3112

REPRINT

Arnaud Adam, Gaëtan Montero, Olivier Finance, Ann Verhetsel and Isabelle Thomas

Revisiting urban models with information and communication technology data ? Some examples from Brussels

Chapter 11 in P.O. Plaut and D. Shach-Pinsly (eds.), *Digital Social Networks and Travel Behaviour in Urban Environments. London: Routledge, 186-202, 2019*

CORE CENTER FOR OPERATIONS RESEARCH AND ECONOMETRICS

CORE
Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve
Tel (32 10) 47 43 04
Email:lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/core/reprints.html

11 Revisiting urban models with information and communication technology data? Some examples from Brussels

Arnaud Adam, Gaëtan Montero and Olivier Finance, Ann Verhetsel and Isabelle Thomas

Introduction

Transport and planning analyses mostly start from simple and available geographic urban data, often driven by administrative considerations. The impact of the delineation of the studied area or the size and shape of chosen basic spatial units is almost always set aside, although their impact on the results of transport models is clearly demonstrated (see Jones, Peeters, &Thomas, 2017; Thomas, Jones, Caruso, & Gerber, 2018). This contribution focuses on the impact of the nature of the data used for the partitioning of Brussels city centre. These partitions are a first step in the process of bridging transport analysis and planning through transport demand revealed by diverse dimensions of the city (e.g. social, morphological, economic, etc.).

Partitioning an urban space into groups of places sharing similar properties is common practice for urban geographers and sociologists since the seminal work of the Chicago school (1930s). It is a way to classify elementary spatial units to better understand their similarities and discrepancies and, above all, to better understand intra-urban spatial structures and dynamics within the city. This is usually done in terms of the socio-economic composition of each place but can also be done in terms of interrelationship between georeferenced populations or that of built-up morphologies: each of these aspects are relevant to figure out the intra-urban homo- or heterogeneity. Identifying parts of a city that are similar is part of a nomothetic approach of geography, as opposed to an idiographic one focusing on the strict unicity of each place. Stressing common properties of places avoids putting too much emphasis on too specific results of the same processes. We take part in this approach here by mobilizing new data and recent innovative methods and by applying them to Brussels.

Quantitative revolution (1960s) has facilitated these analyses methodologically, with the development of factorial and cluster analysis, and faster and faster computing systems, even if available data to describe elementary spatial units were rather scarce (urban factorial ecology). Digital revolution has recently exponentially increased the amount of available data (Floridi, 2012; Kitchin, 2013) and, among them, a lot of localized information. This does not mean that these data did not exist before, yet the detection, the registration, the share and the use of these data reached an unprecedented level since the end of the 20th century. The risk with these new emerging data sciences is that some scientists dive into data before elaborating on the specific goals of the research and could lead to data crunching rather than modelling.

Two general kinds of information are commonly used to classify urban elementary spatial units. First, attributes characterizing each elementary spatial unit can be used to compare places based on these specific properties. These attributes can either describe the population living in the units, the economic activities or the characteristics of the built environment. Uni-, bi- and multivariate analyses are used to consider one or many attributes (X_i) to characterize locations, sometimes followed by clustering places into groups sharing common properties (i.e. clusters). Second, geographers use information about the functional interrelationship (w_{ij}) between couples of elementary spatial units i and j. In that case, graph theory can help geographers to analyse the network of interrelations between places, as well as community detection algorithms. The latter aim at delineating groups of places that are highly connected (i.e. communities). For urban geographers interested since a long time in describing and understanding intra-urban structures and dynamics, the emergence of Information and Communication Technology (ICT) data (Global Positioning System, sensors, etc.) offers new ways to get information about places (X_i) and interrelationships between them (w_{ij}). But what do these new data add to urban geography knowledge?

Limiting urban analyses to ICT data is, for sure, not appropriate nowadays: there are no time series, and the spatial representativeness is often questionable. Some parts of the population can be totally missing in the data coverage.[1] Therefore, censuses and surveys are both still relevant for measuring urban complexity. Censuses have the advantage of being well defined, covering almost the entire population and comparable through time. ICT data measure other aspects, often in a very short time lag and at the individual level (see e.g. Longley, Adnan, & Lansley, 2015; Kitchin, 2013; Miller & Goodchild, 2015).

Cities are complex by definition, and the analysis of their complexity includes, among the already mentioned analyses made on people and interactions, the morphology of their built-up components (see e.g. Thomas & Frankhauser, 2013). Therefore, some other analyses will focus on morphology, by using density, fractal dimension and the natural cities method applied to the footprints of the buildings. This will allow the comparison of characteristics of people and their interrelationships through locations with the morphological features of places. In this chapter, we focus on the Brussels Capital Region (noted as BCR, which is an administrative and political entity on its own), that constitutes the dense urban core of the capital city of Belgium.[2] This allows highlighting intra-urban spatial structures.

Four partitions are here elaborated on a selection of data and methods and are presented in the following sections (by order of appearance): a classical "urban factorial ecology", a community detection based on interrelationships revealed by mobile phone calls and from two different ways of measuring the building footprint: a combination of fractal dimension and the density of buildings, and the "natural city" method. Finally, the two last sections present the conclusions and open the discussion.

Urban factorial ecology: a benchmark

Urban factorial ecology was scientifically very fashionable in early quantitative spatial analysis (1960s; see e.g. Berry & Rees, 1969; Hunter, 1972; Johnston, 1978) and had already been applied to Brussels before (see e.g. Dujardin, Selod, & Thomas, 2008). The aim of these methods is to identify groups of places for which inhabitants have similar characteristics (X_i; socio-economic, ethnic, demographic, etc.) by means of a factorial analysis followed by a clustering method (Pruvot & Weber-Klein, 1984). We here use a principal component analysis and a hierarchical cluster analysis in order to identify spatial clusters and further interpret them in terms of classical "urban models".

In the frame of this chapter, several analyses were performed with different sets of variables, with and without a preliminary component analysis. The results in terms of urban spatial structure are a little sensitive to the used method; here, we develop one example. The data used are provided at the scale of the neighbourhood by the Institut Bruxellois des Statistiques et d'Analyses.[3] The selected attributes (Appendix A) are available online as well as a clear definition of each variable, making the data collection simple and reproducible. Among the 145 neighbourhoods of the BCR, 27 are excluded from the analysis because they are not inhabited (mostly correspond to green spaces or industrial zones). The selection of a large number of variables (48) is inspired by the literature on factorial ecology, in order to catch the socio-economic conditions, as well as the demographic and ethnical conditions of each spatial unit.

As expected, variables are highly correlated and can easily be summarized in four components (see Appendix B for the composition of each component). The score of the two first components are rather organized in concentric structures: the first one around the centre (mainly an age gradient), and the second around the European quarter (mainly a "metropolitan activities" gradient). The third component is organized in sectors and is largely based on income and characteristics of the housing stock: large income and large dwellings in the south-east and low income and small dwellings in the north-west. The fourth component reveals a structure in "donut", highlighting places specialized in residential functions at a certain distance from the city centre.

A ward hierarchical cluster analysis based on the scores of the *Principal Component Analysis* (PCA) is then conducted for grouping the neighbourhoods that share similar profiles. Indicators for finding the best number of clusters (Cubic Clustering Criterion, entropy, Dunn, Calinski-Harabasz, shape of dendrogram) show that the optimal number of clusters turns around five to seven. The solution with six clusters is here illustrated (Figure 11.1-a), six being also a number comparable to the number of mobile phone communities that are extracted and presented in the specific section.

Results (Figure 11.1-a) confirm former analyses conducted (see e.g. Vandermotten & Vermoesen, 1995; Thomas & Zenou, 1999): Brussels combines two structures strongly shaping the city: a concentric and a radial one (Hoyt, 1939). The factorial ecology analysis confirms the existence of two crowns around the Central Business District (CBD). Cluster 2 corresponds to the *CBD* with high values of the scores of the second component, little residential. The *first crown* is mainly residential, characterized by a high population density (positive values on Component 4) and divided into two sectors: Cluster 1 (pink) in the south-east which is mainly residential and better-off (high income, large dwellings) and Cluster 3 (green) in the north and west (low income, small housings). The *second crown* is also divided into two parts even if Cluster 6 (orange) and Cluster 4 (purple) are less marked than those of the first crown. Last, Cluster 5 (blue) corresponds to specific urban structures such as large hospitals, sports halls or cultural centres.

Community detection in ICT data: phone basins

Nowadays, there is no need to wait for official census data to “measure” the city: the emergence of new sources of data now allows to sense its “pulses” in real-time. Contrarily to census data describing attributes of places in terms of socio-economic, demographic and ethnic characteristics of their inhabitants (previous section), ICT data enable to characterize and follow people in space and time. ICT data are particularly useful to monitor daily life and short-term processes by detecting and measuring changes within space (Lee & Lee, 2014). They undoubtedly open new perspectives in interaction analyses and urban geography (Batty et al., 2012). This also allows the capture of intra-urban interrelationships between places (w_{ij}). The availability of such data increases every day due to the multiplication of apps on smartphones that enable following people and their geolocations.

We here limit ourselves to one example: mobile phone calls. They already have been proved to be very useful in monitoring and mapping the de facto population, as well as people’s spatial mobility and social networks (Blondel, Krings, & Thomas, 2010; Griffiths, Hostert, Gruebner, & Van der Linden, 2010; Batty et al., 2012). It enables one to better grasp the interactions between people and their environment (see e.g. Ahas et al., 2015; Deville et al., 2014). With this type of data, it is common to extract groups of people or locations with remarkably high interactions, leading to what is called communities.

If this kind of data are nowadays very attractive for approximating “social networks”, they also appear to have major limitations (see e.g. Calabrese, Ferrari, & Blondel, 2014) including a time-consuming “cleaning” phase. The data are not publicly available, and several operators share the market with no information on market shares. It is quite common that the data are aggregated due to confidentiality issues (no distinction between professional and private use; calls are, in our case, located to the closest antenna and not at the exact location of the call).

We here use one month of phone communications of one of the three major service providers operating in Belgium (April/May 2015). The database includes all mobile phone calls, ignoring their duration, between two phone numbers from the provider, geocoded by antenna at the hour of the call. The data set is here limited to the calls for which both antennas are located in the BCR, that is more than 4.9 million calls. To approximate the coverage of each antenna and to improve the readability of the maps, a Voronoi diagram is designed around each antenna (see Adam et al., 2017).

Methods of community detection based on *modularity* maximization are often adopted to detect communities of nodes that interact in large data sets (Newman, 2004). The inputs of these methods are limited to links between nodes (the weights associated with the edges). If the nodes are geocoded, results can further be mapped, and a spatial partition of the studied area is then obtained. To group together the nodes that are tightly connected and hence detecting communities, these methods search for a compromise between minimizing the connections between nodes classified in different communities (the *cut*) and maximizing the number of communities (the *diversity*; Delvenne, Yaliraki, & Barahona, 2010). The *Louvain Method* is considered as a standard heuristic to maximize the *modularity* value despite its limitations (see e.g. Fortunato, 2009; Traag, 2013; Delvenne, Schaub, Yaliraki, & Barahona, 2013). It is commonly applied because it quickly finds partitions of nodes that maximize the *modularity* without defining a priori central places and/or thresholds (Adam et al., 2018a, Thomas et al, 2017).

Applied to the BCR mobile phone data set, the Louvain Method endogenously detects six communities that can be called “phone basins”. A phone basin corresponds to antennas that have a higher propensity to call other antennas than any antenna classified in other communities. In order to avoid a suboptimal solution, the algorithm is run 1,000 times (Adam et al., 2018b). Hence, for each partition, each node can be characterized by the percentage of runs for which it is associated to the same community and so measuring the stability of the partition.

A striking first feature is that most contiguous Voronoi diagrams belong to the same community, this means that calls are more numerous between antennas that are closely located and that people call more often their closest neighbours than those farther located (Tobler, 1970): geography still matters! Table 11.1 represents the matrix of calls emitted and received between the six communities. For each community, the highest number of calls is observed between antennas belonging to the same community; they correspond to the diagonal of the matrix. The intra-community calls represent more than 50% of the calls emitted or received by each community at the exception of Community 6 (45%). In terms of calls made between communities, Table 11.1 and Figure 11.1-c show different realities if the direction of the calls is taken into account. For instance, on one hand, a high number of calls is made from Community 1 to Community 2 (116,986 calls), and on the other, the antennas belonging to Community 2 were many times in communications with antennas from Community 4 (155,949 calls) as well as of Community 1 (123,096 calls).

Table 11.1 Number of calls between communities

Communities	*1*	*2*	*3*	*4*	*5*	*6*
1	684,066	116,986	63,914	67,722	75,337	38,475
2	123,096	741,082	67,416	155,949	54,281	72,337
3	70,716	71,856	263,209	49,667	40,955	35,133
4	71,806	154,892	45,944	629,775	81,803	83,723
5	75,663	52,409	36,826	78,132	316,415	38,700
6	41,577	72,538	33,734	85,011	40,846	217,392

Figure 11.1-b clearly reveals the spatial organization of calls in the BCR. Two of them are located in the middle of the study area: a community centred on the Pentagon[4] with an extension to the North Railway Station (Community 4) and another one around the EU offices (Community 1). These two communities concentrate 43% of the calls made within the BCR; the other communities are further organized in sectors around these two.

Morphometrics of built-up footprint

The objective of this section is to give an image of the built-up disparities and further partition the city in terms of built-up urban similar morphologies. The literature mentions many indices mainly issued from landscape ecology for measuring and characterizing urban morphologies (see e.g. Medda, Nijkamp, & Rietveld, 1998; Schwarz, 2010; Caruso, Hilal, & Thomas, 2017). We here select and compare two methods to quantitatively grasp the morphological reality of Brussels: fractal dimension combined with density and the “natural cities” (presented in the following sections).

We use the 2009 Cadastral footprint of the buildings as well as the centroids of each building. Every isolated small (less than 20 m^2) building was erased from the database. Analyses are performed on an area larger than the BCR (an area encompassing the entire former province of Brabant that surrounds the BCR) in order to avoid border effects. From this large set, we here isolated and present the results of the BCR. In this analysis, we take care of these border effects because it is well known that fractal dimension and “natural cities” methods are particularly sensitive to these border effects (Montero, Tannier, & Thomas, 2018).

Fractal dimension and density

Built-up urban fabrics with complex geometrical features cannot be described only by simple tools based on Euclidean geometry. Fractal geometry provides an interesting alternative to compare irregular forms, even at different spatial scales (Batty & Kim, 1992). Fractal dimension (D) characterizes the scaling behaviour of fractals, that is the fact that the same structure statically appears on smaller nested scales. For surfaces, D varies between 0 and 2: 2 corresponds to a homogeneous pattern where the mass is distributed uniformly over space, while 0 is a quite unrealistic value corresponding to isolated points without any particular spatial arrangement. The more D is different from 2, the more the patterns show empty areas of different sizes; these empty areas are distributed according to a strong hierarchical law. When $D < 1$, the pattern consists of disconnected, isolated elements concentrated in clusters which are separated by lacunas of different sizes. Hence, the lower D is, the less homogenously the built-up areas are distributed over space. Fractal dimension is independent of the unit of measurement (see e.g. Thomas, Frankhauser, & De Keersmaecker, 2007; Thomas, Cotteels, Jones, & Peeters, 2012).

Fractal dimension can be seen as a proxy of mean density, but the two are not spatially equivalent. Fractal dimension relates to morphology (the internal structure of the built-up areas) while mean density gives a rough idea of the occupation of the area. Geographically weighted fractal analysis is used for computing local fractal dimensions. This method mixes the sandbox multifractal algorithm (see e.g. Vicsek, 2002) and a geographically weighted regression with a kernel to estimate the fractal dimension of cells in a regular grid (Sémécurbe, Tannier, & Roux, 2019). In order to provide results at the finest possible resolution and to avoid estimation problems of the fractal dimension, we use 250-m × 250-m resolution cells.

Density is here simply expressed as the ratio of the surface occupied by the buildings divided by the total surface of the basic spatial unit. The studied area is hence covered by a grid of squared cells for which two values are calculated: density and fractal dimension. A Ward clustering analysis is applied. Results are reported in Figure 11.1-d. The data set is clustered in four crowns in a clear centre–periphery structure. From the outskirts to the centre: (1) low density and non-homogeneous built-up surfaces, (2) average density and average fractal dimension, (3) high density and average homogeneity and (4) very high density and high homogeneity. This structure reminds clearly the classical concentric urban structure model of Alonso-Muth (see e.g. Verhetsel, Thomas, & Beelen, 2010).

From "natural cities" to "natural urban clusters"

A set of papers has recently been initiated by Jiang around the concept of "natural cities" (see e.g. Jia & Jiang, 2011; Jiang, 2013; Jiang & Miao, 2015). The methodology relies on a head/tail division rule to derive "natural" clusters (called "cities" in Jiang's papers), based on the assumption that there are far more smalls things than larger ones. A triangulated irregular network (TIN) is used, made up of individual locations that are considered as *nodes* (initially street nodes and later any location-based social media users such as Twitter or Brightkite, here we adapt the method by using the centroid of each building in order to make the results comparable with the previous section). The method identifies the edges between the nodes (on the TIN) smaller than the average length of the edges as limits of "natural cities". The method relies on the fact that the length of the edges follows a hierarchical distribution where the mean length is used as a threshold for delineating the "cities" (large number of small edges corresponding to cities and small number of large edges corresponding to rural areas). The method appears to be very attractive in a "mechanical" way, but questions remain about its anchoring in urban geography theory and its combination with urban functional issues (see Montero et al., 2018, for a critical and comparative analysis). We here decided to use it for detecting intra-urban clusters within the BCR: Can the BCR be partitioned into "*natural urban clusters*"? Are some parts of the city characterized by tighter networks than others?

If several authors have already supported the use of streets networks for delineating cities (Jia & Jiang, 2011; Arcaute et al., 2015; Masucci, Arcaute, Hatna, Stanilov, & Batty, 2015), Thomas et al. (2012) showed that the spatial organization of buildings within a city is rather different from that of street networks. Hence, in order to compare results with the previous section, we here consider the centroid of each building as the nodes used in the TIN.

The method delineates zones where the spatial proximity of the centroids is higher than the mean length in the study area (48 m). Patches are clearly identified and mapped (Figure 11.1-e). The white space between the patches corresponds to large boulevards, wide infrastructures as railways, green spaces, squares, etc., or simply built-up wards with larger distances between their centroids (larger buildings or larger gardens). Some coloured patches are large and reflect the history and geography of the city (see e.g. Vandermotten & Vermoesen, 1995). Some others are very small, corresponding to specific allotments or pinpoint urban projects. Figure 11.1-e clearly tells another story even if some resemblance exist; there is a clear resemblance between socio-economic realities (Figure 11.1-a) and built-up landscapes (Figure 11.1-e) but the method used for extracting urban built-up landscapes influences the result (Figure 11.1-d and Figure 11.1-e are clearly different).

Complementarity of the four partitions

The four partitions presented in Figure 11.1 result from a unique objective (understanding the city), but the type of data and the methodology, as well as the basic spatial units, differ. Yet some commonalities are observed between the partitions, with some parts of the city being for example encompassed both in a given cluster in the factorial ecology and a specific community into the group of antennas.

An attempt at synthesis is proposed in Table 11.2. The spatial structure of each partition is clearly not similar, at the exception of the specificity of the core area that we found in each of them. Communities centred on the Pentagon and on EU offices are grouped in a socio-economic cluster characterized by a concentration of offices and urban facilities (CBD). The CBD and the first-crown in the factorial ecology are characterized by a dense and homogeneous organization of the built-up environment according to the morphological classification. On the opposite, it is obvious that the comparison of a pure centre–periphery gradient (clustering using the fractal dimension and the built-up density) and a sectoral structure (communities based on cell phone data) will not lead to a strong resemblance between the two partitions.

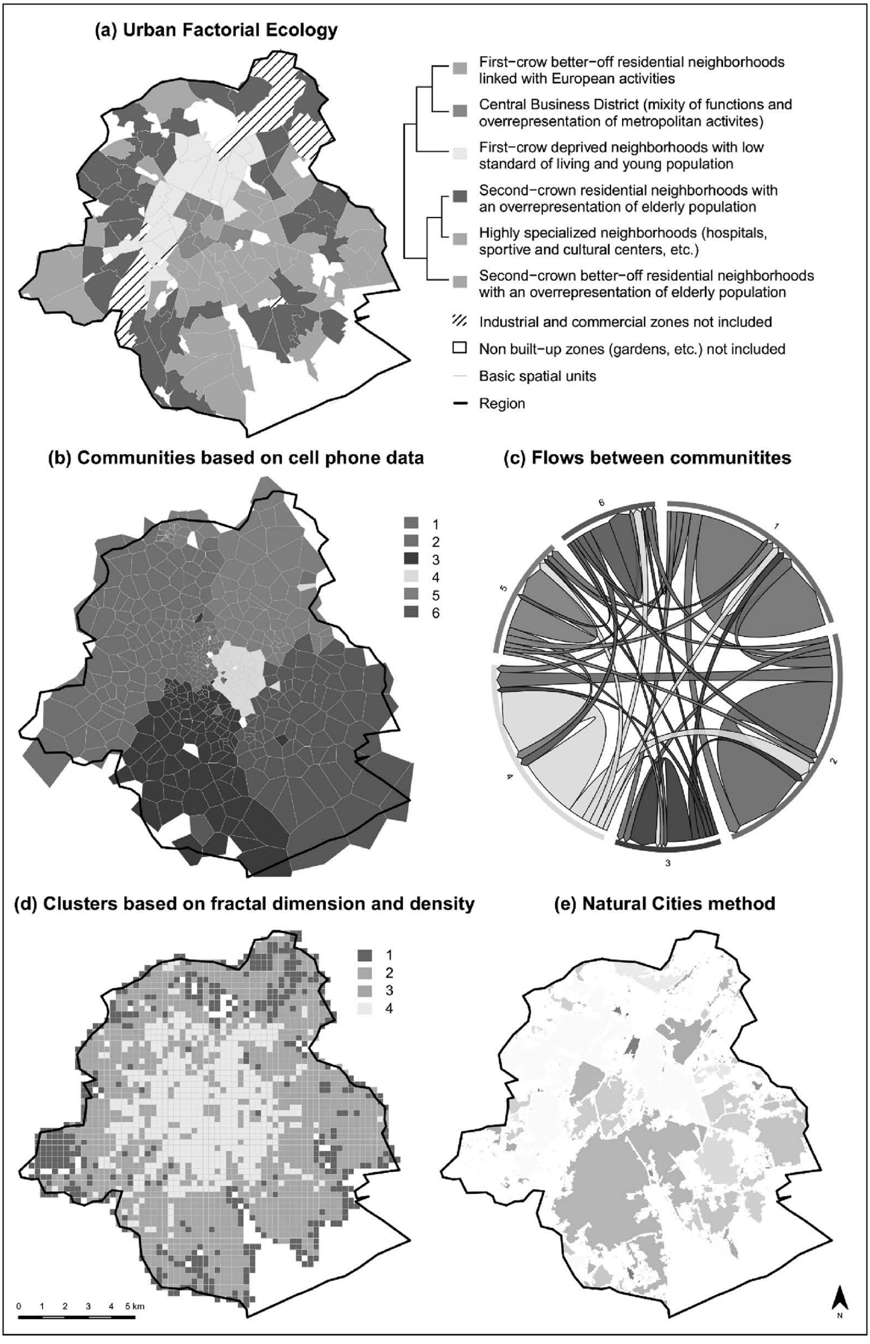

Figure 11.1 Mapping Brussels differently (coloured version of this figure is to be found in https://atlas.brussels/publications/chapter-revisiting-urban-models-coloured-versions-of-figures/)

Table 11.2 Comparison of the four partitions of Brussels

	Type of data	*Basic spatial units*	*Approximative spatial structure of the clustering*
Factorial ecology (PCA + clustering)	Socio-economic, conventional (census)	Statistical neighbourhoods	Combined centre–periphery and sectoral structures
Communities based in cell phone	Relational, unconventional	Voronoi diagrams around antennas	Sectoral structure
Fractal dimension and density	Built-up footprints	Cells grid	Centre–periphery structure
"Natural cities"	Built-up footprints (centroid)	/	Multiple nuclei

The main explanation of the difficulties in comparing the different partitions origins from the fact that the elementary spatial unit is not the same in each analysis. The strict comparison is probably biased by the difference of size, shape and distribution of the basic spatial units (statistical neighbourhoods, Voronoi cells, grid cells) in the city. This is a constraint which we cannot overcome easily because the basic spatial unit is given by the data themselves, and this has to be kept in mind when comparing the different results. It reduces the possibilities of integrating the different kinds of data in a single approach.

Conclusion

With the example of Brussels, this contribution confirms the complexity of the urban structure and the difficulties in fitting data, method and objective. There is no unique way of measuring this complexity: each database and methodology leads to different spatial structures. With this case study, the classical urban models are partly rediscovered by means of new tools and new data. Concentric model (Alonso-Muth/Model of Burgess), sectoral model (Hoyt) and multiple nuclei model (Harris and Ullman) are clearly appearing in Figure 11.1 confirming former studies: Brussels is a superposition of different spatial patterns. Also, for results based on mobile phone data, there is a strong contiguity in the communities: one phones more people who are located closer, reminding one of Toblers's first law of geography.[5]

The definition of the data needs to be controlled, as well as the underlying processes and the research objectives. We cannot let the data speak by themselves, and this is especially true for new ICT data sources. We need to capture the meaning of the data. For ages, geographers have struggled for individual locations to understand processes; ICT data do not change this, as most data are protected by confidentiality reasons (and further, once again, aggregated here by antennas).

Because we did not have the opportunity to conduct all four analyses with the same basic spatial units, the partitions obtained in this contribution are simply juxtaposed and compared between them. Future works will now have to be dedicated to the combination of partitions.

Finally, visualizing data is now current for ICT data, but modelling and linking the results to theory and planning are challenging exercises. There is a clear need to develop and understand methods and link them, first, with spatial theories in a multidisciplinary context before the results can be applied in planning exercises. ICT data are an opportunity to renew quantitative geography, especially as big data enables managing the complexity of interrelationships at local/global scales to add information to conventional data, but big data do not replace them. There is further a clear need for analysis comparing the obtained networks at different scales and for different node definitions (assortativity). Each of the three dimensions we focused on in this chapter brings a distinct point of view and a specific added value. One of the further steps could now be to use methods that mix community detection based on the interactions and cluster analyses based on similarities between places. ICT data are not only a smokescreen: they open new avenues for further dynamic multidisciplinary analyses and for modelling urban realities for the purpose of transport planning.

Acknowledgements – This article was partly written in the framework of the Bru-Net research project financed by Innoviris (Brussels). The authors also wish to thank the telephone provider for the excellent collaboration in the creation of a telephone database and for agreeing to this publication.

Notes

1 For example, there is an over representation of young and male Twitter users in the total resident population in London (Longley et al., 2015).
2 Results shown here are parts of a broader project named Bru-Net that is financed by the Brussels Research Agency and that aims at measuring spatial communities in the whole metropolitan area around Brussels (former province of Brabant). We here limit our analyses to the Brussels Capital Region (BCR), the 19 administrative municipalities of the Brussels core.
3 The "Monitoring des Quartiers" is an interactive tool available online. It has the objective to make available a selection of indicators that characterize disparities and dynamics within the BCR. Not all data are from the same date. Its availability makes the analysis easily reproducible. Retrieved from http://ibsa.brussels/chiffres/chiffres-par-quartier#.WnoZLOjOXyQ.
4 Usual name of the area located inside the former 14th-century walls of the city.
5 "Everything is related to everything else, but near things are more related than distant things" (1970).

References

Adam, A., Charlier, J., Debuisson, M., Duprez, J.-P., Reginster, I., & Thomas, I. (2018a). Bassins résidentiels en Belgique: deux methodes, une réalité ? *L'Espace Géographique*, *1*, 35–50.

Adam, A., Delvenne, J.-C., & Thomas, I. (2017). Cartography of interaction fields in and around Brussels: Commuting, moves and telephone calls. *Brussels Studies*, *118*. Online since December 18, 2017.

Adam, A., Delvenne, J.-C., & Thomas, I. (2018b). On the robustness of the Louvain method to detect communities: Examples on Brussels. *Journal of Geographical Systems*, *20*(4), 363–386.

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., Zook, M. (2015). Everyday space – time geographies: Using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn. *International Journal of Geographical Information Science*, *29*, 2017–2039.

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing cities, deconstructing scaling laws. *Journal of Royal Society Interface*, *12*, 20140745. doi:10.1098/rsif.2014.0745

Batty, M., Axhausen, K., Giannotti, F., Posdnouktov, A., Bazzani, A., Wachowicz, M. . . . Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, *214*, 481–518. doi:10.1140/epjst/e2012–01703–3

Batty, M., & Kim, S. (1992). Form follows function: Reformulating urban population density functions. *Urban Studies*, *29*, 1043–1070.

Berry, B., & Rees, P. (1969). The factorial ecology of Calcutta. *American Journal of Sociology*, *74*, 445–491.

Blondel, V., Krings, G., & Thomas, I. (2010). Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone. *Brussels Studies*. https://doi.org/10.4000/brussels.806

Calabrese, F., Ferrari, L., & Blondel, V. (2014). Urban Sensing using mobile phone network data: A survey of research. *ACM Computing Surveys*, *47*, 1–20.

Caruso, G., Hilal, M., & Thomas, I. (2017). Measuring urban forms from inter-building distances: Combining MST graphs with a Local Index of Spatial Association. *Landscape and Urban Planning*, *163*, 80–89.

Delvenne, J.-C., Schaub, M. T., Yaliraki, S. N., & Barahona, M. (2013). The stability of a graph partition: A dynamics-based framework for community detection. In *Dynamics on and of complex networks* (Vol. 2, pp. 221–242). New York, NY: Springer.

Delvenne, J.-C., Yaliraki, S. N., & Barahona, M. (2010). Stability of graph communities across time scales. *PNAS* (*Proceedings of the National Academy of Sciences of the USA*), *107*, 12755–12760. https://doi.org/10.1073/pnas.0903215107

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F., Gaughan, A. . . . Tatem, A. (2014). Dynamic population mapping using mobile phone data. *PNAS (Proceedings of the National Academy of Sciences of the USA)*, *11*(45), 15888–15893. https://doi.org/10.1073/pnas.1408439111

Dujardin, C., Selod, H., & Thomas, I. (2008). Residential segregation and unemployment: The case of Brussels. *Urban Studies*, *45*(1), 89–113.

Floridi, L. (2012). Big data and their epistemological challenge. *Philosophy & Technology*, *25*, 435–437. https://doi.org/10.1007/s13347-012-0093-4

Fortunato, S. (2009). Community detection in graphs. *Physics Reports*, *486*, 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Griffiths, P., Hostert, P., Gruebner, O., & Van der Linden, S. (2010). Mapping megacity growth with multi-sensor data. *Remote Sensing of Environment*, *114* (2), 426–439. doi:10.1016/j.rse.2009.09.012

Hoyt, H. (1939). *The structure and growth of residential neighborhoods in American cities*. Washington, DC: Government Printing Office.

Hunter, A. (1972). Factorial ecology: A critique and some suggestions. *Demography*, *9*(1), 107–117.

Jia, T., & Jiang, B. (2011). Measuring urban sprawl based on massive street nodes and the novel concept of natural cities. Preprint, arxiv.org/abs/1010.0541

Jiang, B. (2013). Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *The Professional Geographer*, *65*(3), 482–494.

Jiang, B., & Miao, Y. (2015). The evolution of natural cities from the perspective of location-based social media. *The Professional Geographer*, *67*(2), 95–306.

Johnston, R. (1978). Residential area characteristics: Research methods for identifying urban sub-areas analysis and factorial ecology. In D. T. Herbert & R. J. Johnston (Eds.), *Social areas in cities: Processes, patterns and problems* (pp. 175–217). Chichester, West Sussex: John Wiley and Sons.

Jones, J., Peeters, D., & Thomas, I. (2017). Scale effect in a LUTI model of Brussels: Challenges for policy evaluation. *European Journal of Transport and Infrastructure Research (EJTIR)*, *17*(1), 103–131.

Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography*, *3*(3), 262–267.

Lee, J., & Lee, H. (2014). Developing and validating a citizen-centric typology for smart city services. *Government Information Quarterly, ICEGOV*, *2012*(Supplement 31), S93–S105. https://doi.org/10.1016/j.giq.2014.01.010

Longley, P. A., Adnan, M., & Lansley, G. (2015). The geotemporal demographics of twitter usage. *Environment and Planning A: Economy and Space*, *47*, 465–484. https://doi.org/10.1068/a130122p

Masucci, A. P., Arcaute, E., Hatna, E., Stanilov, K., & Batty, M. (2015). On the problem of boundaries and scaling for urban street networks. *Journal of the Royal Society Interface*, *12*(102). http://dx.doi.org/10.1098/rsif.2015.0763

Medda, F., Nijkamp, P., & Rietveld, P. (1998). Recognition and classification of urban shapes. *Geographical Analysis*, *30*(4), 304–314.

Miller, H., & Goodchild, M. (2015). Data-driven geography. *GeoJournal*, *80*(4), 449–461.

Montero, G., Tannier, C., & Thomas, I. (2018). Morphological delineation of an urban space: Critical analysis of three methodologies. *The case study of Brussels* (on going).

Newman, M. E. J. (2004). Detecting community structure in networks. *The European Physical Journal B, 38*, 321–330. https://doi.org/10.1140/epjb/e2004-00124-y

Pruvot, M., & Weber-Klein, C. (1984). Ecologie urbaine factorielle: essai méthodologique et application à Strasbourg. *L'Espace Géographique*, *13*(2), 136–150.

Schwarz, N. (2010). Urban form revisited – Selecting indicators for characterising European cities. *Landscape and Urban Planning*, *96*(1), 29–47.

Sémécurbe, F., Tannier, C., & Roux, S. G. (2019). Applying two fractal methods to characterise the local and global deviations from scale invariance of built patterns throughout mainland. *Journal of Geographical Systems*, *21*, 271. https://doi.org/10.1007/s10109-018-0286-1

Thomas, I., Adam, A., & Verhetsel, A. (2017). Migration and commuting interaction fields: A new geography with a community detection algorithm? Belgeo [En ligne], 4, doi: 10.4000/belgeo.20507Thomas, I., Cotteels, C., Jones, J., & Peeters, D. (2012). Revisiting the extension of the Brussels urban agglomeration: New methods, new data, new results? *E-Belgeo* (on-line). Retrieved from http://belgeo.revues.org/6074

Thomas, I., & Frankhauser, P. (2013). Fractal dimensions of the built-up footprint: Buildings versus roads: Fractal evidence from Antwerp (Belgium). *Environment and Planning B: Planning and Design*, *40*(2), 310–329.

Thomas, I., Frankhauser, P., & De Keersmaecker, M.-L. (2007). Fractal dimension versus density of the built-up surfaces in the periphery of Brussels. *Papers in Regional Science*, *86*(2), 287–307.

Thomas, I., Jones, J., Caruso, G., & Gerber, P. (2018). City delineation in LUTI models: Review and tests. *Transportation Reviews*, *38*(1), 6–32.

Thomas, I., & Zenou, Y. (1999). *Ségrégation urbaine et discrimination sur le marché du travail: le cas de Bruxelles*. In M. Catin, J.-Y. Lesueur, & Y. Zenou (Eds.), *Emploi, concurrence et concentration spatiales* (pp. 105–127). Paris: Economica.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*, 234–240.

Traag, V. (2013). *Algorithms and Dynamical models for communities and reputation in social networks*. Ph.D. at EPL, Université catholique de Louvain, Louvain la Neuve.

Vandermotten, C., & Vermoesen, F. (1995). Structures sociales comparées de l'espace de trois villes européennes: Paris, Bruxelles, Amsterdam. *Espace Populations Sociétés*, *3*, 395–404.

Verhetsel, A., Thomas, I., & Beelen, M. (2010). Commuting in Belgian metropolitan areas: The power of the Alonso-Muth model. *Journal of Transport and Land Use*, 2(3-4), 109–131.

Vicsek, T. (2002). Complexity: The bigger picture. *Nature*, *418*, 131.

Appendix A

48 variables used in the factorial ecology

% children in the neighbourhood and around, enrolled in kindergartens in the neighbourhood.
% children in the neighbourhood and around, enrolled in a primary school in the neighbourhood.
Population density (inhabitants/sq km).
Offices density (sq meters/sq km).
% of buildings with 5 floors and more.
% of impervious surfaces.
% dwellings built before 1961.
% households unsatisfied with the cleanness of the environment around their housing.
% of population aged 18–29.
% of population aged 65–79.
% population aged 80 and more.
Average age (years).
Mobility index (sum of the immigrants and emigrants divided by total population). Sedentariness index (non-migrants divided by the total population).
% of couples with children in total number of private households.
% of EU population (EU15) in the total population.
% of North Africa population in the total population.
% of Latin America population in the total population.
% of Sub-Saharan Africa population in the total population.
% of foreigners in the total number of inhabitants.
% of inhabitants with French nationality.
% people living alone within the 18–29y old.
% people living alone within the 65 and more.
Mean size of private household.
Index of masculinity (number of men *100 divided by the number of women).
% of the 0–17y old. in the total population.
% population aged 65 and more.Economic dependence coefficient (number of 0–17 y and 65+ y divided by the number of 18–64 y).
Ageing coefficient (%) (population of 65+y divided by the population of 0–17y).
Activity rate (%) (number of actives divided by population of 18–64 y).
Application rate (%) (total number of unemployment people divided by population of 18–64 y)
Unemployment rate (%) (total number of unemployment people divided by the active population)
Medium income of the tax report (€).
% of the salaried within the labor force.Employment rate (%).

Absolute difference between male and female activity rates.
% households residing in an apartment.
% dwellings under 55 sq meters.
% dwellings larger than 104 sq meters.
Average area of a dwelling (sq m).
Average area of a dwelling by inhabitant.
% of the social housing (number of social housing for 100 households).
% dwellings occupied by the owner.
Number of renovation subsidies (for 1000 households) (‰).
% of inhabitants living close from a public transportation stop (250 m from the bus, 400m from a tram and 500m from the metro).
Density of private households (number of private households divided by the area of the district)
% of children (less than 18y) living in a household without labour income.
% inhabitants estimating themselves not to be in good health.

Appendix B

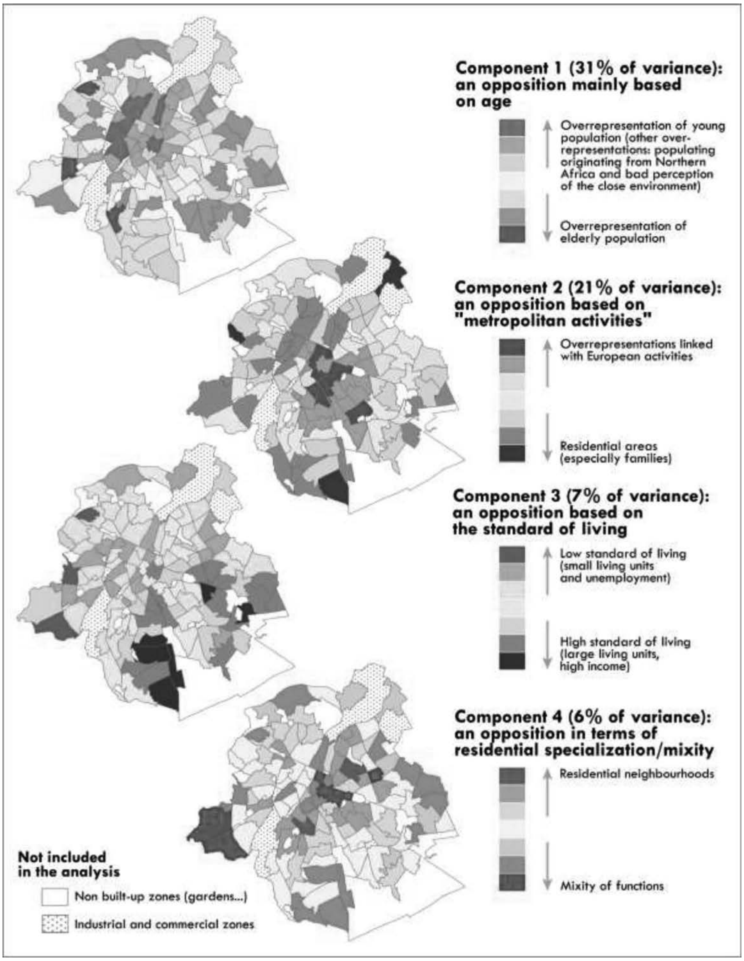

Figure 11.B Four first principal components of the PCA based of variables listed in Appendix A (coloured version of this figure is to be found in https://atlas.brussels/publications/chapter-revisiting-urban-models-coloured-versions-of-figures/)