Variography and Morphometry for Classifying Building Centroids: Protocol, Data and Script

Joan Perez[1(✉)], Alexandre Ornon[1], and Hiroyuki Usui[2]

[1] Université Côte-Azur, CNRS, ESPACE, Nice, France
{joan.perez,alexandre.ornon}@univ-cotedazur.fr
[2] Department of Urban Engineering, Housing and Urban Analysis Laboratory, The University of Tokyo, Tokyo, Japan
usui@ua.t.u-tokyo.ac.jp

Abstract. Different spatial patterns of urban growth exist such as infill, edge-expansion and leapfrog development. This paper presents a methodology, and a corresponding script, that classify new residential buildings as patterns of urban growth. The script performs a combination of variography and morphometry over building centroids on two different dates. The test data is made of the building centroids of 2002 and 2017 for Centre-Var, a region located in southern France. The different bounding regions, yield from series of morphological closings, allow classifying the building centroids that appeared between 2002 and 2017 into different categories of spatial patterns of urban growth. The final classification is made according to the degree of clustering/scattering of new centroids and to their locations regarding existing urban areas. Preliminary results show that this protocol is able to provide useful insights regarding the degree of contribution of each new residential building to the following patterns of urban growth: clustered infill, scattered infill, clustered edge-expansion, scattered edge-expansion, clustered leapfrog and scattered leapfrog. Open access to the script and to the test region data is provided.

Keywords: Classification · Morphology · Geoprocessing · Variography · Geostatistics

1 Introduction

Nowadays, a growing number of sources are directing the attention on the urgency of transforming the way urban spaces are built and managed [1–3]. Amongst these sources, a recurring issue is sprawl, which is deemed has a non-sustainable dynamic regarding environmental issues [2–4]. Yet, the literature on urban sprawl and its different patterns is ambiguous, or even confusing since no consensus is emerging in the academics [5–7]. This is without mentioning other patterns of growth which are not directly related to sprawl, such as redevelopment of existing urban areas, urban scattering, leapfrog urbanization, etc. Numerous denominations of somewhat similar

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12252, pp. 415–424, 2020.
https://doi.org/10.1007/978-3-030-58811-3_30

patterns also exist in the literature such as central-city revival and urban regeneration for urban redevelopment [8]; edge-expansion for continuous sprawl [9], etc. Even if sprawl remain the main scapegoat, other patterns of urban development are also pointed out as hardly sustainable [10]. As suggested by Liu *et al.,* [9], urban growth spatially operates following three main possibilities: infill, edge-expansion and leap-frog. Solely focusing on sprawl, Galster *et al.,* [6] proposes a conceptual definition based on eight dimensions. Among them, one is of particular interest in this research: clustering. It is defined as the degree to which development is tightly bunched together. The degree of clustered development appears not only relevant for edge-expansion patterns, but also for infill and leapfrog. Based on these observations, it should be possible to classify new residential buildings into the aforementioned patterns of urban growth focusing on location and clustered measures of both long-standing and new building centroids. Since certain patterns are deemed more sustainable than others, such as compact or resilient models of development [10, 11], such a classification could provide useful insights regarding into which pattern new building falls. Various methods have been proposed in the literature for measuring, quantifying, and/or delineating urban growth [5, 7, 12–15]. Yet, to the authors' knowledge, no methods are evaluating both urban development and redevelopment from the building scale.

The goal of this paper is to present a protocol, developed within the R platform, that classify spatiotemporal patterns of residential urban growth within a region located in southern France named Centre-Var. The protocol can be described as a location-based morpho-structural approach which combines variography analysis and morphometry (mathematical morphology analysis applied to vector objects). It first detects thresholds of residential building agglomerations in 2002 and 2017 and then performs several morphological closings according to building locations. The bounding regions obtained through the morphological closings then allow classifying the new residential buildings into six different spatial patterns of urban growth: clustered infill, scattered infill, clustered edge-expansion, scattered edge-expansion, clustered leapfrog, and scattered leapfrog. The script makes use of easily accessible data: GIS layers of residential building footprints. Open access to the source code and to the test region data is provided.

This paper is organized as follows. Section 2 introduces the test region and the primary data. Section 3 presents and details each step of the script. Section 4 presents the preliminary results obtained on the test region. Section 5 concludes the paper with a discussion on future applications and developments of the protocol.

2 Test Region and Data Presentation

The region in which the method is applied is the center of the Var department in southern France. The extent of the case study goes from the cities of Brignoles to Le Muy and also includes Draguignan, Vidauban and Le Luc. This area, named

Centre-Var, is located close to three major cities: Marseille, Toulon and Nice. Due to the proximity of these metropolitan areas, this region is sustaining fast processes of urbanization (increase of population, sprawl, urban redevelopment, etc.). According to INSEE[1], Centre-Var population in 1999 was 157.919 inhabitants and 210.071 in 2016, thus showing an impressive growth of 33.02% (Fig. 1).



Fig. 1. Centre-Var presentation. Background: OpenStreetMap, 2020.

The BD TOPO® datasets from the French National Geographic Institute (IGN) are extracted and filtered according to the Centre-Var boundaries. BD TOPO® are GIS layers of building footprints (polygons that represent building shapes on a two-dimensional plane) where buildings are digitized as single-part polygons. Two dates are retained: 2002 and 2017. Buildings not possessing any residential functions are filtered out of the datasets using the specialization attribute included in the primary data. Since these data includes small polygons usually associated to residential buildings (garage, pergola, sheds, etc.), light buildings and small structures below 20 m^2 are filtered out. Data are finally harmonized into a single point feature class (centroids of polygons) possessing two attribute columns indicative of the building presence for each period: *pres2002* and *pres2017* (modalities: "1" for building presence; "0" otherwise). The prepared dataset is ultimately made of 82.249 centroids of residential buildings in 2017, from which 63.238 were also present in 2002. Data are compiled into a GeoPackage file named *CV_PT_Ext_0217*, itself containing two layers: the centroids, named *CV_Building_PT* and the boundaries of Centre-Var, named *Centre_Var_extent*.

[1] French national institute for statistical and economic studies.

3 Script Presentation and Application

The script starts with a subsection that provides information on the R session and packages versions. Necessary packages are then loaded. A second subsection loads the GeoPackage and subsequently divides the centroids into three simple feature geometry objects: the centroids present in 2002 (*Pt_2002.sp*), those present in 2017 (*Pt_2017.sp*) and the new centroids that appeared between 2002 and 2017 (*Pt_2017_new.sp*). The script is then organized in three parts, as follows:

Part one of the script creates a *RasterLayer* object of 50 m of resolution within the case study boundaries using the "raster" function from the *raster* package. The grid, superimposed on the sampling area *Extent.sp*, allows counting the number of building centroids inside each cell for 2002 and 2017 using the "cellFromXY" function. The average number of centroids for non-zero cells is 1.77 in 2002 and 1.87 in 2017, with an increase of occupied cells from 35.722 to 43.829, thus showing both a concentration and an expansion of urban development.

A cross-variogram is calculated for these two variables using the "variogram" function of the *gstat* package [16]. Variography allows exploring the spatial structure inherent to the point distributions by computing the square difference between the values of all the couples of pairs at given distances [17]. The main way to interpret a variogram is to observe where the breaks occur, in order to detect structures at a varying distance. When performed on density grids, the variations of autocorrelation explain a local change in the way that buildings are structured. In a cross-variogram, the first variation is considered as a micro-structure: the first pattern of point agglomeration. It is detected through the local R, which is a local correlation coefficient computed by the cross-variogram values divided by the product of the square roots of both regular variograms [18]. It sets the threshold of the agglomeration distance within the point distribution, which is going to be used in the morphometry section of our protocol. As displayed in Fig. 2, the first micro-structure threshold is identified at 227 m, which is rounded according to the grid resolution at 250 m.

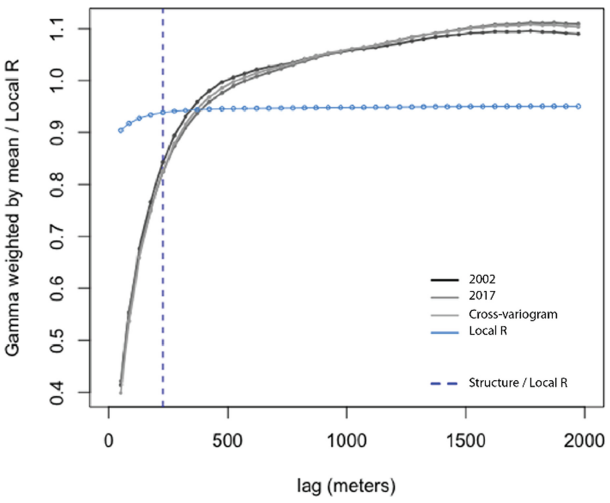

Fig. 2. Cross-Variogram for 2002 and 2017 residential buildings (50 by 50 m)

Offsets around the 2002 and 2017 point features are then created in order to extract built-up areas using the "buffer" function of the *raster* package. These operations, named morphological closings [14], are composed of a dilatation followed by an erosion algorithm. They allow linking close by centroids, ignoring small holes and interstices. The micro-structure threshold of 250 m detected by the analysis of the cross-variogram sets the buffering distance from each point (radius $r = 250/2$). For both closings, a filter is applied in order to remove the small isolated surfaces artificially generated by the dilation algorithm, which are the surfaces below the lower limit of $\pi\ r^2$. The last section of part one performs the difference between the closings of 2017 and 2002 and the difference between the extent of the case study and the closing of 2017 using the "gDifference" function from the package *rgeos*.

Part two of the script divides the new centroids *Pt_2017_new.sp* into three subsets through basic operations of intersect using the "st_intersection" function of the *sf* package. *Point_evo_inside_2002* are the centroids within the 2002 morphological closing, *Point_evo_outside_2017* are the centroids outside the 2017 closing and *Point_evo_diff_0217* are the centroids located both outside 2002 and inside 2017 closings. A downscaling is performed through the creation of a *RasterLayer* object of 25 m of resolution. The number of new building centroids is then counted, using once again "raster" and "cellFromXY" functions. Performing a downscaling is relevant since, at this stage, the focus of the study is no more on the inclusion/exclusion of new buildings within existing urban areas but rather on their clustered/scattered properties. Since each subset is representative of a peculiar trajectory, regular variograms are calculated within each subset (instead of a cross-variogram). As displayed in Fig. 3, they yield thresholds of 113 m for *Point_evo_inside_2002* (rounded at 125 m), 161 m for *Point_evo_diff_0217* (rounded at 150 m), and 238 m for *Point_evo_outside_2017* (rounded at 225 m). Morphological closings are once again performed for each subset using the different identified thresholds as buffering radiuses.

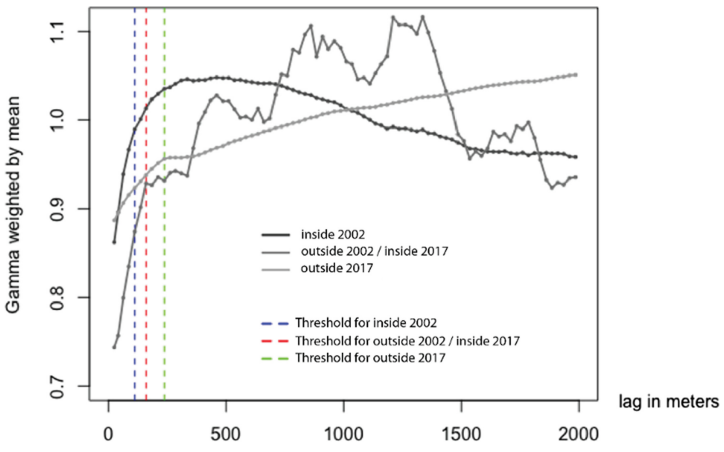

Fig. 3. Variograms for *Point_evo_inside_2002*, *Point_evo_outside_2017* and *Point_evo_diff_0217* (25 by 25 m)

Part three performs the final classification by selecting the building centroids encroaching upon the different buffers, ultimately yielding the following categories: clustered infill, scattered infill, clustered edge-expansion, scattered edge-expansion, clustered leapfrog and scattered leapfrog. The different geometries are also clipped together once again using the function "gDifference". Several plots and maps are displayed and results are recorded, including the variogram samples. Result layers are added within the original GeoPackage file. Figure 4 summarizes the different steps of the protocol. Open access to the script and to the test region data is provided (Appendix).

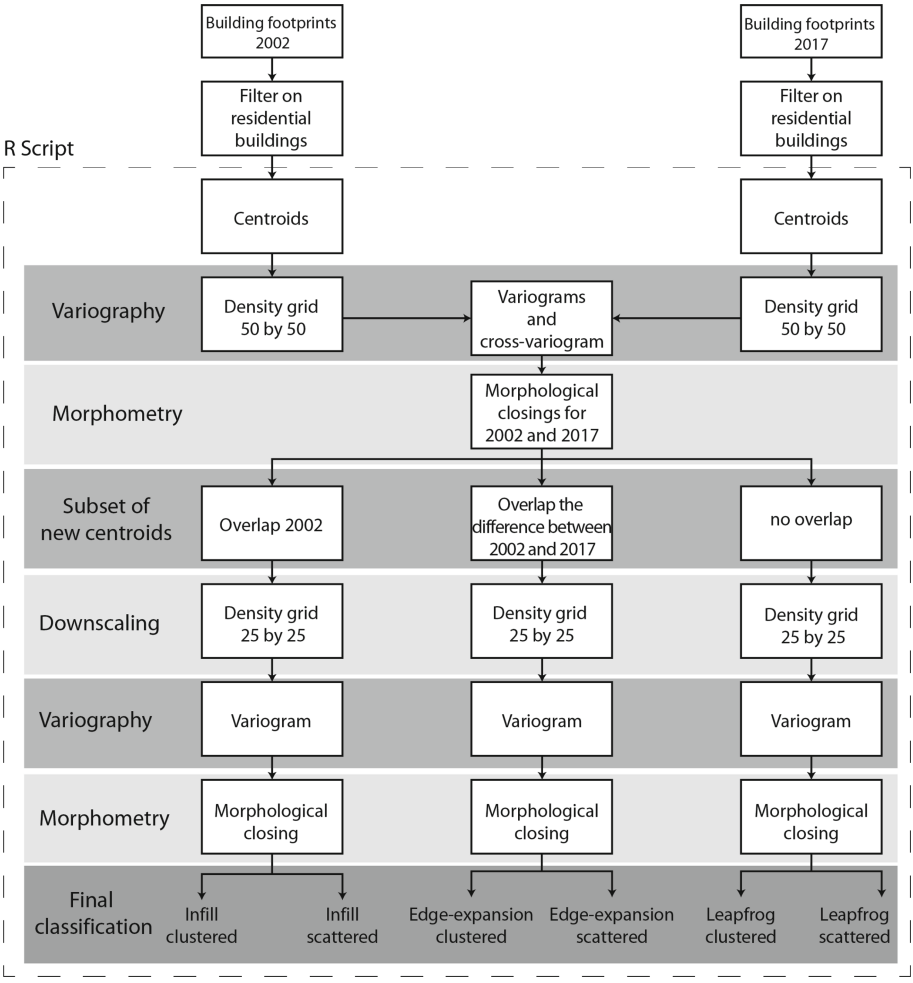

Fig. 4. Flowchart of the protocol

4 Preliminary Results

New residential buildings have been successfully classified into three categories: infill, edge-expansion, and leapfrog; each of the latter further subdivided into clustered and scattered patterns.

As Fig. 5 shows, the three categories, as well as their subdivisions, clearly stand out. Both Brignoles and Le Luc are mostly concerned by infill and edge-expansion patterns. They are both gaining a lot of ground on non-urban spaces from existing urban structures, whereas leapfrogs exist but remain limited comparing to the other patterns.

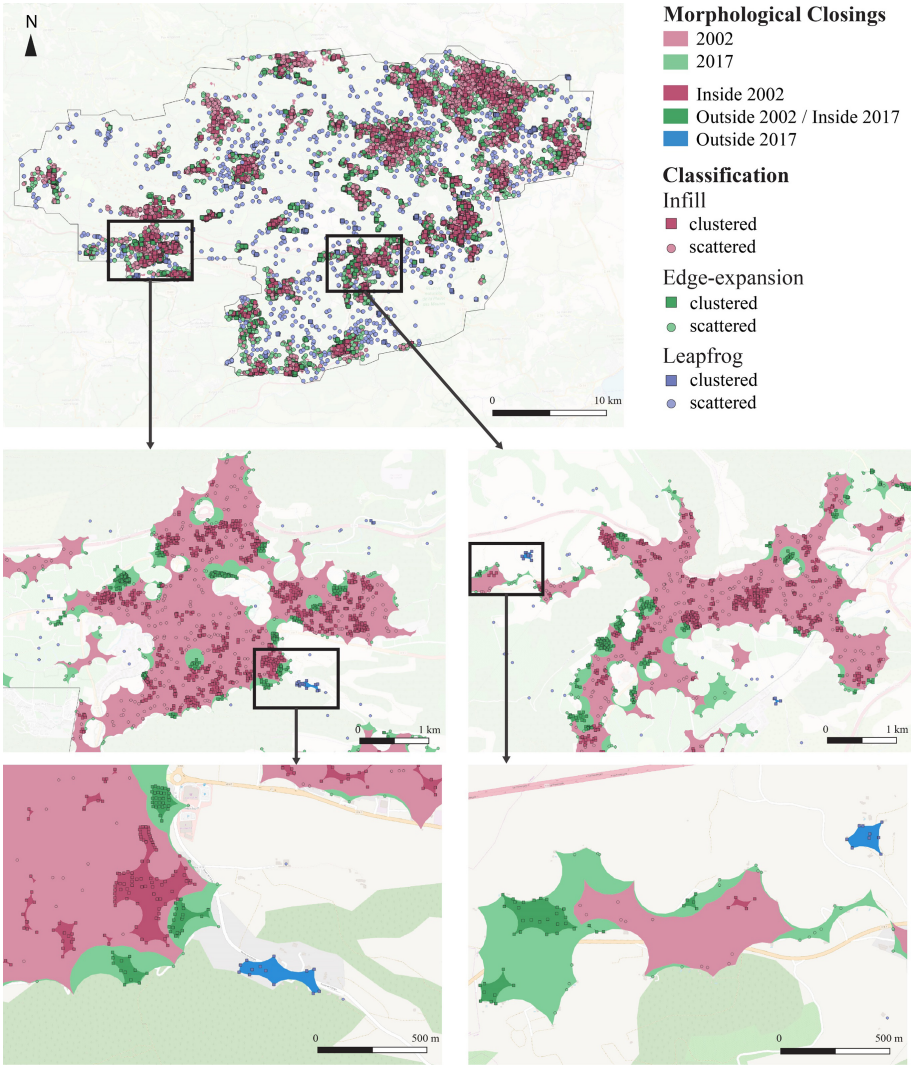

Fig. 5. Morphological closings and classification results. Focus on Brignoles (left) and Le Luc (right). Background: OpenStreetMap, 2020.

66.71% of the new buildings are located within urban structures that were already present in 2002. These buildings are infill patterns of urban growth directly contributing to urban densification. Among this group, the clustered pattern corresponds to the infilling of urban vacuities at the scale of urban blocks or strips of land that were previously not constructed, or at least not fully occupied by residential buildings. This category is also corresponding to various projects of urban redevelopment and regeneration of old urban fabrics. The scattered pattern can however be described as additions of adjacent buildings or infillings of urban vacuities at the scale of the plot, within existing urban blocks. It can be safely assumed that these new buildings are not planned nor coordinated by administrative authorities and promoters, but rather the outcome of individual initiatives.

24.85% of the new buildings, classified as edge-expansion patterns, are directly responsive for urban sprawl. They are the ones gaining ground on non-urban spaces from existing urban structures, thus prolonging the latter in an unbroken fashion. The clustered ones are corresponding to compact new peripheral neighborhoods. Scattered urban development are also contributing to urban expansion. Yet, as compared to the former, this pattern is not compact, thus contributing to low-density urban expansion.

8.44% of the new building centroids are located outside of the 2017 morphological closing, which means that these new buildings are not contributing to urban growth in the sense of continuity. The clustered pattern be described as compact leapfrog urban developments. These patterns are widespread in numerous countries but, as highlighted in Fig. 4, it appears that it is not so much the case in Centre-Var. This phenomenon is often described in the academic literature as compact but yet not continuous urban development [19, 20]. In Brignoles and Le Luc only two small pockets of such emerge. The scattered leapfrog pattern can, for its part, be described as low-density urban development. As compared to the clustered pattern, the hypothesis is that most of these new buildings are single family homes surrounded by large private gardens and villas of complex shapes.

5 Discussion and Future Developments

The protocol presented in this paper allows classifying new residential buildings into six different categories though a combination of variography and morphometry. It yields the following categories: infill, edge-expansion and leapfrog; each of the latter further subdivided into clustered and scattered patterns. As shown in the application to all new residential buildings between 2002 and 2017 in the Centre-Var region, this protocol is able to provide useful insights regarding into which model of development the new residential buildings fall. As highlighted in the literature, some models of development are more sustainable than others. On that basis, 4.726 residential buildings (24.85% of the distribution) are directly responsive for sprawl in the sense of continuity while 1.603 buildings can be described as leapfrog patterns of development (8.44%). The rest of the distribution are characterized by infill densification patterns. Both clustered and scattered patterns are increasing the density of existing urban spaces and are thus in line with compact city theories. Yet, a fieldwork step is required to

validate these preliminary results. The protocol also must be tested in other areas, notably within more urbanized regions such as metropolitan areas.

The script will be improved in several directions. First, alternatives will be sought for some functions that are time-consuming regarding computation, such as "variogram" and "gDifference". Parallel computing is also considered. Second, the script will be simplified to make it reproducible to any area. Since it only requires GIS layers of residential building footprints, it could evolve into a decision tool for the evaluation and quantification of spatial patterns of urban growth.

Appendix: Data and Script

Files

- CV_PT_Ext_0217.gpkg (9.8 MB)
- PT_Classification (v.1.2 - 13.6 kB)
- CV_PT_Ext_0217_RES.gpkg (48.9 MB)
- README.txt (2.9 kB)

Title

Variography and Morphometry for Classifying Building Centroids as Patterns of Urban Growth v1.2

Link

https://zenodo.org/record/3899981

Description

This upload contains a dataset ready to be processed, an R script (v1.2) and a dataset with the associated results of the processing steps. The dataset is a GeoPackage file (EPSG:2154 - RGF93/Lambert-93 - Projected) named "CV_PT_Ext_0217.gpkg" containing two layers (1) "CV_Building_PT": a point feature class made of the centroids of residential buildings in 2002 and 2017 extracted from the French BD TOPO® (National Geographic Institute). This layer possesses two attribute columns indicative of the building presence for each period: "Pres_2002" and "Pres_2017" (modalities: "1" for building presence; "0" otherwise). (2) "Centre_Var_ext": a polygon feature class related to the extent of the case study: a region in southern France named Centre-Var. The script performs a combination of variography analysis and morphological closings over the building centroids. The different bounding regions then allow classifying new residential buildings (the ones that appeared between 2002 and 2017) into different categories of patterns of urban growth according to their degrees of clustering/scattering and to their locations regarding existing urban areas. The GeoPackage associated with the results "CV_PT_Ext_0217_RES.gpkg" contains two additional layers (3) "PT_CLASS" a point feature class made of the new building centroids with an attribute column "cat" related to the classification outputs. The categories are as follows: 1 - clustered infill, 2 - scattered infill, 3 - clustered edge-expansion, 4 - scattered edge-expansion, 5 - clustered leapfrog and 6 - scattered leapfrog. (4) "BR_Clipping": a polygon feature class made of the different bounding regions clipped together.

References

1. United Nations: Transforming our World: The 2030 Agenda for Sustainable Development. United Nations, New York (2015)
2. European Environment Agency. Urban sprawl in Europe, The ignored challenge. European Commission report, p. 60 (2006)
3. Verderber, S.: Sprawling Cities and Our Endangered Public Health. Routledge, London (2012)
4. Brueckner, J.K.: Urban sprawl: diagnosis and remedies. Int. Reg. Sci. Rev. **23**(2), 160–171 (2000)
5. Bhatta, B., Saraswati, S., Bandyopadhyay, D.: Urban sprawl measurement from remote sensing data. Appl. Geogr. **30**(4), 731–740 (2010)
6. Galster, G., Hanson, R., Ratcliffe, M., Wolman, H., Coleman, S., Freihage, J.: Wrestling sprawl to the ground: defining and measuring an elusive concept. Hous. Policy Debate **12**(4), 681–717 (2001)
7. Bhatta, B.: Analysis of Urban Growth and Sprawl from Remote Sensing Data. Advances in Geographic Information Science, p. 172. Springer, Berlin (2010). https://doi.org/10.1007/978-3-642-05299-6
8. Smith, N.: The New Urban Frontier, Gentrification and the Revanchist City. Routledge, New York (1996)
9. Liu, Z., He, C., Wu, J.: General spatiotemporal patterns of urbanization: an examination of 16 world cities. Sustainability **8**, 41 (2016)
10. Jenks, M., Burton, E., Williams, K.: The Compact City: A Sustainable Urban Form? Spon Press, London, p. 360 (1996)
11. Burdett, R., Philipp, R.: Shaping Cities in an Urban Age, p. 448. Phaidon Press, New York (2018)
12. Tannier, C., Thomas, I., Vuidel, G., Frankhauser, P.: A fractal approach to identifying urban boundaries. Geogr. Anal. **43**(2), 211–227 (2011)
13. Chaudhry, O., Mackaness, W.A.: Automatic identification of urban settlement boundaries for multiple representation databases. Comput. Environ. Urban Syst. **32**(2), 95–109 (2008)
14. Perez, J., Fusco, G., Moriconi-Ebrard, F.: Identification and quantification of urban space in India: defining urban macro-structures. Urban Stud. **56**(10), 1988–2004 (2019)
15. Usui, H.: A bottom-up approach for delineating urban areas minimizing the connection cost of built clusters: comparison with top-down-based densely inhabited districts. Comput. Environ. Urban Syst. **77**, 101363 (2019)
16. Pebesma, E.: Multivariable geostatistics in S: the GSTAT package. Comput. Geosci. **30**, 683–691 (2004)
17. Armstrong, M.: Experimental variograms. In: Basic Linear Geostatistics, pp. 47–58. Springer, Berlin, Heidelberg (1998). https://doi.org/10.1007/978-3-642-58727-6_4
18. Dauphiné, A., Voiron Canicio, C.: Variogrammes et structures spatiales, Collection Reclus Modes d'Emploi no 12, GIP Reclus (1988)
19. Mills, D.E.: Growth, speculation and sprawl in a monocentric city. J. Urban Econ. **10**(2), 201–226 (1981)
20. Zang, W., Wrenn, D.H., Irwin, E.G.: Spatial heterogeneity, accessibility, and zoning: an empirical investigation of leapfrog development. J. Econ. Geogr. **17**(3), 547–570 (2017)