MONTERO G., TANNIER C., THOMAS I. (2021), Delineation of cities based on scaling properties of urban patterns: a comparison of three methods, *International Journal of Geographical Information Science*, DOI: 10.1080/13658816.2020.1817462

Delineation of cities based on scaling properties of urban patterns: a comparison of three methods

Abstract

Identifying urban boundaries involves analysing both the functional and the morphological aspects of urban systems. In this paper, we adopt a purely morphological approach and compare three methods for the morphological delineation of cities. Each method avoids using any predefined quantified threshold (size, distance, built density, etc.) to detect crucial discontinuities in space. The first method identifies *Natural Cities* by clustering points. The other two are the fractal-based *MorphoLim* method and the *Hierarchical Percolation*; both involve transforming the data using a step-by-step dilation process. The three methods are critically compared and illustrated by applications to theoretical urban patterns. We further apply each method to the urban agglomeration of Brussels, the Belgian capital, using different data (building footprints, building centroids and street nodes) and considering two study areas, namely the former province of Brabant and the entire country of Belgium. The results show that it is impossible to draw an unambiguous morphological boundary for an urban agglomeration. Consequently, it is crucial to relate the data used, the size of the study area and the method chosen to the objectives of the delineation.

Keywords: urban boundary; urban patterns; scaling laws; fractal; percolation.

1. Introduction

Defining city boundaries is invariably an enthralling challenge for planners and geographers. Planners need to know where cities start and end in order to govern and manage them efficiently (count populations, levy taxes, organise transport systems, etc.); geographers need to know the city limits in order to map urban growth, characterise land-use changes, compare cities (their areaand population), etc. Sotomayor-Gómez and Samaniego (2020) quote a United Nations World Cities report, which indicates that countries' urban proportions face multiple non-comparable definitions. As an example, the city of Toronto in Canada may be considered to have between 2.6 and 5.1

million inhabitants, depending on which of the three definitions of city is used to calculate its size. According to Sotomayor-Gómez and Samaniego (2020), such variable figures reflect annual growth rates ranging from 0.9% to 1.8%, which may clearly lead to significant complications for efficient planning efforts. Moreover, in most disciplines, assessing the effect of system boundaries is important to provide certainty and robustness in model findings. This has long been an issue for geographers, from either a theoretical or an empirical perspective as explained in Thomas *et al.* (2018).

Even though the "bounding problem" was clearly defined by Haggett (1972) early on, nowadays there is no consensus on the most appropriate way to delineate urban agglomerations either in terms of methods or in terms of criteria or thresholds, and hence results do not converge (see Cottineau *et al.*, 2018; Ferreira *et al.*, 2010; Rozenfeld *et al.*, 2011; Thomas *et al.*, 2018, for examples and reviews).

A first way to delineate urban areas is to consider morphological criteria alone based on physical data (buildings, road networks, etc.) (e.g. Medda *et al.*, 1998; Thomas *et al.*, 2018). A second way is to consider functional criteria based on variables characterising either places and people living in these places, or exchanges (flows) between places (Blondel *et al.*, 2010; Guérois *et al.*, 2014; Kauffmann, 2012; Sotomayor-Gómez and Samaniego, 2020). Here, we have decided to adopt a morphological approach because the criteria used are often easily comparable internationally (Fang and Yu, 2017; Gisbert *et al.*, 2018; Parr, 2007; Weber, 2001). In comparison, functional approaches are less inclusive and capture only part of the urban functioning, depending on the variables considered (Thomas *et al.*, 2018).

Methods used to delineate cities by morphological criteria can be grouped into four categories:

1. Density measures – a density threshold is set, above which a place is considered "urban". This threshold can be applied on discrete maps, where elementary spatial units are administrative units or cells of a grid, or on smoothed maps using a kernel density estimation (Borruso, 2003; Oliveira *et al.*, 2014; Zhou, 2015; Zhou & Guo, 2018).

2. Classification of remote sensing images or aerial photographs, so that similar spatial units (pixels) can be grouped together. Usually a criterion of contiguity—generally involving a distance threshold—is added (Abed & Kaysi, 2003; Nicolau & Cavaco, 2018; Weber, 2001).

3. Clustering methods that group elements (usually buildings) according to their proximity or other predefined common characteristics (Caruso *et al.*, 2017; Chaudhry & Mackaness, 2008; Deng *et al.*, 2011; Maisonobe *et al.*, 2018; Rozenfeld *et al.*, 2008, 2011; Usui, 2019; Wu *et al.*, 2017).

4. Analyses of a statistical distribution across scales — for example, distribution of distances between buildings, street blocks, road intersections or points of interest — and identification of a threshold in this scaling distribution (Arcaute *et al.*, 2016; Jiang, 2013, 2015; Jiang & Jia, 2011; Jiang & Liu, 2012; Jiang & Miao, 2015; Liu *et al.*, 2019; Long, 2016; Masucci *et al.*, 2015; Tannier *et al.*, 2011; Tannier & Thomas, 2013).

The present paper is anchored in the fourth category of methods that avoids using any predefined quantified threshold (size, distance, built density, etc.) to detect crucial discontinuities in space. Thresholds are identified a posteriori; they are specific to each area under study and cities characterised by similar global densities may exhibit different thresholds. Avoiding ex ante thresholds is especially valuable, as it makes it possible to compare cities of varying sizes and shapes (Le Gléau *et al.*, 1997). The first method under consideration derives *Natural Cities* by clustering points (street nodes or building centroids). It combines several methodological processes proposed in a series of five papers (Jiang & Jia, 2011, Jiang, 2013, 2015, 2019; Jiang & Miao, 2015). Two thresholds are successively identified: first, the mean distance between the points of interest, which enables us to create clusters of points; second, the mean area of the clusters of points calculated via the recursive *head/tail breaks* method applied to the rank-size distribution of the clusters. The *Natural Cities* identified on this basis are the clusters whose area is larger than the nested means. The other two methods under consideration are the *Hierarchical Percolation* (Arcaute *et al.*, 2016; Masucci *et al.*, 2015) and fractal-based *MorphoLim* methods (Tannier *et al.*, 2011; Tannier & Thomas, 2013). Both involve transforming the data (buildings mapped in 2D or road intersections) using a step-by-step dilation. The resulting dilation curves represent the size of the dilation buffer on the *x*-axis and the change in the number of clusters (in the case of *MorphoLim*) or the number of road intersections in the biggest cluster (in the case of *Hierarchical Percolation*) on the *y*-axis. A crucial threshold is then identified on the dilation curve. This threshold corresponds to a break point in the fractal organisation of the built pattern (*MorphoLim*) or to the condensation threshold of a logistic function (*Hierarchical Percolation*).

Few papers broach the comparison of the morphological limits of cities resulting from different methods (Borruso, 2003; Cetinkaya, 2015; Zhou, 2015) and none tackle the comparison of the three methods cited above. This is the research objective of this paper. It is well-known that case studies in geography are highly dependent on the size, shape and position of the spatial units used for the analysis (MAUP: Modifiable Areal Unit Problem) (Briant *et al.*, 2010; Openshaw, 1983), as well as the spatial extent of the study area (Goodchild, 2016). Therefore, we seek to determine the impact of these effects on the delineation of morphological agglomerations for each of the three methods. Data used are of two types: the footprint of buildings (2D) or their centroid (1D), and the road intersections (1D). Such data are easy to collect, freely available in many countries and often easy to handle.

This paper is structured as follows: a critical analysis of each method based on the analysis of theoretical urban patterns (Section 2), the application of the three methods to delineate the morphological agglomeration of Brussels (Belgium) (Section 3), and a conclusion discussing the impossibility of defining a single unambiguous morphological limit of a city (Section 4). The theoretical and empirical urban patterns analysed in sections 2 and 3 are freely available in the form of shapefiles (see section entitled “Data and codes availability statement” at the end of this paper).

2. Description and comparison of three delineation methods based on the analysis of theoretical urban patterns

2.1 Natural Cities

The data used to identify *Natural Cities* are street nodes including intersections and end points (Jiang & Jia, 2011) (Figure 1a), although other types of points could be used, e.g. social media users’ locations (Jiang & Miao, 2015). The methodological process adopted to delineate *Natural Cities* begins with the creation of a *Triangular Irregular Network* (TIN) as proposed by Jiang & Miao (2015) (Figure 1b). On this basis, we plot the rank-size distribution of the edges of the TIN according to their length and we verify that this distribution is heavily right-skewed, with a minority of large values in the head and a majority of small values in the tail. If the condition is confirmed, we apply the head-tail division rule set out by Jiang & Liu (2011): “given a variable X, if its values x follow a heavy tailed distribution, then the mean (m) of the values divides all the values into two parts: a high percentage in the tail, and a low percentage in the head” (Figure 1c). Thus, edges shorter than the mean length are selected (Figure 1d). We then transform the sets of contiguous edges into polygons, i.e. morphological clusters. We plot the rank-size distribution of the clusters according to their area. On the condition that this distribution is heavily right-skewed, we apply the *head/tail breaks* method proposed by Jiang (2013, 2015, 2019). The values of the distribution are separated into two groups: those above the mean and those below the mean; the partitioning for the values above the mean then continues iteratively until the head part of the distribution is no longer heavy-tailed, i.e. more than 40% of the clusters have an area higher than the mean (Jiang, 2019). By the end of this process, the largest clusters across scales (i.e. the *Natural Cities*) are identified.

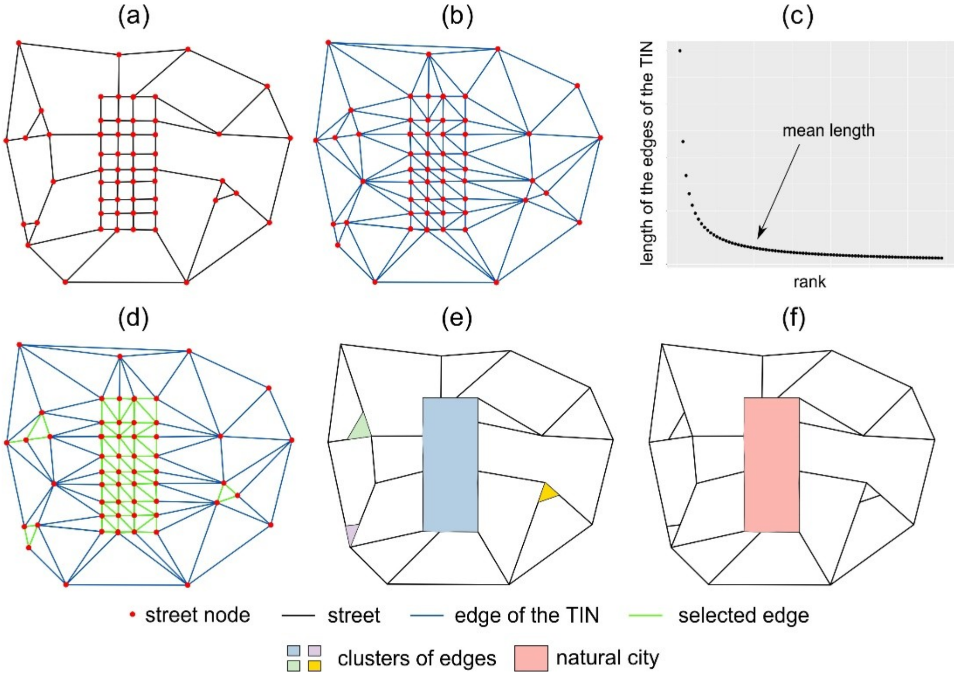

Figure 1. Methodological process adopted to delineate *Natural Cities*: (a) identification of the street nodes; (b) creation of a *Triangular Irregular Network* (TIN); (c) rank-size distribution of the edges of the TIN according to their length; (d) selection of the edges shorter than the mean; (e) transformation of the selected edges into polygons; (f) selection of the largest clusters (i.e. the *Natural Cities*).

In order to select the shortest edges of the TIN, we apply the non-recursive head-tail division rule and not the recursive *head/tail breaks* method because the rank-size distribution of the edges of the TIN is short-tailed and not long-tailed for the two types of data used and the two study areas considered in Section 3 below. This distinction between short-tailed and long-tailed distributions is pointed out in Jiang (2019): "*if the notion of far more smalls than larges recurs on at least twice, the data are long-tailed (...); otherwise, they are short-tailed.*" In other words, if a distribution is no longer heavy-tailed after the first *head/tail break*, this means that it is short-tailed. Thus, we use the mean to select the shortest edges of the TIN. For right-skewed data distributions, the median is substantially smaller than the mean. Incidentally, using the median instead of the mean would lead to the creation of smaller *Natural Cities*. In Figure 2, it can be seen that the delineation of the *Natural Cities* is more relevant with the median than with the mean. Nevertheless, this observation should be confirmed in the case of real-world urban patterns (see Section 3.5).

The delineation of *Natural Cities* also crucially depends on the geographical extent of the study area. Expanding the study area may lead to the introduction of supplementary "rural" street nodes being more distant from each other than "urban" street nodes: this mechanically increases the mean length of the edges of the TIN (Figure 3b). Conversely, the introduction of new "urban" street nodes shortens the mean length of the edges of the TIN (Figure 3c).

Finally, concerning the selection of the largest clusters across scales (i.e. the *Natural Cities*), it is worth noticing that Wang *et al.* (2015) and Liu *et al.* (2019) propose to use the maximal entropy instead of the *head/tail breaks* to identify the optimal size of clusters. This alternative is not considered in this paper but could be worth testing further.

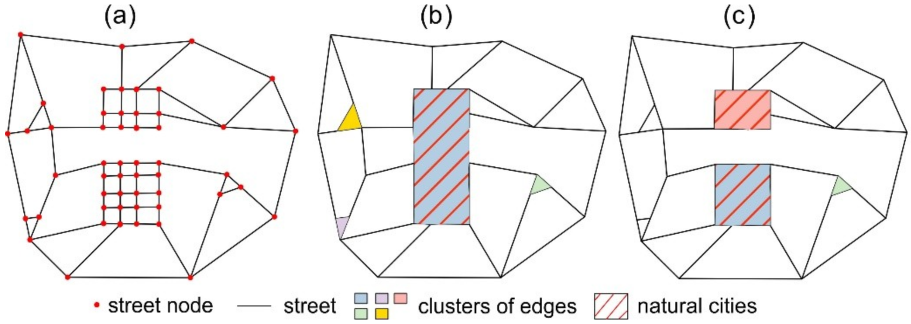

Figure 2. A road network (a) and two variants of the *Natural Cities* method: delineation based on the mean length of the edges of the TIN (b); delineation based on the median length of the edges of the TIN (c).

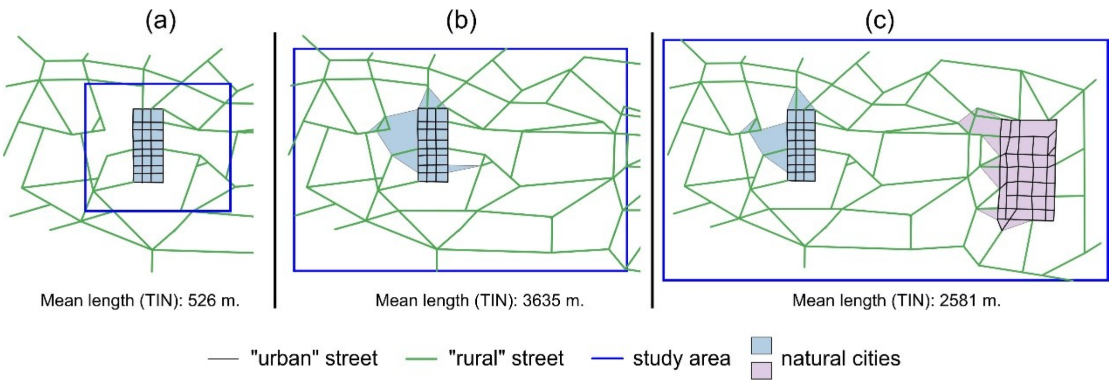

Figure 3. Effect of the geographical extent of the study area on the delineation of *Natural Cities*.

2.2 *MorphoLim*

Data used to delineate morphological agglomerations with *MorphoLim* (https://sourcesup.renater.fr/www/morpholim/) are building footprints mapped in 2D. A step-by-step dilation is applied to each building on the map; the number of built clusters is counted after each dilation step and the results are presented on a log–log plot, where the *x*-axis represents the width of the dilation buffer and the *y*-axis the corresponding number of built clusters (Figure 4a). If a pattern is fractal, then the number of clusters is related to their size by a power-law function corresponding to a linear relationship on a log–log plot and the slope of the curve is the fractal dimension (Tannier *et al.*, 2011). If the pattern is not purely fractal or even multifractal, the dilation curve is not linear and the point at which the dilation curve deviates most from a straight line is a crucial distance threshold. On the curvature function of the dilation curve, which measures how far the curve deviates from a straight line at each point (Lowe, 1989), it is possible to identify the point characterised by the maximum curvature (Figure 4b). The maximum curvature of the dilation curve reveals a major spatial discontinuity across scales. The corresponding distance threshold separates two morphological spatial subsets that are distinct in fractal terms: below that threshold, built elements are organised according to the same spatial logic and belong to the same morphological subset. Mapping the urban boundaries then consists in applying to the building map a buffer with a diameter equal to the distance threshold. On this map, the largest built cluster(s) is (are) identified by visual analysis of the rank-size distribution of all built clusters. The largest built cluster(s) is (are) considered to form the urban morphological agglomeration (Tannier & Thomas, 2013).

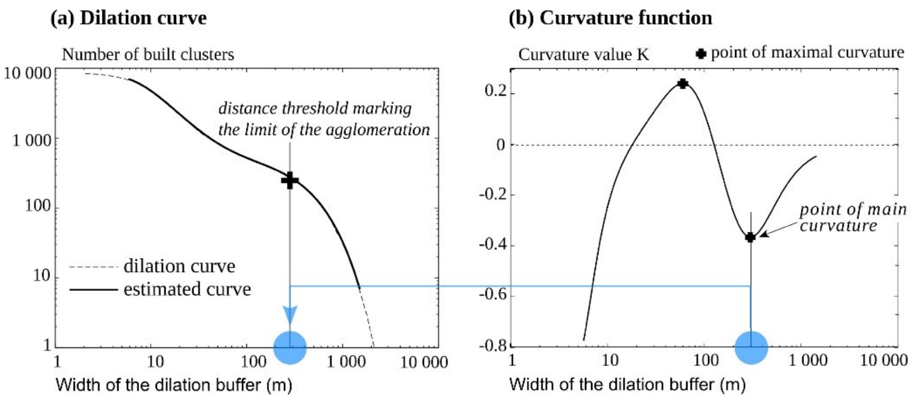

Figure 4. *MorphoLim*: identification of the dilation threshold marking the limit of the morphological agglomeration (blue disks).

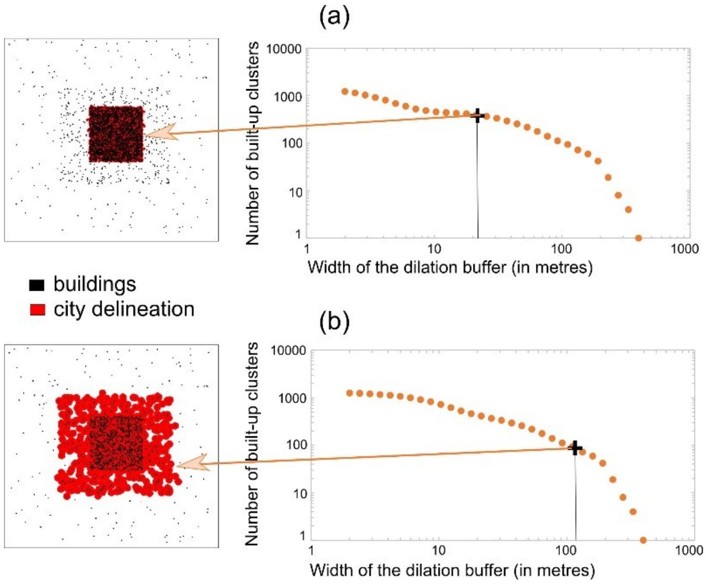

Figure 5. Analysis of two theoretical built patterns using *MorphoLim*. The urban core of pattern (a) is denser than the core of pattern (b).

Yet the separation of two morphological spatial subsets that are distinct in fractal terms, which is indicated by the maximum curvature of the dilation curve, does not always correspond to the separation between the city and its periphery, as Figure 5 shows. Figures 5a and 5b represent two theoretical built patterns with a densely built-up urban core surrounded by a less dense suburban area that is itself surrounded by an even less dense rural area. The two Figures are really similar except that the urban core in Figure 5a is denser than that in Figure 5b. As a consequence, the maximum curvature in Figure 5a indicates the boundary of the urban core, while in Figure 5b it indicates the boundary of the suburban area. In Figure 5a, the built shape of the suburban area is actually more similar to the built shape of the rural area than to the shape of the urban core. In other words, the strongest morphological difference concerns the urban core with respect to its suburban and rural periphery. Thus, in this case, the delineation of the city excludes the suburban area. In contrast, in Figure 5b, the built shape of the suburban area is closer to that of the urban core. Consequently, the delineation of the city includes the suburban area.

The delineation of a morphological agglomeration using *MorphoLim* can depend on the geographical extent of the study area. Figure 6 illustrates this phenomenon. If we look at Figures 6a and 6b, the delineation of the morphological agglomeration does not change with the expansion of the study area: the morphological agglomeration comprises the dense urban core and the suburban area surrounding it. This is because the expansion of the study area in Figure 6b with respect to Figure 6a consists in the addition of supplementary buildings in the rural area. Since the spatial distribution of those supplementary buildings conforms with the spatial distribution of buildings located in the rural area of Figure 6a, the morphological difference between the morphological agglomeration and its "rural" periphery is the same in both cases and their delineation is identical. In contrast, the expanded study area in Figure 6c, compared with Figure 6b, contains a second dense urban core. Thus, the strongest morphological difference is now between, on the one hand, the dense urban cores and, on the other hand, their suburban and rural peripheries. As a result, the suburban area, which surrounds the urban core located on the left of Figure 6c, is no longer included in the morphological agglomeration.

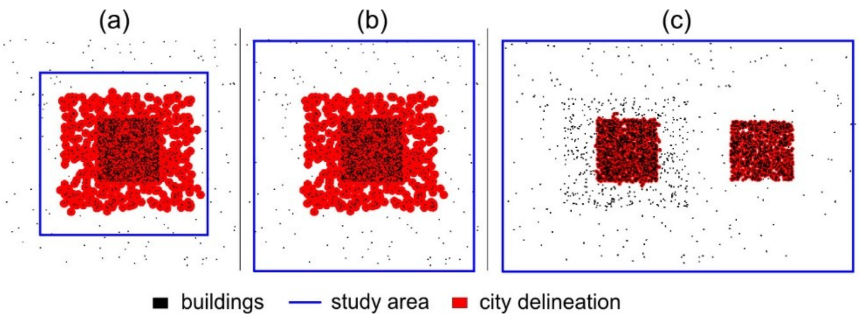

Figure 6. Effect of the geographical extent of the study area on the identification of city boundaries using *MorphoLim*.

2.3 Hierarchical Percolation

Hierarchical Percolation applies to street intersections (Arcaute *et al.*, 2016; Masucci *et al.*, 2015) (Figure 7). A street network is defined as a planar graph where the street intersections are the vertices and the street segments are the edges. Thus, intersection points are connected only if there is a street connecting them. Two street intersections belong to the same morphological cluster if the distance between them is below a given distance threshold. Progressively increasing the distance threshold via a dilation process enlarges the size of the clusters, until eventually a giant component appears, spanning the entire street network. The size of the largest cluster $N_{max}(\tau)$, measured in terms of the number of street intersections, increases with respect to the dilation distance (τ) according to a logistic growth function, where C is the carrying capacity, r the growth rate and τ_0 the inflection point (equation 1). The R code written to calculate the carrying capacity is provided (see section entitled “Data and codes availability statement”).

$$N_{max}(\tau)=\frac{C}{1+e^{-r(\tau-\tau_0)}} \qquad (1)$$

The city condensation threshold is defined as the threshold where the maximum cluster size $N_{max}(\tau)$ intersects the carrying capacity of the logistic function. The largest cluster identified at this condensation threshold forms the city.

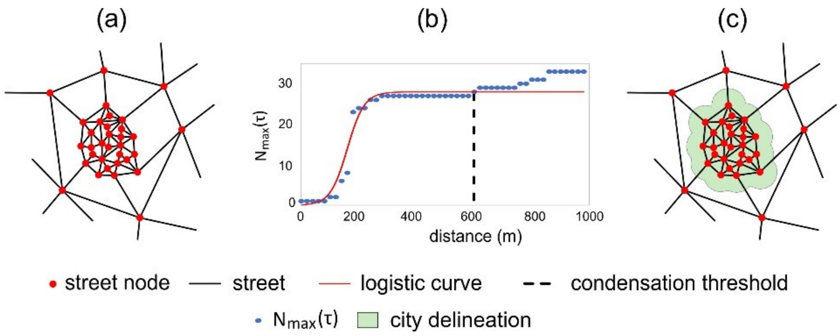

Figure 7. *Hierarchical Percolation* method applied to a theoretical street network. (a) identification of street intersections; (b) fitted logistic function (*x*-axis: size of the dilation buffer; *y*-axis: number of street intersections in the biggest cluster); (c) city delineation.

The geographical extent of the study area is critical when delineating cities by *Hierarchical Percolation*, as the expansion of the study area may involve the introduction of supplementary cities. In the example shown in Figure 8, we see that a first condensation threshold is reached at a 500 m dilation distance. A second condensation threshold then appears at about 1000 m. Hence, as mentioned by Masucci *et al.* (2015), identification of the largest clusters (i.e. the cities) within a pattern containing several cities cannot be completely automated and requires visual inspection.

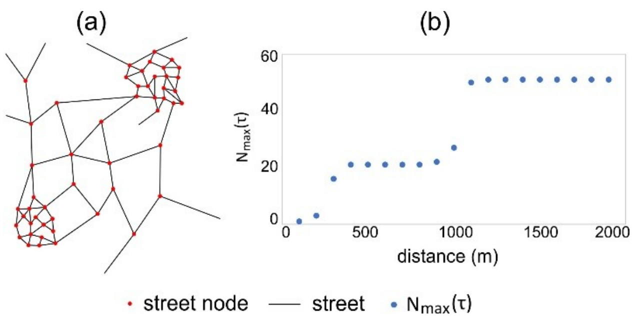

Figure 8. A polycentric street pattern (two centres) (a), and a double condensation threshold in the logistic growth function of the size of the largest cluster (b).

2.4 Comparison of the methods

All three methods involve the use of fully disaggregated data and follow a bottom-up approach. Thus, they avoid all biases associated with the MAUP. Nevertheless, their results are influenced by the existence of natural barriers, especially wide rivers or forests, or major transport routes. In the case of *Natural Cities*, the distances between street nodes separated by any such barrier can be above the mean (Figure 9a). In the case of *MorphoLim*, a barrier can separate a morphological agglomeration into two clusters. In Figure 9b, the two clusters have almost the same size. Yet if one of these clusters were smaller, it would not appear at the top of the rank-size distribution of built clusters and so would not be identified as a part of the morphological agglomeration. In the case of *Hierarchical Percolation*, the barriers are reflected as jumps in the logistic function (Figure 9c). To overcome this limitation, Masucci *et al.* (2015) have artificially added street intersections in the middle of the bridges on the River Thames in London (UK). Interestingly, city delineation obtained using the three methods can take into account ribbons of buildings or street nodes linking two cities (Figure 10).

All three methods are also sensitive to the geographical extent of the study area. Yet the underlying mechanism is different. For *MorphoLim* and *Hierarchical Percolation*, the extent of the study area does not change the dilation threshold or the percolation threshold unless a new built structure appears with the expansion of the study area. In contrast, city delineations obtained using the *Natural Cities* method vary with respect to the expansion of the study area even if no new built structure is included.

In the case of *Natural Cities*, the distances separating street nodes are measured on a Triangular Irregular Network (TIN). In the case of *Hierarchical Percolation* and *MorphoLim*, distances separating buildings are Euclidean distances. City delineation with *Natural Cities* results from direct analysis of the distribution of distances that separate street nodes. With *MorphoLim* and *Hierarchical Percolation*, the distribution of distances is not directly analysed: city delineation appears with a jump in the size of the largest cluster of street intersections in the course of a dilation process (*Hierarchical Percolation*) and a maximum deviation from scale invariance of the number of built clusters in the course of dilation (*MorphoLim*). Although the criteria (measurement and analysis of distances) are different between the three methods, they all look for heterogeneity, i.e. spatial differentiation. Consequently, if buildings or street intersections are uniformly distributed in space, i.e. if distances separating buildings do not vary, no morphological agglomeration can be identified, whatever the method chosen. Besides, with *Hierarchical Percolation* and *MorphoLim*, no morphological agglomeration can be detected if local variations in distances separating the elements (buildings or street intersections) are always the same across scales.

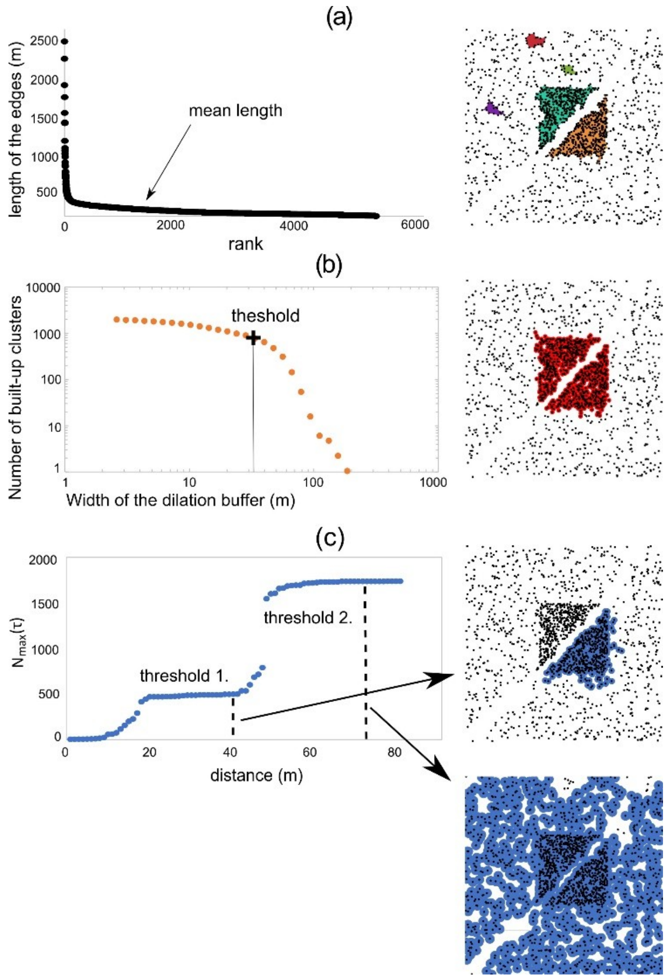

Figure 9. Morphological agglomeration within a theoretical urban pattern crossed by a large non built area: (a) *Natural Cities*; (b) *MorphoLim*; (c) *Hierarchical Percolation.*

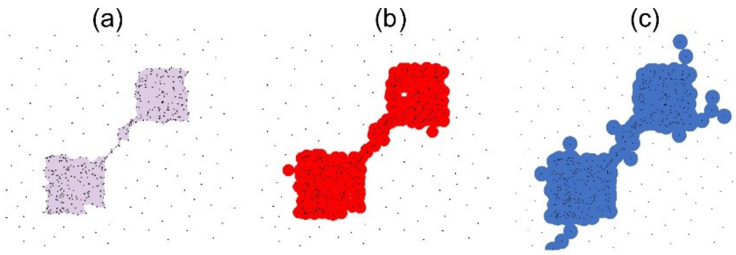

Figure 10. Morphological agglomeration of a theoretical urban pattern where a built ribbon links two urban centres by means of (a) *Natural Cities*, (b) *MorphoLim*, and (c) *Hierarchical Percolation*

Finally, each method involves some arbitrariness. As for *Natural Cities*, the *head/tail breaks* are defined here as more than 40% of the clusters with an area greater than the mean, but another percentage (20% or 30%) could be applied. As for *Hierarchical Percolation*, the arbitrariness refers to the visual inspection of the curve representing the growth of the size of the largest cluster in the course of dilation to identify the condensation thresholds. As for *MorphoLim*, the arbitrariness refers to the choice of the appropriate polynomial to fit the empirical dilation curve before calculating the curvature function.

3. Real case study: delineation of the morphological agglomeration of Brussels (capital city of Belgium)

The three methods under consideration have been applied to delineate the morphological agglomeration of the Brussels metropolitan area. While Brussels is undoubtedly the name of a city, its spatial delineation is not clear. From an administrative and political standpoint, Brussels is referred to as Brussels Capital Region (BCR) (Figure 11a). Yet the city extends over a much larger area across the borders of BCR into the Flemish and Walloon regions. The boundaries of this large metropolitan area are not officially or scientifically fixed: indeed, there is no governance at the scale of the metropolitan area and scientists do not concur about its limits (see, for example, Dujardin *et al.*, 2007; Boussauw *et al.*, 2012; Thomas *et al.*, 2012; Vandermotten, 2016; Van Hecke *et al.*, 2007). Hence the delineation of the morphological agglomeration of Brussels raises not only interesting political issues but also the question of statistical homogeneity throughout the three administrative regions concerned.

3.1 Data and study areas

The *Hierarchical Percolation* and *MorphoLim methods* necessitate adopting a large study area comprising at least one urban agglomeration and its hinterland (i.e. suburban or rural areas under the influence of the urban core). The *Natural Cities* method works better with big data, hence a large study area, as the statistical distribution of the data is more likely to be heavy-tailed (Jiang 2015). Additionally, we have shown that all three methods are sensitive to the geographical extent of the study area. For those reasons, we delineate the morphological agglomeration of Brussels within two different study areas: (1) the former province of Brabant (Adam *et al.*, 2017) and (2) the entire country (whole of Belgium). The former province of Brabant (3357 km²) was created in 1815 and is today divided into three entities: BCR (1.2 million inhabitants), the province of Flemish

Brabant in the north (1.1 million inhabitants) and the province of Walloon Brabant in the south (0.4 million inhabitants). It includes about ten regional cities (Figure 11b). The entire country of Belgium (30689 km²) has 11.46 million inhabitants (2019).

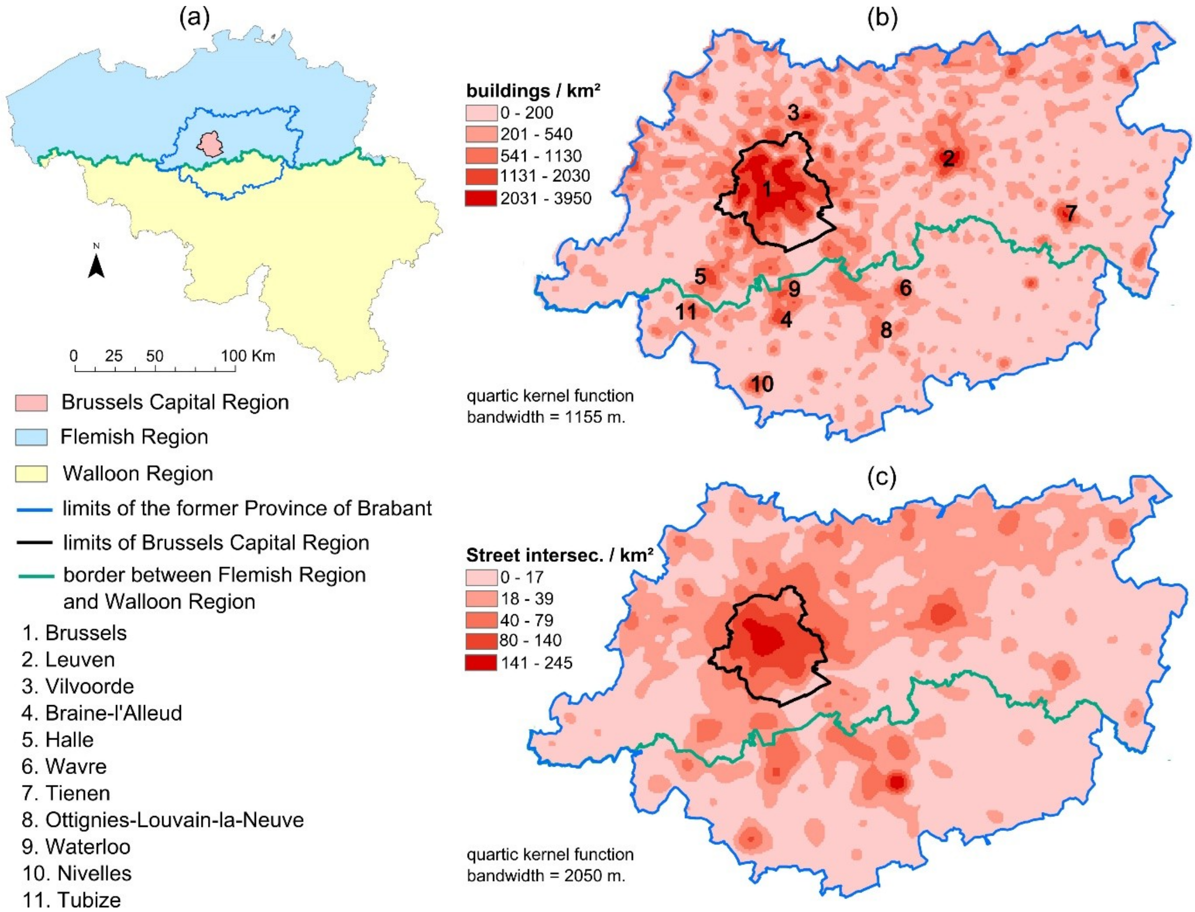

Figure 11. Belgium: location of the former province of Brabant around the Brussels Capital Region (a). Smoothed densities of building centroids (b) and road intersections (c) in the province of Brabant. Note: the bandwidths correspond to the optimal values proposed by default in the ArcGis software application.

Data about buildings come from the Land Registry Administration of Belgium (© 2009 *Administration Générale de la Documentation Patrimoniale*). In this registry, all buildings with a footprint of more than 4 m^2 are mapped in 2D (polygons), regardless of their function (e.g. residential, commercial or industrial). For our study, to avoid noise due to small buildings such as garages or garden sheds, all entities under 12 m², which represent less than 5% of all buildings in the whole of Belgium, were removed from the database. Road network data come from the Geofabrik platform of OpenStreetMap (http://download.geofabrik.de accessed 08/21/2018). OSM data have been chosen because of their large availability across the world. All cycle paths and car parks were removed from the database and all streets with the same name were merged to avoid their over-representation. Within cities, it can happen that street segments are duplicated two or three times in OSM data. Yet a visual inspection of the data has shown that this is rarely the case in Belgium. Street segments have been transformed into street nodes according to the tutorial by Ren (2018). All nodes of degree two were also removed.

3.2 Morphological agglomerations of Brussels obtained using the *Natural Cities* method

The *Natural Cities* method is successively applied to street nodes, which include street junctions and end points, and to building centroids (see Appendix 1). The basic condition for applying the *Natural Cities* method is that the rank-size distribution of the edges of the TIN according to their length is heavily right skewed. This condition is fulfilled both for the province of Brabant and for Belgium.

3.2.1 Deriving *Natural Cities* from street nodes

When considering both the province of Brabant and the whole of Belgium, a single *Natural City* forms the morphological agglomeration of Brussels (Figure 12). This *Natural City* is centred on Brussels Capital Region. Its size is close to that of BCR but its shape differs. The *Natural City* is slightly larger for the whole of Belgium than for the province of Brabant (148 km² versus 111 km²; see Table 1).

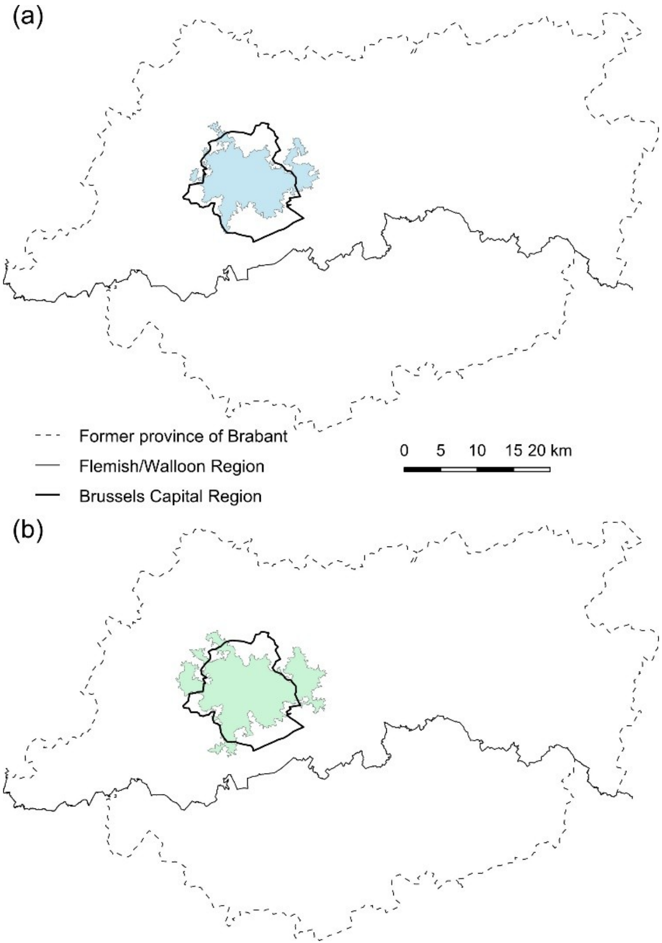

Figure 12. Morphological agglomerations of Brussels obtained using *Natural Cities* computed on street nodes (a) for the province of Brabant and (b) for the whole of Belgium.

3.2.2 Deriving *Natural Cities* from building centroids

When considering the province of Brabant, four small *Natural Cities* make up the morphological agglomeration of Brussels. Three are located within Brussels Capital Region and the fourth covers the regional cities of Braine-l'Alleud and Waterloo (see Figure 11b and Figure 13a). When considering the whole of Belgium, only one *Natural City* located within Brussels Capital Region makes up the morphological agglomeration of Brussels (Figure 13b).

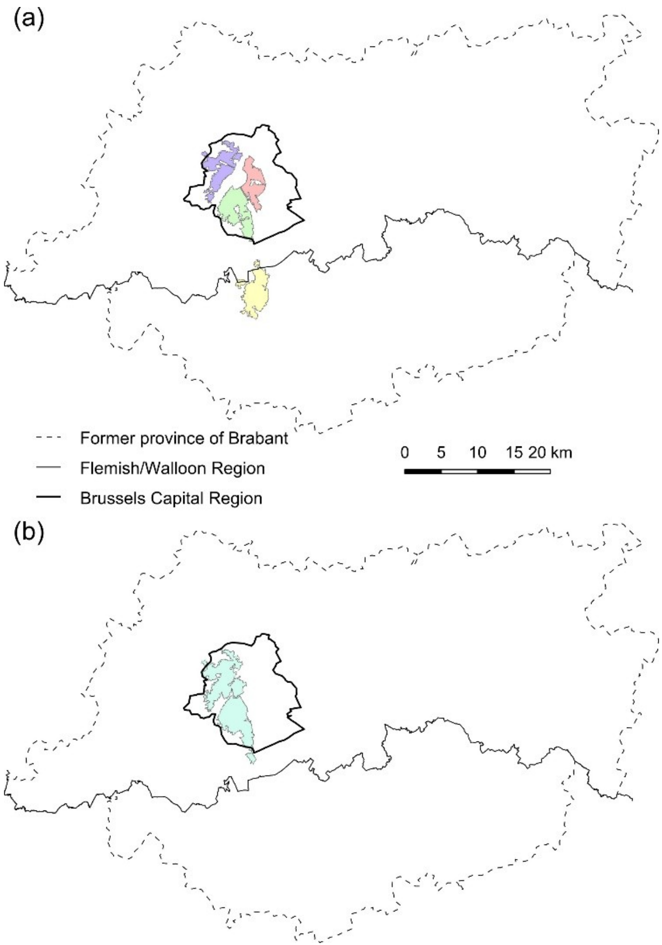

Figure 13. Morphological agglomeration of Brussels obtained using *Natural Cities* on the basis of building centroids: (a) for the province of Brabant and (b) for the whole of Belgium.

3.3 Morphological agglomerations of Brussels obtained using *MorphoLim*

Data used here are buildings mapped in 2D. The analysis was first performed on the province of Brabant. Figure 14a shows that the dilation curve is quite close to a straight line. The maximum curvature value is low (-0.087). The spatial distribution of buildings in the cluster at the top of the rank-size distribution (Figure 14b) therefore differs little from the spatial distribution of buildings in the rest of the province of Brabant. This largest cluster covers a major part of the study area (Figure 15a) and incorporates all regional cities in Figure 11b apart from Tienen and Nivelles. Indeed, Nivelles is quite confined within itself and Tienen is connected by a few built ribbons to some nearby villages but not to the other regional cities.

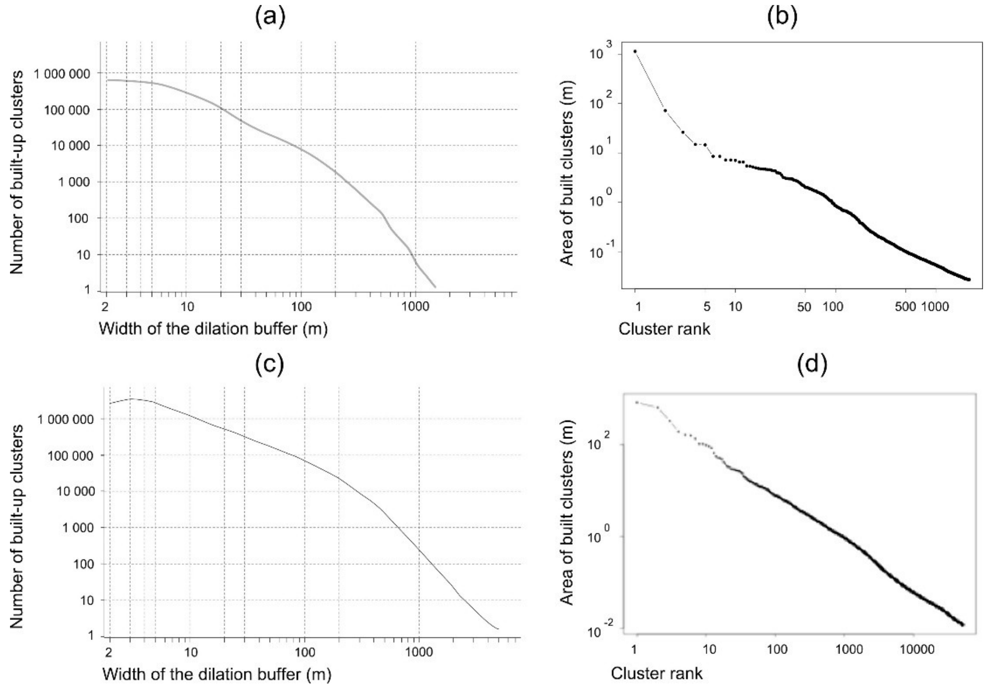

Figure 14. Dilation curve and rank-size distribution of built clusters obtained using *MorphoLim* for the province of Brabant (a & b) and for the whole of Belgium (c & d).

The results are markedly different when the study area is the whole of Belgium: the dilation curve exhibits a stronger main curvature (equal to -0.11) (Figure 14c); the cluster at the top of the rank-size distribution is not so large (Figure 14d); the morphological agglomeration spreads towards the south (Walloon Brabant) and towards the north-west (Flemish Brabant) beyond the border of the province where it stops at the edge of the city of Ghent (Figure 15b).

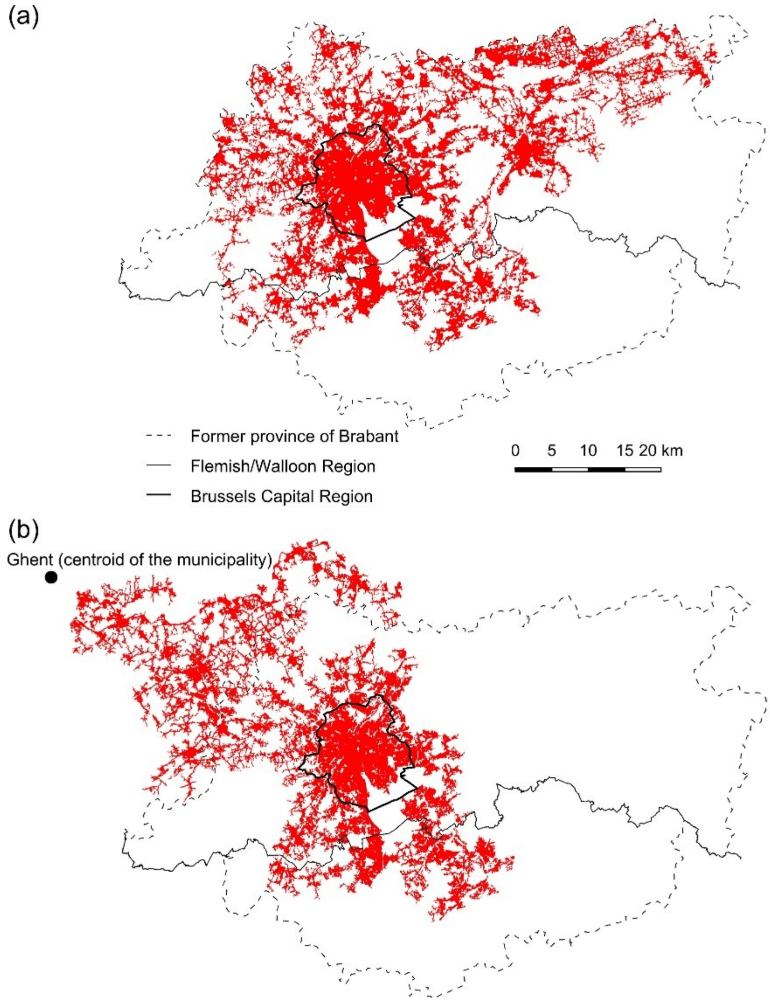

Figure 15: Morphological agglomerations of Brussels obtained using *MorphoLim*; (a) for the province of Brabant; (b) for the whole of Belgium.

3.4 Morphological agglomerations of Brussels obtained using *Hierarchical Percolation*

Hierarchical Percolation applies to road intersections (without the street end points unlike the *Natural Cities* method). Adaptation of the method enables us to apply it to building centroids too (Behnisch *et al.*, 2019).

3.4.1 *Hierarchical Percolation* applied to road intersections

In the case of the province of Brabant, the first condensation threshold appears at 140 m. The morphological agglomeration of Brussels extends to the close periphery of Brussels Capital Region and a little to the south to Braine-l'Alleud (Figure 16a). With the second threshold (170 m), the agglomeration extends beyond the administrative limits of BCR in all directions and encompasses Braine-l'Alleud in the south. The results obtained for Belgium are very similar (Figure 16b): a first percolation threshold appears at 140 m and the morphological agglomeration of Brussels extends to the close periphery of BCR and to Braine-l’Alleud.

Turning now to the largest condensation thresholds, the morphological agglomeration of Brussels is no longer identifiable. In the case of the province of Brabant (threshold: 220 m), the resulting morphological agglomeration covers almost the entire study area. In the case of Belgium (threshold: 440 m), the resulting morphological agglomeration covers almost the entire country.

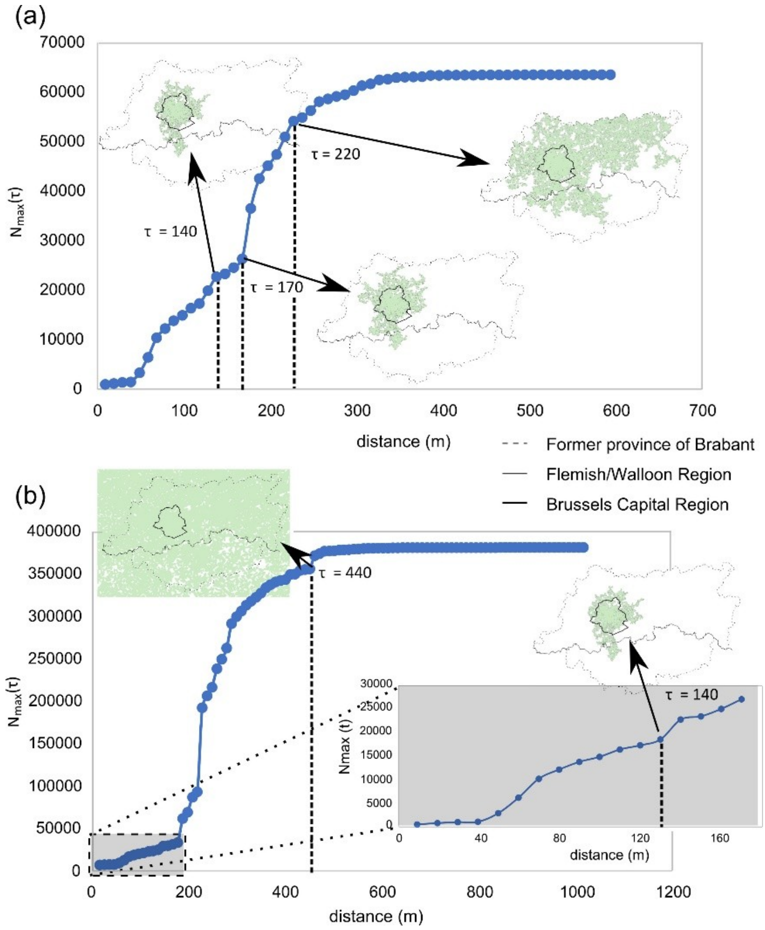

Figure 16. Logistic functions and morphological agglomerations of Brussels obtained using *Hierarchical Percolation* on the basis of road intersections; (a) for the province of Brabant; (b) for the whole of Belgium. *x*-axis: width of the dilation buffer; *y*-axis: number of road intersections in the largest cluster; τ: condensation threshold.

3.4.2 *Hierarchical Percolation* applied to building centroids

Similarly to the above-mentioned results, several percolation thresholds emerge at different dilation steps. For the province of Brabant, three thresholds are identified (Figure 17a). If we consider the first one (30 m), the agglomeration of Brussels is almost entirely limited to BCR and only extends beyond it to the south where it encompasses Braine-l'Alleud. Considering the second threshold (90 m), the agglomeration of Brussels is much larger and covers almost all of the western half of the study area. With the third threshold (100 m), the two largest cities in the study area (Brussels and Leuven) merge and the morphological agglomeration covers all regional cities apart from Tienen and Nivelles.

With the extension of the study area to the whole of Belgium, the first percolation threshold appears at 60 m (Figure 17b). At this stage, Brussels covers the western half of the study area and encompasses all regional cities apart from Leuven, Tienen and Nivelles. With the second threshold (70 m), the agglomeration of Brussels encompasses the city of Ghent but still not Leuven. Finally, with the next threshold (160 m), all of Flanders is attached to the agglomeration of Brussels because of the existence of built ribbons that link the Flemish cities.

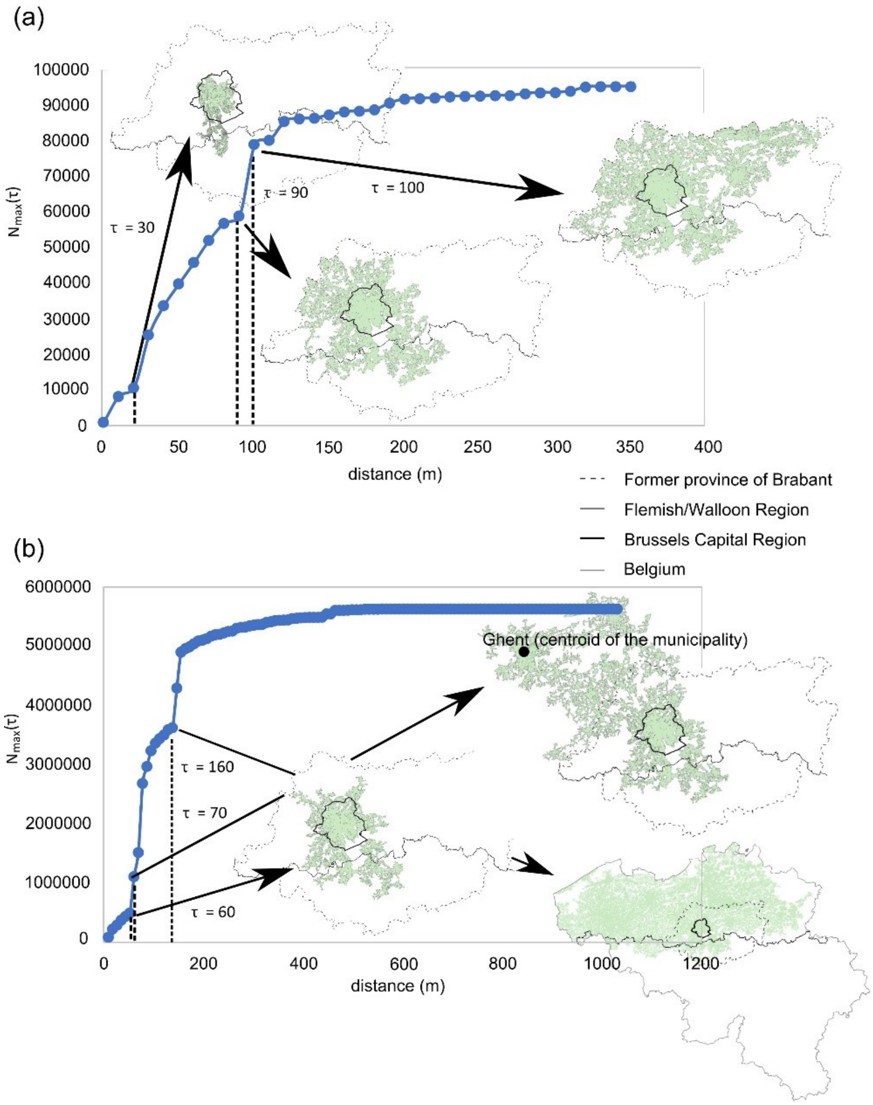

Figure 17. Logistic functions and morphological agglomerations of Brussels obtained using *Hierarchical Percolation* on the basis of building centroids; (a) for the province of Brabant; (b) for Belgium. *x*-axis: width

of the dilation buffer; *y*-axis: number of building centroids in the largest cluster; τ: condensation threshold.

Table 1. Summary of results obtained using the three delineation methods. (*) Mean length of the TIN; (+) distance at which the dilation curve deviates most from a straight line; (#) percolation threshold.

Method	**Data**	**Study area**	**Threshold** (in m)	**Area of Brussels' agglomeration** (in km^2)	**Perimeter of Brussels' agglomeration** (in km^2)
Natural Cities	Road intersections	Province of Brabant	212 (*)	110	169
		Whole of Belgium	250 (*)	148	204
	Building centroids	Province of Brabant	48 (*)	67	329
		Whole of Belgium	54 (*)	46	187
MorphoLim	Building footprints	Province of Brabant	181 (+)	653	5299
		Whole of Belgium	121 (+)	502	5186
Hierarchical percolation	Road intersections	Province of Brabant	140 (#)	271	1038
			170 (#)	395	1262
			220 (#)	1395	4165
		Whole of Belgium	140 (#)	269	1030
			440 (#)	21580	20876
	Building centroids	Province of Brabant	30 (#)	128	1261
			90 (#)	710	3053
			100 (#)	1176	5093
		Whole of Belgium	60 (#)	395	2073
			70 (#)	1158	6818
			160 (#)	8779	25463

3.5 Discussion

By using three methods for delineating the morphological agglomeration of Brussels, we obtain a series of morphological limits that differ markedly. With *Hierarchical Percolation*, *MorphoLim* and *Natural Cities* applied to street nodes, the morphological agglomeration is made up of a single morphological cluster. In contrast, the application of the *Natural Cities* method to building centroids results in the identification of several small *Natural Cities*. In general, all *Natural Cities* are quite small. It would therefore be irrelevant to use the median instead of the mean, as considered in Section 2.1.

With *MorphoLim*, the morphological agglomeration of Brussels is smaller if the study area is Belgium (Figure 15b) than if it is the province of Brabant (Figure 15a); see Table 1. This is because a crucial change in distances separating buildings occurs at a lower distance threshold when considering all of Belgium. The opposite is found for the *Hierarchical Percolation* method applied to building centroids when the first percolation threshold is considered (see Figures 17a and b, and Table 1). Yet in the case of street intersections, the percolation thresholds and the area of the morphological agglomerations are the same whichever study area is under consideration (province of Brabant or whole of Belgium) when the first percolation threshold is considered (see Figures 16a and b, and Table 1). For *Natural Cities*, the effect of the expansion of the study area also depends on the data used. With the street intersections, we observe an increase in the mean length of the edges of the TIN and a decrease in the minimal area of the clusters. With the building centroids, we again observe an increase in the mean length of the edges of the TIN, but this time an increase in the minimal area of the clusters (see Figures 12 and 13, and Table 1).

Among all delineations of the agglomeration of Brussels, some are undoubtedly irrelevant. First, *Natural Cities* derived from building centroids are much too small (Figure 13). They delineate large neighbourhoods within BCR as well as the morphological agglomeration of two regional cities (Braine-l'Alleud and Waterloo), but not one inclusive morphological agglomeration. Second, the boundaries of the morphological agglomeration are sometimes artificially truncated by the border of the province of Brabant when this study area is chosen. This is the case with the delineation obtained using *MorphoLim* (Figure 15a) as well as the application of the largest percolation thresholds of *Hierarchical Percolation* (Figure 16a - threshold 220 m, Figure 17a - threshold 100 m). In those cases, the province of Brabant is too small with respect to the chosen method or parameters and a larger study area (e.g. the whole of Belgium) would be required. Third, when the study area is the whole of Belgium, the largest percolation thresholds do not enable us to delineate the morphological agglomeration of Brussels because the morphological cluster being identified covers either all of Flanders (for building centroids) or almost all of Belgium (for street intersections).

Finally, nine delineations of the morphological agglomeration of Brussels appear to be geographically relevant: the *Natural Cities* identified on the basis of street nodes for the two study areas (province of Brabant and Belgium), the delineation obtained using *MorphoLim* applied to the whole of Belgium, the delineations obtained using the *Hierarchical Percolation* method applied to road intersections (province of Brabant - thresholds 140 m and 170 m, and Belgium - threshold

140 m) and to the centroids of buildings (province of Brabant - threshold 30 m, and Belgium - thresholds 60 m and 70 m). In six cases out of nine, the limit of the morphological agglomeration of Brussels remains in the close periphery of BCR (*Natural Cities*, Figure 12a and b), or spreads a little into the Walloon region to Braine-l'Alleud (*Hierarchical Percolation*, Figure 16a - threshold 140 m and 170m, Figure 16b - threshold 140 m, Figure 17a - threshold 30 m). In another case, the morphological agglomeration of Brussels is larger and extends clearly over the borders of BCR into the Flemish and Walloon regions (Figure 17b - threshold 60 m). It encompasses seven regional centres: Vilvoorde, Halle, Tubize, Waterloo, Braine-l'Alleud, Wavre and Ottignies-Louvain-la-Neuve. In the last two cases, the agglomeration of Brussels is even much larger. It spills over the border of the province of Brabant and reaches the edge of the city of Ghent in the north-west (*MorphoLim*, Figure 15b) or even encompasses the city of Ghent (*Hierarchical Percolation*, Figure 17b - threshold 70 m).

Finally, two opposing views of the settlement pattern of Belgium are supported by the results obtained. On the one hand, the limit of the morphological agglomeration of Brussels remains in the close periphery of BCR or spreads a little into the Walloon region to Braine-l'Alleud (six relevant delineations of Brussels out of nine). This confirms the idea that Belgian cities are still separate entities (Van Criekingen *et al.*, 2007). In contrast, in three delineations out of nine, the morphological agglomeration of Brussels extends clearly over the borders of BCR into the Flemish and Walloon regions, or spills over the border of the province of Brabant and reaches the city of Ghent in the north-west. Moreover, the morphological agglomeration even covers either all of Flanders or almost all of Belgium, with the largest percolation thresholds. This supports the idea that Belgium is one large conurbation (Vandermotten *et al.*, 2008) in opposition to Van Criekingen *et al.* (2007). In actual fact, Belgian cities are separate entities but they have strong morphological ties that take the form of built ribbons, especially in Flanders.

4 Conclusion

Conventional methods that integrate a fixed threshold set ex ante have the advantage of being easy to apply and their results are easy to handle. Yet such methods overlook the specific features of the patterns under consideration that appear across scales and that are revealed by the three methods compared in this paper. Each of these methods has a specific way of detecting the urban boundary — the nested mean value in a *head/tail breaks* distribution for *Natural Cities*, a maximum deviation from a hyperbolic distribution for *MorphoLim* and a condensation threshold in a logistic growth function for *Hierarchical Percolation*. Cities with the same built density or the same density of street nodes may display different dilation thresholds (*MorphoLim*) or percolation thresholds (*Hierarchical Percolation*), or even mean edge lengths (*Natural Cities*). All three methods involve the use of fully disaggregated data and follow a bottom-up approach. Consequently they avoid all biases of the MAUP. Nevertheless, they are all sensitive to the geographical extent of the study area.

The three methods were applied to two study areas (the province of Brabant and the whole of Belgium) with two types of data (buildings and street intersections). The resulting delineations of the morphological agglomeration of Brussels differ markedly. Accordingly, it is impossible to draw a single unambiguous morphological boundary for an urban agglomeration in general, and for Brussels in particular. This is consistent with the observations by Chen (2016): "Urban form has no characteristic scale, and thus an urban boundary cannot be identified exactly" and by Jiang (2018): "(...) things are not measurable, or measurement depends on the measuring scales because of the fractal nature of geographic phenomena (e.g. Goodchild and Mark 1987, Batty and Longley 1994, Chen 2011)". Indeed, if a city is a boundless fractal, none of the three methods can enable its boundary to be identified. As suggested by Chen *et al.* (2017), "(...) geographical phenomena fall into two groups: one is with characteristic scale (scaleful group), and the other, without characteristic scale (scale-free group). The former can be described with characteristic length such as average value, standard deviation, and eigenvalue, while the latter should be described with scaling exponent such as fractal dimension." Nevertheless, the three methods make it possible to delineate a morphological agglomeration of Brussels because local deviations from scale invariance actually exist (Chen *et al.*, 2017; Sémécurbe *et al.*, 2019). As stated by Jiang and Ren (2018), geography has three fundamental issues regarding geographical space: how it looks, how it works, and what it ought to be. In the present paper, we compare three methods with a view to seeing how they describe the built-up realities on the basis of the example of the morphological agglomeration of Brussels. However, we are not comparing these methods with a view to finding out how a morphological agglomeration should be, nor are we trying to determine whether or not results comply with norms or ideas of goodness or optimality, as done for instance in Chen (2016) and Jiang (2019). With the example of Belgium, we here compare methodologically and empirically three methods for delineating morphological agglomerations; Belgium is sufficiently diverse in terms of urban textures for trusting the present results. Comparing with case studies in other countries/continents opens a new avenue for further research in human and urban geography.

Our study has shown that street nodes and building centroids do not convey the same information about the spatial organization of built patterns. This is very clear in the case of the *Natural Cities* method, as the *Natural Cities* delineated on street nodes are much larger than the *Natural Cities* delineated on building centroids. Distances that separate street intersections are actually above the mean in between individual cities, whereas distances that separate building centroids are above the mean everywhere except in a few very dense intra-urban built clusters. Moreover, the nature of the data used exerts a joint effect with the size of the study area on city delineations. This joint effect varies with the method under consideration. With this in mind, the crucial interest of the *Natural Cities* method applied to street nodes is that it enables clear identification of the morphological limits of individual cities, in this case Brussels but it could be other cities as well. Considering the Belgian case, changing the study area has a minor impact on the resulting delineations. Indeed, expanding the study area expands Brussels' *Natural City* just a little bit. In contrast, the *MorphoLim* and *Hierarchical Percolation* methods applied to the building data for the whole of Belgium enable the identification of an extensive morphological agglomeration of Brussels that includes many regional cities, among which Brussels, linked by built ribbons. However, the application of those two methods to building data considering only the province of Brabant is not relevant because the boundaries of the morphological agglomeration are truncated by the border of the study area. Finally, the *Hierarchical Percolation* method applied to street intersections is interesting, as it enables the identification of three percolation thresholds, and thus three nested morphological agglomerations of growing size that show how the peripheries of Brussels are more or less closely linked to the core agglomeration from a morphological point of view. Conversely, this method does not provide a single clear limit of the morphological agglomeration of Brussels.

To conclude, geographers and planners should be more than ever conscious of the importance of their choices about the data used, the size of the study area and the delineation method. They all have to be carefully selected with respect to both the objectives of each delineation and the specific features of the area under study.

Data and codes availability statement

The data and codes that support the findings of this study are available on *figshare.com*: https://doi.org/10.6084/m9.figshare.11405598

References

Abel, J., Kaysi, I.A., 2003. Identifying urban boundaries: Application of remote sensing and geographic information system technologies. *Canadian Journal of Civil Engineering* 30(6), 992–999. https://doi.org/10.1139/l03-051

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., Batty, M., 2014. Constructing cities, deconstructing scaling laws. *Journal of The Royal Society Interface* 12, 20140745–20140745. https://doi.org/10.1098/rsif.2014.0745

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*. https://doi.org/10.4000/brussels.1592

Batty M., Longley P., 1994. Fractal Cities: A geometry of form and function, Academic Press:London.

Behnisch, M., Schorcht, M., Kriewald, S., Rybski, D., 2019. Settlement percolation: A study of building connectivity and poles of inaccessibility. *Landscape and Urban Planning* 191, 103631.

Blondel, V., Krings, G., Thomas, I., 2010. Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone. *Brussels Studies* 42 (general collection, online: https://journals.openedition.org/brussels/806

Borruso, G., 2003. Network Density and the Delimitation of Urban Areas. *Transactions in GIS* 7, 177–191.

Boussauw, K., Neutens, T., Witlox, F., 2012. Relationship between Spatial Proximity and Travel-to-Work Distance: The Effect of the Compact City. *Regional Studies* 46, 687–706.

Briant, A., Combes, P.-P., Lafoucarde, M., 2010. Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics* 67(3), 287–302.

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.

Cetinkaya, S., Basaraner, M., Burghardt, D., 2015. Proximity-based grouping of buildings in urban blocks: a comparison of four algorithms. *Geocarto International* 30, 618–632.

Chaudhry, O., Mackaness, W.A., 2008. Automatic identification of urban settlement boundaries for multiple representation databases. *Computers, Environment and Urban Systems* 32, 95–109.

Chen, Y., 2016. Defining urban and rural regions by multifractal spectrums of urbanization. *Fractals* 24, 1650004.

Chen, Y., 2011. Derivation of the functional relations between fractal dimension of and shape indices of urban form. Computers, Environment and Urban Systems, 35, 442–451.

Chen Y. , Wang J. , Long Y. , Zhang X. , Liu X. , Li K., 2017. Defining urban boundaries by characteristic scales. arXiv:1710.01869 [physics.soc-ph]

Chaudhry, O., Mackaness, W.A., 2008. Automatic identification of urban settlement boundaries for multiple representation databases. *Computers, Environment and Urban Systems* 32, 95–109.

Cottineau C., Finance O., Hatna E., Arcaute E., Batty M., 2018. Defining urban clusters to detect agglomeration economies, Environment and Planning B: Urban Analytics and City Science, 1-16

Deng, M., Liu, Q., Cheng, T., Shi, Y., 2011. An adaptive spatial clustering algorithm based on delaunay triangulation. *Computers, Environment and Urban Systems* 35, 320–332.

Dujardin, C., Thomas, I., Tulkens, H., 2007. Quelles frontières pour Bruxelles ? Une mise à jour. *Reflets et perspectives de la vie économique* XLVI, 155.

Fang C., Yu D., 2017. Urban agglomeration: an evolving concept of an emerging phenomenon. Landscape and Urban Planning 162, 126-136

Ferreira, J.A., Condessa, B., Castro e Almeida, J., Pinto, P., 2010. Urban settlements delimitation in low-density areas—An application to the municipality of Tomar (Portugal). *Landscape and Urban Planning* 97, 156–167.

Gisbert, F., Marti, I., Gielen, E., 2017. Clustering cities through urban metrics analysis, *Journal of Urban Design*, 22:5, 689-708

Goodchild, M., 2016. Laws in Geography. Conférence au Collège International des Sciences du Territoire, Paris.

Goodchild, M., D. Mark. (1987). The Fractal Nature of Geographic Phenomena. Annals of the Association of American Geographers 77(2), 265–278.

Guérois, M., Bretagnolle, A., Mathian, H., Pavard, A., 2014. Functional Urban Areas (FUA) and European harmonization. A feasibility study from the comparison of two approaches: commuting flows and accessibility isochrones (Technical Report, Espon 2013 Database). Paris: Union Européenne. (p. 35).

Haggett P., 1972. Geography: a modern synthesis. Harper & Row

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

Jiang, B., 2015. Head/tail breaks for visualization of city structure and dynamics. Cities, 43, 69–77.

Jiang, B., 2018. A topological representation for taking cities as a coherent whole. Geographical Analysis 50, 298–313.

Jiang, B., 2019. A recursive definition of goodness of space for bridging the concepts of space and place for sustainability. Sustainability, 11(15), 4091.

Jiang, B., Jia, T., 2011. Zipf's law for all the natural cities in the United States: a geospatial perspective. *International Journal of Geographical Information Science* 25, 1269–1281.

Jiang, B., Liu, X., 2012. Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. International Journal of Geographical Information Science 26, 215–229.

Jiang, B., Miao, Y., 2015. The evolution of natural cities from the perspective of location-based social media. *The Professional Geographer* 67, 295–306.

Jiang, B., Ren, Z., 2018. Geographic space as a living structure for predicting human activities using big data. Int. J. Geogr. Inf. Sci. 50, 298–313.

Jones, J., Peeters, D., Thomas, I., 2017. Scale effect in a LUTI model of Brussels: challenges for policy evaluation. EJTIR 17, 103–131.

Kauffmann, A., 2012. Delineation of City Regions Based on Commuting Interrelations: The Example of Large Cities in Germany (IWH Discussion Paper No. 4). Halle Institute for Economic Research.

Le Gléau, J., Pumain, D., Saint-Julien, T., 1997. Towns of Europe: To Each Country Its Definition. *INSEE Studies* 6, 1–14.

Liu, L., Xia, B., Wu, H., Zhao, J., Peng, Z., Yu, Y., 2019. Delimitating the Natural City with Points of Interests Based on Service Area and Maximum Entropy Method. *Entropy* 21, 458.

Long, Y., 2016. Redefining Chinese city system with emerging new data. *Applied Geography* 75, 36–48.

Lowe, D.G., 1989. Organization of smooth image curves at multiple scales. *International Journal of Computer Vision* 3, 119–130.

Maisonobe, M., Jégou, L., Eckert, D., 2018. Delineating urban agglomerations across the world: a dataset for studying the spatial distribution of academic research at city level. *Cybergeo: European Journal of Geography* 871, online: https://journals.openedition.org/cybergeo/29637.

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, 20150763.

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

Nicolau, R., Cavaco, C., 2018. Automated delimitation of urban areas comprising small-sized towns – Comparison of two methodologies applied to mainland Portugal. *Environment and Planning B: Urban Analytics and City Science* 45, 180–201.

Oliveira, E.A., Andrade, J.S., Makse, H.A., 2015. Large cities are less green. *Scientific Reports* 4, 4235.

Openshaw, S., 1983. The Modifiable Areal Unit Problem. Concepts and Techniques in Modern Geography, Geo Books, Norwich (Headley Brothers Ltd. Kent). 38.

Parr, J.B., 2007. Spatial Definitions of the City: Four Perspectives. *Urban Studies* 44, 381–392.

Ren, Z., 2018. Tutorials for Extracting Street Nodes from OSM Data and Creating Natural Cities from Street Nodes Using ArcGIS Model, A Workshop on Big Data Analytics: Topological and scaling perspective for better understanding and (re-)making sustainable cities, 12 June 2018, Lund, Sweden, in conjunction with AGILE 2018.

Rozenfeld, H.D., Rybski, D., Andrade Jr., D.S., Batty, M., Stanley, H.E., Makse, H.A., 2008. Laws of population growth. PNAS, 105(48): 18702-18707

Rozenfeld, H.D., Rybski, D., Gabaix, X, Makse, H.A., 2011. The area and population of cities: New insights from a different perspective on cities. American Economic Review, 101(5): 2205–2225

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 France, Journal of Geographical Systems, 21(2), pp. 271-293.

Sotomayor-Gómez, B., Samaniego, H. (2020). City limits in the age of smartphones and urban scaling. Computers, Environment and Urban Systems 79, 101423

Tannier, C., Thomas, I., Vuidel, G., Frankhauser, P., 2011. A Fractal Approach to Identifying Urban Boundaries. *Geographical Analysis* 43, 211–227.

Tannier, C., Thomas, I., 2013. Defining and characterizing urban boundaries: A fractal analysis of theoretical cities and Belgian cities. *Computers, Environment and Urban Systems* 41, 234–248.

Thomas, I., Cotteels, C., Jones, J., Peeters, D., 2012. Revisiting the extension of the Brussels urban agglomeration: new methods, new data … new results? *Belgeo* 2–13.

Thomas, I., Jones, J., Caruso, G., Gerber, P., 2018. City delineation in European applications of LUTI models: review and tests. *Transport Reviews* 38, 6–32.

Usui, H., 2019. A bottom-up approach for delineating urban areas minimizing the connection cost of built clusters: Comparison with top-down-based densely inhabited districts. *Computers, Environment and Urban Systems* 77, 101363.

Vandermotten, C., 2016. Adéquations et inadéquations du découpage territorial aux contextes politiques et économiques : le cas de la Belgique. *EchoGéo* 35, online: http://journals.openedition.org/echogeo/14490.

Vandermotten, C., Halbert, L., Roelandts, M., Cornut, P., 2008. European Planning and the Polycentric Consensus: Wishful Thinking? *Regional Studies* 42, 1205–1217.

Van Criekingen, M., Cornut, P., Luyten, S., 2007. Brussels: polycentricity as ‘images on the map’, not in reality, in: Cattan N. (Ed), Cities and Networks in Europe. A Critical Approach of Polycentrism, Paris, John Libbey Eurotext. pp. 105–112.

Van Hecke, E., Halleux, J.-M., Decroly, J.-M., Merenne-Schoumaker, B., 2009. Noyaux d’habitat et Régions urbaines dans une Belgique urbanisée. Monographie 9 : L’Urbanisation. SPF Economie, P.M.E., Classes moyennes et Energie, Bruxelles.

Wang, S., Wang, D., Li, C., Li, Y., 2015. Comment on “Clustering by fast search and find of density peaks.” arXiv:1501.04267 [cs].

Weber, C. 2001. Urban Agglomeration Delimitation Using Remote Sensing. In: Remote Sensing and Urban Analysis, 145–59, edited by Donnay, J., Barnsley, M. and Longley, P., London: Taylor & Francis.

Wu, B., Yu, B., Wu, Q., Chen, Z., Yao, S., Huang, Y., Wu, J., 2018. An Extended Minimum Spanning Tree method for characterizing local urban patterns. *International Journal of Geographical Information Science* 32, 450–475.

Zhou, Q., 2015. Comparative Study of Approaches to Delineating Built-Up Areas Using Road Network Data: Delineating Built-Up Areas Using Road Network Data. *Transactions in GIS* 19, 848–876.

Zhou, Q., Guo, L., 2018. Empirical approach to threshold determination for the delineation of built-up areas with road network data. *PLoS ONE* 13, e0194806.

Appendix 1: Head/tail breaks method applied to the rank-size distribution of morphological clusters, i.e. polygons formed by the continuous edges of the TIN whose length is shorter than the mean edge length. Grey line in the Table: end of the head/tail breaks process. After this step, the percentage of clusters in the head of the distribution is equal to 1 or higher than ~40%.

Street nodes of the province of Brabant

Step of the head/tail breaks	Total number of morphological clusters	Mean area of the clusters (in m^2)	Number of clusters in the head of the distribution	Number of clusters in the tail of the distribution	% of clusters in the head of the distribution	% of clusters in the tail of the distribution
1	4153	0.0854	375	3778	9%	91%
2	375	0.8105	48	327	13%	87%
3	48	4.7974	7	41	15%	85%
4	7	22.6925	1	6	14%	86%

Street nodes of the whole of Belgium

Step of the head/tail breaks	Total number of morphological clusters	Mean area of the clusters (in m^2)	Number of clusters in the head of the distribution	Number of clusters in the tail of the distribution	% of clusters in the head of the distribution	% of clusters in the tail of the distribution
1	23770	0.1082	2435	21335	10%	90%
2	2435	0.8980	402	2033	17%	83%
3	402	4.0189	75	327	19%	81%
4	75	13.8951	17	58	23%	77%
5	17	39.2771	5	12	29%	71%
6	5	81.3112	2	3	40%	60%
7	2	118.8025	1	1	50%	50%

Building centroids of the province of Brabant

Step of the head/tail breaks	Total number of morphological clusters	Mean area of the clusters (in km^2)	Number of clusters in the head of the distribution	Number of clusters in the tail of the distribution	% of clusters in the head of the distribution	% of clusters in the tail of the distribution
1	22629	0.0191	1517	21112	7%	93%
2	1517	0.2575	250	1267	16%	84%
3	250	1.2446	63	187	25%	75%
4	63	3.3218	14	49	22%	78%
5	14	8.2630	4	10	29%	71%
6	4	16.7524	3	1	75%	25%

Building centroids of the whole of Belgium

Step of the head/tail breaks	Total number of morphological clusters	Mean area of the clusters (in km^2)	Number of clusters in the head of the distribution	Number of clusters in the tail of the distribution	% of clusters in the head of the distribution	% of clusters in the tail of the distribution
1	137445	0.0199	8815	128630	6%	94%
2	8815	0.2793	1372	7443	16%	84%
3	1372	1.4354	369	1003	27%	73%
4	369	3.6304	104	265	28%	72%
5	104	7.2101	36	68	35%	65%
6	36	11.6104	9	27	25%	75%
7	9	19.9756	2	7	22%	78%
8	2	35.9760	1	1	50%	50%