Journal of Transport Geography 82 (2020) 102565

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Combining travel and population data through a bivariate spatial analysis to define Functional Urban Regions

Larissa Limongi Aguiar[a], Gustavo Garcia Manzato[b,*], Antônio Nélson Rodrigues da Silva[c]

[a] São Paulo State University – UNESP, School of Engineering, Bauru, Brazil
[b] São Paulo State University - UNESP, School of Engineering, Bauru, Avenida Engenheiro Luiz Edmundo Carrijo Coube, 14-01, 17033-360.Bauru, SP, Brazil
[c] University of São Paulo – USP, São Carlos School of Engineering, Brazil

ABSTRACT

The aim of this paper is to provide support to the issue of defining Functional Urban Regions (FURs). We carried out an exploratory study in which both commuting trips and population data were used in a combined setting. To do this, Bivariate Exploratory Spatial Data Analysis was used, enabling permuted analyses of the variables considered. The municipalities of São Paulo state, Brazil, were used in the case study investigated in this research. We observed that this combined analysis has more satisfactory results when compared to the analyses of variables in an isolated way. There is also a greater correlation of the results obtained from bivariate analysis with the official delimitations. Considering the possible permuted analyses, which are travel data versus population or population versus travel data, the first showed more reliable results in terms of identifying municipalities with strong correlations between the two variables. This can be observed both in areas within official FURs and throughout the territory, highlighting the areas of regional influence.

1. Introduction

Over the past half century, the world became more urban. Data from the United Nations (UN, 2019) show that urban areas house 55% of the world's population, but this figure may reach 68% within the next 30 years. This shift from rural to urban has remarkably been seen in developing countries (Cohen, 2006). For example, the United States took nearly 90 years compared to what Korea and Brazil took, respectively, 20 and 30 years (Henderson, 2002). Tokyo, New Delhi, Shanghai, Mexico City and São Paulo are the world's largest cities, all with more than 22 million inhabitants within their respective metropolitan areas. In addition, Cairo, Mumbai, Beijing and Dhaka also rank among the most populous cities, all comprising almost 20 million inhabitants (UN, 2019). Except from Tokyo, these cities share the same fact: they all belong to developing countries.

For those countries, increasing urbanization brings challenges in managing the demands for housing, jobs, basic infrastructures and services. Such challenges are tied up with integrated policies for the sustainable development of urban areas, which in most cases are not formed by a single city. This fact introduces an additional task to be faced by urban planners and managers, regarding the analysis of the various processes emerging from metropolization. In turn, this task is also associated with the issue of achieving greater success in terms of solving urban problems at this scale. The so-called Functional Urban Regions (FURs) are usually established in these cases, comprising the figure of a metropolitan government to manage the interests of some of the actors involved (population, local governments, stakeholders, etc.). The issue, however, relies on how to define (or delimit) such FURs.

In the United States of America, for example, a Metropolitan Statistical Area (MSA, which corresponds to the term FUR used here) is defined by two factors, which are: an urban area with more than 50,000 inhabitants, as well as adjacent counties in which at least 25% of the residents go on commuting trips to that urban area (Office of Management and Budget, 2010). Similarly, FURs in Europe (also known as Functional Urban Areas) comprise at least one main urbanized area containing 20,000 (or more) jobs and any other adjacent region that is classified as NUTS3 from where more commuting to this main urban area is observed than to any other area (Cheshire and Hay, 1989). NUTS stands for the Nomenclature of Territorial Units for Statistics, which is a geographical nomenclature subdividing the economic territory of the European Union into regions according to Eurostat, the statistical office of the European Union. NUTS3 corresponds to the third subdivision level, which is less aggregated than NUTS1 and NUTS2. The Organization for Economic Co-operation and Development (OECD) also worked on a common definition of metropolitan areas in order to assess national urban policies in OECD countries, so as to make international comparisons of economic, social and environmental performances. Together with the European Union, the OECD developed a "*harmonized definition of urban areas*" called "*Functional Economic Units*", which avoid restrictions related to an administrative-based definition (OECD,

* Corresponding author.
E-mail address: gustavo.manzato@unesp.br (G.G. Manzato).

https://doi.org/10.1016/j.jtrangeo.2019.102565
Received 19 June 2018; Received in revised form 14 October 2019; Accepted 14 October 2019
Available online 02 December 2019
0966-6923/

Table 1
Main indicators employed for the definition of FURs and their respective authors.

Main indicator	Authors
Political-administrative relations	Arellano Ríos (2014a, 2014b, 2015)
Remote sensing features	Weber (2001); Huang et al. (2014); Niemeyer et al. (2014); Sahoo and Pekkat (2014)
Land price	Bode (2008)
Distribution of jobs	Coombes and Overman (2004)
Urban form and morphology	Bereitschaft and Debbage (2014)
Population density	Ramos and Rodrigues da Silva (2003, 2007); Ramos et al. (2004); Manzato et al. (2007)
Transport infrastructures and services	Boarnet and Haughwout (2000); Manley (2014); Manzato et al. (2015); Guzik et al. (2017); Moisés et al. (2017)
Combination of demographic and transport infrastructure supply indicators.	Manzato and Rodrigues da Silva (2010); Pereira and Rodrigues da Silva (2010); Ajauskas et al. (2012); Rodrigues da Silva et al. (2014); Oliveira Junior et al. (2017)
Transport flows	Konjar et al. (2010); Kauffmann (2012); Williams et al. (2012); Horňák and Kraft (2015); Aguiar et al. (2017); Kraft and Marada (2017); Soares et al. (2017)

2012). The method is also based on the distribution of the population and on commuting data (OECD, 2016).

In fact, flows of people and goods have generally been used as a measure of functional relationships between municipalities, which seem reasonable for the definition of FURs (Kauffmann, 2012; Williams et al., 2012; Klapka and Halás, 2016). However, various approaches can be found in the literature concerning the definition of FURs, some including the intrinsic relationship with transport flows. Table 1 summarizes the indicators used and the respective authors involved.

The methods for defining FURs established over the last decades have analyzed different variables and sought to represent the complex relationships in these areas. This complexity and the variety of approaches are supported by Breitung (2011), arguing that the analysis of the delimitation of FURs should be done according to five different, though interrelated aspects: political and administrative matters; physical or morphological characteristics; socio-spatial and cultural issues; psychological and behavioral perspectives; and functional relations and cooperation networks. However, according to Kourtit et al. (2015), these aspects might be difficult to measure due to data availability, lack of skilled labor force to perform the analyses, etc. or may provide ambiguous or inaccurate definitions. These issues are particularly evident in developing countries (Cohen, 2006).

In Brazil, although the installation of FURs is foreseen in the Federal Constitution of 1988, there are no minimum criteria, mainly quantitative, established by the Federal government for its definition. The Brazilian states are responsible and, consequently, have the autonomy to install their own FURs, whose definition criteria are established at the state level. Obviously, if we compare the definitions among states, wide disparities can be observed. As states aim to meet their needs by seeking federal resources by institutionalizing new FURs, this also results in a considerable number of FURs that actually have a low-profile metropolitan dynamic (Moura and Carvalho, 2012). Thus, it is assumed that the definitions adopted in the country are predominantly based on political-administrative relations. However, in reality, this method may not be able to express the necessary loyalty to the planning and urban management processes of the whole area resulting from the union of municipalities that comprise a FUR.

Previous studies carried out in Brazil (Manzato et al., 2007; Manzato and Rodrigues da Silva, 2010; Pereira and Rodrigues da Silva, 2010; Ajauskas et al., 2012; Rodrigues da Silva et al., 2014; Manzato et al., 2015; Oliveira Junior et al., 2017) attempted to contribute to a more quantitative method for defining FURs, using data about population distribution and highway infrastructure supply. Population distribution is a well-established indicator to measure urbanization (Office of Management and Budget, 1998; UN, 2019) as well as it is available in virtually every country. Transportation supply indicators have been employed (see Table 1) to analyze the definition of FURs, in some cases as a proxy variable to transport flows. A remarkable feature observed in those studies carried out in Brazil was the application of spatial analysis techniques, such as spatial statistics and spatial modeling as part of the methods. That is, the studies brought new insights to the issue of FURs definition, which in fact were initiated by Ramos and Rodrigues da Silva (2003, 2007) and Ramos et al. (2004), employing techniques to take into account the spatial dependency within the data.

The aforementioned Brazilian studies were developed considering the fact that data about transport flows were not easily available (as it is the case of most developing countries). However, an effort of the Brazilian Institute for Applied Economic Research (IPEA) made available data about commuting flows. Based on these data, Castello Branco et al. (2013) provided a discussion about the delimitation of the Brazilian Metropolitan Regions, highlighting the existing heterogeneity in approaches across the country, and arguing about the difficulties in evaluating their pros and cons. In that sense, the authors worked on a unified methodology for the entire country, based on the distribution of the population and commuting flows. More recently, Aguiar et al. (2017) also investigated the data on commuting flows provided by IPEA, setting up various scenarios with a combination of population levels of a given municipality, and the percentage levels of residents in other municipalities traveling to work in that municipality. This study was an attempt to analyze the definitions of FURs in the São Paulo state, Brazil, based on the method used in the United States of America.

Despite what has been increasingly developed to provide approaches for the definition of FURs, there is still room for new insights in this field. Developing countries still lack for studies regarding the effects of urbanization, especially metropolization. In line with what Aguiar et al. (2017) proposed, we also explored a combination of the distribution of the population and commuting flows, but we tried to analyze such variables with the application of spatial statistics techniques. More specifically, we propose an approach to define FURs based on the Bivariate Exploratory Spatial Data Analysis (ESDA) technique applied to those variables. To our knowledge, the use of this technique to the issue of analyzing FURs has not been approached yet, except for the univariate version as described before (which, in fact, led to promising results according to the respective authors). Another contribution of this study relies on the fact that the commuting flows are considered in terms of incoming and outgoing trips. On one hand, the cross-border flows are not taken into account, which in fact could represent a limitation in our study. On the other hand, the cross-border flows may be unavailable in some developing countries, where only the statistics about the totals of incoming and outgoing trips might be obtained. Therefore, the proposed study here tries to cover this gap as well.

The case study was carried for São Paulo state. Regarding the description of the study developed, this article is organized into sections. Section 2 contains the theoretical reference and details of the variables and procedures used. Afterwards, the results of the approach are presented and analyzed in Section 3. Finally, conclusions regarding the contribution of this indicator are discussed in Section 4.

2. Methodology

The methodology described in this paper is divided into two parts: the first provides an overview of the theoretical framework of spatial analysis tools; and the second part describes the case study.

2.1. Theoretical framework

The methodology developed in this study investigated population data and commuting travel data using ESDA (Anselin, 1995, 1998). The ESDA technique is based on the characterization of spatial dependence, designating how the variables are spatially correlated. Depending on the number of variables to be analyzed, it can be classified as univariate or multivariate analysis.

The univariate ESDA technique classifies each observation using two parameters. The first, Z, corresponds to a vector that represents the calculation of the difference between the value of a given variable in each observation and the overall average of that variable considering all the observations. The second, W_Z, corresponds to the difference between the average value of such a variable computed for the neighboring observations of a given observation and the overall average of that variable considering all observations. The Moran's index (I) is obtained from these parameters, as given by Eq. (1), in which the exponent t defines the transposed vector.

$$I = \frac{Z^t.\ W_Z}{Z^t.\ Z} \tag{1}$$

The results of this classification can also be observed in the Moran scatterplot, in which four quadrants are identified as described hereafter. Quadrant HH (for High-High) depicts the observations whose value of the variable is higher than the overall average and the average value of the neighboring observations is also higher than the overall average. In sequence, quadrant LL (for Low-Low) depicts the observations whose value of the variable is lower than the overall average and the average value of the neighboring observations is also lower than the overall average. Next, quadrant LH (for Low-High) depicts the observations whose value of the variable is lower than the overall average, but the average value of the neighboring observations is higher than the overall average. Finally, quadrant HL (for High-Low) depicts observations whose value of the variable is higher than the overall average, but the average value of the neighboring observations is lower than the overall average. In addition to the Moran scatterplot, the observations classified in one of the quadrants described can be also seen in the thematic maps, the so-called “Box Maps”.

A method developed by Anselin et al. (2002) tests multivariate spatial autocorrelation in lattice data. This concept of autocorrelation represents the degree of association between the values of a variable (z_k) observed at a given location and the values assumed by a second variable (z_l) visualized in neighboring locations. More specifically, it is a linear and systematic association (more than under spatial randomness) between the variable z_k of a location i (${z_k}^i$) and the spatial lag for the other variable, that is, ${W_{Z_l}}^i$. To relate these variables, a multivariate coefficient of spatial autocorrelation is defined, according to Eq. (2).

$$m_{kl} = z_k^t.\ \mathrm{W}^S.\ z_l \tag{2}$$

where:

$$z_k = \frac{[\mathrm{x_k} - \bar{\mathrm{x}}_\mathrm{k}]}{\sigma_\mathrm{k}} \text{ and } z_l = \frac{[\mathrm{x_l} - \bar{\mathrm{x}}_\mathrm{l}]}{\sigma_\mathrm{l}}$$

The variables z_k and z_l are standardized, so that the mean is zero and the standard deviation is one. The vector Z_k provides the deviations between the values of the variable z_k in the observed locations and the overall average. The other vector, Z_l, incorporates the deviations between the average values for the z_l variable in the neighboring areas of the observed locations and the overall average. As for W^s, this is a spatial weights matrix that identifies the set of neighbors of each location, assigning non-zero values to the neighbors and zero to the others. Moreover, zero is assigned to the main diagonal as locations are not considered adjacent to themselves.

The Moran's index characteristic of the multivariate analysis, in turn, can be obtained from Eq. (3), which makes use of the basic elements already mentioned, where Z_k is the vector of deviations for the evaluated variable in the observed locations, Z_l is the vector of deviations for the variable analyzed in adjacent locations, W is the spatial matrix of weights with normalized lines and the exponent t defines transposed vector.

$$I_{kl} = \frac{z_k^t.\ W.\ z_l}{z_k^t.\ z_k} \tag{3}$$

Considering n as the number of observations, once the z-variables are standardized, we can consider that the sum of squares, contained in the denominator of Eq. (3), is equal to n, regardless of the variable values. By doing this, Eq. (4) is obtained.

$$I_{kl} = \frac{z_k^t.\ W.\ z_l}{n} \tag{4}$$

The multivariate ESDA technique also evaluates the results using the Moran scatterplot and Box Maps, following the same reasoning explained before. In quadrant HH, the value of variable z_k is higher than the overall average and its neighbors' average for variable z_l is also higher than the overall average. In quadrant LL, the value of variable z_k is lower than the overall average and its neighbors' average for variable z_l is also below the overall average. Thus, a positive spatial autocorrelation is visualized for the observations classified in these two quadrants. In quadrant LH, the value of variable z_k is lower than the overall average and the average of variable z_l in adjacent observations is higher than the overall average. Finally, in quadrant HL, the value of variable z_k is higher than the overall average and its neighbors' average for variable z_l is lower than the overall average. It is understood, therefore, that the spatial autocorrelation between the observations classified in these two quadrants is negative.

2.2. Case study

In this study, data with information about population and commuting travel regarding the year 2010 were explored. These data were obtained from two main sources, i.e., the Brazilian Institute of Geography and Statistics (IBGE, 2010) and the Observatory of the Metropolis (2013). They were organized, manipulated and presented at the level of municipalities, considering São Paulo state as our case study. More specifically, three variables comprised the analyses, including: 1) *total number of people making outgoing trips from the municipality*; 2) *total number of people making incoming trips to the municipality*; and 3) *total population in the municipality*.

A first exploratory analysis was carried out with the variables independently, i.e., performing a univariate ESDA. In sequence, the two variables related to the commuting trips were analyzed together with population using the multivariate ESDA technique. Initially, the variable z_k was associated with commuting trips data and the variable z_l was associated with population. Afterwards, these variables were permuted, so that population became the variable z_k and commuting trips became the variable z_l.

The software packages used were Maptitude, which is a geographic information system (GIS) for data manipulation generating Box Maps, and GeoDa, which was used for applying the ESDA technique.

3. Results

Initially, the variables related to commuting travel as well as population were explored using the univariate ESDA technique, so that the analysis of each variable generated a separate Box Map, as displayed in Fig. 1. It is worth noting that both total population and

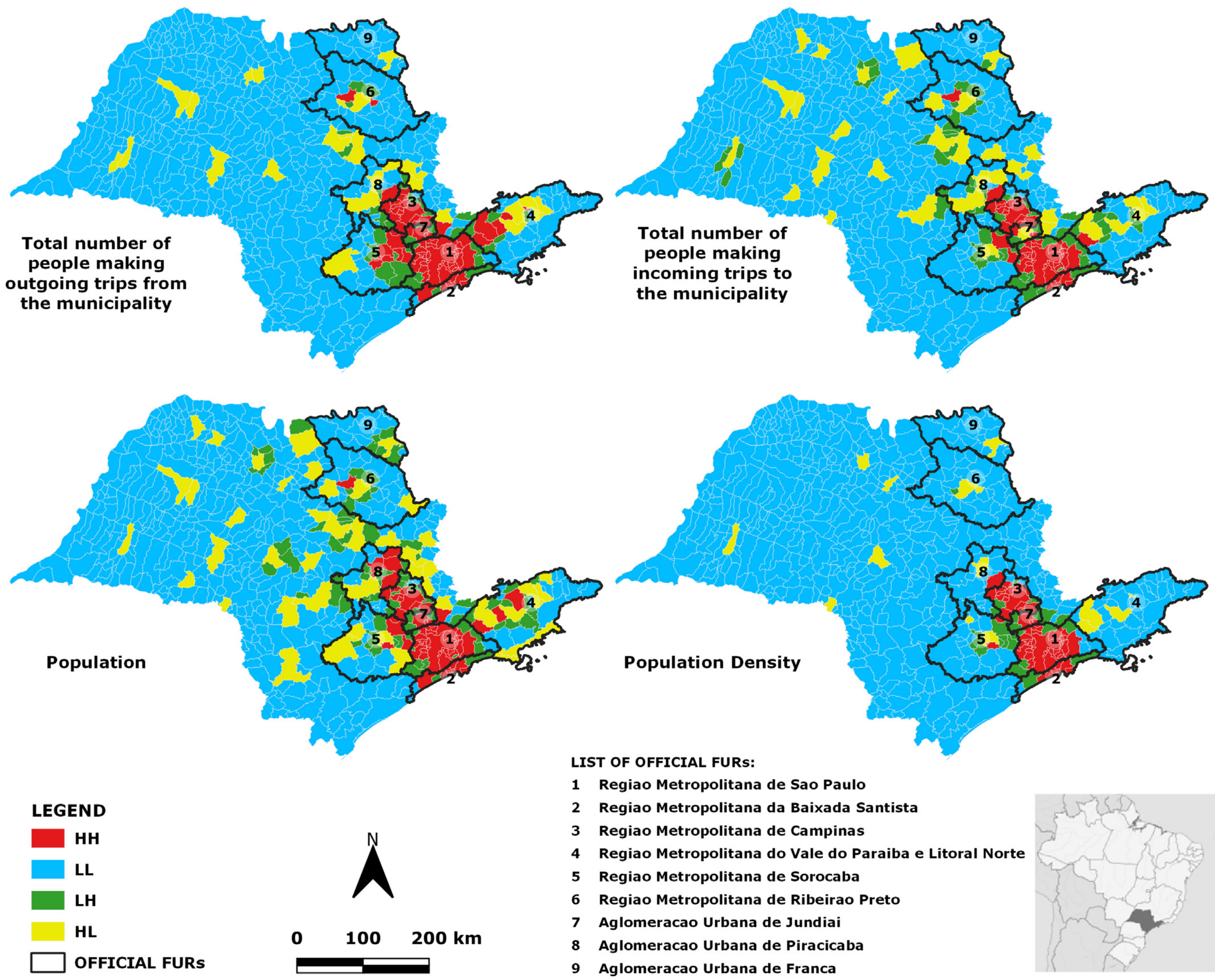

Fig. 1. Box maps representing the total number of people making outgoing trips from the municipality (top left), the total number of people making incoming trips to the municipality (top right), population (bottom left) and the population density (bottom right) - Source: Created by the authors.

population density are presented because the first follows part of the criteria employed in the United States of America and in Europe whereas the latter is commonly found in studies in which spatial statistics techniques are applied. Although the application of such techniques is one of the key points of this study, which might lead us to use population density, we chose the variable total population for the analyses. We did so to follow the same criteria that are based on the combination of travel and population data (e.g., Office of Management and Budget, 2010; OECD, 2016; and the specific study carried out in the São Paulo state by Aguiar et al., 2017).

Concerning the classification of municipalities in each quadrant, an explanation of the interpretations is presented as follows, taking as an example the variable "*Total number of people making incoming trips to the municipality*". For the other variables, the same rationale was followed. Therefore, municipalities classified in quadrant HH have, as well as their neighbors, a value of the *total number of people making incoming trips to the municipality* which was higher than the overall state average. Thus, municipalities classified in this quadrant expect a strong correlation with official FURs due to the high number of commuting trips. In contrast, municipalities classified in quadrant LL have, as well as their neighbors, a value of the *total number of people making incoming trips to the municipality* lower than the overall state average. In these cases, we expected to identify the municipalities with a low number of commuting trips, usually located away from the official FURs.

Considering the interpretations for the quadrant LH, the municipalities classified in this quadrant have a value of the *total number of people making incoming trips to the municipality* lower than the overall average of the state, while their neighbors display a higher value of the related variable compared to the overall state average. Thus, the municipalities located in the vicinity of those classified as HH and, consequently, adjacent to the official FURs are generally identified in quadrant LH. Finally, the municipalities classified in the quadrant HL present a value of the *total number of people making incoming trips to the municipality* higher than the overall average of the state, while their neighbors had a lower value of the related variable compared to the overall state average. These cases occur commonly away from official FURs, but they highlight municipalities that generally exert a regional influence in the state.

Based on the *Box Map* regarding the variable "*Total number of people making incoming trips to the municipality*" (Fig. 1), we observe that the municipalities in quadrant HH, which demonstrate a high intensity of commuting trips, are located in the official FURs. It can also be observed that the municipalities belonging to quadrant LL are distributed across most of the territory, considering that these municipalities (as well as their neighbors) have a low number of incoming trips, i.e., they are not attractive areas.

In addition to identifying the FURs, there is also evidence of the regional centers represented by the quadrant HL, which comprise municipalities with high incoming trip values, since they are very attractive areas. Around these areas, as well as those surrounding the quadrant HH, there are some municipalities located in the quadrant LH. These municipalities do not show high incoming trips, but their neighbors have intense inflows. This is, however, a relevant fact, as they have the potential to be incorporated into the HH or HL quadrants because of their proximity to areas whose incoming travel values are high.

The analyses of the variable "*Total number of people making outgoing trips from the municipality*" shown in the respective *Box Map* of Fig. 1 basically follows the same rationale described for the incoming trips. We highlight, though, that a visual inspection reveals a larger amount of municipalities classified in the quadrant HH located within the official FURs. In addition, we observe that the municipalities classified in the quadrant HL are less scattered across the territory and there are also less municipalities classified in the quadrant LH.

Finally, comparing the results obtained with the data on commuting trips and the variable "*population density*", we note that there is a high correlation among the municipalities classified in the HH quadrant, although there are less municipalities in this quadrant when using the population density. The analysis with the official FURs also shows reasonable results, as we can see that the municipalities in the HH quadrant are within the official boundaries. There is, however, a larger amount of municipalities classified in the LL quadrant when using the population density criterion. This results in less municipalities classified in the HH quadrant and also in the HL and LH quadrants, revealing a low correlation on these quadrants when comparing both population and commuting trip criteria. By this, we can observe that the variable population density neither highlights regional centers (quadrant HL) nor stresses FURs (quadrants HH and LH) when compared to the results obtained by means of the variables related to commuting trips.

On the other hand, the *Box Map* representing the "*total population"* shows a high correlation between this variable and travel data. More specifically, most municipalities in the HH quadrant regarding the total population are coincident with the municipalities in this quadrant regarding travel data. They are also located within the official FURs. A visual inspection of total population and travel data maps still reveals more coincident municipalities in both HL and LH quadrants, although the total population displays more scattered observations across the territory in such quadrants, especially HL.

Up to this point, we have demonstrated and discussed the application of the univariate ESDA technique for the definition of FURs using population-based criteria and commuting travel separately. However, there is evidence for a joint analysis taking into account such variables (Cheshire and Hay, 1989; Office of Management and Budget, 2010; OECD, 2016), given that employing a multivariate ESDA setting (or, more specifically, a bivariate ESDA as the case studied here) has not been explored yet. Four combinations in total were then investigated, as follows:

- Combination A: total number of people making outgoing trips (z_k) and total population (z_l);
- Combination B: total population (z_k) and total number of people making outgoing trips (z_l);
- Combination C: total number of people making incoming trips (z_k) and total population (z_l);
- Combination D: total population (z_k) and total number of people making incoming trips (z_l);

For each combined pair listed above, the bivariate ESDA technique was applied and a Box Map was obtained illustrating the results. To completely understand the Box Maps, it is important to consider the convention adopted in the legends of the figures. Note that the title of each legend consists of one variable versus a second variable. Therefore, the first variable to appear in the title corresponds to the variable related to the given location (z_k) and the second variable to appear in the title corresponds to the variable associated with the neighboring areas (z_l) of the given location.

An explanation of the interpretations resulting from the classification of municipalities in each of the quadrants is presented below, taking as an example the variables "*Total population*", associated with the given location (z_k), and "*Total number of people making outgoing trips*", associated with the adjacent areas (z_l). For the other combination scenarios, the same rationale can be followed. Therefore, the municipalities classified in quadrant HH have a value of the *total population* higher than the overall state average. Moreover, their neighbors also present a value of the *total number of people making outgoing trips* which is higher than the overall state average. Thus, for municipalities classified in this quadrant, a strong correlation with official FURs is expected due to their large population and the high intensity of outgoing trips from adjacent areas. In contrast, the municipalities classified in quadrant LL have a value of the *total population* which is lower than the overall state average and their neighbors also present a value of the *total number of people making outgoing trips* which is lower than the overall state average. In these cases, it is expected to identify the various municipalities distributed with small territorial dimensions and with a low intensity of commuting, usually located away from official FURs.

Considering the interpretations for quadrant LH, the municipalities have a value of the *total population* which is lower than the global average, while their neighbors present values of the *total number of people making outgoing trips* which is higher than the global average. Thus, the municipalities located in the vicinity of those classified as HH and, consequently, adjacent to the official FURs are generally identified in this quadrant. Municipalities classified as LH in the vicinity of municipalities classified as HL may also be observed. Finally, regarding the interpretations for the municipalities classified in quadrant HL, it is understood that they present a value of the *total population* which is higher than the overall average, whereas its neighbors present a value of the *total number of people making outgoing trips* lower than the global average. These cases occur commonly away from official FURs, but they highlight municipalities that generally exert a regional influence, despite receiving less intense flows from their neighbors.

With the application of the bivariate ESDA considering both the total number of people making outgoing trips and the total population, the Box Maps in Fig. 2 were obtained. Based on these results, it could be observed that the municipalities in quadrant HH along with quadrant LH are located in the official FURs. It is important to point out that in the metropolitan areas that have been consolidated for a longer time, such as São Paulo and Campinas, the municipalities in quadrant HH cover significant portions of these official limits, with a reduced number of municipalities in other quadrants.

Regarding the municipalities belonging to quadrant LL, many cases can be observed across the territory. In general, these municipalities are characterized by a low number of inhabitants and by the existence of adjacent areas with minimal influence and not very significant flows. This is because these municipalities are not attractive areas and, similarly, are far from the attractive areas of employment. However, some of these municipalities can be found inside the official functional urban regions consolidated recently, which may indicate their future growth, due to the closer socioeconomic relations with the influential cities that are developing in these regions.

Another relevant aspect is the evidence of municipalities with regional influence. In the Box Map that represents combination B, a large number of municipalities in the countryside are classified in quadrant HL, which indicates that these municipalities have large *population* values and their adjacent municipalities have a low *number of people making outgoing trips*. On the other hand, combination A has fewer (compared to B) municipalities classified as HL (high *number of people making outgoing trips* and low *total population* at the neighboring

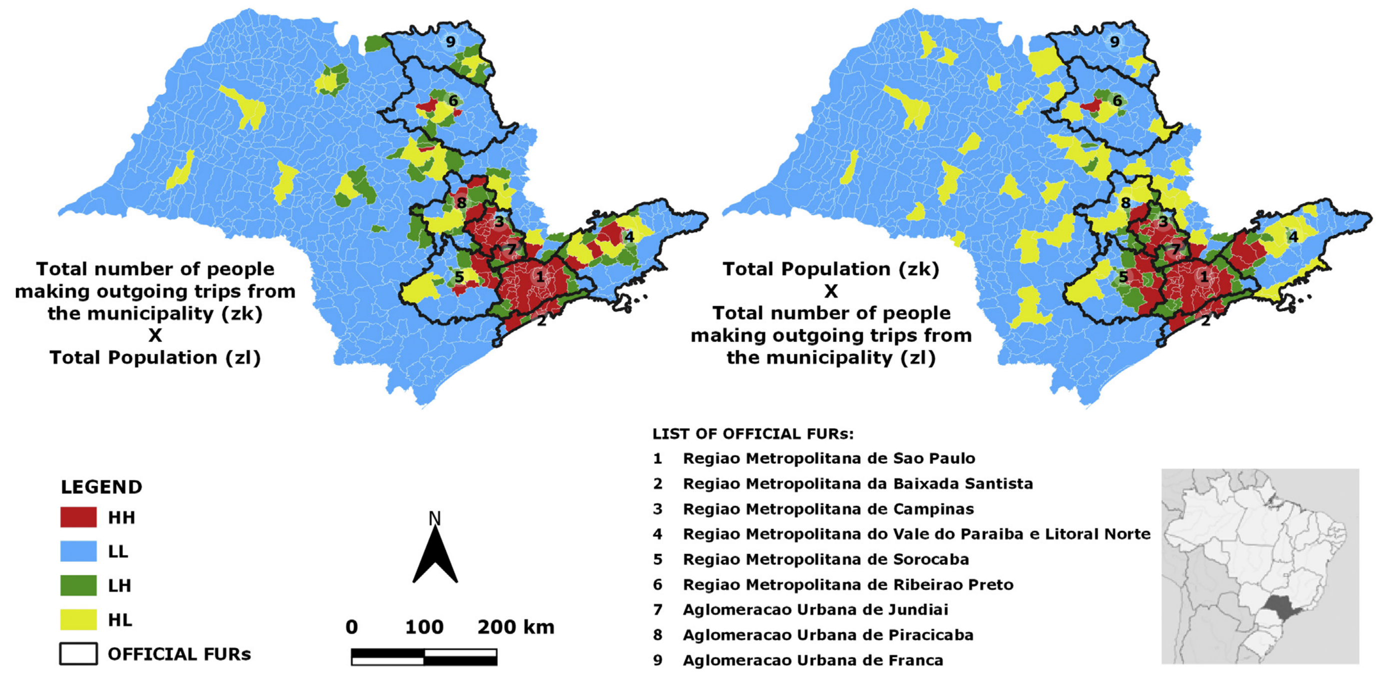

Fig. 2. Box maps representing the total number of people making outgoing trips versus the total population (combination A) and the total population versus the total number of people making outgoing trips (combination B) - Source: Created by the authors.

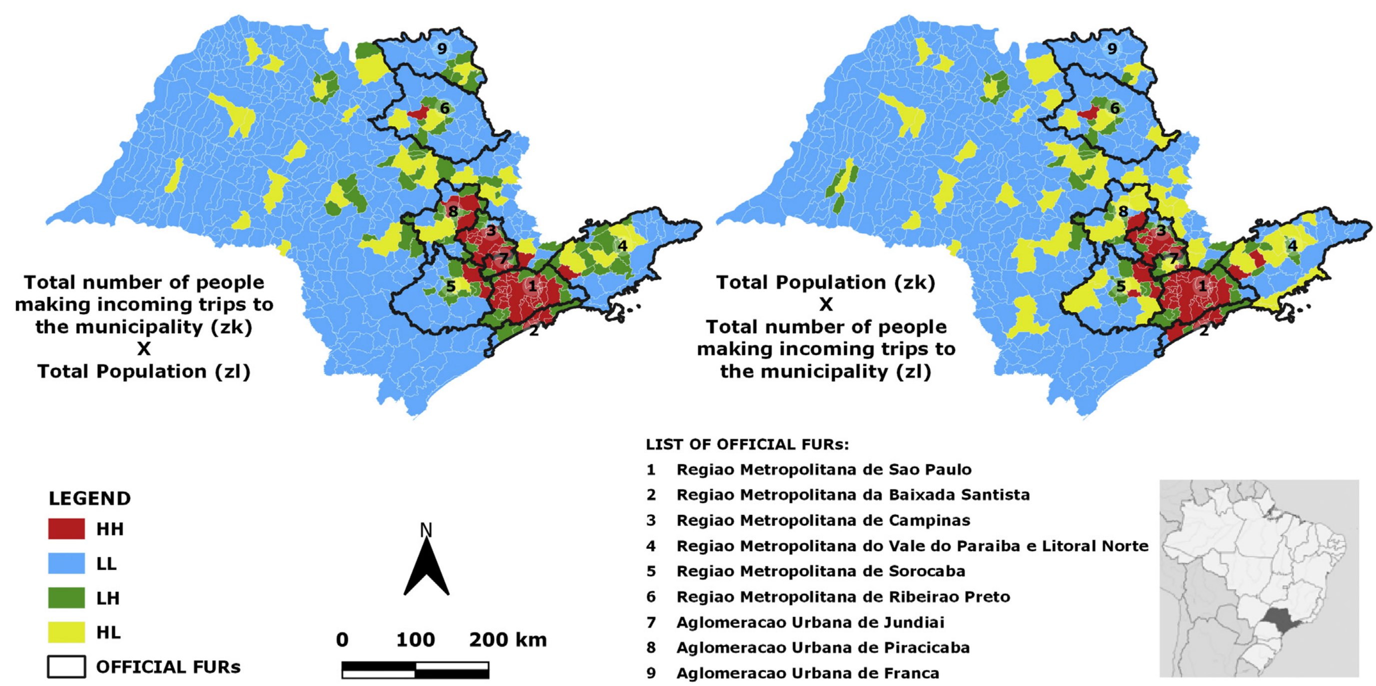

Fig. 3. Box maps representing the total number of people making incoming trips versus the total population (combination C) and the total population versus the total number of people making incoming trips (combination D) - Source: Created by the authors.

municipalities). However, they are less scattered and are mostly found in groups with municipalities classified in quadrant LH. Thus, these results may show some evidence that the use of the travel data as the variable associated with a given location (z_k) and the total population as the variable associated with adjacent areas (z_l) produces more reliable results.

The results presented in Fig. 3, which show the combinations between the *variables total number of people making incoming trips* and *total population* verify similar patterns to what was described previously. That is, municipalities in quadrant HH along with municipalities in quadrant LH are mostly within the official FURs; municipalities in quadrant HL point to areas of regional influence; and, most importantly, the combination of the *total number of people making incoming trips* as the variable at the given location (z_k) and the *total population* as the variable of the adjacent areas (z_l) also shows more satisfactory results than the opposite combination.

4. Conclusions

Based on a bivariate ESDA technique combining commuting travel and population data, in a case study conducted in the state of São Paulo, we investigated whether such an approach demonstrated promising outcomes to the issue of delimitating FURs.

Analyzing the distribution of the municipalities in the various Box Maps generated, we concluded that the combinations in which the commuting trips were associated with the given location (z_k) and population in related adjacent areas (z_l) presented more satisfactory results. For these cases, the attractive HH and HL areas are mostly surrounded by LH areas. Thus, these combinations reproduced the movement of people from peripheral regions to central areas in a better way, which generally correspond to populated municipalities.

Considering the comparison of the FURs outlined in the approach developed here with the official boundaries, we noticed that the metropolitan areas of São Paulo, Campinas and Baixada Santista were well characterized, demonstrating a high degree of coincidence. In the FURs consolidated recently (i.e., all others except São Paulo, Campinas and Baixada Santista), although there are municipalities classified in the four quadrants, we can observe some predominance of municipalities in quadrants HL and LH, which indicates the emergence of important relationships in a regional scenario.

Finally, in addition to the characterization of FURs, the approach also represented relatively well the influence of the municipalities in the interior of the state, highlighted by the observations classified in the HL quadrant. Although not characterized as functional urban regions, these areas display large populations and have local influence. Thus, it is also important to consider them as emerging municipalities, since they may comprise centers of urban agglomerations in the future.

Declaration of Competing Interest

None.

Acknowledgments

The authors thank the Brazilian agencies FAPESP (São Paulo Research Foundation – Grant 2014/26903-0) and CNPq (Brazilian National Council for Scientific and Technological Development – Grants 471397/2013-0; 302177/2015-9; 400617/2016-1) for their support. The authors are also grateful to Caliper Corporation for the educational license grant of the Maptitude software used in this study.

References

Aguiar, L.L., Manzato, G.G., Rodrigues da Silva, A.N., 2017. Patterns of commuting flows for delimitating Functional Urban Regions in the state of São Paulo, Brazil. In: Proceedings of the 15th International Conference on Computers in Urban Planning and Urban Management, Adelaide, Australia.

Ajauskas, R., Manzato, G.G., Rodrigues da Silva, A.N., 2012. The definition of Functional Urban Regions: validation of a set of spatial models with recent census data and analysis of an additional model specification. In: Proceedings of CAMUSS, the International Symposium on Cellular Automata Modeling for Urban Spatial Systems, Porto, Portugal, pp. 91–104.

Anselin, L., 1995. Local indicators of spatial association – LISA. Geogr. Anal. 27, 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Anselin, L., 1998. In: Longley, P.A., Brooks, S.M., McDonnell, R., Macmillian, W. (Eds.), Exploratory Spatial Data Analysis in a Geocomputational Environment. Geocomputation: A Primer. Wiley and Sons, New York, pp. 77–94.

Anselin, L.; Syabri, I; Smirnov, O. (2002) Visualizing multivariate spatial correlation with dynamically linked windows. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi = 10.1.1.118.7163. (Accessed on March 14th, 2018).

Arellano Ríos, A., 2014a. La definición jurídica del fenómeno metropolitano en el ámbito subnacional mexicano. Opinión Jurídica 13 (26), 91–108.

Arellano Ríos, A., 2014b. La coordinación metropolitana en el ámbito subnacional mexicano: un análisis institucional. Documentos y Aportes en Administración Pública y Gestión Estatal (DAAPGE) 14 (23), 33–70.

Arellano Ríos, A., 2015. Metropolitan coordination in Mexico. Curr. Urban Stud. 3, 11–17. https://doi.org/10.4236/cus.2015.31002.

Bereitschaft, B., Debbage, K., 2014. Regional variations in urban fragmentation among U.S. metropolitan and megapolitan areas. Applied Spatial Analysis and Policy 7, 119–147. https://doi.org/10.1007/s12061-013-9092-9.

Boarnet, M.G., Haughwout, A.F., 2000. Do Highways Matter? Evidence and Policy Implications of Highways' Influence on Metropolitan Development. In: The Brookings Institution Center on Urban and Metropolitan Policy (Discussion Paper), . http://www.brookings.edu/research/reports/2000/08/highways-boarnet>.

Bode, E., 2008. Delineating metropolitan areas using land prices. J. Reg. Sci. 48 (1), 131–163. https://doi.org/10.1111/j.1467-9787.2008.00544.x.

Breitung, W., 2011. Borders and the city: intra-urban boundaries in Guangzhou (China). Quaestiones Geographicae 30 (4), 55–67. https://doi.org/10.2478/v10117-011-0038-5.

Castello Branco, M.L.G., Pereira, R.H.M., Nadalin, V.G., 2013. Rediscutindo a Delimitação das Regiões Metropolitanas no Brasil: Um Exercício a Partir dos Critérios da Década de 1970. Texto para discussão TD. Instituto de Pesquisa Econômica Aplicada (IPEA), Rio de Janeiro, pp. 1860.

Cheshire, P.C., Hay, D.G., 1989. Urban Problems in Western Europe: An Economic Analysis. Unwin Hyman, London.

Cohen, B., 2006. Urbanization in developing countries: current trends, future projections, and key challenges for sustainability. Technol. Soc. 28 (1–2), 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005.

Coombes, P.P., Overman, H.G., 2004. In: Henderson, J.V., Thisse, E.J.F. (Eds.), The spatial Distribution of Economic Activities in the European Union. 4. Handbook of Urban and Regional Economics: Cities and Geography, North Holland, Amsterdam, pp. 2845–2909.

Guzik, R., Kołoś, A., Gwosdz, K., 2017. Interconnections in public transport as a method for delimiting urban functional areas and the settlement hierarchy in Poland. Regional Statistics 7 (1), 063–077. https://doi.org/10.15196/RS07104.

Henderson, V., 2002. Urbanization in developing countries. World Bank Res. Obs. 17 (1), 89–112. https://doi.org/10.1093/wbro/17.1.89.

Horňák, M., Kraft, S., 2015. Functional transport regions in Slovakia defined by passenger-car traffic flows. Mitteilungen der Osterreichischen Geographischen Gesellschaft 157 (1), 109–128. https://doi.org/10.1553/moegg157s109.

Huang, X., Lu, Q., Zhang, L., 2014. A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas. ISPRS J. Photogramm. Remote Sens. 90, 36–48. https://doi.org/10.1016/j.isprsjprs.2014.01.008.

IBGE (Instituto Brasileiro de Geografia e Estatística), 2010. Censo 2010 (Available at). http://www.censo2010.ibge.gov.br/.

Kauffmann, A., 2012. Delineation of city regions based on commuting interrelations: the example of large cities in Germany. IWH-Diskussionspapiere 2012 (4) (Available at). http://nbn-resolving.de/urn:nbn:de:101:1-201212176768.

Klapka, P., Halás, M., 2016. Conceptualising patterns of spatial flows: five decades of advances in the definition and use of functional regions. Moravian Geographical Reports 24 (2), 2–11. https://doi.org/10.1515/mgr-2016-0006.

Konjar, M., Lisec, A., Drobne, S., 2010. Methods for delineation of functional regions using data on commuters. In: Proceedings of the 13th AGILE International Conference on Geographic Information Science. Guimarães, Portugal.

Kourtit, K., Nijkamp, P., Partridge, M.D., 2015. Challenges of the new urban world. Applied Spatial Analysis and Policy 8 (3), 199–215. https://doi.org/10.1007/s12061-015-9155-1.

Kraft, S., Marada, M., 2017. Delimitation of functional transport regions: understanding the transport flows patterns at the micro-regional level. Geografiska Annaler: Series B, Human Geography 99 (1), 79–93. https://doi.org/10.1080/04353684.2017.1291741.

Manley, E., 2014. Identifying Functional Urban Regions within traffic flow. Reg. Stud. Reg. Sci. 1 (1), 40–42. https://doi.org/10.1080/21681376.2014.891649.

Manzato, G.G., Rodrigues da Silva, A.N., 2010. Spatial-temporal combination of variables for monitoring changes in metropolitan areas. Appl. Spatial Analysis and Policy 3 (1), 25–44. https://doi.org/10.1007/s12061-009-9028-6.

Manzato, G.G., Baria, I., Rodrigues da Silva, A.N., 2007. A GIS-based comparison of methodologies for the definition of metropolitan areas in a developing country. In: Proceedings of 10th International Conference on Computers in Urban Planning and Urban Management, Iguazu Falls, Brazil.

Manzato, G.G., Dias, R.S., Rodrigues da Silva, A.N., 2015. Capacity of roadway infrastructure and its relation with Functional Urban Regions. In: Proceedings of 14th International Conference on Computers in Urban Planning and Urban Management, Cambridge MA, USA.

Moisés, O.A., Royuela, V., Xavier, V., 2017. Computing Functional Urban Areas Using a Hierarchical Travel Time Approach: An Applied Case to Ecuador. University of Barcelona, Research Institute of Applied Economics. http://diposit.ub.edu/dspace/bitstream/2445/110563/1/IR17-005-Obaco_Computing.pdf.

Moura, R., Carvalho, I., 2012. Estatuto da Metrópole: onde está a região metropolitana? In: Observatório das Metrópoles, Available at. http://observatoriodasmetropoles.net/index.php?option = com_k2&view = item&id = 455%3Aestatuto-da-metr%C3%B3pole-onde-est%C3%A1-a-regi%C3%A3o-metropolitana%3F&Itemid = 165&lang = pt (Acessed April 15, 2014).

Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. ISPRS J. Photogramm. Remote Sens. 87, 152–165. https://doi.org/10.1016/j.isprsjprs.2013.11.001.

Observatory of the Metropolis, 2013. Banco de Dados de Movimento Pendular - Municípios Brasileiros. Available at. http://www.observatoriodasmetropoles.net/index.php?option = com_content&view = article&id = 152&Itemid = 155&lang = pt Accessed: January 14, 2014.

OECD (Organisation for Economic Co-operation and Development), 2012. Redefining "Urban": A New Way to Measure Metropolitan Areas. OECD Publishing, Paris. https://doi.org/10.1787/9789264174108-en.

OECD (Organisation for Economic Co-operation and Development), 2016. OECD Regions at a Glance 2016. OECD Publishing, Paris. https://doi.org/10.1787/reg_glance-2016-en.

Office of Management and Budget, 1998. Alternative approaches to defining metropolitan and non-metropolitan areas. Federal Register 63 (244) (December 21, 1998).

Office of Management and Budget, 2010. 2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas; Notice. Federal Register 75 (123) (June 28, 2010).
Oliveira Junior, M.A., Matiolli, J.A.C., Manzato, G.G., 2017. Incorporating recent census and transportation supply data into cellular automata-based models for the delimitation of Functional Urban Regions. In: Proceedings of 96th Annual Meeting of the Transportation Research Board (TRB). Washington D.C., USA.
Pereira, H.T.S., Rodrigues da Silva, A.N., 2010. Comparing spatial analysis methods for the definition of Functional Urban Regions - the case of Bahia, Brazil. In: Proceedings of 10th International Conference on Design and Decision Support Systems in Architecture and Urban Planning, Eindhoven, The Netherlands.
Ramos, R.A.R., Rodrigues da Silva, A.N., Miranda, V.P., 2004. A comparison of two methods for the definition of regional metropolitan areas through an application in the North of Portugal. In: Proceedings of 44th European Congress of the European Regional Science Association, Porto, Portugal.
Ramos, R.A.R., Rodrigues da Silva, A.N., 2003. A data-driven approach for the definition of metropolitan regions. In: Proceedings of 8th International Conference on Computers in Urban Planning and Urban Management, Sendai, Japan.
Ramos, R.A.R., Rodrigues da Silva, A.N., 2007. A spatial analysis approach for the definition of metropolitan regions - the case of Portugal. Environment and Planning B: Planning and Design 34 (1), 171–185. https://doi.org/10.1068/b31117.
Rodrigues da Silva, A.N., Manzato, G.G., Pereira, H.T.S., 2014. Defining Functional Urban Regions in Bahia, Brazil, using roadway coverage and population density variables. J. Transp. Geogr. 36, 79–88. https://doi.org/10.1016/j.jtrangeo.2014.03.001.
Sahoo, S.N., Pekkat, S., 2014. Determination of urbanization based on imperviousness. Urban Design and Planning 167 (DP2), 49–57. https://doi.org/10.1680/udap.13.00027.
Soares, E., Figueiredo, R., Vala, F., 2017. Defining labour market areas and its relevance from a statistical perspective: the Portuguese case. Stat. J. IAOS 33 (3), 615–625. https://doi.org/10.3233/SJI-170381.
UN (United Nations), 2019. World Urbanization Prospects: The 2018 Revision. Department of Economic and Social Affairs, Population Division (ST/ESA/SER.A/420), New York. https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.
Weber, C., 2001. In: Donnay, J.P., Barnsley, M.J., Longley, E.P.A. (Eds.), Urban Agglomeration Delimitation Using Remote Sensing data. Remote sensing and urban analysis, Taylor & Francis, London, pp. 145–159.
Williams, A.M., Foord, J., Mooney, J., 2012. Human mobility in Functional Urban Regions: understanding the diversity of mobilities. Int. Rev. Sociol. 22 (2), 191–209. https://doi.org/10.1080/03906701.2012.696961.