Computers, Environment and Urban Systems 89 (2021) 101675

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Exploring night and day socio-spatial segregation based on mobile phone data: The case of Medellin (Colombia)

Borja Moya-Gómez [a], Marcin Stępniak [a,*], Juan Carlos García-Palomares [a], Enrique Frías-Martínez [b], Javier Gutiérrez [a]

[a] Complutense University of Madrid, Spain
[b] Telefonica Research, Madrid, Spain

ARTICLE INFO

Keywords:
Social segregation
Mobile phone data
Spatial statistics
Medellin (Colombia)

ABSTRACT

Social segregation research has a long tradition in urban studies. Usually, these studies focus on the residential dimension, using official registries (e.g., census data), which show population distribution at night. Nevertheless, these studies disregard the fact that the population in cities is highly mobile, and its spatial distribution dramatically changes between night and day. The emergence of new data sources (Big Data) creates perfect conditions to consider segregation as a process, by providing the opportunity to dynamically analyse temporal changes in social segregation.

This study uses mobile phone data to analyse changes in social segregation between night and day. Our case study is Medellin (Colombia), a highly socially-segregated, South American city, where social integration policies are being developed, targeting the population in the most disadvantaged neighbourhoods. We use several complementary indicators of social segregation, supplementing them with mobility indicators that help explain changes in spatial segregation between night and day.

The main conclusion is that daily mobility reduces the concentration of a particular group within neighbourhoods and increases the degree of social mixing (exposure) in local settings. This greater social exposure softens local contrasts (outliers) and increases the extension of spatial clusters (positive spatial autocorrelation), so general clustering trends emerge more clearly. The study also makes clear that increased exposure during the day mainly occurs due to the mobility of the low-income population, who are the most likely to leave their neighbourhood during the day and who travel the greatest distances to the most diverse set of destinations.

1. Introduction

Socio-spatial segregation research is well grounded in urban studies, focusing mostly on residential segregation. This is because the main source used (the census) provides data on the characteristics of the population based on its place of residence. However, urban inhabitants do not spend their entire time in their place of residence, but conduct activities across different parts of a city. In consequence, their mobility modifies the city's socio-spatial segregation patterns throughout the day. Moreover, spending time outside the residential neighbourhood may play a significant role in determining segregation experiences. Accordingly, individuals residing in the same area may be exposed to very different social and economic contexts, depending on the nature of their daily mobility and the type of residential area.

Mobile-phone data offers new opportunities for segregation studies, thanks to their high temporal resolution. As people carry out their activities in a city, they leave behind a digital fingerprint that can be used to analyse the population's segregation patterns at different times of the day (Östh, Shuttleworth, & Niedomysl, 2018). Mobile-phone data offer less complete socioeconomic data than population censuses. However, they bear some noteworthy advantages: samples are usually very large and are already being stored by mobile network operators, rendering the cost of data collection much lower than that of a survey with an equivalent sample size. Besides, these data present fine high spatial and temporal resolution, and enable accurate and dynamic analysis (García-Albertos, Picornell, Salas-Olmedo, & Gutiérrez, 2018).

Many authors argue that segregation is not only the result of the conditions of the home neighbourhood, but also of daily mobility of

* Corresponding author at: tGIS, Department of Geography, Complutense University of Madrid, C/ Profesor Aranguren, s/n, Ciudad Universitaria, 28040 Madrid, Spain.

E-mail addresses: marcinstepniak@ucm.es (M. Stępniak), javiergutierrez@ghis.ucm.es (J. Gutiérrez).

https://doi.org/10.1016/j.compenvurbsys.2021.101675
Received 28 January 2021; Received in revised form 12 April 2021; Accepted 6 June 2021
Available online 14 June 2021
0198-9715/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

individuals (Heringa, Bolt, Dijst, & van Kempen, 2014; Kwan & Weber, 2003; Östh et al., 2018; Wong & Shaw, 2011). The underlying theoretical basis for research of segregation outside the residential context is that one's exposure during their daily activities might have an impact on their personal experiences of segregation (Östh et al., 2018). It is assumed that the probability of cross-group interactions and integration depend on the presence of people who belong to different groups in one's activity space (Zhou, Chen, Yeh, & Yue, 2019). As Östh et al. (2018) state, it needs to be emphasised that we do not suggest that increased levels of spatial proximity would necessarily reduce or abolish the negative and discriminatory practices commonly associated with residential segregation. However, it is well known that segregation in its the most extreme state ultimately leads to the complete spatial separation of individuals and groups. This, in turn bounds integration of different groups, in at least some areas of daily life, and limits opportunity to share the same spaces (e.g. Dixon, Durrheim, & Tredoux, 2005). This statement is rooted in a multitude of studies calling for the dispersal of members of segregated groups in space as a mean to enhance social integration (Musterd and Ostendorf, 2013).

Further, we identify a research gap on changes in socio-spatial segregation patterns between night and day in cities of the Global South, even though these cities are usually assessed as highly segregated. One of the main reasons is related to the limited access to necessary data, which can be at least partly addressed by the application of mobile phone data. Nevertheless, the research on temporal socio-spatial segregation based on mobile phone data has focused on European cities (see, for example, Silm & Ahas, 2014a, 2014b; Silm, Ahas, & Mooses, 2018; Östh et al., 2018), with an exception of Santiago de Chile study by Dannemann, Sotomayor-Gómez, and Samaniego (2018). At the same time, mobile phone data are particularly useful to generate new knowledge about cities of the Global South, partly because they cover some of the gaps in official statistics and partly, because they allow to "enter" to neighbourhoods where interviewers have limited access, e.g. due to high criminality. Our study area, the city of Medellin (Colombia), is of special interest for this kind of research, since it is a highly segregated city, suffering severe problems of social exclusion and insecurity.

Our main hypothesis is that urban daily mobility mitigates socio-spatial segregation. The main contribution of the presented paper is that we test our hypothesis in the understudied context of the Global South, applying the multidimensional approach to segregation studies which combines a broad range of complementary segregation and mobility indicators, and showing that even very basic mobile phone data can extend our understanding of urban segregation. First, we evaluate an income urban segregation covering its three complementary dimensions, namely concentration, clustering and exposure. Moreover, we introduce a distance-decay Shannon Index, a modification of the Shannon Index, extending measurement of social exposure (social mixing) beyond a particular neighbourhood in order to consider broader local settings. These local settings more realistically reflect potential for social interaction between different groups, independently of socially meaningless "administrative borders" between neighbourhoods (or census tracks). Further, we implement mobility indicators and the concept of socio-spatial-mix trips into segregation studies, using them to explain changes of the social mix between night and day. Finally, we show how to take advantage of very basic Call Detail Record (CDR) data in order to provide insightful conclusions on urban socio-spatial segregation. In consequence, the paper contributes to fill the research gap in knowledge on night and day socio-spatial segregation in cities in less developed countries exploring the potential of mobile-phone data for this purpose and providing new knowledge for future comparative analysis.

The remainder of this paper is structured as follows. Section 2 presents a literature review on socio-spatial segregation. Section 3 describes the study area. Sections 4 and 5 report the data and the methodology, respectively. Section 6 describes results, and finally Section 7 presents the conclusions.

2. Socio-spatial segregation

The roots of research on residential segregation are based on studies that focus on differences in housing and living conditions experienced by inhabitants from different ethnic groups in urban America (Massey & Denton, 1993), however, ethnic segregation has been studied extensively all over the world (Johnston, Poulsen, & Forrest, 2007; Musterd & Van Kempen, 2009; Sabatini, 2006). Currently, the main line of residential segregation studies does not limit themselves to ethnic issues and deals with broader phenomena, such as socioeconomic segregation (Musterd & Ostendorf, 2013; Tammaru, 2015), including examples from the Global South (e.g. Molinatti, 2015; Monkkonen, 2012).

The most important reasons that underlie residential segregation are related to voluntary or involuntary housing choices. Better-off households express their economic power through their housing preferences, while more vulnerable ones have limited options, as their housing decisions are constrained by a lack of resources. In consequence, even though the former demonstrate freedom of choice, while the latter express limitations, both processes equally lead to spatial segregation between particular socioeconomic strata (Maloutas, 2016). In the background behind these processes, one can observe gentrification or suburbanisation, which are the two sides of the same phenomenon: colonisation of preferred urban areas by high-income groups, and the pushing of low-income groups toward urban margins (Musterd, Marcińczak, van Ham, & Tammaru, 2017).

Thus, the most important reason for residential segregation is income-related social inequality (Musterd & Ostendorf, 2013; Tammaru, Marcińczak, Aunap, & van Ham, 2017). Several studies show that the growth of income disparities is associated with an increase in socioeconomic segregation in both, the US (e.g. Reardon & Bischoff, 2011) and Europe (Tammaru, 2015). Nevertheless, the recent comparative study shows that this relationship is not straightforward, underlying differences between paths followed by individual cities (Tammaru et al., 2017). This path dependency proves the need for case-specific segregation studies, in particular in different socioeconomic contexts. This is even more important in case of cities in less developed countries, as they experience the most radical growth in socioeconomic inequalities (Wissink, Schwanen, & van Kempen, 2016).

Segregation studies became major sources of information for the formulation of social integration policies based on the assumption of direct associations between the spatial and the social dimensions of segregation (Schnell, Diab, & Benenson, 2015). In the classical urban geography, spatial distance, as expressed by residential segregation, is associated with social distance (Musterd et al., 2017). Importantly, the coexistence of both distances causes that consequences of residential segregation go far beyond housing conditions or different levels of access to particular types of public services. The spatial separation of particular groups provokes social division and decreases trust and increases tensions between them (Östh, Andersson, & Malmberg, 2013).

Several studies link the social stratum of a household's residential area with the spatial range of their daily behaviour (cf. Kamruzzaman, Baker, Washington, & Turrell, 2014). Despite increasing interest in spatial separation between individuals during their non-employment activities (Silm & Ahas, 2014a), time spent at home and at work locations constitutes a significant fraction of an individual's time, and as such, constrains social interactions. Thus, there is a growing need to understand the degree of workplace segregation, as this can exaggerate the social separation of particular groups or, on the contrary, increase social exposure to other social groups. For example, for a better understanding of immigrant integration into the local labour market, Ellis, Wright, and Parks (2009) show the need to extend the meaning of segregation beyond its residential manifestation in order to include segregation at the workplace into the picture.

Previous studies suggest an interrelation between residential and workplace segregation, by showing that living in less-segregated neighbourhoods reduces segregation at the workplace level

(Strömgren et al., 2014). The segregation study shows that that residential and workplace locations are similarly segregated (Toomet, Silm, Saluveer, & Tammaru, 2012); however, the study at hand is based on a particular (European) context. On the other hand, Ellis et al. (2009) show that, in the US context, residential segregation is an order of magnitude more considerable than segregation at the work location, providing one more argument that supports the well-grounded statement that segregation is a highly context-specific issue (Tammaru, 2015; van Ham & Tammaru, 2016). Thus, Silm and Ahas (2014b) argue that a single-domain segregation study is likely to misjudge individuals' social interaction, as these interactions may occur at different locations, during different activities. In consequence, the contemporary critique of the "classical" segregation indices call for a multidimensional approach to socio-spatial segregation. It underlines interdependency between residential location, social networks, activity schedules, and other social and cultural aspects of daily life. In the modern, highly mobile urban life, socio-spatial interactions become weakly, if ever determined by the individual's residential location (Schnell et al., 2015).

Census data perform reasonably well when one must capture the residential dimension of segregation, but they hardly contain information that covers other segregation domains. However, several new data sources may contribute to filling this gap. A tailor-made mobile phone application may collect very detailed spatial and temporal data on activity spaces of individuals, supplemented by socioeconomic data on individuals (Yip, Forrest, & Xian, 2016). Nevertheless, even though it would automate data collection, it shares similar constraints with more "traditional" travel diaries (see Li & Wang, 2017) related to an extreme effort in contacting participants and a high data-collection cost.

Mobile phone data offers an excellent opportunity to analyse socio-spatial segregation to different social groups in their activity places beyond their place of residence, since it allows us to know the location of people over time. It is collected passively and continuously, generating very large samples, much larger than those obtained in household mobility surveys. Ironically most of the socio-spatial segregation studies based on mobile phone data have been performed in cities with relatively detailed and regularly collected census data, while they have been hardly applied in highly segregated cities of the Global South, where the access to data is limited. Pioneering research conducted in Estonia uses mobile phone data to define activity spaces and their interrelation with segregation during non-employment activities (Silm & Ahas, 2014a), the temporal fluctuation of segregation levels (Silm & Ahas, 2014b) and ethnic segregation among the young population (Silm et al., 2018). The main finding of the Estonian studies is that daily mobility decreases residential segregation, however not uniformly, with significant differences between social groups and spaces. Similarly, Östh et al. (2018) use a large dataset of more than 1.2 million Swedish mobile phone users and conclude that even though mobility alleviates segregation for some individuals, many areas remain highly segregated even, when daily mobility is taken into account. Zhou et al. (2019) attempted to elucidate workplace segregation using cellular network data with a special focus on the segregation of rural migrants. They show that rural migrants who worked in manufacturing industries and lived in suburban areas suffered from higher workplace segregation from other social groups compared with those who worked in service jobs and resided in the central-city areas. Beyond daily variation of segregation, other authors focus their attention on segregation during holidays. For example, the Estonian study applies mobile phone positioning data to show that during holidays, ethnic segregation is significantly higher than during regular, labour days, in particular outside the capital city (Mooses, Silm, & Ahas, 2016).

The main obstacle of the application of mobile phone data for segregation studies, is related to the (un) availability of social characteristics of users. These data can be obtained from phone operators or estimated through data enrichment processes using auxiliary data sources (Kwan, 2016; Park & Kwan, 2018). Mobile network operators store useful information about their clients, for example age, gender, address and others. In the case of the aforementioned Estonian studies, data on call locations were supplemented by the information obtained from phone operators, which distinguished the language selected by the user (used to identify ethnicity), age and gender. Östh et al. (2018) crossed mobile phone data with information on urban socio-spatial structures. Similarly, Zhou et al. (2019) used the same approach in order to identify migrant enclaves and measure their exposure to other social groups in their workplaces. However, this information is usually not available for researchers due to privacy concerns. Moreover, many mobile phone contracts are linked to companies rather than to individuals. In addition, some customers do not update their personal data. In case the socioeconomic data is not available or not reliable, it has to be estimated. The application of GIS (Geographic Information System) enables cross mobile phone data with another sources, like e.g. national census. In this case, the socioeconomic characteristics of users is identified with the dominant characteristic of residents of the area where they live, however it is not possible to infer the individual characteristic of each user.

3. Study area

Medellin is the second-most populated city in Colombia and the capital of the Department of Antioquia, with approximately 2.5 million inhabitants. Medellín's population is relatively young: 24.9% is under the age of 15, 68.5% between the age of 15 and 65, and only 6.7% over the age of 65. The official unemployment rate is 10.2%. The use of mobile phone services is very widespread among the adult population. The subscriber ratio of the country is 109.5 per inhabitant. Most of them (79%) are prepaid subscribers (Ministerio de Tecnologías de la Información y las Comunicaciones, 2014). Medellin is especially appropriate for conducting a socio-spatial segregation study, since this is a city with great social contrasts and a very violent recent past. However, important political actions have been recently conducted, aiming to promote social integration in the most disadvantaged neighbourhoods.

Medellin is spread throughout the valley that runs north-south (Fig. 1a). The main areas of economic activity, together with the main transport infrastructures are located at the bottom of the valley. The most disadvantaged residential areas (stratum 1 in the range 1–6; Table 1) are found in the most elevated, peripheral zones, especially in the northern part of the municipality (Fig. 1b). Middle income strata neighbourhoods predominate in the centre (strata 3 and 4). The highest income group (stratum 6) concentrates in the southeast where residential buildings alternate with office spaces and shopping centres.

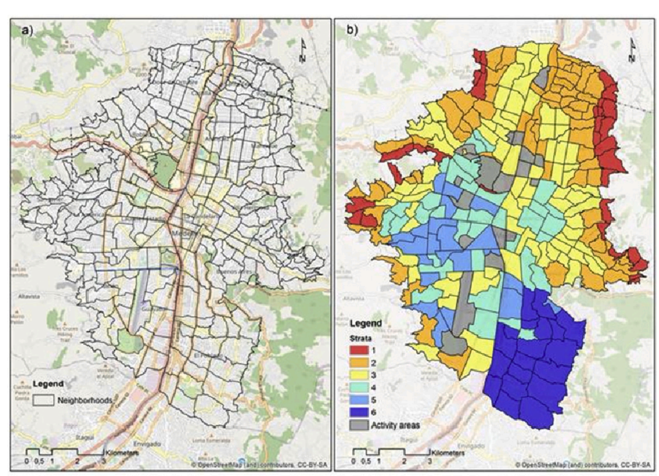

Fig. 1. Reference map (a) and predominant income stratum in each neighbourhood (b).
Source: Medellin Town Hall.

Table 1
Neighbourhoods of Medellin according to predominant income stratum (2014).

Predominant income stratum	Number of neighbourhoods	Households 2014 (thousands)	Percentage	Population 2014 (thousands)	Percentage
Activity areas (*)	19	0	0	0	0
1 (very low)	32	80.0	12.4	280.1	13.0
2	66	211.3	32.9	741.6	34.5
3	76	188.0	29.3	637.4	29.7
4	35	76.1	11.9	229.1	10.7
5	22	56.4	8.8	167.1	7.8
6 (very high)	19	30.3	4.7	93.2	4.3
Total	269	641.8	100.0	2148.4	100.0

[*] Non-residential areas occupied by offices, universities, hospitals, etc.

During the 1980s, Medellin emerged as the epicentre of Colombia's burgeoning illegal drug trade, and in the 90s it became the world's most violent city, as many areas of the city were effectively ungovernable. The settlements on the hillsides of the city were severely impoverished and abandoned. Both public and private investment in these neighbourhoods were progressively phased out as insecurity increased. Public funding and economic activity were concentrated in the socially peaceful and orderly areas of the city (Brand & Davila, 2011).

The 2001–2003 Medellin Development Plan (Plan de Desarrollo 2001-2003; Pérez Gutiérrez, 2001) marked a turning point for the city. It implemented an urban-regeneration model of socially responsible investments in infrastructure, which sought to repay the city's historical debt to the most impoverished sectors in the periphery (Dávila, 2009). Metrocable, a gondola lift system, was central to this social integration strategy, since, firstly, it gives the population who live in the outlying areas convenient access to the city centre, thereby also increasing their employment opportunities, and, secondly, it facilitates the involvement of public institutions in these areas. Thus, over the last decade, the city districts connected by Metrocable have benefited from a significant increase in public funding for the provision and improvement of infrastructure and public services, such as libraries, schools, parks, bank branches, etc. (Brand & Davila, 2011).

The metro system, operating since 1995, consists of two elevated rail lines and three cable lines. In terms of organisation and fares, these cable lines are treated equally as other metro lines. More recently, two Bus Rapid Transit (BRT) corridors were opened, the so-called Metroplus (Heinrichs & Bernet, 2014). Nevertheless, Metrocable has been a key factor in the social integration of the population in deprived neighbourhoods in the periphery. Even though Metrocable is not, on its own, enough to combat crime, it has supported the creation of a climate of peaceful co-existence in which there has been a significant decline in crime rates, as it has been accompanied by a complete package of measures to promote urban development and social integration. The social urbanism projects developed in Medellin in relation to Metrocable have proved effective as a crime-prevention tool (Bea, 2016).

4. Data sources and data pre-processing

This study combines different data sources, both traditional ones and aggregated, anonymised mobile phone data. The latter serves to characterize population density during day and night, and model mobility between particular neighbourhoods. Additionally, we apply data provided by Medellin Town Hall, which contains the distribution of population, structured by income strata at neighbourhood level, as well as transport infrastructure and land use.

4.1. Mobile phone data

Mobile phone networks are built using a set of phone towers (Base Transceiver Stations; BTS) that connect mobile phones to the network. Each BTS covers an area called a cell. Depending on population density, BTS coverage typically ranges from less than 1 km^2 in dense urban areas, to more than 4 km^2, in rural ones. For the sake of simplicity, it is commonly assumed that each BTS covers a 2-dimensional non-overlapping polygon. Hence, we use the Voronoi tessellation generated from known BTS points to define each cell.

We use a 5-month (December 1, 2013, through April 30, 2014), anonymized and aggregated Call Detail Record (CDR) dataset from one of the top three biggest mobile phone companies in Colombia (Ministerio de Tecnologías de la Información y las Comunicaciones, 2014). The whole dataset consists of around 2.2 billion records, stored in a MongoDB database. CDRs are generated when a phone calls or uses basic messaging service (like SMS, MMS, among others).

In the dataset, every CDR is limited to three fields: an ad-hoc user id anonymized for this study (every mobile phone number has a unique id in the whole dataset), the time and date when the interaction started, and the BTS's id that the user was connected to at the beginning of the interaction. Note that no information about the exact position of a user within the BTS's cell is known. Also, if the phone is not actively using mobile network services, there would be no CDR record. No other data from the provider, such as a contract or demographic data, are available for this study. In summary, more than 358,000 users interacted with any BTS of Medellin. These users created over 112,000,000 CDRs in Colombia, out of which roughly 35,000,000 in Medellin.

The pre-processing data task consists of cleaning (deleting inconsistent records, such as unknown BTS's coordinates) and filtering of users with at least one call in Medellin during the analysed time window. Then, a home detection algorithm is applied. Home detection algorithms are a specific kind of a wider group of algorithms used to identify meaningful places (like home, work, or others) from mobility information, applying some criteria to define time slots for particular types of activities. Ahas, Silm, Järv, Saluveer, and Tiru (2010) use an anchor-point model to identify home and work and validated it with the actual geography of the population. Frias-Martinez, Virseda, Rubio, and Frias-Martinez (2010) propose a genetic algorithm approach to identify the time slot that had to be used for more precise identification of home and work locations. Kang, Welbourne, Stewart, and Borriello (2005) and Isaacman et al. (2011) use a similar approach consisting of clustering towers (or active points) to identify home and work, but while Kang et al. does it for WiFi signals, Isaacman et al. apply it to mobile phone records.

Considering the previous approaches, we implement a simplified version of Isaacman et al. (2011) approach. By considering the activity of a user during night-time hours (18:00 to 06:00), we identify the place of residence as the area (Voronoi polygon of a BTS) where the user has the highest number (mode) of weeks with at least one CDR, considering only those users with at least 5 weeks with CDRs, in order to eliminate those who, represent sporadic overnight stays. In case of results with more than one weekly-mode, the cell with the highest number of days of activity is selected as the residence, and if still tied, the cell with the highest number of CDRs is selected. If a user's data does not fulfil these conditions, their records were discarded from a dataset.

The same procedure is followed to determine the most likely work/study area, using 09:00 to 12:00, and 14:00 to 16:00 time windows. The

12:00 to 14:00 time window is not considered to avoid activities other than work or study usually undertaken during this period. Then, we create a matrix of trips between home and work or study areas for all users whose area of residence and work are located within the city of Medellin.

Due to the criteria described above, in the analysis we use data for the third part of users with any CDR located within Medellin. In case of the others the home and work locations are identified elsewhere in Colombia or no home / work location is possible to identify when the five-different-weeks-rule is to be applied. Eventually, we identified approximately 80,000 users with home and 87,500 with work locations in Medellin, out of which 68,000 were identified with both, home and work location in Medellin.

Once home and work locations are identified within BTS Voronoi polygons, they are transferred to the layer of city's neighbourhoods, applying the areal weighting method (Chakraborty & Armstrong, 1997; García-Palomares, Gutiérrez, & Cardozo, 2013; Tu et al., 2020) and the dasymetric technique (using the city neighbourhood layer and a land-use map as ancillary maps). Finally, expansion factors are used to estimate home and work location as well home-to-work mobility patterns of the entire population of Medellin using data for users with both locations identified.

4.2. Socio-economic data

We apply data provided by Medellin Town Hall, which contains estimated neighbourhood population in 2010 and 2013 (DANE and Alcaldía de Medellín, n.d.) and housing quality strata (in the scale from 1 to 6) at neighbourhood level in 2010 (Alcaldía de Medellín, n.d.). In our study, in the scope of missing any other data which describes socio-economic status of residents, we use the latter as an approximation of the socio-economic status. Moreover, we assume the latter is constant between 2010 and 2013.

Then, we use a commuting matrix between home and work/study location in order to redistribute population by stratum to destination neighbourhoods during daytime, in proportion to the total number of residents from each stratum by origin neighbourhood (night-time hours). In result, we obtained a population distribution in each neighbourhood based on strata both during night and daytime hours, which is essential to conduct a temporal segregation analysis of the city.

5. Segregation indicators

We apply a set of indicators which provides complementary information on different dimensions of the segregation: concentration, clustering and exposure. We use a Geographic Information System (ArcGIS 10.6) in order to intersect official information on the neighbourhoods (population and income strata) with information obtained from the analysis of mobile phone data (population presence), always differentiating between night and day to analyse the city's dynamic.

Additionally, three mobility-based indicators are calculated to measure a degree of openness: the intensity of the openness, the variety of destinations and the distance travelled. These indicators aim to capture the response of the population from the most disadvantaged neighbourhoods to employment and service deficits in these neighbourhoods, analysing the extent to which the population takes advantage of opportunities offered to them by the city as a whole.

5.1. Concentration

The concentration of particular socioeconomic groups (distinguished by income strata) at day and night is measured with the location quotient (LQ). It is a widely-used indicator in residential segregation studies (see Bauder & Sharpe, 2002), calculated as follows:

$$LQ_{si} = \frac{x_{si}}{X_i} \cdot \frac{Y_s}{Y}$$

where LQ_{si} is the location quotient of the social group s in the neighbourhood i, x_{si} is the number of individuals of the group s in the neighbourhood i, X_i is the total number of individuals in the neighbourhood i, Y_s is the total number of individuals of the social group s in the city, and Y is the total number of individuals in the city.

The indicator compares the share of population of one group in a given neighbourhood with the share of that group within the city as a whole. LQ higher than 1 indicates that the group is over-represented in a given neighbourhood, while LQ below 1 - under-represented. The more extreme values are reported per group (very high in certain neighbourhoods and very low in others), the more concentrated, and therefore the more segregated, that group is.

To obtain a global degree of spatial concentration for each group, its corresponding coefficient of variation is calculated. Higher disparities, i.e. very high values in some neighbourhoods and very low in others, result in high values of coefficient of variation indicating a high degree of concentration of that group in the space. On the contrary, an even share (i.e. close to the average) of a particular social group among neighbourhoods means a low level of segregation.

5.2. Clustering

The LQ and its coefficient of variation analyses concentration levels in each of the neighbourhoods separately, disregarding whether over or under-representation of a particular group is spatially clustered. The Global Moran's Index evaluates spatial autocorrelation indicating whether the pattern is clustered, dispersed, or random. Additionally, the z-score and *p*-value evaluate the significance of the results (Mitchel, 2005). A positive and statistically significant z-score indicates positive spatial autocorrelation (i.e. that high values and/or low values are spatially clustered), while a negative and statistically significant z-score value means that high and low values are spatially dispersed (negative spatial autocorrelation). Finally, a non-significant p-value indicates that social groups are randomly distributed among neighbourhoods. When calculating the index, we consider the spatial interaction between observations within a 1000 m radius (an equivalent of a typical 15-min walk), with a weight inversely proportional to the distance. This distance limit aims to reproduce the local setting in each neighbourhood, meaning the space where inhabitants experience their “neighbourhood life”.

Additionally, Anselin's Local Moran's Index (LISA statistic) is used to identify spatial concentrations of high or low values and spatial outliers (Anselin, 2010), providing z-scores and *p*-values. The cluster/outlier type distinguishes between a statistically significant cluster of high values (HH), a cluster of low values (LL), an outlier in which a high value is surrounded primarily by low values (HL), and outlier in which a low value is surrounded primarily by high values (LH; Mitchel, 2005). This analysis is performed for each social group with a cut-off distance of 1000 m for inverse distance procedure.

Location quotient and LISA statistics measure different dimensions of segregation. The first one measures concentration of particular socio-economic groups in each of the neighbourhoods separately, without considering neighbouring areas, while the second one is focused on spatial relations between neighbourhoods (see Fig. 2). Thus, a spatial distribution can be concentrated and clustered (Fig. 2d) or concentrated and random (Fig. 2b) at the same time.

5.3. Exposure

Exposure refers to the likelihood of potential interaction between individuals from different strata groups. This likelihood is greater in spaces with higher diversity of groups or strata. In order to evaluate the

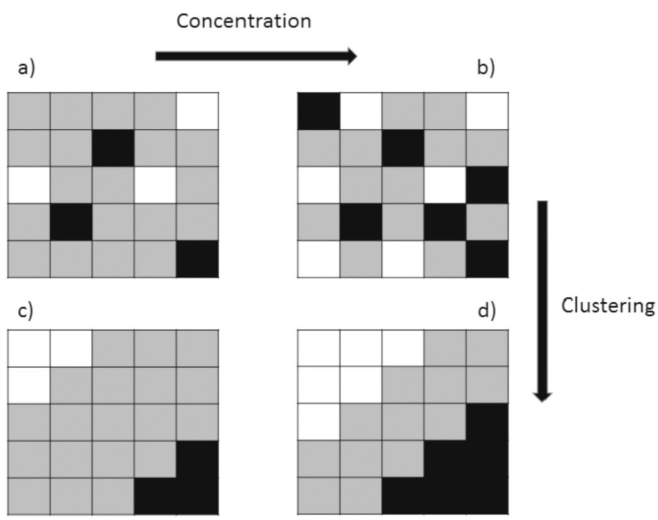

Fig. 2. Concentration and clustering as different dimensions of socio-spatial segregation. The colours of the cells represent the values of the analysed variable: white (low), grey (medium) and black (high). The direction of arrows expresses increasing of concentration and clustering, respectively.

degree of exposure in each neighbourhood, during night-time and day-time hours, we use the Shannon Index (Shannon, 1948). This index is commonly used to analyse diversity in ecological studies (cf. Nagendra, 2002). Within the field of social studies, it has rarely been used, with a few exceptions (Borruso, 2009; Sarma & Das, 2015; Sheskin & Hartman, 2019; Warf & Vincent, 2007). The Shannon Index is calculated according to:

$$H_i = -\sum_{s=1}^{S} p_{si}\cdot log(p_{si})$$

where:H_i is the Shannon Index for the neighbourhood i,p_{si} is the proportion of individuals belonging to the selected social group s in neighbourhood i,and S is the set of social groups.

The index ranges from 0 to infinity, where 0 represents a perfectly homogeneous population, while higher values indicate higher heterogeneity.

The Shannon Index is calculated for particular neighbourhoods, but in reality borders of neighbourhoods are not impregnable for population. On the contrary, "neighbourhood life" spreads into nearby neighbourhoods. Thus, we extend the application of the Shannon Index beyond a particular neighbourhood in order to reflect real activity spaces of population. To do so, we calculate the Shannon Index considering a radius of 1000 m instead of limiting it to administrative units (neighbourhoods). Additionally, we apply distance decay, putting greater weight on the population from nearer units (with which social interaction is more likely) and lower on the farther units (with which interaction is less likely). The proposed approach has the additional advantage of mitigating MAUP (Modifiable Areal Unit Problem; Openshaw & Taylor, 1981), as results are not dependent on the size and shape of applied areal units.

Thus, we calculate the Distance-decay Shannon Index (DDH), separately at night and at day, and then aggregate the results by income strata. Formally:

$$DDH_i = -\sum_{s=1}^{S} \widetilde{p}_{si}\cdot log\left(\widetilde{p}_{si}\right)$$

$$\widetilde{p}_{si} = \sum_{j}^{N} p_{sj}\cdot f\left(d_{ij}\right)$$

$$f\left(d_{ij}\right) = \left\{1-\frac{d_{ij}}{1000}\ if\ d_{ij}\in\left[0,1000\right)\ 0\ if\ d_{ij}\geq 1000\right.$$

where:DDH_i is the distance-decay Shannon Index for the neighbourhood i,$\widetilde{p}_{si}$ is the weighted proportion of individuals belonging to the selected social group s in neighbourhood i and its surroundings.p_{sj} is the share of individuals belonging to the selected social group s in the neighbourhood j.$d_{ij}$is the distance between neighbourhood i and neighbourhood j,$f(d_{ij})$ is the weighting function with a 1000 m threshold.N is the set of neighbourhoods, and S is the set of social groups.

5.4. *Mobility indicators: socio-spatial-mix trips*

Unlike previous indicators, which reflect a static situation by nature, indicators that consider segregation as a process should reflect daily mobility of population. The exposure of a population from one income stratum to another stratum (and, therefore, the likelihood of interaction between members of different strata) largely depends on daily mobility patterns.

A mobility-based indicator is calculated using the trips matrix derived from mobile phone data (see Section 4). It shows the percentage of trips with origin and destination in neighbourhoods of different strata (for example, from a neighbourhood of stratum 1 to stratum 4). The trips can be considered as "socio-spatial-mix trips", since they connect neighbourhoods of different socio-economic characteristics. Therefore, the indicator expresses a level of increase of social mix due to daily mobility of individuals.

Additionally, we calculate an average distance (as a straight line) covered by "socio-spatial-mix trips". We aggregate results by neighbourhoods and income strata and use them as a proxy of a cost that has to be covered by residents of disadvantaged neighbourhoods in order to offset limited employment opportunities and service provision in their neighbourhood.

6. Results

The population distribution during night and day at neighbourhood level looks very different (Fig. 3). During the night (Fig. 3a), the population is concentrated in peripheral neighbourhoods, especially in the north of the municipality, while central areas are almost empty. On the other hand, during a day (Fig. 3b), the population leaves peripheral neighbourhoods and concentrates in central areas in the bottom of the valley and the southern part of the city. The map of differences between day and night (Fig. 3c) identifies the net trip generation and attraction areas, the former located in the periphery and the latter in the centre and in the south.

6.1. *Concentration*

The left column of Fig. 4 presents concentration of particular income strata at night showing two overlapping patterns. The income strata tend to spread over the space in successive periphery-centre aureoles, from the most elevated hillside areas (where the most disadvantaged strata are located) to central areas in the bottom of the valley (where middle and upper classes live). Simultaneously, a north-south gradient is visible, with an increase of presence of income strata from north to south. The stratum with the lowest income level (stratum 1) is concentrated on the northwest and northeast edges of the municipality. Stratum 2 is predominant in a second ring toward the centre, while stratum 3 is located in areas near the bottom of the valley in the northern half, and a few south-western neighbourhoods. Population from stratum 4 is predominant in centre neighbourhoods. Finally, the wealthiest neighbourhoods (stratum 6) are concentrated in the southeast of the municipality, while stratum 5 is located both in the centre-west and in the south-east.

The right column of Fig. 4 shows changes in the location quotient between night and day, according to the dominant income strata. In general, we observe how the population relocates to nearest

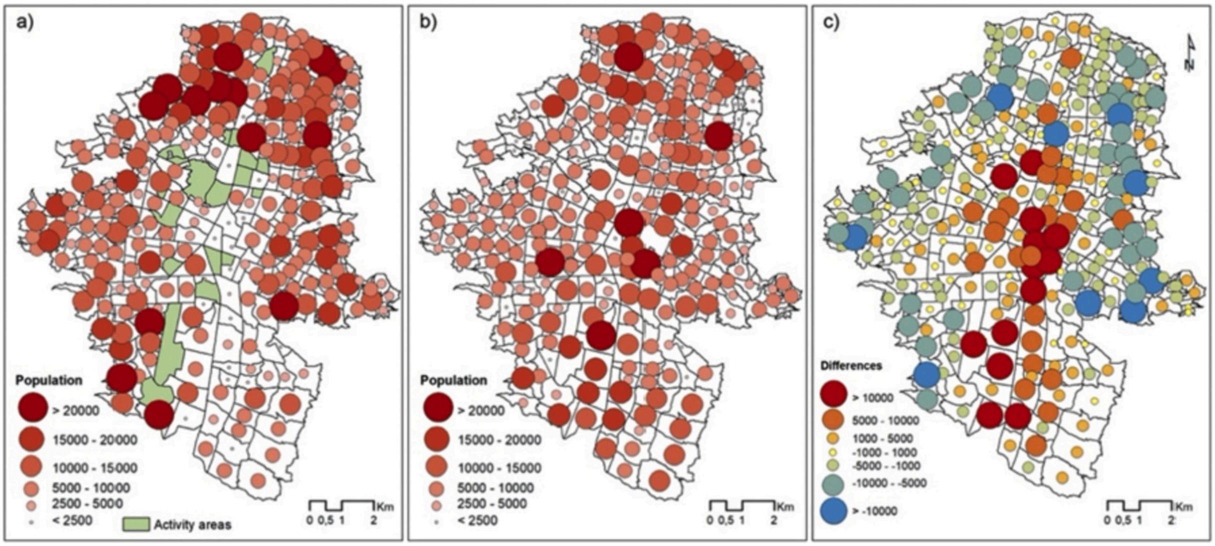

Fig. 3. Population at night (a), during a day (b) and differences between night and day (c).

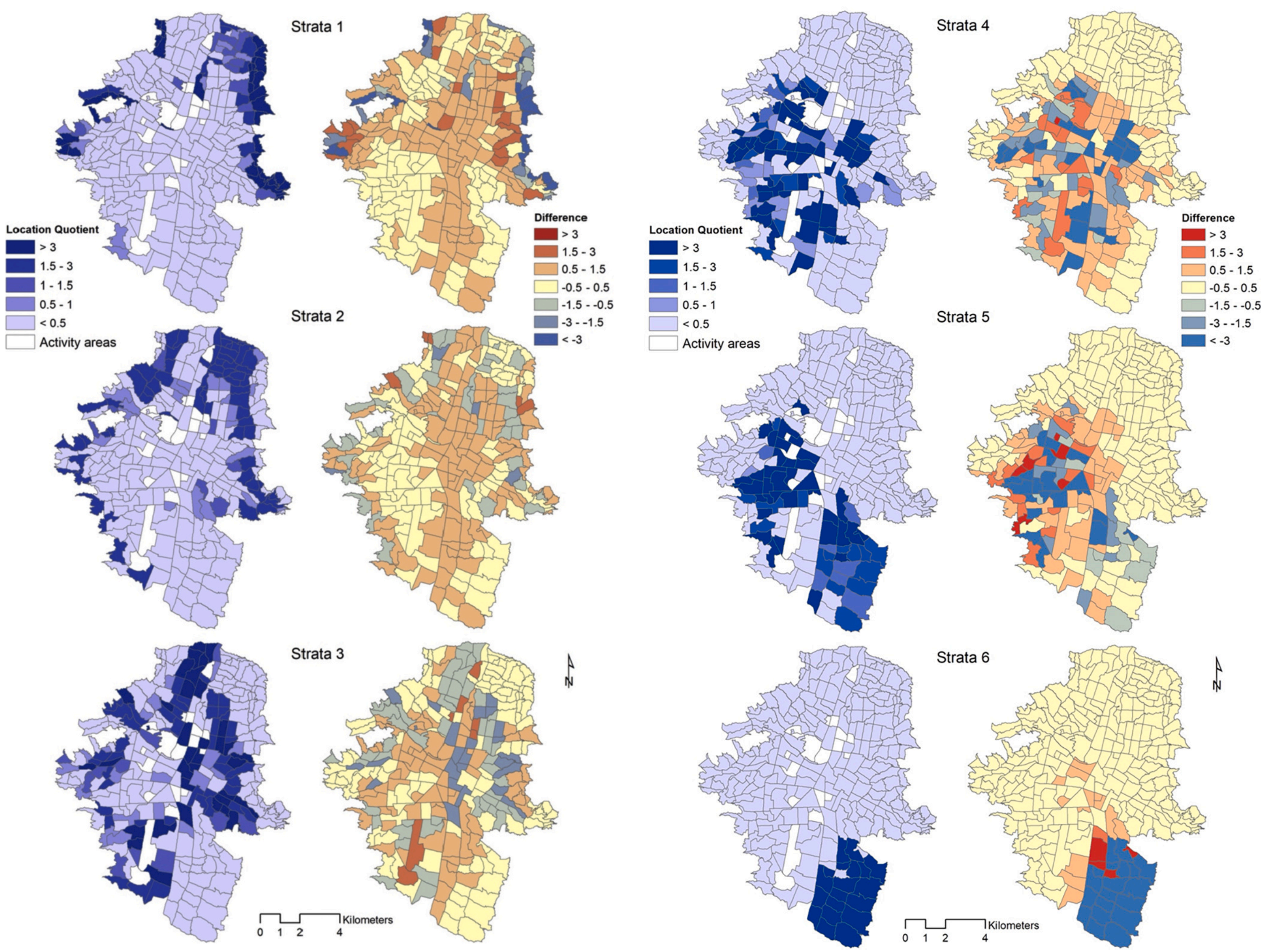

Fig. 4. a. Location quotient of the population for each stratum (rows) in each neighbourhood at night (left) and changes between night and day (right): strata 1–3. b. Location quotient of the population for each stratum (rows) in each neighbourhood at night (left) and changes between night and day (right): strata 4–6.

neighbourhoods and to the neighbourhoods at the bottom of the valley. However, while strata 1 and 2 tend to relocate especially in the northern part of the municipality, stratum 3 does so practically all around the municipality and the rest of the strata hardly consider the northern neighbourhoods as spaces of activity. Finally, the highest income stratum relocates only in a reduced number of neighbourhoods in the centre and south of the city.

Coefficients of variation of the LQ for each stratum at night and during a day (Table 2) reveal that the highest-income strata (5 and 6) are the most spatially concentrated, both during day and at night. Simultaneously, the degree of spatial concentration of residents of all income strata decreases during a day, in case of low and middle-low income groups in particular.

6.2. Clustering

Results from the previous analysis show that higher income strata bear greater spatial concentration, and that spatial concentration decreases when going from night to day. However, this initial analysis addresses each neighbourhood independently, so it does not provide information on whether neighbourhoods with high and/or low concentration level are spatially grouped, or they are arranged randomly in the space.

As expected, Global Moran's Index indicates a positive spatial autocorrelation in all strata (both at night and day), since all z-score values are positive and statistically significant (Tables 3 and 4). The stratum 1 (the lowest income level) has the highest z-score at night, indicating the highest degree of spatial clustering (greatest segregation) among all strata. Moreover, spatial clustering is uniformly higher during a day (Table 4) than at night (Table 3).

Fig. 5 supports this statement, as all maps show a larger area of HH and LL clusters (positive spatial autocorrelation) and reduced area of HL and LH outliers (negative spatial autocorrelation) during a day. The increase of clustering during a day (clustering) is compatible with decrease of concentration presented in the previous section (see Fig. 2), since different segregation aspects are measured: spatial autocorrelation analyses relationship between values of neighbouring units (applying 1000 m radius) and concentration considers each neighbourhood separately.

Change of population distribution between night and day reduces the number of outliers (Fig. 5). For example, two low-income strata enclaves, (HL outliers) are visible in the centre-north of the city, but only at night (Strata 1 at Fig. 5). As both are enclaves surrounded by higher-income strata neighbourhoods, their inhabitants have greater potential for social interaction than the HH clusters located on the northeast and northwest periphery. In result, both enclaves are not visible during a day anymore, as their location facilitates mixing with surrounding neighbourhoods. The number of LH outliers is also reduced, due to population mobility which leads to social mixing in these neighbourhoods and those around them. Spatial patterns during a day are much clearer than those at night, with an aureole of HH clusters in the northeast and northwest periphery and an LL cluster in the activity zone of the valley's left bank. Similar trends are observed in case of stratum 2, although neighbourhoods included in HH clusters have a bit less peripheral location. Similarly, more of the middle and middle-high income strata (strata 3, 4 and 5) outliers (especially LH ones) disappear during a day, simplifying spatial structures, particularly in strata 4 and 5. Finally, stratum 6 shows the simplest spatial structures, both at night and during a day, with the large HH cluster in the southeast.

6.3. Exposure

The analyses of differences of the level of concentration during night and day show a scale of changes in population distribution (Fig. 4). We have yet to discover how this mobility influences social mixing: does the population travel to areas where higher exposure occurs? Or, do the people go to areas similar to where they reside, thus reinforcing the segregation recorded during night-time?

Spatial exposure patterns at night (Fig. 6a) are fairly complex, although we do observe that neighbourhoods with less exposure (with values below 1 and even below 0.75) tend to be predominant both in the north (low income strata) and in the south (high income strata). Simultaneously, more socially diverse neighbourhoods are located in the central and western part of the city, where middle or middle-high income strata predominate. During a day (Fig. 6b), the level of distance-decay Shannon Index is much higher than at night, revealing a general increase in exposure, and the emerging spatial pattern is clearer and easier to recognize. The centre and south of the city and at the bottom of the valley in particular, concentrate more jobs and facilities, where inhabitants from different strata and neighbourhoods go to work or to conduct other activities. Thus, they become spaces where it is more likely that people interact with representatives of different strata.

In all neighbourhoods, exposure is higher during a day than at night, although the increase is spatially diverse. The greatest increase of distance-decay Shannon Index takes place in the areas which attract more trips, i.e. in the centre and south of the city (cf. Fig. 2).

Table 5 shows weighted average values of distance-decay Shannon Index, aggregated by dominant income stratum, using population size at a given time as a weighting factor. It confirms that exposure to interaction with other income strata increases considerably when going from night to day. It also reveals that the lowest and highest strata are the least exposed during the night. Simultaneously, low and middle-low strata (1, 2 and 3) bear less exposure during the day. Stratum 6 experiences the greatest contrast between night and day, since the neighbourhoods dominated by this stratum receives during a day a significant mobility inflow from other types of neighbourhoods.

6.4. Mobility-based indicators: socio-spatial-mix trips

Fig. 7a shows that the percentage of trips to neighbourhoods of a different income stratum is higher in areas on the hillsides, contributing to a higher social mix (exposure) at the destination, while residents of wealthy neighbourhoods in the southeast tend to move to areas of the same stratum. As expected, the average distance travelled by socio-spatial-mix trips (Fig. 7b) is greater from peripheral neighbourhoods than from central ones. Moreover, due to the deficit of jobs and other activities in the northern hillsides, trips from these areas are disproportionately larger.

Table 6 shows mobility indicators aggregated by income stratum. It is noteworthy that, as income stratum decreases, the percentage of "socio-spatial-mix trips" increases, as does the average of distance travelled per inhabitant. Residents of high income stratum areas find more opportunities for their daily activities (including employment) in their neighbourhood or nearby, in neighbourhoods of the same stratum.

Table 2
Coefficient of variation of the spatial concentration, according to income strata and neighbourhoods at night and during a day (269 neighbourhoods).

Strata	1	2	3	4	5	6	Weighted Average
Night	211.7	140.9	131.5	198.3	245.4	365.8	171.7
Day	97.6	51.0	60.2	101.7	138.7	291.5	82.6
Difference	−114.1	−89.9	−71.3	−96.6	−106.7	−74.3	−89.1
Difference (%)	−53.9	−63.8	−54.2	−48.7	−43.5	−20.3	−54,5

Table 3
Global Moran Index according to income strata: night.

Strata	1	2	3	4	5	6	Weighted Average
Moran's Index:	0.594260	0.477933	0.371989	0.286071	0.400216	0.365052	0.430215
Expected Index:	−0.003731	−0.003731	−0.003731	−0.003731	−0.003731	−0.003731	-0.003731
Variance:	0.001118	0.001133	0.001135	0.001118	0.001103	0.001088	0.001126
z-score:	17.881416	14.312412	11.150828	8.668499	12.165428	11.178735	13.029628
p-value:	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Table 4
Global Moran Index according to income strata: day.

Strata	1	2	3	4	5	6	Weighted Average
Moran's Index:	0.827137	0.831332	0.732913	0.834656	0.823340	0.403601	0.783936
Expected Index:	−0.003731	−0.003731	−0.003731	−0.003731	−0.003731	−0.003731	−0.003731
Variance:	0.001111	0.001131	0.001135	0.001132	0.001119	0.001080	0.001127
z-score:	24.922183	24.833590	21.864572	24.918737	24.726716	12.396457	23.460938
p-value:	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

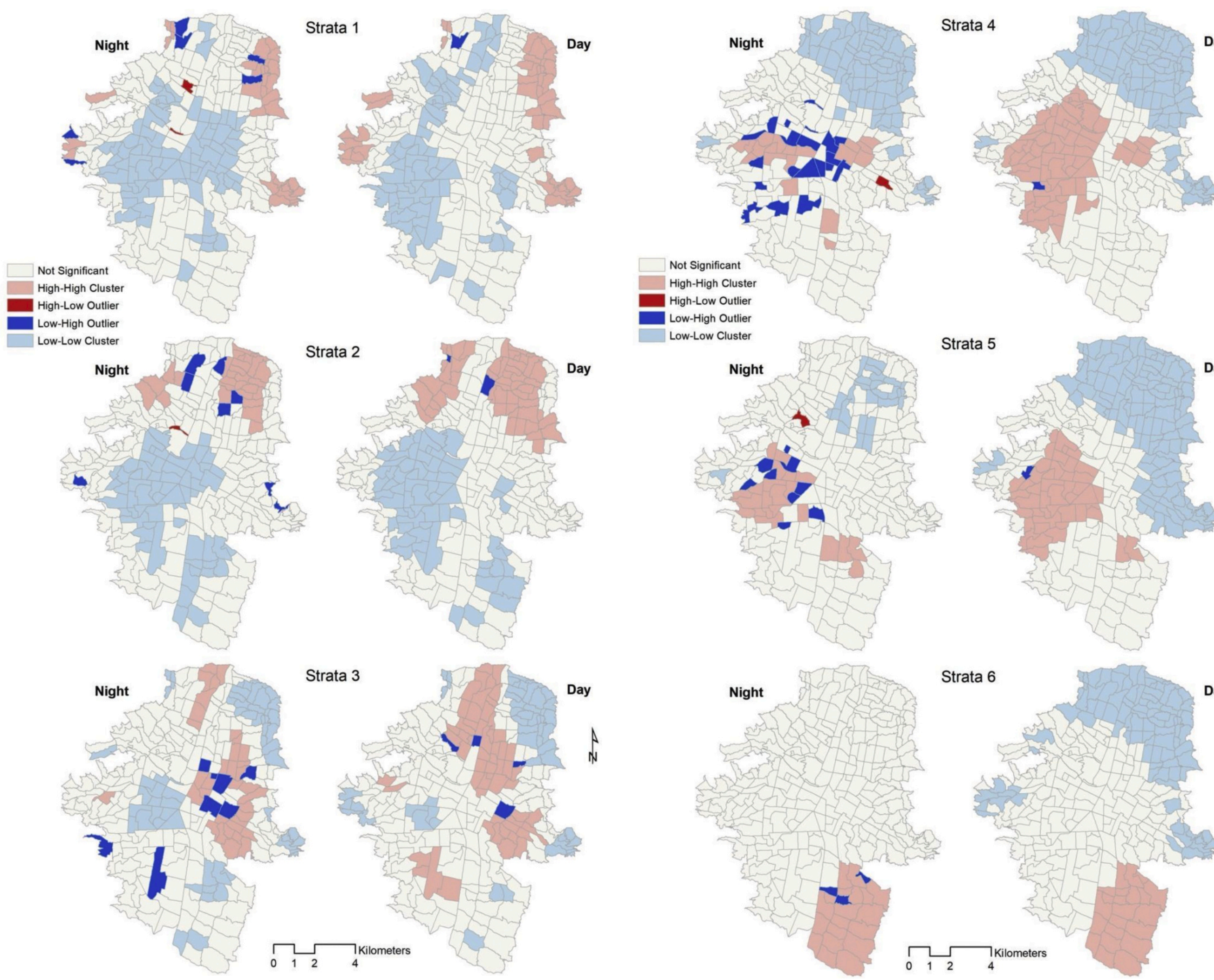

Fig. 5. a. LISA during the night and day: Strata 1–3 in rows. b. LISA during the night and day: Strata 4–6 in rows.

On the other hand, low-income strata neighbourhoods have fewer jobs and facilities, thus during a day their inhabitants need to travel to meet their needs. Their destinations are usually farther, in neighbourhoods of higher income strata. Thus, we argue that the greater social mix during a day, as measured with the distance-decay Shannon Index, is mainly a result of the low-income strata population's mobility.

7. Conclusions

This study addresses the variations between night and day in income

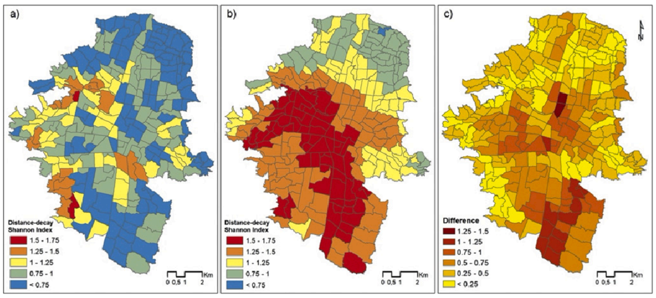

Fig. 6. Distance-decay Shannon Index according to neighbourhoods at night (a), day (b) and changes between night and day (c), calculated with a 1000 m radius, weighted by the inverse of distance.

spatial segregation in Medellin (Colombia), providing new angles in the Europe-centric research field. It relies on data derived from official registries supplemented by mobile phone data. Segregation indicators are calculated based on income spatial patterns of the population during day and night. We show that even a limited amount of details derived from mobile phone data provides important insights on population mobility and helps to enrich our understanding of segregation's differences between day and night.

In the presented study, the multidimensional segregation analysis which evaluates concentration, clustering and exposure is supplemented by mobility-based indicators which provide comprehensive understanding of changes of segregation patterns between night and day. In our approach concentration is evaluated for each neighbourhood separately, while clustering and exposure indicators apply 1000 m radius in order to consider the local setting of each neighbourhood. The proposed distance-decay Shannon Index permits to go beyond individual neighbourhoods in exposure analysis. It covers areas within a walking distance in order to better reflect real-life experience of inhabitants and provide complementary information on the complex phenomenon of spatial segregation. In result, our approach is less sensitive to the modifiable areal unit problem (MAUP) as it does not operate on administratively defined spatial units.

The analysis of the concentration indicator reveals strong spatial contrasts at night between north and south, on the one hand, and the hillsides and the bottom of the valley, on the other. Both, the neighbourhoods of low strata, concentrated in the northern hillsides as well as the high strata in the south-eastern, form spatial clusters and exhibit very low exposure to other groups. The highest exposure values are recorded in central neighbourhoods, inhabited by a population of middle-income strata.

The dynamic of concentration levels of particular strata between night and day suggests that, during a day the population of the disadvantaged neighbourhoods tends to move toward the bottom of the valley, while the higher-income population from the south-east tends to stay in the area. The middle-income strata (strata 4 and 5) tend to move toward the south and the bottom of the valley as well. In result, neighbourhoods in the centre and south receive a large amount of population from neighbourhoods of different strata, reaching a very high level of exposure during a day, while those in the hillsides in the north continue exhibiting very low levels of exposure. We can hypothesise, that insufficient job opportunities and limited activities in the disadvantaged areas (services, leisure etc.) do not attract residents of other areas limiting potential social interaction of their inhabitants.

Furthermore, the analysis of the coefficients of variation of the LQ and the weighted average values of Shannon index shows that during night the extreme strata (highest-income and lowest-income strata) are the most spatially concentrated and the least exposed. Moreover, the results of LISA statistics reveal that the lowest income stratum exhibits the highest autocorrelation value (clustering). Nevertheless, the degree of spatial concentration for all income strata decreases during the day, in particular for low and middle-low income groups. This is consistent with previous studies (Ellis et al., 2009), including those that apply mobile phone data (Östh et al., 2018). Interestingly, the decrease of concentration levels during the day is compatible with an increase of clustering. Income strata are spatially less concentrated, but show a greater spatial order, with less outliers and more extended spatial clusters, softening local contrasts and reinforcing general trends. Moreover, all strata notice higher exposure during the day, especially the highest income stratum. Interestingly, neighbourhoods dominated by the highest stratum experienced the greatest contrast between night and day, since they receive a significant, daytime mobility inflow from other types of neighbourhoods.

Finally, an increased exposure during the day is mostly caused by the mobility of low-income strata. This observation is in line of Xu, Belyi, Bojic, and Ratti (2018) findings for Singapore but contradictory to conclusions derived from the same study for Boston. In Medellin, the lower the income strata level, the larger share of socio-spatial-mix trips

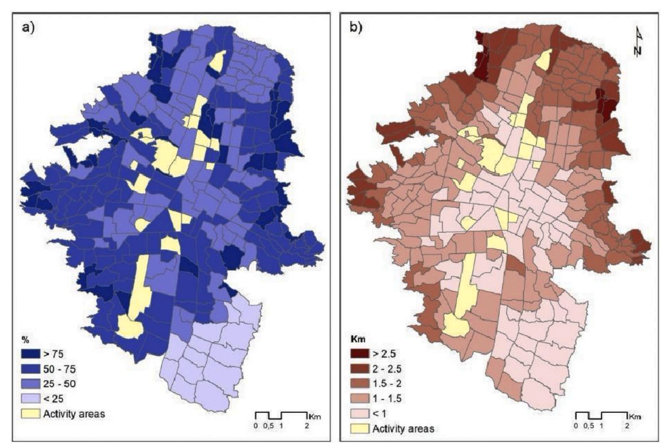

Fig. 7. Mobility indicators: Percentage of "socio-spatial-mix trips" (a) and average distance (km) travelled per inhabitant in "socio-spatial-mix trips" (b).

Table 6
Mobility indicators.

Strata	1	2	3	4	5	6	Weighted Average
Socio-spatial-mix trips (%)	78.4	59.3	47.9	53.1	59.0	23.3	56.2
Average distance travelled (km) per inhabitant in socio-spatial-mix trips	2.31	1.82	1.24	1.03	1.19	0.85	1.51

Table 5
Distance-decay Shannon Index according to income strata.

Strata	1	2	3	4	5	6	Weighted Average
Night	0.775	0.783	0.808	1.006	0.922	0.656	0.819
Day	1.210	1.193	1.295	1.473	1.508	1.549	1.296
Difference	0.435	0.410	0.487	0.467	0.586	0.893	0.477
Difference %	56.1	52.4	60.3	46.4	63.6	136.1	58.2

(trips with origin and destination in neighbourhoods of different strata), increasing social mix during the day. The crucial observation is that daily mobility mitigates the spatial segregation observed at night. However, it is related to a great effort of inhabitants from low strata areas, who need to travel to work and conduct activities in other neighbourhoods, predominantly of higher income stratum, covering larger distances in socio-spatial-mix trips.

We agree that public transport is crucial to improve the integration of the population from disadvantaged areas. Public transport and spatial mismatch strategies should reduce socio-spatial segregation during the day. However, they are not sufficient to address socio-spatial segregation during the night. From a public policy perspective, we must insist on the need to increase spatial mismatch in the city and on the promotion of services and commerce in the most disadvantaged neighbourhoods. These are in line with the strategy adopted by the Medellin Town Hall.

In general, social mix policies are widely accepted and promoted as an important strategy to overcome problems related to residential segregation. Strategies for developing mixed communities can be based, among others, on the spatial dispersion of low income population, the regeneration of neighbourhoods and the regulation of new developments. These strategies should encourage social interaction between different groups, reduce criminality and anti-social behaviour, improve neighbourhood reputation, and enhance services and infrastructure. Studies evaluating social mixing policies indicate that the expected benefits are only partially achieved. However, it is generally accepted that social mixing policies can at least contribute to mitigate the negative effects of residential segregation, in particular if they are combined with a good quality public transport.

The presented analysis may be significantly enriched if more detailed mobile phone data is available. Greater temporal and spatial resolution would provide more robust conclusions about the actual interaction between representatives of different socioeconomic strata. Nevertheless, even though this kind of data is becoming more available for researchers, it is still inaccessible in many locations all over the world. We believe that our approach shows how even a limited scope of mobile phone data can enhance our understanding of urban segregation.

Declarations of interest

None.

Acknowledgments

The authors gratefully acknowledge funding from the Spanish Ministry of Science and Innovation and the European Regional Development Fund (Project DynMobility, RTI2018-098402-B-I00). M. Stępniak would like to thank the Spanish Agencia Estatal de Investigación (Ministerio de Ciencia, Innovación y Universidades) and the European Social Fund for their financial support through the Ramón y Cajal Programme.

References

Ahas, R., Silm, S., Järv, O., Saluveer, E., & Tiru, M. (2010). Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of Urban Technology, 17*, 3–27. https://doi.org/10.1080/10630731003597306.

Alcaldía de Medellín. (n.d.). Viviendas residenciales estratificadas por comuna y barrio, según estrato 2010. Medellín (Colombia). Retrieved from https://web.archive.org/web/20150525164554/https://www.medellin.gov.co/irj/go/km/docs/wpccontent/Sites/SubportaldelCiudadano/PlaneaciónMunicipal/Secciones/IndicadoresyEstadísticas/Documentos/Estratificación/ViviendasEstratificadasporComunayBarrio2010.pdf.

Anselin, L. (2010). Local indicators of spatial association-LISA. *Geographical Analysis, 27*, 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Bauder, H., & Sharpe, B. (2002). Residential segregation of visible minorities in Canada's gateway cities. *The Canadian Geographer, 46*, 204–222. https://doi.org/10.1111/j.1541-0064.2002.tb00741.x.

Bea, D. C. (2016). Transport engineering and reduction in crime: the Medellín case. *Transportation research procedia, 18*, 88–2292. https://doi.org/10.1016/j.trpro.2016.12.012.

Borruso, G. (2009). Geographical analysis of foreign immigration and spatial patterns in urban areas: Density estimation, spatial segregation and diversity analysis. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (pp. 301–323). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-10649-1_18.

Brand, P., & Davila, J. (2011). *Aerial cable-car systems for public transport in low-income urban areas: lessons from Medellin, Colombia*. https://doi.org/10.35648/20.500.12413/11781/ii006.

Chakraborty, J., & Armstrong, M. P. (1997). Exploring the use of buffer analysis for the identification of impacted areas in environmental equity assessment. *Cartography and Geographic Information Science, 24*, 145–157. https://doi.org/10.1559/152304097782476951.

DANE, & Alcaldía de Medellín. (n.d.). Perfil Sociodemográfico 2005–2015 por comuna. Medellín (Colombia).

Dannemann, T., Sotomayor-Gómez, B., & Samaniego, H. (2018). The time geography of segregation during working hours. *Royal Society Open Science, 5*, 180749. https://doi.org/10.1098/rsos.180749.

Dávila, J. D. (2009). Being a mayor: The view from four Colombian cities. *Environment and Urbanization, 21*, 37–57. https://doi.org/10.1177/0956247809103003.

Dixon, J., Durrheim, K., & Tredoux, C. (2005). Beyond the optimal contact strategy: A reality check for the contact hypothesis. *The American Psychologist, 60*, 697–711.

Ellis, M., Wright, R., & Parks, V. (2009). Geography and the immigrant division of labor. *Economic Geography, 83*, 255–281. https://doi.org/10.1111/j.1944-8287.2007.tb00354.x.

Frias-Martinez, V., Virseda, J., Rubio, A., & Frias-Martinez, E. (2010). Towards large scale technology impact analyses. In *Proceedings of the 4th ACM/IEEE international conference on information and communication technologies and development - ICTD '10* (pp. 1–10). New York, New York, USA: ACM Press. https://doi.org/10.1145/2369220.2369230.

García-Albertos, P., Picornell, M., Salas-Olmedo, M.H., Gutiérrez, J., 2018. Exploring the potential of mobile phone records and online route planners for dynamic accessibility analysis. Transportation Research Part A: Policy and Practice 0–1. doi: https://doi.org/10.1016/j.tra.2018.02.008.

García-Palomares, J. C., Gutiérrez, J., & Cardozo, O. D. (2013). Walking accessibility to public transport: An analysis based on microdata and GIS. *Environment and Planning. B, Planning & Design, 40*, 1087–1102. https://doi.org/10.1068/b39008.

van Ham, M., & Tammaru, T. (2016). New perspectives on ethnic segregation over time and space. A domains approach. *Urban Geography, 37*, 953–962. https://doi.org/10.1080/02723638.2016.1142152.

Heinrichs, D., & Bernet, J. S. (2014). Public transport and accessibility in informal settlements: Aerial cable cars in Medellín, Colombia. *Transportation Research Procedia, 4*(1), 55–67.

Heringa, A., Bolt, G., Dijst, M., & van Kempen, R. (2014). Individual activity patterns and the meaning of residential environments for inter-ethnic contact. *Tijdschrift voor Economische en Sociale Geografie, 105*, 64–78. https://doi.org/10.1111/tesg.12046.

Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Martonosi, M., Rowland, J., & Varshavsky, A. (2011). Identifying important places in people's lives from cellular network data. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (pp. 133–151). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-21726-5_9.

Johnston, R., Poulsen, M., & Forrest, J. (2007). The geography of ethnic residential segregation: A comparative study of five countries. *Annals of the Association of American Geographers, 97*, 713–738. https://doi.org/10.1111/j.1467-8306.2007.00579.x.

Kamruzzaman, M., Baker, D., Washington, S., & Turrell, G. (2014). Advance transit oriented development typology: Case study in Brisbane, Australia. *Journal of Transport Geography, 34*, 54–70. https://doi.org/10.1016/j.jtrangeo.2013.11.002.

Kang, J. H., Welbourne, W., Stewart, B., & Borriello, G. (2005). Extracting places from traces of locations. *ACM SIGMOBILE Mob. Comput. Commun. Rev., 9*, 58–68. https://doi.org/10.1145/1094549.1094558.

Kwan, M. P. (2016). Algorithmic geographies: Big data, algorithmic uncertainty, and the production of geographic knowledge. *Annals of the American Association of Geographers, 106*, 274–282. https://doi.org/10.1080/00045608.2015.1117937.

Kwan, M.-P., & Weber, J. (2003). Individual accessibility revisited: Implications for geographical analysis in the twenty-first century. *Geographical Analysis, 35*, 341–353. https://doi.org/10.1111/j.1538-4632.2003.tb01119.x.

Li, F., & Wang, D. (2017). Measuring urban segregation based on individuals' daily activity patterns: A multidimensional approach. *Environment & Planning A, 49*, 467–486. https://doi.org/10.1177/0308518X16673213.

Maloutas, T. (2016). Socio-economic segregation in Athens at the beginning of the twenty-first century. In *Socio-economic segregation in European capital cities* (pp. 180–209). London and New York: East Meets West. Routledge. https://doi.org/10.4324/9781315758879-15.

Massey, D. S., & Denton, N. A. (1993). *American apartheid. Segregation and the making of the underclass*. Harvard University Press.

Ministerio de Tecnologías de la Información y las Comunicaciones. (2014). Boletín trimestral de las TIC - Cifras primer trimestre de 2014. Bogotá (Colombia). Retrieved from https://colombiatic.mintic.gov.co/679/w3-article-6276.html.

Mitchel, A. (2005). The ESRI guide to GIS analysis, volume 2: Spartial measurements and statistics. In *ESRI Guid. to GIS Anal*.

Molinatti, F. (2015). Cambios en los patrones de segregación residencial socioeconómica en la ciudad de Córdoba. Años 1991, 2001 y 2008. *International Migration, 9*. https://doi.org/10.17428/rmi.v9i33.51.

Monkkonen, P. (2012). Residential segregation in urban Mexico: Levels and patterns. *Eure, 38*, 125–146. https://doi.org/10.4067/S0250-71612012000200005.

Mooses, V., Silm, S., & Ahas, R. (2016). Ethnic segregation during public and national holidays: A study using mobile phone data. *Geografiska Annaler: Series B, Human Geography, 98*, 205–219. https://doi.org/10.1111/geob.12100.

Musterd, S., Marcińczak, S., van Ham, M., & Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban Geography, 38*, 1062–1083. https://doi.org/10.1080/02723638.2016.1228371.

Musterd, S., & Ostendorf, W. (2013). Urban segregation and the welfare state: Inequality and exclusion in western cities. In *Urban segregation and the welfare state: Inequality and exclusion in western cities*. Taylor and Francis. https://doi.org/10.4324/9780203448533.

Musterd, S., & Van Kempen, R. (2009). Segregation and housing of minority ethnic groups in Western European cities. *Tijdschrift voor Economische en Sociale Geografie, 100*, 559–566. https://doi.org/10.1111/j.1467-9663.2009.00558.x.

Nagendra, H. (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography, 22*, 175–186. https://doi.org/10.1016/S0143-6228(02)00002-4.

Openshaw, S., & Taylor, P. J. (1981). The modifiable areal unit problem. In N. Wrigley, & R. Bennett (Eds.), *Quantitative geography: A British view* (pp. 60–69). London (the United Kingdom): Routledge and Kegan Paul.

Östh, J., Andersson, E., & Malmberg, B. (2013). School choice and increasing performance difference: A counterfactual approach. *Urban Studies, 50*, 407–425. https://doi.org/10.1177/0042098012452322.

Östh, J., Shuttleworth, I., & Niedomysl, T. (2018). Spatial and temporal patterns of economic segregation in Sweden's metropolitan areas: A mobility approach. *Environment & Planning A, 50*, 809–825. https://doi.org/10.1177/0308518X18763167.

Park, Y. M., & Kwan, M.-P. (2018). Beyond residential segregation: A spatiotemporal approach to examining multi-contextual segregation. *Computers, Environment and Urban Systems, 71*, 98–108. https://doi.org/10.1016/j.compenvurbsys.2018.05.001.

Pérez Gutiérrez, L. (2001). *Medellin competitiva : plan de desarrollo 2001-2003 : hacia una revolucion de la cultura ciudadana*. Medellín: Alcaldia de Medellín.

Plan de Desarrollo. (2001–2003). *Medellín Competitiva, 2001*. Medellín.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *The American Journal of Sociology, 116*, 1092–1153. https://doi.org/10.1086/657114.

Sabatini, F. (2006). *La segregación social del espacio en las ciudades de América Latina*.

Sarma, P., & Das, D. (2015). Application of Shannon's index to study diversity with reference to census data of Assam. *Asian Journal of Management Research, 5*, 620–628.

Schnell, I., Diab, A. A. B., & Benenson, I. (2015). A global index for measuring socio-spatial segregation versus integration. *Applied Geography, 58*, 179–188.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Sheskin, I. M., & Hartman, H. J. (2019). Religious diversity and religious participation in U.S. Jewish communities. *The Professional Geographer, 71*, 39–51. https://doi.org/10.1080/00330124.2018.1455520.

Silm, S., & Ahas, R. (2014a). Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Annals of the Association of American Geographers, 104*, 542–559. https://doi.org/10.1080/00045608.2014.892362.

Silm, S., & Ahas, R. (2014b). The temporal variation of ethnic segregation in a city: Evidence from a mobile phone use dataset. *Social Science Research, 47*, 30–43. https://doi.org/10.1016/j.ssresearch.2014.03.011.

Silm, S., Ahas, R., & Mooses, V. (2018). Are younger age groups less segregated? Measuring ethnic segregation in activity spaces using mobile phone data. *Journal of Ethnic and Migration Studies, 44*, 1797–1817. https://doi.org/10.1080/1369183X.2017.1400425.

Strömgren, M., Tammaru, T., Danzer, A. M., van Ham, M., Marcińczak, S., Stjernström, O., & Lindgren, U. (2014). Factors shaping workplace segregation between natives and immigrants. *Demography*. https://doi.org/10.1007/s13524-013-0271-8.

Tammaru, T. (2015). *Socio-economic segregation in European capital cities*. Routledge. https://doi.org/10.4324/9781315758879.

Tammaru, T., Marcińczak, S., Aunap, R., & van Ham, M. (2017). *Inequalities and segregation across the long-term economic cycle: An analysis of south and north European cities*.

Toomet, O., Silm, S., Saluveer, E., & Tammaru, T. (2012). *Where do ethnic groups meet? Copresence at places of residence, work, and free – time segregation across full activity space: background* (pp. 1–27).

Tu, W., Zhu, T., Xia, J., Zhou, Y., Lai, Y., Jiang, J., & Li, Q. (2020). Portraying the spatial dynamics of urban vibrancy using multisource urban big data. *Computers, Environment and Urban Systems, 80*, 101428. https://doi.org/10.1016/j.compenvurbsys.2019.101428.

Warf, B., & Vincent, P. (2007). Religious diversity across the globe: A geographic exploration. *Social and Cultural Geography, 8*, 597–613. https://doi.org/10.1080/14649360701529857.

Wissink, B., Schwanen, T., & van Kempen, R. (2016). Beyond residential segregation: Introduction. *Cities*. https://doi.org/10.1016/j.cities.2016.08.010.

Wong, D. W. S., & Shaw, S.-L. (2011). Measuring segregation: An activity space approach. *Journal of Geographical Systems, 13*, 127–145. https://doi.org/10.1007/s10109-010-0112-x.

Xu, Y., Belyi, A., Bojic, I., & Ratti, C. (2018). Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Computers, Environment and Urban Systems, 72*, 51–67. https://doi.org/10.1016/j.compenvurbsys.2018.04.001.

Yip, N. M., Forrest, R., & Xian, S. (2016). Exploring segregation and mobilities: Application of an activity tracking app on mobile phone. *Cities, 59*, 156–163. https://doi.org/10.1016/j.cities.2016.02.003.

Zhou, X., Chen, Z., Yeh, A. G., & Yue, Y. (2019). Workplace segregation of rural migrants in urban China: A case study of Shenzhen using cellphone big data. *Environment and Planning B: Urban Analytics and City Science*. https://doi.org/10.1177/2399808319846903 (online first).