- 29 Let us go back to Figure 1 that gives the delineation urban agglomerations and the importance of the city, and so implying that not all municipalities are allocated to a centre (white on the map). The objective is hence different from that developed in this paper. Hence, in this paper all municipalities are allocated to a community and no centrality is a priori defined. Hatched surfaces reveal places for which the allocation to a community is less certain and it is a posteriori possible to define centres (see Beckers *et al.*, 2017 or Adam *et al.*, 2017). The objective of the map reported in Figure 1 is hence different from that pursued in this paper, but both bring an interesting and complementary viewpoint on the geography of Belgium.

Concluding comments

- 30 A “new” geography of Belgium is here proposed, relying on two OD matrices (commuting and migrations) and a network based method (the Louvain method) that optimally clusters places without predefining central places. It here proved to be quite useful and informative, especially when adapting the objective function with recent methodological developments (Delvenne *et al.*, 2013). With such methodological choices, one can easily state that “data speak by themselves”: we indeed obtain partitions without fixing any a

priori thresholds, not even including geocoding in the method. We however showed that behind this idealistic methodological viewpoint, there are still some choices to be made by the researcher and hence there is no unique and indisputable solution for defining optimality (choice of ρ , stability of the allocation of a node to a community). Optimality and robustness are two different but mixed concepts that should be weighted when providing partitions to the decision maker, and results have always to be interpreted with a geographical knowledge of the study area confirming a recent paper of Miller (2017).

- 31 Both datasets provide a different mosaic of basins reminding how important it is to control data definition in any quantitative analysis. **Commuting** data reveal a few large interaction fields confirming that commuters travel daily a long distance starting from their place of residence (communities are always made of contiguous municipalities). 10 job basins are designed. Politics in Belgium that allow and encourage the reimbursement of home-work travel expenses reinforce people to search for a job within a larger range of distances. Results obtained at the level of the statistical sectors are exactly the same as those obtained at the level of municipalities. Results are however different when working on **migrations** data: a much larger number of communities is optimally detected, each covering a small number of contiguous municipalities. This refers to “the Belgian way of life”: households are only willing to move within a small distance, where relatives and other social relations are prominent, and therefore being prepared to commute daily over relative large distances for an appropriate job.
- 32 Two common features in these results are worth mentioning as concluding comments. First, distance still matters nowadays in a small country such as Belgium: surprisingly enough, most communities are made out of **contiguous municipalities** strengthening Tobler’s Law in geography, and/or the idea that relationships are more intense between places that are close to each other than often admitted in spatial economics (see e.g. Gaspar and Glaeser (1998) or Fujita and Thisse (2013)). Second, it is quite unexpected to retrieve **provincial borders** when dealing with 2011 data, because the origins of the provinces date back to the end of the 18th century, before industrialisation. After decades of advocating “Functional urban regions” and minimising the impact of administrative borders by geographers, provinces are clearly present on all the maps presented here. This is especially true for the provinces West-Vlaanderen, Oost-Vlaanderen, Limburg, Liège and Luxembourg that appear to be prominent interaction fields. Note that – quite expectedly – the province Antwerp integrates Waasland, but only for commuting and not for migration communities. The community centred on Brussels includes the province of Brabant Wallon and the western part Flemish Brabant, for commuting as well as for migration. Of course these two areas were part of the historical Province of Brabant till 1995, but it is worth mentioning that in the eastern part of Brabant Leuven forms now an independent community. The province of Namur is linked through commuting and migration movements with the area of Charleroi, leaving the western part of Hainaut around Mons and Tournai as a community on its own. We clearly see that there is a juxtaposition of provinces, urban and transportation networks for explaining these maps. The prominence of the provinces in our results is quite unexpected, especially as the political and administrative dimensions of the provinces are losing in importance as spatial units of the organization of economic interaction over the past decades in recent years. In the Flemish Region, it has indeed been agreed that from 2018 most of the provincial “person oriented” services will be transferred to the regional level. In the Walloon Region, the deletion of the provincial level pops up regularly in medias, like a

Loch Ness monster. Resistance to change is anyway multifaceted; some institutional competences are still exerted at the provincial level today, explaining probably some of the observed behavioural realities by means of commuting and migration data. Napoleons' decisions still impact current (behavioural) realities!

- 33 Data Sciences and Big Data open nowadays new avenues for further research in geography (new data, finer spatial and time scales), but one needs to remain very cautious about the meaning and representativeness of these new data, the limits of the method(s) used and especially the geographical interpretation of the obtained results.

BIBLIOGRAPHY

- ADAM A., DELVENNE J.C. & THOMAS I. (2017a), *On the robustness of the Louvain Method to detect communities: examples on Brussels* (submitted to *Journal of Geographical Systems*).
- ADAM A., CHARLIER J., DEBUISSON M., DUPREZ J.P., REGINSTER I. & THOMAS I. (2017b), *Bassins migratoires en Belgique : deux méthodes, une réalité ?* (submitted to *L'Espace géographique*).
- ADAM A., DELVENNE J.C. & THOMAS I. (2017), « Cartographies des champs d'interaction dans et autour de Bruxelles : navettes, déménagements et appels téléphoniques », *Brussels Studies*, 118.
- BECKERS J., THOMAS I., VANOUTRIVE T. & VERHETSEL A. (2017), *Logistics clusters, including inter-firm relations through community detection* (submitted).
- BLONDEL V., KRINGS G. & THOMAS I. (2010), « Régions et frontières de téléphonie mobile en Belgique et dans l'aire métropolitaine bruxelloise », *Brussels Studies*, 42, 12 p. (translations in Dutch and English).
- BLONDEL V., GUILLAUME J.L., LAMBIOTTE R. & LEFEBVRE E. (2008), "Fast unfolding of communities in large networks", *Journal of Statistical Mechanics: Theory and Experiment*, 10, P10008.
- CHARLIER J., DEBUISSON M., DUPREZ J.P. & REGINSTER I. (2016), « Mouvements résidentiels en Wallonie (1994-2014) », Namur, *Working paper de l'IWEPS*, 21, 82 p.
- DE WASSEIGE Y., LAFFUT M., RUYTERS C. & SCHLEIPER P. (2000), « Bassins d'emploi et régions fonctionnelles : méthodologie et définition des bassins d'emploi wallons », *Discussion Paper du SES n°2005*, Namur.
- DECUYPER A., CLOQUET C., THOMAS I. & DELVENNE J.C. (2017), *On the issue of aggregation for community detection in social networks* (to be submitted).
- DELVENNE J.C., SCHAUB M.T., YALIRAKI S.N. & BARAHON M. (2013), "The stability of a graph partition: A dynamics-based framework for community detection", in *Dynamics On and Of Complex Networks*, 2, pp. 221, Complex Networks.
- FORTUNATO S. (2009), "Community detection in graphs", *Physics Reports*, 486, pp. 7-174.
- FUJITA M., THISSE J.F. (2013), "Economics of agglomeration. Cities, Industrial Location, and Globalization", *Cambridge University Press*, 528 p.
- GASPAR J., GLAESER E. (1998), "Information technology and the future of cities", *Journal of Urban Economics*, 43, pp. 136-156.

- GRIMMEAU J.P., DECROLY J.M. & WERTZ I. (2012), « La démographie des communes belges de 1980 à 2010 », *Courrier hebdomadaire du CRISP*, 37, pp. 2162-2163.
- JONES J., CLOQUET C., ADAM A., DECUYPER A. & THOMAS I. (2016), “Belgium through the lens of rail travel requests: does geography still matter?”, *International Journal of Geo-Information*, 5, 11, p. 216.
- KITCHIN R. (2013), “Big data and human geography: Opportunities, challenges and risks”, *Dialogues in Human Geography* 3, 3, pp. 262-267.
- LEEMANS S., PATTYN M., ROUSSEAU S. & VAN DER HAEGEN H. (1989), “De Belgische stadsgewesten 1981”, *Statistische Studieën 1989*, Nationaal Instituut voor de Statistiek, Brussel.
- LUYTEN S., VAN HECKE E. (2007), *De Belgische Stadsgewesten 2001 Brussels: Statistics Belgium*, 84 p.
- MARISSAL P., MEDINA LOCKHART P., VANDERMOTTEN C. & VAN HAMME G. (2006), *Les structures socio-économiques de l'espace belge. Une exploitation des données d'emploi de l'enquête socio-économique de 2001*, Enquête socio-économique 2001, Monographie, 133 p.
- MILLER H. (2017), “Geographic information science II: Mesogeography: Social physics, GIScience and the quest for geographic knowledge”, *Prog. Hum. Geogr.* doi:10.1177/0309132517712154.
- NEWMAN M. (2004), “Detecting community structure in networks”, *Eur. Phys. J.B., The European Physical Journal B*, 3, 82, pp. 321-330.
- THOMAS I., VERHETSEL A. & LORANT V. (2009), « Le recensement de la population : un requiem ? », *Regards Economiques*, 67, 8 p.
- TRAAG V., KRINGS G. & VAN DOOREN P. (2013), “Significant Scales in Community Structure”, *Scientific Reports*, 3, Article 2930.
- VAN DER HAEGEN H., PATTYN M. (1979), “De Belgische stadsgewesten”, *Statistische en Econometrische Studieën*, 59, Nationaal Instituut voor de Statistiek, Brussel.
- VAN DER HAEGEN H., PATTYN M. (1980), “An operationalization of the concept of City Region in West-European perspective: the Belgian City regions”, *Tijdschrift voor Economische en Sociale Geografie*, 71, 2, pp. 70-77.
- VAN DER HAEGEN H., VAN HECKE E. & JUCHTMANS G. (1996), “De Belgische stadsgewesten 1991”, *Statistische Studieën*, 104, Nationaal Instituut voor de Statistiek, Ministerie van Economische zaken.
- VAN DER HAEGEN H., VAN HECKE E. & SAVENBERG S. (2000), “Belgians on the move”, *Belgeo*, pp. 173-188.
- VAN HAMME G., WERTS I. & BIOT V. (2011), « La croissance économique sans le progrès social : l'état des lieux à Bruxelles », *Brussels Studies*, 48.

ABSTRACTS

The objective is to refresh the geography of Belgium using interactions between places by means of a community detection algorithm (Louvain Method) inspired by Complex theory and Data Sciences. Places that are tightly related are optimally clustered into communities, leading to a new and optimal partition of Belgium. Migrations and commuting movements (*Census11*) are here analysed. We obtain a mosaic of “interaction fields” that are here interpreted in terms of methodological choices, human and urban geography as well as Belgian political dilemmas. They give the opportunity to remind that researchers have to control the impact of their

methodological choices and that each type of data leads to a different geographical partitioning, with one major unexpected common spatial feature in Belgium: the pre-eminence of the provincial borders. This perfectly fits with current political questioning.

L'objectif est d'apporter un regard neuf sur la géographie de la Belgique à l'aide de données relationnelles et d'un algorithme de détection de communautés (Méthode de Louvain) inspiré de l'ingénierie mathématique et des sciences de données. Les lieux qui sont fortement liés en termes d'échanges sont ici classés de façon optimale en « communautés », conduisant à une partition innovante de la Belgique. Nous nous limitons ici aux mouvements de navettes et aux déménagements (migrations) issus du dernier recensement (*Census11*).

Les différentes partitions de la Belgique sont discutées et interprétées d'un point de vue géographique et méthodologique. Ces résultats nous donnent l'occasion de rappeler qu'il importe de maîtriser les choix méthodologiques et que chaque type de données conduit à une partition différente qu'il convient d'interpréter en fonction de la théorie. De manière surprenante, une tendance spatiale inattendue apparaît à travers les résultats obtenus : la prééminence des frontières provinciales. Ce résultat interpelle au vu des questionnements politiques actuels en Belgique.

INDEX

Keywords: community detection, interaction fields, migration, commuting, provinces, Belgium, Census11

Mots-clés: détection de communautés, champs d'interactions, migration, navettes, provinces, Belgique, Census11

AUTHORS

ISABELLE THOMAS

Centre for Operations Research and Econometrics (CORE), Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, isabelle.thomas@uclouvain.be

ARNAUD ADAM

Centre for Operations Research and Econometrics (CORE), Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, a.adam@uclouvain.be

ANN VERHETSEL

Department of Transport and Regional Economics, University of Antwerp, Belgium, ann.verhetsel@uantwerpen.be