University of Tartu

Faculty of Science and Technology

Institute of Ecology and Earth Sciences

Department of Geography

Master thesis in Geoinformatics for Urbanized Society (30 ECTS)

Comparing activity-space based segregation methods: a study with GPS data

Eliana Solange Ortiz Gavela

Supervisors: Siiri Silm, PhD

Veronika Mooses, MSc

Tartu, 2020

Annotation

Comparing activity-space based segregation methods: a study with GPS data

Spatial segregation is a complex multidimensional process that affects different spheres of the society and people's life. GPS data has allowed the development of more accurate measures to understand spatial segregation. Thus, the aim of the thesis is to compare activity space methods for studying segregation and interaction with GPS data. Groups under comparison are staff and students, and the activity space indicators are calculated for weekdays and weekends.

Three methods have been used for the comparison: minimum convex polygon (MCP), buffer and grid. For the research GPS data has been obtained through MobilityLog smartphone application. The dataset consists of 50 staff members and 50 students. The month covered by the data is March for the years 2015, 2017 and 2018. The data consisted of pre – calculated stops based on aggregated sequential GPS points and the social characteristic of the respondents: academic role.

Each method, according to its complexity, depicted different results. Students registered larger activity spaces than staff. Buffer and grid methods provided more detailed and realistic insights of the activity spaces compared to MCP. MCP produces overgeneralised results which seems to be unsuitable for GPS data. Buffer and grid based calculations are more suitable for GPS data. With the spatially precise data, there is the opportunity to measure the overlap of activity spaces (i.e. interaction spaces). Interaction spaces showed that student's spatial behaviour is more variable than staff both during weekdays and weekends. Results indicated that the used indicators did not provide sufficient evidence to confirm segregation between staff members and students. However, the existence of interaction spaces suggests that segregation might be reduced between staff and students, especially within the university environment.

Key words: spatial segregation, measuring segregation, activity space, comparison of methods, GPS data.

CERS code: S230 – Social geography

Annotatsioon

Tegevusruumipõhiste segregatsioonimeetodite võrdlus GPS andmetel

Ruumiline segregatsioon on keerukas ja mitmemõõtmeline protsess, mis mõjutab nii ühiskonna kui ka inimeste elu mitmesuguseid valdkondi. Üha täpsemad andmed nagu GPSi põhised liikuvusandmed võimaldavad inimeste tegevusruumi täpsemalt mõõta ja seeläbi segregatsiooni protsessi paremini mõista. Käesoleva magistritöö eesmärk on võrrelda tegevusruumi meetodeid segregatsiooni ja interaktsiooni uurimiseks GPS andmetel. Vaadeldavad grupid on Tartu Ülikooli töötajad ja üliõpilased ning tegevusruumi indikaatorid leitakse tööpäevade ja nädalavahetuse jaoks eraldi.

Töötajate ja üliõpilaste tegevusruume mõõdetakse kolme meetodi abil: minimaalne kumer hulknurk (*minimum convex polygon*), puhver (*buffer*) ja võre (*grid*). Kolme meetodi põhjal arvutatud tegevusruumi indikaatoreid võrreldakse ja antakse hinnang, kui sobilikud need meetodid on GPS-andmete põhiseks segregatsiooni uurimiseks. GPS-andmed on kogutud MobilityLog nutitelefoni rakenduse abil. Järjestikku paiknevad GPS punktid on eelnevalt agregeeritud peatusteks. Töös kasutatav andmestik koosneb 50 töötaja ja 50 üliõpilase 2015, 2017 ja 2018 aasta märtsikuu peatustest.

Sõltuvalt oma keerukusest andis iga meetod erinevaid tulemusi. Üliõpilaste tegevusruumid olid suuremad kui töötajate omad. Puhvri ja võre meetodid andsid detailsema ja realistlikuma arusaama tegevusruumide ulatusest võrreldes minimaalse kumera hulknurga meetodiga. Viimane annab liialt üldistatud pildi tegevusruumide ulatusest, mis tundub olevat GPS-andmete jaoks ebasobiv. Puhvri- ja võrepõhised arvutused on GPS-andmete jaoks sobivamad. Ruumiliselt täpsete andmetega on võimalik mõõta ka tegevusruumide kattumist, nn interaktsiooniruume. Interaktsiooniruumide põhised arvutused näitasid, et üliõpilaste ruumikasutus on töötajatega võrreldes muutlikum nii tööpäevadel kui ka nädalavahetusel. Töös kasutatud näitajad ei ole piisavad, et kinnitada töötajate ja üliõpilaste vahelist segregatsiooni. Interaktsiooniruumide olemasolu viitab sellele, et kahe grupi vaheline eraldatus võib olla väiksem ja seda eriti ülikooli keskkonnas.

Võtmesõnad: ruumiline segregatsioon, segregatsiooni mõõtmine, tegevusruum, meetodite

võrdlus, GPS andmed

CERS kood: S230 – sotsiaalgeograafia (social geography)

Table of contents

Introduction

The concept and measurement of segregation has evolved through time. Since the early spatial segregation studies, residence places have been the main spatial focus to resume people's daily activity spaces and measure segregation between racial groups (Park & Kwan, 2018; Shen, 2019; Wissink et al., 2016). However, researchers have suggested that other aspects of segregation should be incorporated (Wang et al., 2012), as it has been proven that the segregation notion extends from the residential space to other socio-geographical spaces that an individual experiences with other population groups across daily life (Park & Kwan, 2018). Thus, spatial segregation not only is the separation of groups of population in geographical areas, but is also a complex process that includes human interactions and collective constructions, as the result of the social interaction of the people across the multiple socio-geographical spaces within their daily activity spaces (Raanan & Shoval, 2014).

Wong (2002) stated that the nature of spatial segregation is the lack of interaction among population groups provided by the geographical settings and the locations of people. The choice of people's activity locations generates spatial processes in which people interact with their environment as well as with other people (Ren, 2016). Cities, transportation, communications and other systems shape human activities by altering relationships between space and time in human interaction (Miller, 2004). Therefore, the interaction between people is bound to the urban space (Shen, 2019). In this context, interaction spaces are those places that potentially could bring members of two social groups together in one geographical location, where they may have higher chances of interaction. Moreover, interaction spaces have the potential to be an ideal setting within integration and social interaction can be enhanced, or otherwise a potential space of social conflict (Rose-redwood & Rose-redwood, 2013).

Activity space was introduced as a multidimensional (spatial, temporal and cognitive) measure to analyse segregation beyond the residential sphere (Wang et al., 2012). Since it displays activity patterns from locations in which people perform daily activities, activity space reflects an individual experience of segregation with other population groups across daily life activities (Park & Kwan, 2018; Yao et al., 2018). This is known as activity space segregation.

Several studies have used different available data such as census data, interviews, travel surveys, travel diaries, and call detail records (CDR), to measure activity space segregation at individual level (Kestens et al., 2012; Sadahiro, 2019; Shen, 2019; Silm et al., 2018; Susilo & Kitamura, 2005; Tao et al., 2020). However, due to the quality and detail level of data, these studies have raised new questions to address methodological and conceptualization problems (Jones & Pebley, 2014; Wong & Shaw, 2011)

Due to the digital technological advances, novel digital data sources have surged such: call detail records (CDR) and global positioning data (GPS) (Ahas et al., 2007; Järv et al., 2017; Jones & Pebley, 2014). However, GPS data provides higher spatial accuracy compared to CDR data (Ahas et al., 2007). GPS data provides individual-centered and continuous location data that has helped to identify true activity locations in time and space (Jones & Pebley, 2014). And owing to the precision and space-time granularity of this data more accurate measures of activity spaces to evaluate segregation levels have been developed (Cornwell & Cagney, 2017; Wong & Shaw, 2011). The range of possibilities is from activity space methods widely used in segregation literature because of its easiness to compute and interpret (i.e. standard deviation ellipse (SDE), minimum convex hull polygon (MCP)), to methods that require advanced methodologies for computation and interpretation (i.e. daily path areas (DPA), kernel densities). Both imply methodological and computational limitations. Consequently, each method produces different results that are related to the availability, collection and type of data, and spatial settings of the area of analysis (Vich et al., 2017).

Thus, the aim of the thesis is to compare activity space methods for studying segregation and interaction with GPS data. Groups under comparison are staff and students, and the activity space indicators are calculated for weekdays and weekends. To achieve the aim of the research, following research questions are stated:

1) How does the activity space of staff and students vary across different methods?

2) How do the characteristics of interaction space of staff and students vary across different methods?

Data used for the research has been provided by the Department of Geography consisting of members of University of Tartu. The data is obtained with the tracking application for smartphones: MobilityLog. The data set used for the research consist of 50 staff members and 50 students randomly selected. Due to data availability, the month covered by the data is March for the years 2015, 2017 and 2018. The study area is Tartu city including a 5 km buffer from Tartu city administrative borders. The data set included the pre-calculated stops based on GPS data and social characteristics of the respondents (university role and place of work/study).

1. Theoretical overview

1.1 Spatial segregation

Segregation has been studied since the beginning of the Chicago School (Yip et al., 2016), and it has been recognised as a critical issue in many Western cities (Tan et al., 2017). For instance, African American, Latino and Asian experiences in U.S industrial cities (Kaplan & Holloway, 2001; Wissink et al., 2016) or the accelerated immigration in Europe from Africa, Asia, and the Caribbean as Europe offered a larger job opportunities and quality of life, and as the result of the impact of post-colonialism (Kaplan & Woodhouse, 2004; Kivisto, 2016; Musterd, 2005).

Segregation has not only been studied in the racial-ethnic sphere (Ellis et al., 2004; Rumberger & Willms, 1992; van Kempen & Ozuekren, 1998) but also in occupational segregation by nativity and race (Hanson & Pratt, 1991; Harrison & Lloyd, 2013; Xu & Leffler, 1992), and sex-gender segregation at workplace (Bender et al., 2005; Blair & Lichter, 1991; Blau & Hendricks, 1979; Grusky, 2001); income segregation (Owens et al., 2016; Reardon & Bischoff, 2011; Wessel, 2000); religious segregation (Brimicombe, 2007; Field et al., 2008; Nelson, 2010; Shdema et al., 2018; Smith, 2001); and class segregation (Dwyer, 2010; Schnore, 1965). In the mid-1940s, the spatial dimension of segregation was incorporated into the traditional socio – anthropological segregation studies for the first time, by conceptually classifying measures with respect to distinct dimensions of spatial variations (Jahn et al., 1947; Sadahiro, 2019).

In this context, spatial segregation is a complex spatial process, multidimensional and specific phenomenon (Kaplan & Holloway, 2001; Massey & Denton, 1988), consisting of the interaction of social, spatial and temporal dimensions (Järv et al., 2015), which refers to the spatial separation of social groups of a population in a geographical area (Silm & Ahas, 2014b). Since its complex and multidimensional nature, understanding spatial segregation requires assessment from various perspectives and spatial arrangements (Sadahiro, 2019; Yao et al., 2018). For instance, black and Latino children (as minority social groups) in the United States are highly exposed to environmental lead contamination than white children (as majority social group), and there is evidence that this is the result of geographic location (Palmer, 2013).

1.2 Methods for measuring spatial segregation

Since the early spatial segregation studies, the tradition grounds residence places as the main spatial focus to measure segregation between racial groups (Park & Kwan, 2018; Shen, 2019; Wissink et al., 2016). Various indices to measure residential segregation have been developed in many quantitative segregation studies (Park & Kwan, 2018).

The earliest studies during the 40s incorporated and developed spatial segregation (Jahn et al., 1947). Duncan & Duncan (1995) developed the dissimilarity index, which has been one of the most popular and long standard measures of residential segregation because of its easiness to compute and interpret (Li & Wang, 2017; Sadahiro, 2019). Later in 1988, Massey & Denton delineated five index-based dimensions of residential segregation: evenness, exposure, concentration, centralization and clustering.

However, indexes of residential segregation have been criticised as non spatial measures because they do not account for the spatial relationships among locations, and they share a number of significant limitations (Li & Wang, 2017; Reardon & O'Sullivan, 2004). First, they focus on a city or region as a whole, assuming spatial relations are consistent. Therefore, there is no urban context. Second, they discard a meaningful amount of spatial information, not considering the spatial relations of the units within the urban system. Third, they consider the residential space as the anchor points of people's mobility without looking out to the varied locations people visit across the course of a day (Tao et al., 2020; Yao et al., 2018). The methodological flaws that encompass these limitations are described hereafter.

Two methodological flaws of the index-based residential segregation methods have been identified. The first is known as the checkerboard problem. It relies on the fact that segregation measures focus on the racial composition of neighbourhoods rather than the spatial proximity of neighbourhoods within an entire area. Figure 1a. illustrates the checkerboard problem, where each square represents exclusively a black or a white neighbourhood. If all the black squares were moved to one side of the board, and the white squares to the other side, we would expect an increase in segregation to be registered, since in that scenario neighborhoods are not only composed by the same racial group but also surrounded by the same racial group. Index-based measures do not distinguish the difference between the aforementioned patterns since in each case the racial composition of the neighborhoods is the same (Borjas, 1998; Oakes & Kaufman, 2017; Park & Kwan, 2018; Reardon & O'Sullivan, 2004).

Second, the modifiable areal unit problem (MAUP), arose in residential segregation because of how data (residential population, census, travel diary data, etc.) is collected, aggregated and reported for spatial units, which have not correspondence with social/spatial divisions, ignoring the spatial relationships between units and the real scale of areal units (Massey & Denton, 1988; Reardon & O'Sullivan, 2004). Figure 1b shows two aspects of MAUP: aggregation effects, which result in measured segregation if the subareas used to compute it are different sized; and zoning effects, which result in measured segregation if the boundaries of the same subareas are shifted, even if the size and number of the sub-areas remain the same (Oakes & Kaufman, 2017).

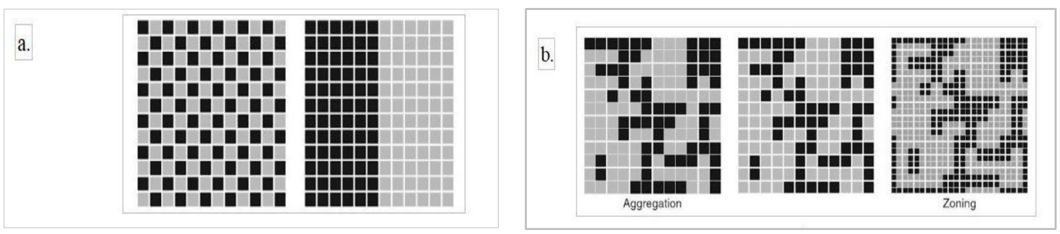

Figure 1. 1a. the checkerboard problem 1b. Modifiable areal unit problem (MAUP).

Source: Oakes & Kaufman (2017)

Although residential segregation has been an important dimension to study segregation, researchers have suggested that other aspects of segregation should be incorporated to avoid the neglect of other potentially important dimensions and sites of social interaction (Wang et al., 2012; Yip et al., 2016). People also experience segregation outside their residential neighbourhoods, as a result of their daily lives, for instance, workplaces or sites for social and recreational activities (Kwan, 2013). Thus, spatial segregation not only is the separation of groups of population in geographical areas, but is also a complex process that includes human interactions and collective constructions as the result of the social interaction of the people across the multiple socio-geographical spaces within their daily lives(Raanan & Shoval, 2014).

Park & Kwan, (2018) stated that the segregation notion extends from the residential space to other socio-economic spaces that an individual experiences with other population groups across daily life activities. Every human activity occurs at a particular place in a defined period of time, thus space and time are inseparable dimensions when studying people's mobility (Yao et al., 2018).

Accordingly, new approaches to measure and analyse segregation beyond residential – based sphere have surged (Shdema et al., 2018), focusing on activity patterns from locations in which people perform daily activities and reflect an individual experience of segregation (Park & Kwan, 2018; Yao et al., 2018). This is known as segregation in space and time or activity- space segregation (Wang & Li, 2016).

1.3 The activity space concept

Activity space is a relatively old concept that was first mentioned by behavioural geographers around the 1970s (Kwan, 1999; Patterson & Farber, 2015). Generally speaking, activity spaces display the use of space based on the daily travel and activities of an individual, to understand geographic mobility patterns (Tao et al., 2020). However, several authors have developed different terms and methods to describe the activity space concept (Patterson & Farber, 2015).

Activity space is an indicator of an individual's spatial behaviour (Patterson & Farber, 2015), represented as a two-dimensional geometric area form of the locations within action spaces during a period of time (Rai et al., 2007; Schönfelder & Axhausen, 2003). These action spaces can be considered portions of the environment constituted by the spatial distribution of those locations an individual has personal experience with, as a result of daily-life activities (Horton & Reynolds, 1971; Kestens et al., 2010; Schönfelder & Axhausen, 2003). Activity spaces can also be an indicator for aggregated spatial behaviour. On an aggregated level they can show recurrent locations in a city, main transport routes and temporal and spatial differences in the use of space (Lee et al., 2016).

Each daily life activity occurs at a particular place in a defined period of time (Yao et al., 2018). Therefore each activity can be encompassed in a different domain (i.e. personal care and enjoyment, family and friend's interaction, work, home, grocery shopping, recreation) (Susilo & Kitamura, 2005; van Ham & Tammaru, 2016). The ability of individuals to be at a certain place in a specific time depends on the resources available to them. Thus, institutional, social, environmental and transportation network conditions are some of the factors that shape and influence the set of places an individual is capable to visit during a normal day (Susilo & Kitamura, 2005). In this sense, activity space is a concept that include spatial, temporal and cognitive dimensions (Wang et al., 2012; Wunderlich, 2008), and according to Horton & Reynolds (1971), it is defined as the subset of all locations an individual visit as the result of daily life activities.

1.4 Activity space based segregation

The activity space based segregation approach is based on the idea of seeking to understand social relations in their spatial context (Palmer, 2013). In this sense it incorporates an additional aspect to the spatial segregation studies, which is the differences in the physical properties, social relations and activity locations within people's space – time spaces themselves (Palmer, 2013; Toomet et al., 2015).

The spatial dimension of activity spaces has been the main focus of many segregation studies, paying attention to the geographical properties of the daily activity spaces (Mooses et al., 2016; Silm et al., 2018). The differences in the extent of activity spaces are influenced by constraints, needs, preferences or resources available to individuals (Mooses et al., 2016; Sherman et al., 2005), that can be categorised in two groups: based on their relation to an individual, or based on environmental characteristics (Vich et al., 2017).

The first group relates to social and demographic characteristics of the individual. For instance, university role (staff members and students) is believed to be a strong determinant of the size of the activity spaces, because it is considered a proxy of social status that reflects age and income differences (Vich et al., 2017). In terms of age, previous literature shows that mature adults with higher incomes and family responsibilities (generally staff members) tend to have larger activity spaces than younger adults (generally students) (Buliung & Kanaroglou, 2006; Ta et al., 2016). But younger adults tend to be more spatially active compared to mature adults (Masso et al., 2019). In terms of income, households with higher income have shown to have larger activity spaces (Vich et al., 2017).

The second group of factors relates to the environmental characteristics. Individual's locations in space and time depend on the resources available to them (Susilo & Kitamura, 2005), so institutional, social, environmental and transportation networks shape human activities by altering relationships between space and time (Miller, 2004; Susilo & Kitamura, 2005). For instance, individuals experiencing long commutes are forced to use motorized transport modes. Long commuters have shown to have larger activity spaces, mainly because the time invested in commuting is prioritized with respect to the time invested in other activities. Otherwise, non commuters, due to proximity dynamics, tend to have short trips which are linked to smaller activity spaces (Vich et al., 2017). Additionally, a mobility-related factor is the trip frequency to activity locations, which have shown to denote differences in the spatial extent of activity spaces (Sherman et al., 2005). Activity locations can be distinguished in two groups. Routine activity locations that are repetitive and visited during weekdays, normally related to home and work places; and the discretionary activity locations (less or randomly visited), generally during weekends, which might be connected to socializing and leisure activities (Silm & Ahas, 2014a; Susilo & Kitamura, 2005). Both types of activity locations have proven to influence the size of activity spaces (Buliung et al., 2008; Sherman et al., 2005).

The temporal dimension of activity space is the less frequently explored in segregation studies (Järv et al., 2014; Silm & Ahas, 2014b). Everyday life entails mandatory activities such: eating, sleeping, working or going to school. Based on this, individuals will decide the destinations of their daily activities, which will determine the characteristics of their individual spatial experience. For instance, people experiencing long working hours are forced to reduce the time spent in other activities, this is reflected in the size of their activity space which tends to be reduced (Vich et al., 2017).

Susilo & Kitamura (2005) studied the temporal variation of individual's activity spaces. Their study demonstrated that the type of activities (Routine activities on weekdays and discretionary activities on weekends) influence the size of the activity spaces. On weekdays, activity locations and spatial behaviour tend to be repetitive. Whereas, on weekends activity locations and spatial behaviour are more variable and tend to be more random and less repetitive. For workers and students on weekdays, the spread of activity locations are more stable than on weekends. However, as workers tend to have more rigid activity schedules they will have less variable activity spaces. Similarly, Buliung et al., (2008) explored the temporal variation in activity space characteristics. Their findings revealed the existent variation in the use of space on weekdays and weekends. The variation in spatial behaviour is higher during the week than on weekends. Weekdays activities of respondents appear to be performed in a wider geographical area when compared to weekends which are carried out over smaller areas.

Spatial segregation studies have shown the multi-dimensions and adaptability of activity spaces because they capture the relevant interaction spaces an individual may experience as the result of daily activities (Wong & Shaw, 2011). Wong, (2002), stated that the nature of spatial segregation is the lack of interaction among population groups provided by the geographical settings and the locations of people. However, the main social spaces that segregation studies have focused on are residential, school and work, ignoring the fact that an individual may also experience segregation across the interaction spaces, where these social spaces overlap. For instance, individuals at home are likely to interact not only with their neighbours, but also with the workers from workplaces in the same neighbourhoods (Wong & Shaw, 2011). Schönfelder & Axhausen (2003), assumed that interaction takes place between stops, therefore when people travel between locations they may interact or be exposed to other groups.

In this context, interaction spaces are those places that potentially could bring members of two social groups together in one geographical location, where they may have higher chances of interaction. For instance, a university environment could be consider a potential interaction space because it brings together local and international students and academics in one geographic location, increasing the chances of interaction between staff and students. Interaction spaces have the potential to be an ideal setting within integration and social interaction can be enhanced, or otherwise a potential space of social conflict (Rose-redwood & Rose-redwood, 2013). Therefore, they could provide insights about segregation from a different perspective that could be relevant for the newer segregation studies that try to tackle how to break the vicious circle of segregation (van Ham et al., 2018).

1.5 Comparison of activity space based segregation methods

Activity spaces have been represented with a variety of indicators that capture their geometry, size and inherent structure (see Table 1) (Schönfelder & Axhausen, 2002, 2003). The range of possibilities is from methods that are easy to compute and represent (i.e. standard deviation ellipse (SDE), minimum convex hull polygon (MCP), but overgeneralise the actual utilised space, to methods that require advanced methodologies to compute and have demonstrated to represent a more realistic and accurate representation of activity space (i.e. daily path areas (DPA), kernel densities) (Vich et al., 2017).

One of the factors that strongly influence the accuracy of the results provided by each method is the data mining methods on human motion and activities (Ahas et al., 2007). For instance, data such travel diaries, travel surveys, questionnaires, interviews, and census data are still in active use but do not give accurate characteristics of the movements of individuals (Ahas et al., 2007; Kestens et al., 2010; Sadahiro, 2019; Shen, 2019; Susilo & Kitamura, 2005; Tao et al., 2020). Due to the digital technological advances, novel digital data sources have surged such: call detail records (CDR) and global positioning data (GPS) (Ahas et al., 2007; Järv et al., 2017; Jones & Pebley, 2014; Silm et al., 2018). However, GPS data provides higher spatial accuracy compared to CDR data (Ahas et al., 2007).

GPS data provides individual-centered and continuous location data that has helped to identify true activity locations in time and space (Jones & Pebley, 2014). And owing to the precision and space-time granularity of this data more accurate methods to represent activity spaces have been used, improving the spatial temporal resolution in activity space based segregation studies (Cornwell & Cagney, 2017; Hirsch et al., 2014; Shoval & Isaacson, 2007; Wong & Shaw, 2011).

The MCP has a long tradition in activity space research and segregation studies (Palmer, 2013). By definition a MCP is the smallest convex polygon that encompasses all activity locations for an individual, or a group of individuals, over the course of a day or longer periods (Buliung et al., 2008; Patterson & Farber, 2015). MCP is a generalised representation of the shape of activity spaces that captures large information about interactions between people and their activity locations (Jones & Pebley, 2014).

MCP has been used to describe the geographical extent and the dispersion characteristics of an individual's activity patterns (Buliung et al., 2008). MCP has been useful in activity space based segregation studies because it describes the range of movement between activity locations, is easy to compute and conceptually an intuitive approach for the analysis and geographic visualization of human activity patterns (Buliung et al., 2008; Vich et al., 2017). However, literature is consistent about the disadvantages of MCP. First, it overgeneralises the space individuals are in contact with (Palmer, 2013; Vich et al., 2017), therefore results include those areas that individuals are not actually visiting (Patterson & Farber, 2015). Second, MCP method is sensitive to spatial outliers, for instance the presence of unusual activity locations (spatial outliers) in the set to calculate MCP can provide very large areas covered for daily activities, when in the reality the majority of locations are clustered only in a particular area (Buliung et al., 2008; Patterson & Farber, 2015). Finally, MCP cannot be calculated if the activity locations are collinear, an example can be when an individual's daily activity locations are performed along a transport or commercial corridor (Buliung et al., 2008).

Table 1. Summary of calculation methods for activity space.

Methods	Type and shape of measure	Advantages	Disadvantages	Applied by	Type of data used
Minimum convex polygon	Euclidean measure Vector image Geometrical shape	Describes the range of movement Easy calculation Captures all activity destination	Overgeneralises the space individuals are in contact with	1.Chan et al., (2014) 2.Vich et al., (2017) 3.Zenk et al., (2018) 4.Zhang et al., (2020)	1.Interviews 2.Travel diaries and GPS 3. GPS 4. GPS
Standard deviation ellipse	Euclidean measure Vector image Ellipsoid shape	Shows the orientation of activities Possibility to weight by frequency	Does not capture all activities A minimum of three unique activities required Overgeneralises the space individuals are in contact with	1.Sherman et al., (2005) 2.Rai et al., (2007) 3.Järv et al., (2015) 4.Silm et al., (2018)	1. Interviews 2. Travel surveys and GPS 3. CDR 4. CDR
Kernel density map	Density map Raster image	Captures all activity destinations Shows the orientation of activities Possibility to weight by frequency	Not comparable with other methods Does not capture the connections between activities	1.Schönfelder & Axhausen, (2002) 2.O'Sullivan & Wong, (2007)	1. Travel surveys 2. Census data
Shortest route	Network-based measure Vector image Network-derived polygon	Captures all activity destinations Shows the orientation of activities Approximation of routes followed	May be too restricted for predictive purposes	Schönfelder & Axhausen, (2003)	Travel surveys
Daily path areas	Network-based measure Vector image Network-derived polygon	Captures all activity destinations Shows the orientation of activities Shows the actual routes followed	The occasional loss of signal of navigation devices might misrepresent the actual route	1.Weber, (2003) 2.Sherman et al., (2005) 3.Zenk et al., (2011) 4.Vich et al., (2017)	1. Travel diary surveys 2. Interviews 3. GPS 4. Travel diaries and GPS

Source: Vich et al., (2017) Modified by the author of the thesis.

Daily path area (DPA) is a relatively new measure that has emerged as a more realistic and accurate method that represents activity spaces by creating buffers around daily activity location routes of individuals (Vich et al., 2017). Several variations of DPA have been used in segregation studies, but within all these variations the use of buffers have remained for the analysis (Kwan, 1999;

Patterson & Farber, 2015; Schönfelder & Axhausen, 2003; Sherman et al., 2005; Zenk et al., 2011), this is mainly because buffers delimit the area containing all the activity location and the potential use of space linked to urban opportunities (Kwan, 1999). For example, Kwan (1999) by analysing the sizes of DPA suggested that access to urban opportunities for women is significantly less than men. However, the size of DPA is not enough to provide a complete and in-depth overview of the results. Given the complexity of the segregation phenomenon recommended other indicators of time-budget and space-time constraints to be included.

DPA is a measure that explicitly measures where people travelled. However, the accuracy of DPA depends on two factors. First, the data quality of the road networks, low quality and missing roads in the data may need a degree of arbitrary decision making and consequently generating biased results (Zenk et al., 2018). Second, the accuracy of the data, in the case of analysis with GPS data, the occasional loss of signal of the navigation devices might not capture all the activity locations hence misrepresent the actual activity space. The buffer widths for the analysis should be chosen considering the maximum accuracy error recorded by GPS network signals and the spatial settings of the study area (Lee et al., 2016; Vich et al., 2017; Zenk et al., 2018).

Vich et al., (2017), compared three activity space methods: minimum convex polygon (MCP), standard deviation ellipse (SDE) and daily path area (DPA), to explore the spatial extent of daily mobility of staff members and students by analysing their activity spaces. Methodological results of the study indicate that DPA produces lower area estimates of activity spaces than MCP and SDE and the influence of the different factors on the area of activity spaces varies across the three different methods. Additionally, DPA area estimates are more appropriate and accurate and minimise spatial overgeneralization. Regarding the socio-demographic results, university role is the most statistically significant determinant of the size of activity space. Student's activity spaces are larger than staff members. Respondents investing more time in work/study at the university tended to reduce the number of other activities and this resulted in smaller activity spaces (Vich et al., 2017).

In the last decades, there have been considerable methodological advances to measure segregation. However, the recently developed measures have not been taken into account, mostly because of their complex calculation and lack of efficient geoprocessing tools (Hong et al., 2014). Grids have been associated with residential segregation indexes and the methodological flaws related to them

(O'Sullivan & Wong, 2007). Hong et al., (2014), implemented grid as surface-based segregation measure, where population counts are not agglomerated into arbitrarily defined geographic areas (i.e. census tracts, neighbourhoods), therefore it do not require the use of aggregate spatial units and therefore avoids the modifiable areal unit problem (MAUP) (described in section 1.2). They stated that the grid approach does not eliminate all the methodological errors, but it could reveal important patterns that would not be possible to find using conventional measures. Also, the results of grid are data dependent and finer grids tend to produce more accurate estimates of segregation. The main advantage of using the grid to measure segregation is consistent geographic scale for analysis (Hong et al., 2014). The main limitation of the grid method is the accuracy needed to determine the spatial resolution of the grid, to display a more realistic measure of segregation the spatial and functional logic of the analysed area should be examined (Hillier et al., 1993). Additionally, careful attention should be paid when reducing the spatial resolution of the grid cells because it could depict totally different results (Wong, 2009).

The activity space concept has been used widely providing new insights of how people's spatial mobility can contribute to measure different dimensions of segregation. Several researchers have compared some of the aforementioned methods to test the impact of the applied method to identify segregation. For instance, Wong, (2018) compared different activity space methods (standard deviation ellipse, network buff er, and potential path area) to demonstrate the experience of accessibility of people with visual impairment. Wong's, (2018) research showed that each method produces different results not only in terms of area size but also the spaces delineated as accessible and inaccessible.

Schönfelder & Axhausen, (2003) explored the extent of activity spaces based on three methods: confidence ellipse, kernel density estimates, and shortest paths networks to identify people at risk of social exclusion. Their analysis indicated that the main driver for activity space size is the overall number of unique locations visited by the individuals, rather than their socio-demographic characteristics. However, with the proposed measures they could not identify social exclusion within groups. Their results indicate that the study design, sample and selected methods were not suitable to address social exclusion highlighting the importance of the purpose of the study and data to select the fitting activity space measures. Measures with different methodological approaches might be more suitable to analyse segregation.

2. Data and methods

2.1 Description of the study area

The study area is Tartu city, Estonia. Tartu is the second largest city in Estonia with around 97 000 inhabitants at the beginning of 2019 (Statistics Estonia, 2019). It is considered a student city and an international established research town being the home of more than 15 000 students and scholars (University of Tartu, 2019). The spatial units under observation are Tartu city including its 17 districts and a 5 km buffer from Tartu city administrative borders (see Figure 2). For the purposes of the research Tartu borders before the 2017 administrative reform are considered.

Figure 2. Study area. Tartu city borders before 2017 administrative reform. Highlighted districts: Kesklinn (city centre) and Maarjamõisa as those are the districts where the respondents work or study.

2.2 Data

GPS data was collected with the tracking application for smartphones: MobilityLog, initially called YouSense. The application was developed by the Computer Laboratory of the University of Cambridge and the Mobility Lab of the University of Tartu, it is designed for long-term mobility tracking and social-network-related experiments tracking location with GPS, WiFi access points, accelerometer, and mobile network cells (Ahas et al., 2017). The dataset is provided by the Department of Geography and consists of members of University of Tartu.

For privacy and data protection, no personal information was provided nor any type of identification of the respondents. Each respondent has a unique user ID, which is a random number. The participation of the respondents is voluntary and signing an agreement was required.

The dataset included pre-calculated stops. The stops were calculated by aggregating the sequential GPS points collected with MobilityLog smartphone application. The data about stops provided includes the following information (Table 2):

- user ID which identifies each respondent;
- start and end time for each stop in milliseconds;
- duration of the stops in millisecond which is the difference between the start time and end time in milliseconds;
- start and end timestamps for each stop;
- geometry type: point of each stop.

In addition to the stops calculated based on GPS data, social characteristics of each respondent such as their academic role and work/study place was provided.

Table 2. Example of the data structure of the stops.

	user_id	millis_start	millis_end	duration_millis	time_start	time_end	geom
1	20	1.422761e+12	1.422766e+12	4676000	2015-02-01 05:22:28+02	2015-02-01 06:40:24+02	0101000020E50C0000F0A14FAD5D32244136C0F4EF05B25...
2	20	1.422766e+12	1.422779e+12	13154000	2015-02-01 06:45:01+02	2015-02-01 10:24:15+02	0101000020E50C0000BEDF72366F32244130467CD606B25...
3	20	1.422782e+12	1.422784e+12	2261000	2015-02-01 11:12:53+02	2015-02-01 11:50:34+02	0101000020E50C0000FB022AFE84322441A35C51A206B25...
4	20	1.422784e+12	1.422785e+12	343000	2015-02-01 11:54:04+02	2015-02-01 11:59:47+02	0101000020E50C00000EDC7F1596322441B1699CB405B25...
5	20	1.422790e+12	1.422859e+12	69240000	2015-02-01 13:30:28+02	2015-02-02 08:44:28+02	0101000020E50C000066F8DF81713224417F8F41BA03B25...
6	20	1.422860e+12	1.422868e+12	8404000	2015-02-02 08:54:14+02	2015-02-02 11:14:18+02	0101000020E50C0000651FA3938F322441CA8CB64D08B25...

Both the data of stops and the social characteristic: academic role of each respondent are used for the analysis. The month covered by the data is March for the years 2015, 2017 and 2018, notice that respondents have different starting and ending times of the data sequence. The number of data collection days varied individually, this means that each individual had a different number of days in which data was collected.

The initial dataset is composed of 177 respondents from which 124 were students and 53 staff. Due to the higher number of students than staff members, to avoid possible methodological flaws and biased results 50 students and 50 staff members were randomly selected by their user ID. The random sample for students and staff members was obtained with a formula for random sampling in Excel which provided random user IDs from each of the groups.

From the initial dataset, complementary variables are calculated for the analysis (Figure 3). First the social characteristics were joined by the user ID to the stops data to have one dataset with stops calculated based on GPS points and social characteristics.

In the next step new variables for the analysis were calculated, resulting in 13 variables in the end:

- duration of the stop in minutes, calculated from duration of stops in milliseconds;
- hour when the stop started calculated from the start timestamp;
- number and name of the day (i.e. 1, Monday) when the stop started calculated from the start timestamp

In the calculation of the new variables the combination of software PostgreSQL and RStudio was used.

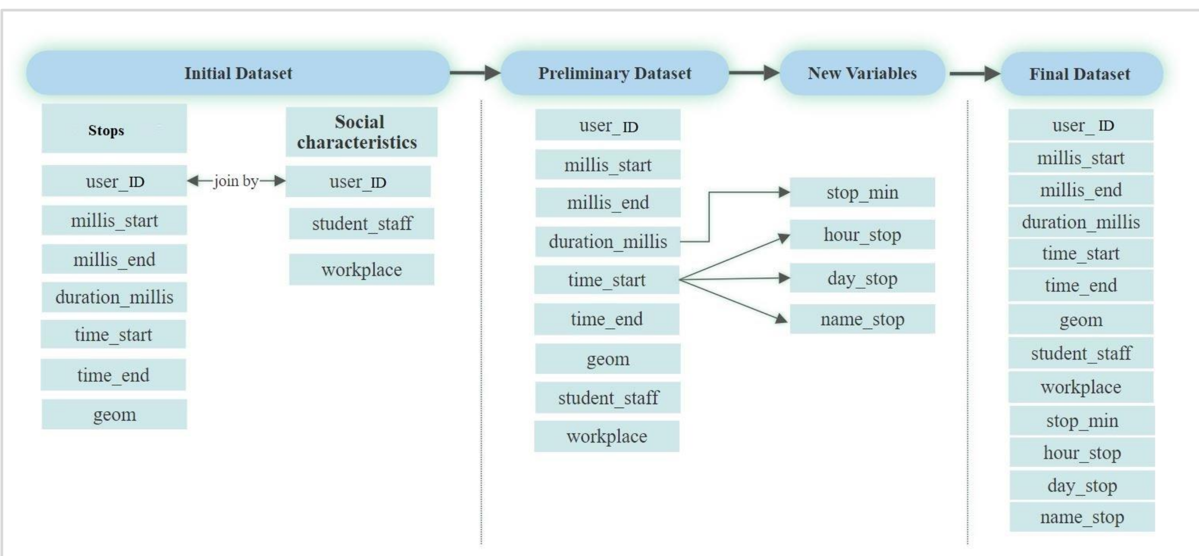

Figure 3. Dataset pre-processing and structure.

2.3 Methodological outline

The research was conducted in three major phases to achieve the desired objectives. The temporal units are type of days, dividing them into weekdays and weekends, and the spatial unit is the extent of the activity spaces depending on the method: minimum convex polygon (MCP), buffer and grid.

In the first phase the activity spaces and interaction spaces with each method were calculated. The second phase focused on generating the activity space indicators. The last phase concentrated on the analysis of the results providing a comparison of the results obtained by each method.

All the calculations are done by using a combination of ArcGIS, PostGIS (geoprocessing) and RStudio (statistical analysis) software.

2.3.1 Calculation of activity spaces

Minimum convex hull (MCP)

By definition the MCP of a set of points is the smallest convex set that contains those points (Barber et al., 1996). The MCP for each user is calculated with the *Minimum Bounding Geometry tool* from the ArcGIS *Data Management package*. The MCP for each respondent was aggregated with the *Dissolve tool*, by academic role and type of day. As a result the MCP activity spaces for staff and students for weekdays and weekends were obtained.

As student's activity spaces are the result of aggregating the MCP of each student, these activity spaces contain only the stops corresponding to students. The same applies to staff activity spaces.

Buffer

The buffer approach was inspired by the daily path area method (DPA) (see section 1.5). Since the data covered the activity locations of March for the years 2015, 2017, 2018 and the analysis is performed for all the years (aggregated), it is not possible to include the routes based on sequential activity locations. Therefore, only the buffer component from DPA is included.

To calculate the activity space a 50 meters buffer is generated around each stop. This calculation was performed using the *Buffer tool* from ArcGIS *Data Management package*. Then the buffers were aggregated with the *Dissolve tool*, by academic role and type of day, obtaining the buffer activity spaces for staff and students for weekdays and weekends. Student's activity spaces are those buffer areas containing only student's stops. Staff activity spaces are those buffer areas containing only staff stops.

Grid

To calculate the grid based activity spaces a 50x50 meters grid is created covering the study area. The grid was created with the create *Fishnet tool* from the ArGIS *Data Management* toolbox. Then the layer containing all the stops was subset by weekdays and weekends. To define the grid cells as activity spaces each subset of stops (weekday, weekend) was intersected with the grid (*Intersect tool* from the *Analysis toolbox*). Then the resulting layers were exported to PostGIS where queries were ran to find the grid cells containing only staff or student stops, separately. The resulting table was then exported back to the ArcGIS environment and joined with the grid, using the feature *Join Field* from the *Data Management toolbox*. As a result two activity spaces were obtained: student's activity spaces which are the grid cells containing only student's stops, and staff activity spaces which are the grid cells containing only staff stops.

2.3.2 Calculation of interaction spaces

The interaction spaces were calculated following the same dynamic with MCP and buffer method. To obtain the interaction spaces the staff and student activity spaces were intersected with the *Intersect tool* from the *Analysis toolbox*.

The calculations to obtain interaction spaces with the grid method were more complex. Since the activity spaces are represented by a grid cell and determined by the stops within each grid cell, staff/student activity spaces are the grid cells that specifically contain staff/student stops. Whereas, interaction spaces are the grid cells that contain at least one staff and one student stop. Thus, interaction spaces are not defined by the intersection of areas but the stops within each grid cell.

To calculate grid interaction spaces the following methodology was applied. After each layer containing the stops by type of day (weekday, weekend) was intersected with the grid covering the study area, the resulting layers were exported to PostGIS where queries were ran to find the number of unique staff and students per grid cell, separately. Each layer was exported back to the ArcGIS environment and joined with the grid by the cell ID, using the feature *Join Field* from the *Data Management toolbox*. In the settings of the join field tool it should be specified that only the matching records should be kept, in that way, only the grid cells that contain at least one staff stop and one student stop are displayed.

2.3.3 Calculation of activity space indicators

In order to characterise activity spaces and interaction spaces four indicators are calculated for each method:

- extent of the activity spaces and interaction spaces (km^2);
- unique number of staff and students number per interaction spaces;
- distribution of interaction spaces across Tartu city districts;
- number of stops within interaction spaces.

2.3.3.1 Extent of activity spaces and interaction spaces

The extent of activity and interaction spaces was obtained by calculating the areas in km^2 of the geometric shapes of the activity and interaction spaces. The same methodology was used to calculate the areas of activity and interaction spaces across the three activity space methods: using the tool *Calculate Geometry* from the *Data Management toolbox*. As a result the staff and student activity spaces, and interaction spaces areas for weekdays and weekends were obtained. It is important to notice that based on how MCP activity spaces were calculated (section 2.3.1), it was only possible to calculate the areas and mean areas for MCP activity space based on academic role and type of day.

The total areas of staff and students activity spaces included the areas of interaction spaces. Therefore, to obtain only the staff and student activity space areas, the area of the interaction spaces was excluded. For the MCP and grid method this was performed by using the geoprocessing tool: *Select by Location*. On the other hand, with the buffer method the dynamic was different. Buffer activity spaces were not calculated per respondent but based on the stops. Hence to obtain only the staff and student activity space areas the interaction area was subtracted from the total staff and student activity space areas, separately for weekdays and weekends, since interaction areas are different for each type of day.

2.3.3.2 Unique number of staff and students per interaction spaces

The unique number of staff and student per interaction spaces was calculated for the three methods following the same methodology. First, the layer containing the stops of all respondents was intersected with the layer containing the interaction spaces, obtaining the stops within each interaction space. Second, the resulting layer was exported to PostGIS where queries were ran to find the number of unique staff and students per interaction space, separately. Finally, the layers containing the number of unique staff and unique students per interaction space were exported back to the ArcGIS environment and joined with the layer containing the interaction spaces by each interaction space ID. As a result a layer with the interaction spaces containing the number of unique staff and students was obtained. Then the difference between unique staff and student number per interaction spaces is calculated.

The difference between unique staff and student number per interaction spaces is calculated following the same dynamic for the MCP and buffer interaction spaces. It is obtained by subtracting the unique staff number and unique student number per interaction space. The subtraction (Sb) is as follows:

$$Sb = unique\ number\ of\ staff - unique\ number\ of\ students$$

As a result a difference map is presented with 3 activity space categories that should be interpreted as follows:

- The neutral interaction spaces: there is the same number of staff and students.
- Staff dominated interaction spaces: there is a higher number of staff.
- Student dominated interaction spaces: there is a higher number of students.

The dynamic to calculate the difference between unique staff and student number per interaction spaces is different with the grid method. The staff and student activity spaces are converted to raster layers (*Feature to Raster tool from Conversion toolbox*). The pixel size is 50m and the value of each pixel is defined by the count of unique students and staff, respectively. Therefore, each pixel of the new raster layers will contain the number of unique students or staff.

The subtraction of the student raster layer from the staff raster layer is performed with map algebra (*Raster Calculator tool from Spatial Analyst toolbox*) to follow the same criteria as with the other methods (unique number of staff – unique number of students). The result is a new raster layer containing the values of this subtraction.

Then the raster layer is reclassified (*Reclassify tool from Spatial Analyst toolbox*) to identify and classify the staff dominated, student dominated and neutral interaction spaces with the criteria in Table 3. As a result a difference map is presented with 3 activity space categories described above.

Table 3. Criteria for raster reclassification.

Grid cell value	Type of activity space	Reclassification value
0	Neutral	0
Positive	Staff dominated	1
Negative	Student dominated	2

2.3.3.3 Distribution of interaction spaces across Tartu city districts

The distribution of interaction spaces across the study area is analysed by calculating the percentage of interaction space area per district *(% ISD)*, the same methodology was used for the buffer and grid method (see explanation in section 3.2).

First the interaction spaces layer was intersected with the Tartu city districts layer (*Intersect tool from the Analysis toolbox*), obtaining the distribution of the interaction spaces within Tartu city districts. Then the area in km^2 of interaction spaces within each district is calculated (*Calculate Geometry tool from the Data Management toolbox*). Finally, with the *field calculator*, the percentage of interaction space per district with respect to the total area of the interaction space was computed. The total areas of interaction spaces are different during the weekdays and weekends. Therefore, the percentage of interaction spaces per district are calculated separately based on the total area of interaction spaces per weekdays and weekends as follows:

$$\%\boldsymbol{ISD} = \frac{weekdays\ ISD\ (km^2)}{total\ weekdays\ ISD\ (km^2)} * 100 \qquad \%\boldsymbol{ISD} = \frac{weekends\ ISD\ (km^2)}{total\ weekends\ ISD\ (km^2)} * 100$$

2.3.3.4 Number of stops per respondent within interaction spaces

The number of stops per interaction spaces is calculated following the same dynamic for the three methods. All the stops were intersected with each interaction space (per method, separately), obtaining only the stops within the interaction space. Then the number of stops was quantified based on academic role (student, staff) and type of day (weekday, weekend).

The distribution of stops per respondent within interaction spaces is calculated following the same dynamic for the three methods. First, all the stops were intersected with each interaction space (per method, separately), obtaining only the stops within the interaction space. Second, the resulting layer was exported to PostGIS. In PostGIS by writing queries, the number of stops per respondent were obtained based on academic role (student, staff) and type of day (weekday, weekend). No outliers are removed because that would mean eliminating a respondent from the analysis and therefore bias in the results.

2.3.4 Statistical analysis

Statistical analysis was performed to explore whether there was a statistically significant difference:

1. between weekdays and weekends in the *extent of staff and student activity spaces* calculated by MCP method
2. between student and staff median values in *number of stops per respondent within interaction spaces* calculated by MCP, buffer and grid methods

1. First, the distribution of the areas per respondent based on type of day is tested with the Shapiro-Wilk normality test. The Shapiro-Wilk normality test allows to test the normality of the data to select the correct statistical test for further analysis (parametric and non-parametric) (Karadimitriou, 2019). The data resulted not normally distributed (weekdays: $p = 1.46e\text{-}10$; weekends: $p = 2.12e\text{-}11$). Second, the difference between staff and students' activity space mean areas per respondent during weekdays and weekends is analysed by comparing the medians of the activity space areas. Mann-Whitney non-parametric test is used to examine the significance of the difference of the mean areas per respondent based on type in type of day and academic role. Mann-Whitney test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed (Lund Research Ltd, 2018).
2. To explore the variation of the spatial behavior between staff and students during weekdays and weekends within interaction spaces, the analysis of outliers, interquartile range (IQR) and

medians is performed. Outliers are objects that have very different behavioural attribute values from those of their surrounding spatial neighbors (Aggarwal, 2016). They could be good indicators to analyse how respondents' spatial behaviour varies based on the academic role during weekdays and weekends. The interquartile range (IQR) is a measure that expresses how spread is the data with respect to median value and it is mostly used for comparison purposes (Day, 2018). The interquartile range could provide insights into repetition across activity types (Buliung et al., 2008). Finally, to test the significance of the difference in median values of number of stops per respondent the Mann-Whitney test is used.

3. Results

3.1 Comparison of the activity spaces by different methods

Extent of activity spaces

Each method provided a different value of the extent of the activity spaces: the extent of the activity space was the largest for MCP, followed by buffer, and grid (Figure 4). MCP depicts overgeneralised activity space areas: they are approximately 19 times larger than the extent calculated by the buffer method and approximately 46 times larger when calculated with the grid. This is due to the fact that MCP calculation depends on the distribution of the stops across the study area rather than the concentration or dispersion of the stops, thus the further a stop is located the larger area the MCP will capture.

Compared to MCP, buffer based activity spaces are considerably smaller. Finally, grid activity spaces are the smallest but similar with buffer. They represent more detailed area estimates of activity spaces because it captures and defines an activity space based on the stops in a 50x50 meters grid cell, and includes the areas where people actually visited.

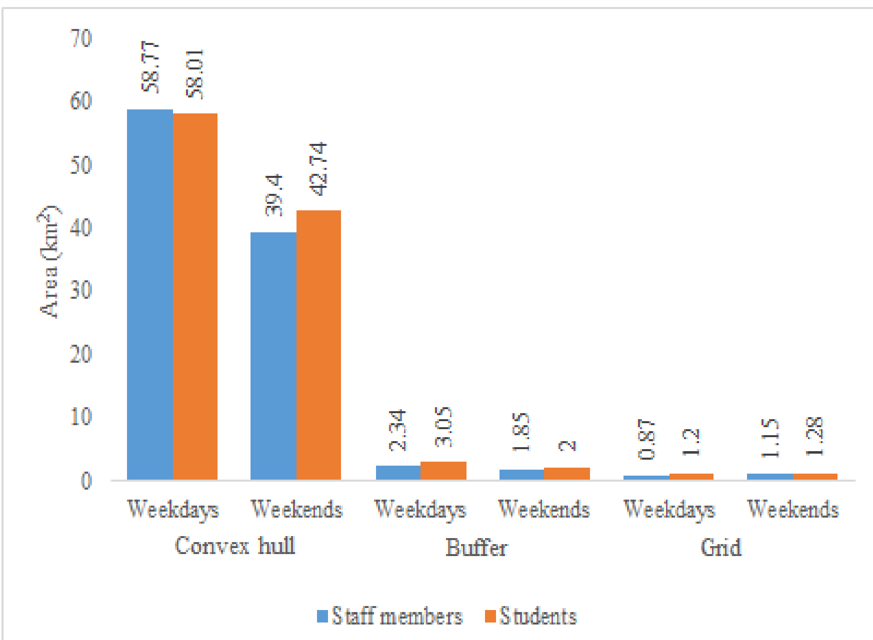

Figure 4. Extent of staff and student activity spaces by method.

Regarding the activity space for staff and students during weekdays, similar results were obtained with buffer and grid methods (Figure 4): student’s activity spaces are larger than staff activity spaces during the weekdays. MCP based calculations were exceptional in this regard: staff activity spaces are larger than student activity spaces but the difference is quite small.

During the weekends, student's activity spaces are larger than staff activity spaces across the three methods. The difference in the extent between student and staff activity spaces were similar for the buffer and grid method (0.15 and 0.13 km^2, respectively). On the other hand, for MCP calculations indicated that students' activity spaces are 3.34 km2 larger than staff.

Even though the area sizes are different, MCP and buffer methods indicate similar results across weekdays: the extent of activity space is larger during the weekday than weekend. Interestingly, calculations based on grid method, indicate the opposite: the extent of activity spaces is larger during weekends both for the staff and students.

For the MCP method it was possible to perform a statistical analysis to find the significance of the difference between the extent of staff and student activity spaces during weekdays and weekends. Results showed that the differences were not statistically significant (weekdays:$p = 0.3027$; weekends: $p = 0.1105$). Differently, with buffer and grid it was not possible to perform a comparative statistical analysis, as discussed in section 2.3.2.1.

To conclude, with buffer and grid methods the spatial units that represent activity spaces (50 meter buffer and 50x50 meters grid cell) are smaller than MCP providing a more detailed and realistic insight of the activity spaces (Figure 4).

Use and distribution of space across Tartu city

To determine the differences of the use of space, two types of activity spaces were calculated: student and staff activity spaces for each method. Student and staff activity spaces are the areas that contain only staff and students stops as described in section 2.3.1. Therefore these are the areas where segregation may increase due to the lack of interaction between staff and students. The interaction areas are included in figures to better identify the staff and student segregated areas. The difference of the use of space is displayed differently by each method.

MCP activity spaces are depicted as large polygons covering various districts from Tartu city and districts from other municipalities both during the weekdays and weekends. Since MCP calculations are overgeneralised, staff and student activity spaces do not reflect the segregated areas within Tartu city (Figure 5).

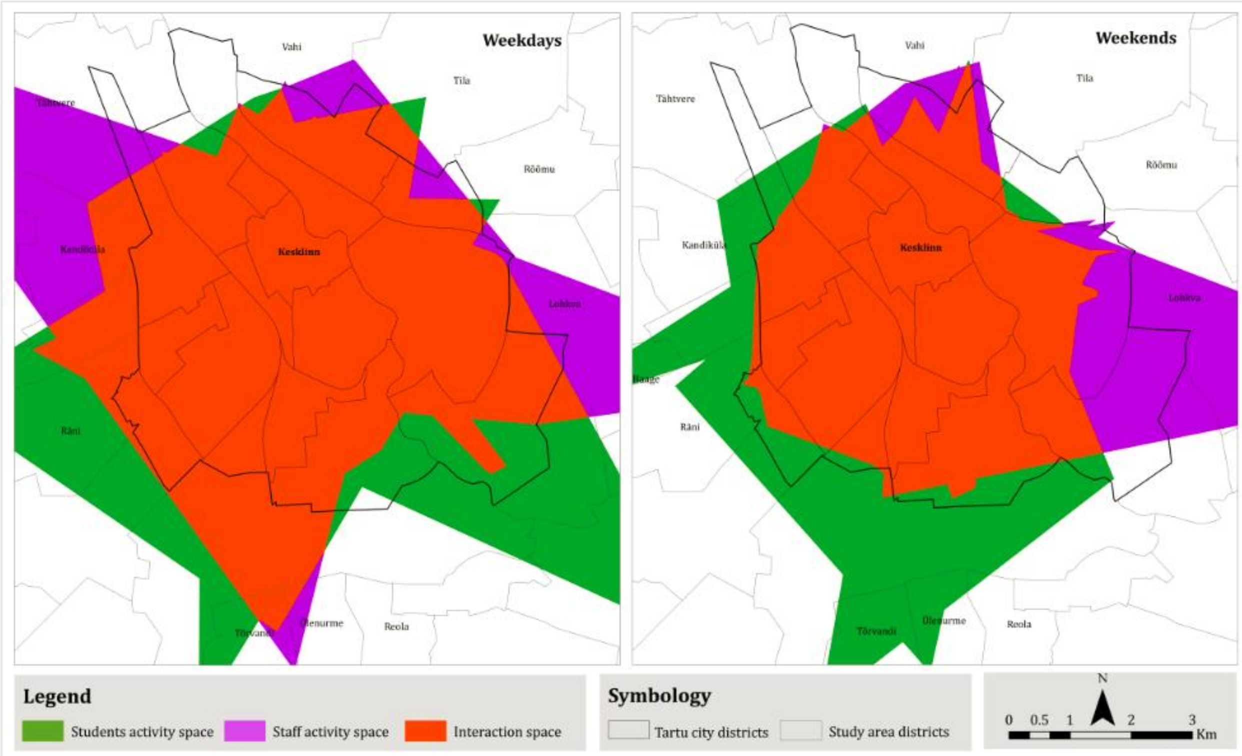

Figure 5. Aggregated MCP activity spaces based on the type of day and academic role. Note: interaction spaces included to identify staff and student segregated areas.

Compared to MCP, buffer (Figure 6) and grid (Figure 7) activity spaces indicate more realistic results, where segregated places are identifiable. Both methods indicate that during the weekdays and weekends, Kesklinn, Ülejõe and Annelinn are the districts where students have higher concentration of activity spaces compared to staff. Ülejõe and Annelinn might be related to home location. In Ülejõe district university dorms are located and in Annelinn the prices of renting a flat could be more accessible to students compared to other districts. Kesklinn concentrates facilities such as university buildings, banks, shopping malls or coffee places that in terms of distance might be more accessible to students. On the other hand, Veeriku, Supilinn and Ränilinn are the districts where staff have higher concentration of activity spaces compared to students, both during weekdays and weekends. This could be related to home location: in these districts housing prices might be less accessible to students. From Figure 6 and 7, it is visible that students tend to move within districts that are near to Kesklinn (city centre), whereas staff tend to move in further districts from Kesklinn. This could be for several reasons, for instance: housing price, as mentioned before, but also transportation mode, staff might have higher chances to own a car compared to students.

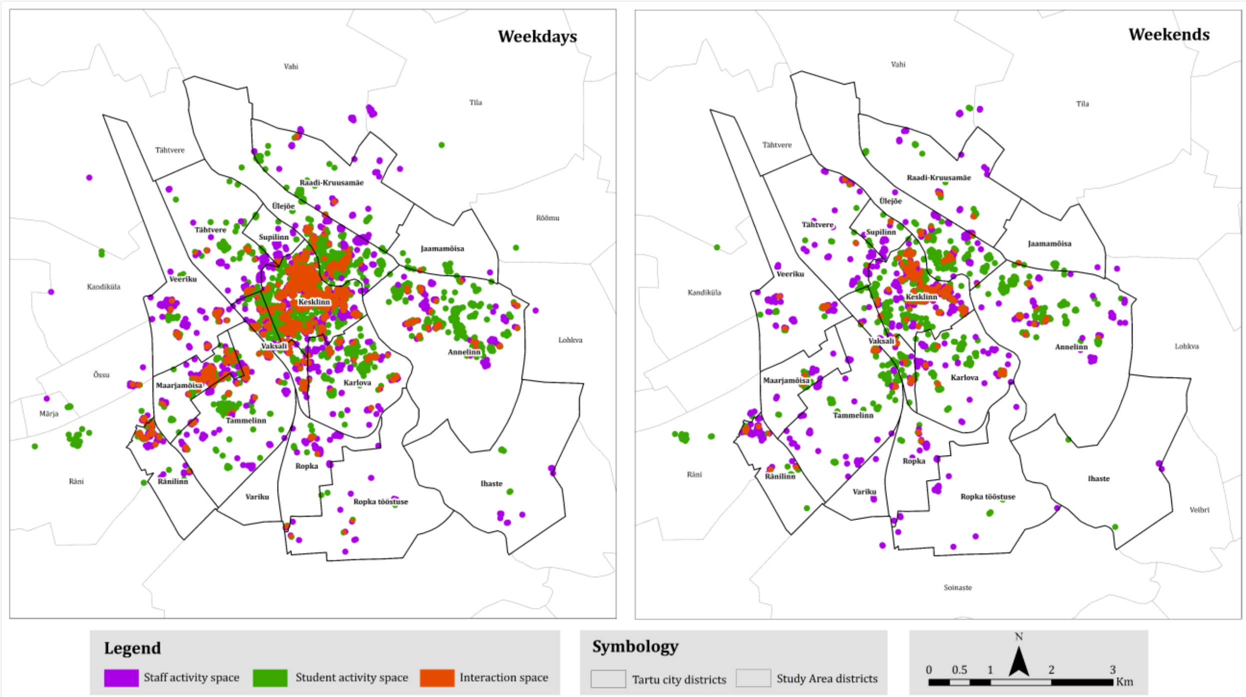

Figure 6. Aggregated buffer activity spaces based on the type of day and academic role. Note: interaction spaces included to identify staff and student segregated areas.

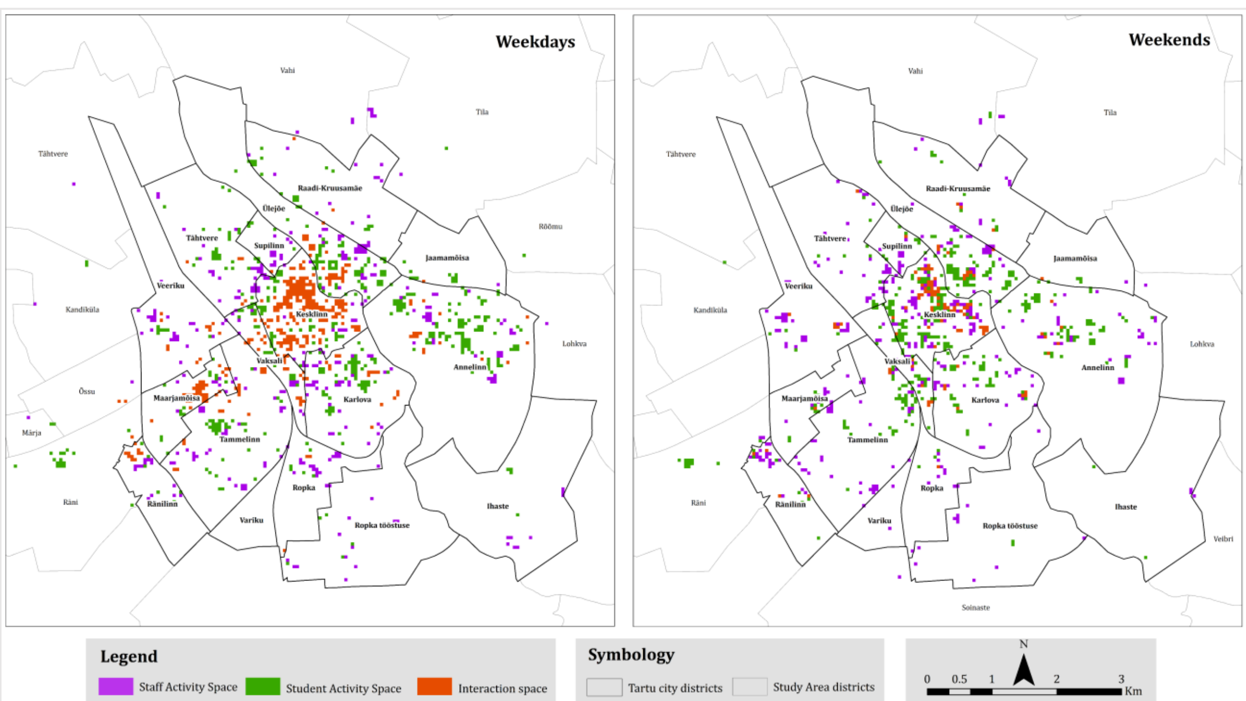

Figure 7. Aggregated grid activity spaces based on the type of day and academic role. Note: interaction spaces included to identify staff and student segregated areas.

Even though, buffer and grid depicted similar results about the segregated districts, grid activity spaces are more localised presenting a more detailed distribution of the space across Tartu city (Figure 7). The level of detail that the grid activity spaces provide, makes the segregated areas easier to identify not only across the city of Tartu but also within each district. Therefore, grid might be the most suitable method to measure segregation because it displays realistic activity spaces that might indicate localised segregated areas at city and district level.

3.2 Comparison of the interaction spaces by different methods

Extent of interaction spaces

The extent of interaction spaces is different for each method: MCP provides the largest extent of interaction space (66.73 km^2), followed by buffer (2.95 km^2), and grid (1.19 km^2) (Figure 8). MCP interaction spaces are overgeneralised: they are approximately 23 times larger than the extent of buffer interaction spaces, and approximately 56 times larger than the extent of grid interaction spaces.

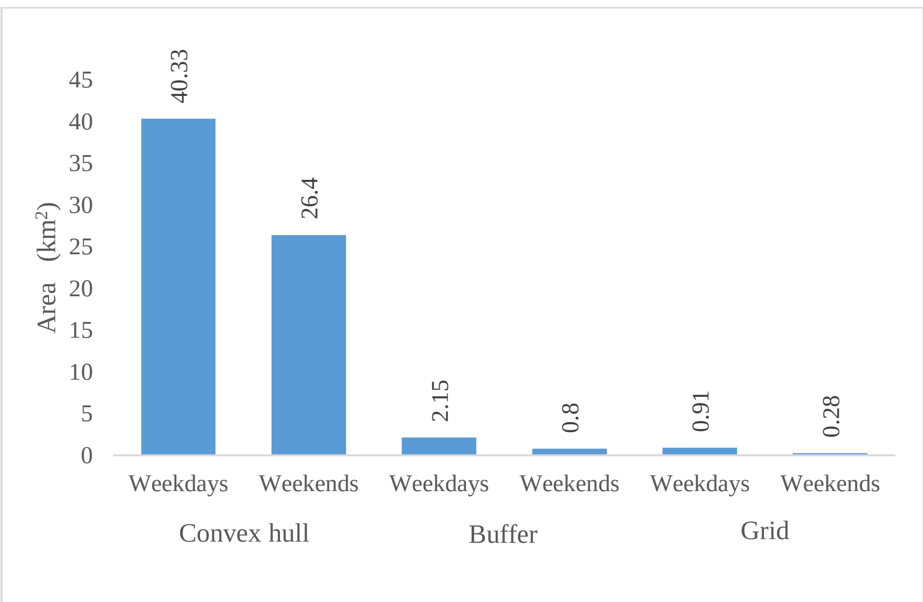

Figure 8. Extent of interaction spaces by method.

Since interaction spaces are the areas where there are higher opportunities that staff and students encounter, MCP calculations indicate non-realistic results because they suggest that staff and students have the same chances to meet across all Tartu city. Buffer interaction spaces are smaller and indicate more realistic interaction spaces across Tartu city, compared to MCP. However, grid calculations represent the most realistic and detailed results because they indicate more localised and smaller areas of interaction spaces, and they represent the spaces where there is actually at least one staff and one student in a 50x50 meter grid cell.

Regarding the interaction spaces during weekdays and weekends, similar results were obtained for the three methods: interaction spaces during the weekdays are larger than interaction spaces during weekends. This suggests that staff and students have higher opportunities to encounter across Tartu city during weekdays.

To conclude, each method provided different results to represent the extent of interaction spaces. However, grid calculations indicate the most realistic and localised interaction spaces across Tartu city compared to buffer and MCP calculations.

Use and distribution of interaction space across Tartu city

To determine the differences of the use of space, three categories of interaction spaces were calculated: neutral, staff and student dominant areas for each method. However, the difference of the use of space is identified only with the buffer and grid methods. Due to the overgeneralised interaction spaces that MCP depicted, it is not possible to identify differences in the use of space with this method (Figure 9).

MCP interaction spaces are depicted as whole polygons that cover almost all the city of Tartu, making no distinction in the differences of the use of space between staff and students within interaction spaces. MCP calculations showed that all the student and staff stops are included in the MCP interaction spaces, making no differentiation in the use of space within interaction spaces, therefore interaction spaces have been defined as neutral space because there are the same number of staff and students (see Figure 9). Although MCP interaction spaces during weekdays and weekends are overgeneralised, they could be considered as the extreme boundaries where staff and students are more likely to encounter. Outside the MCP boundary the chances that staff and students are in the same space is almost null.

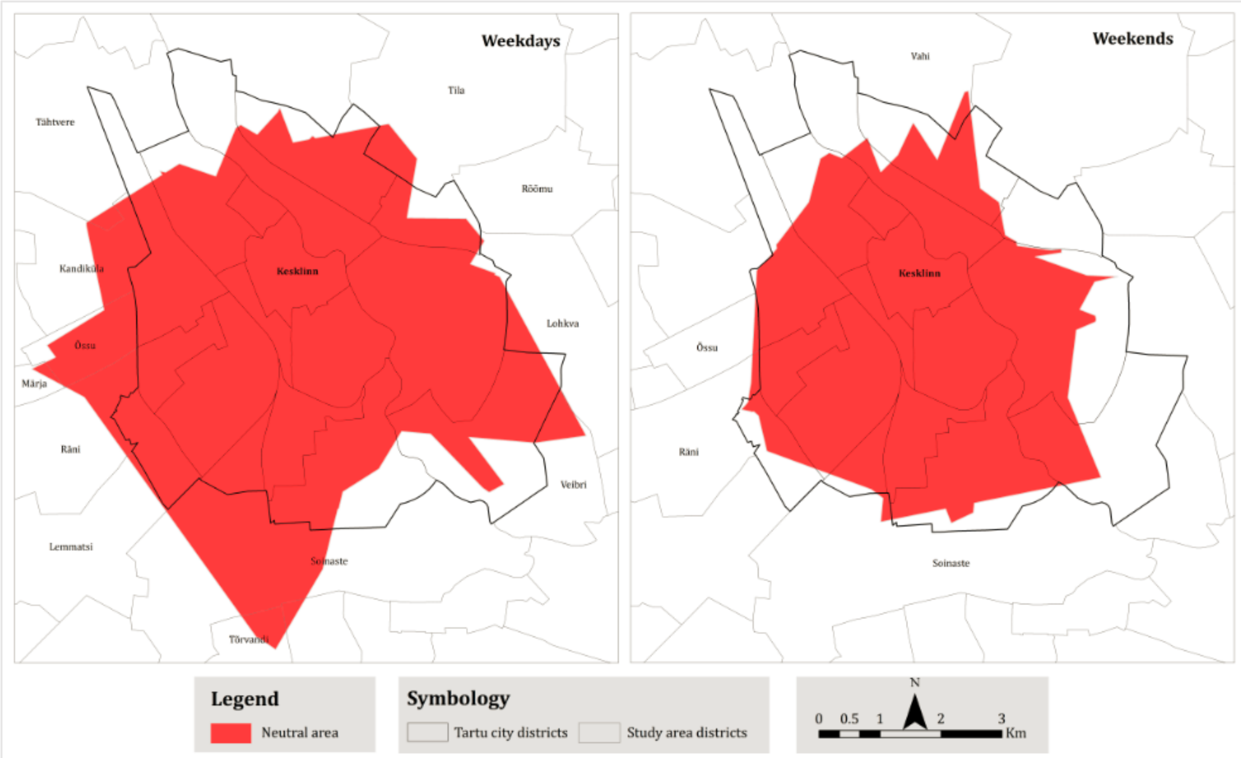

Figure 9. Differences in the use of space within MCP interaction spaces by type of day.

Compared to grid results, buffer interaction spaces are easier to identify (Figure 10). Grid interaction spaces are harder to identify because each grid interaction space has the same area: 0.0025 km2 (50x50 meter grid cell) (Figure 11). However both methods indicate that during weekdays, interaction spaces are more spread across Tartu city compared to weekends. Additionally, student dominated areas are clustered to certain places across Tartu city, compared to staff dominated areas which are more scattered (Figure 10, 11).

During the weekdays, the buffer method provides 93 interaction spaces, whereas grid method 364. During the weekends, the buffer method provides 55 interaction spaces, whereas grid method 115. This indicates that compared to the grid method, the buffer provides more generalised results about the difference of use of space between staff and students across Tartu city both during weekdays and weekends. The level of detail that the grid method provides is higher than the buffer method.

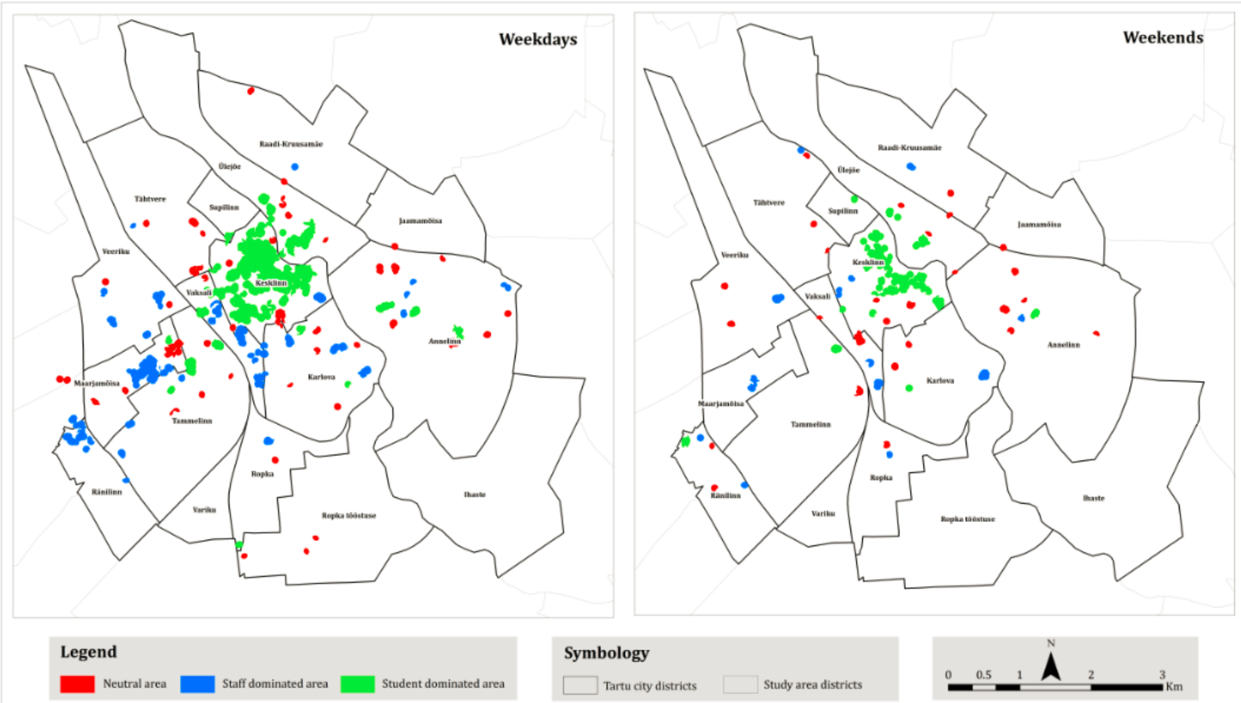

Figure 10. Differences in the use of space between staff and students within buffer interaction spaces by type of day.

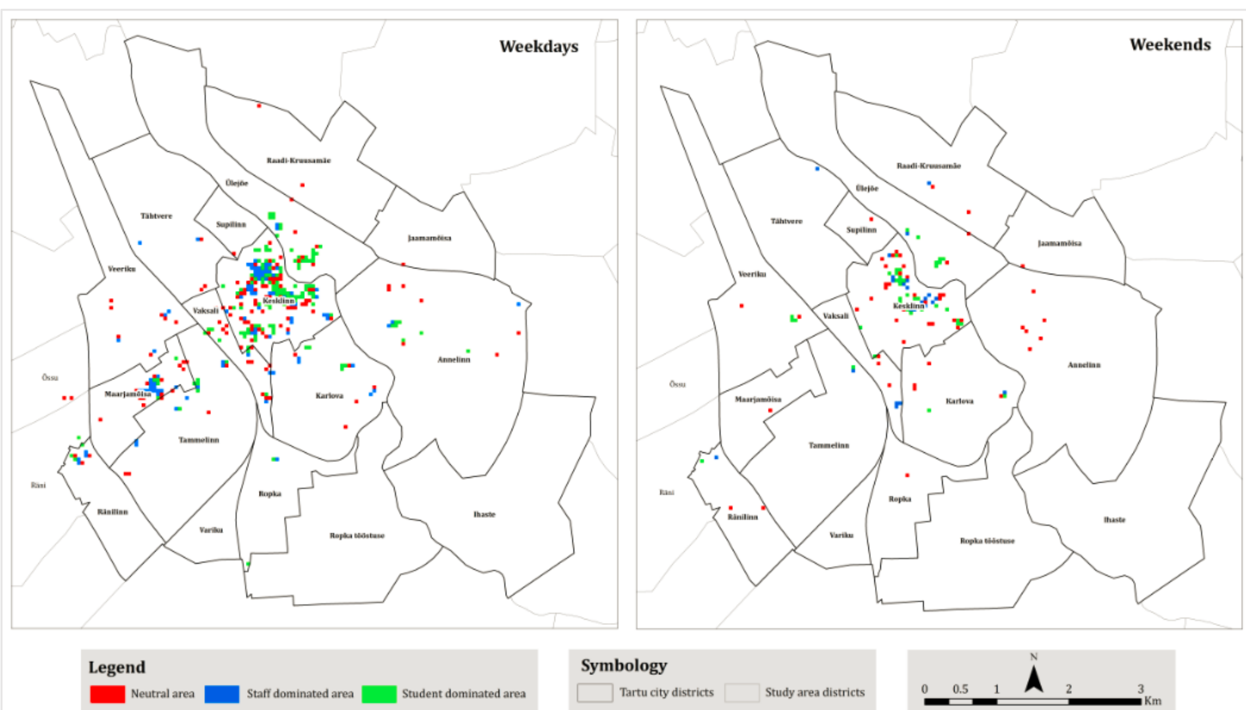

Figure 11. Differences in the use of space between staff and students within grid interaction spaces by type of day.

During the weekdays buffer calculations indicate that the student dominant interaction space with the highest number of students (15 students more than staff) is located in Ülejõe district. This might be mainly because in this area of the district some student facilities, such as dormitories, are located. There are two staff dominant interaction spaces with the highest number of staff (12 staff more than students) located in Maarjamõisa and Veeriku districts. This might be mainly because in this area of the district new Tartu University buildings are located (Figure 10). On the other hand, grid calculations indicate three student dominant interaction spaces with the highest number of students (9 students more than staff) during weekdays all located in Kesklinn district, this might be due to the presence of several university buildings and other places like grocery stores or shopping malls. There are six staff dominant interaction spaces with the highest number of staff in Kesklinn (3), Maarjamõisa (2) and Vaksali (1) districts (Figure 11).

During the weekends, buffer calculations indicate one student dominant interaction space with the highest number of students (10 more students than staff) which is located in Kesklinn district, this could be due to the presence of pubs, night clubs and restaurants that might be more frequented particularly by students during the weekends compared to staff (Figure 10). On the other hand, one staff dominant interaction space with the highest number of staff (4 staff more than students) is located in Ränilinn district. Grid calculations, on the other hand, showed that during weekends there is one student dominant interaction space with the highest number of students (4 students more than staff) is located in Ülejõe district. There are two staff dominant interaction spaces with the highest number of staff (5 students more than staff) are located in located Kesklinn and Karlova districts (Figure 11).

Although the buffer and grid method showed that staff and student dominated areas are located in different districts, both methods indicate that the number of interaction spaces during the weekdays is higher compared to weekends. This suggests that during the weekdays the chances that staff and students interact are higher than during the weekends. However, compared to the buffer method, grid calculations indicate more localised results of those spaces where staff and students may interact across Tartu city.

The percentage of interaction spaces per district is calculated for the buffer and grid methods to explore the distribution of the interaction spaces across Tartu city, and to determine the most visited districts during the weekdays and weekends.

For the buffer and grid method (see Table 4), during weekdays Kesklinn (buffer: 40.97%; grid: 50.99%) is the most visited district followed by Maarjamõisa (buffer: 8.74%; grid: 9.93%) and Ülejõe (buffer: 8.03%; grid: 8.16%). As the most visited districts during the weekdays this indicates that most of the activities could be related to routine activities such as work or studies, and residential purposes.

Table 4. Top three most visited districts by type of day with buffer and grid interaction spaces during weekdays. Note: Percentage is calculated separately based on total area of interaction spaces per weekdays per method.

	Weekdays	
	Percentage	
District	**Buffer**	**Grid**
Kesklinn	40.97	50.99
Maarjamõisa	8.74	9.93
Ülejõe	8.03	8.16

During weekends, both methods showed that Kesklinn (buffer: 49.09%; grid: 57.05%) is also the most visited district followed by Ülejõe (buffer: 8.03%; grid: 8.69%). However, the buffer method depicted that the third most visited district is Annelinn (7.61%) and the grid method Karlova (6.25%) (Table 5). This difference could be due to the different level of detail that each method provides. As the most visited districts during the weekends this indicates that most of the activities could be related to residential and leisure purposes. For instance, Annelinn district has the largest residential area in Tartu city formed by Soviet apartment buildings, and Karlova district has become a trendy district due to the several different emerging coffee places, pubs and the presence of parks as well as schools.

Buffer and grid method depicted an interesting result: while Maarjamõisa district is the second most visited district during weekdays (buffer: 8.74%; grid: 9.93%), it is the least visited during weekends (buffer: 2.56%; grid: 1.79%). This reveals that most of the activities within Maarjamõisa district could be related to study and work purposes.

Table 5. Top three most visited districts by type of day with buffer and grid interaction spaces during weekends. Note: Percentage is calculated separately based on total area of interaction spaces per weekdays per method

	Weekends	
	Percentage	
District	**Buffer**	**Grid**
Kesklinn	49.09	57.05
Ülejõe	8.03	8.69
Annelinn	7.61	-
Karlova	-	6.25

Additionally, buffer and grid calculations during weekdays and weekends indicate that interaction spaces are concentrated within Tartu city, with the exception of a small area during weekdays located in Räni and Õssu districts outside Tartu city borders.

To conclude, the percentage of interaction spaces per district showed that although the total areas of interaction space vary with buffer and grid methods, the most visited districts are the same during the weekdays and weekends. During the weekends there is a small variation with the third most visited district.

Number of stops per respondent within interaction spaces

The number of stops in relation to the interaction space per method is reviewed. Results showed that the number of stops vary across different methods: the larger the interaction space extent the more stops it will capture. Overall MCP, as expected, captures a higher number of stops (17729), followed by the buffer (8174) and grid method (5573) (Table 6).

Even though the differences in the number of stops during weekdays and weekends per method: MCP depicted 10511 more stops during the weekdays than weekends, followed by buffer and grid (6096 and 4505, respectively), the stops captured by the three methods during the weekdays are larger than in weekends suggesting that mobility during weekdays is higher than in the weekend.

Table 6. Number of stops within interaction spaces per method.

	Number of stops		
	Weekdays	Weekends	Total
MCP	14120	3609	17729
Buffer	7135	1039	8174
Grid	5039	534	5573

The smaller the interaction space extent, the smaller difference of the median values, outliers and interquartile ranges of the distribution between staff and students. Therefore the number and distribution of stops have a directly proportional relationship with the extent of the interaction spaces. Gradually the number and distribution of the stops per respondent decrease with each method (Table 7), this impacts the level of detail and type of the results that each method provides.

Each method provided different results (see Table 7). MCP results were the easiest and straightforward to interpret compared to the other two methods: MCP outliers, median and IQR values indicated that within interaction spaces, students have a greater variation in spatial behaviour compared to staff members, and the difference is statistically significant both during the weekdays and weekends.

Table 7. Number of stops per respondent difference across three methods. [a] Interquartile range (IQR); [b] p-value from Mann-Whitney test; * significant p-value

		Weekdays				*Weekends*			
		Median	Outliers	IQR[a]	p^b	Median	Outliers	IQR[a]	p^b
MCP	Staff	82	1	76.87	4.921e-07*	19	1	31.87	0.0005*
	Students	163	0	265.5		32.50	2	90.75	
Buffer	Staff	50	1	57.75	0.09	6	5	9	-
	Students	58.50	8	13.5		8	5	13.5	
Grid	Staff	42	1	48.75	0.63	4	5	4.5	-
	Students	38	5	52.5		4	4	5.25	

Finally, grid results showed a different perspective of the spatial behavior between staff and students within interaction spaces (Table 7): during weekdays, median values suggest that staff mobility is slightly higher than students, but was not proven to be statistically significant. However, outliers and IQR indicates that student spatial behaviour is more variable than staff. This means that although staff mobility is slightly higher, students are visiting more places compared to staff.

During weekends, median values indicated no variation in the student and staff mobility (Table 7). However, outliers and IQR indicates that student spatial behaviour is slightly more variable than staff. As with the buffer method, significance was not possible to test because only the stops from 43 staff and 47 students were identified. This indicates that during the weekends, the 7 staff and 3 students may have not interacted at all with the rest of the respondents of the sample, which suggests that the opportunities that staff and students encounter are lower during the weekends.

To conclude, each method provided different insights about the spatial mobility of staff and students within interaction spaces. The number of stops is highly related to the extent of the interaction spaces that each method provided, therefore the level of detail of the results is different for each method. MCP provided very general results about the spatial behavior of the respondents, compared to the buffer and the grid method. The grid method provided results that were more detailed and bringing out additional insights about spatial mobility within interaction spaces.

4. Discussion and conclusions

The comparison of activity space and interaction space measures showed that each method, according to its complexity, depicted different results, corroborating the underlined literature about how different methods produce different results not only in terms of area size but also the characteristics of the spaces (Vich et al., 2017; Wong, 2018).

Minimum convex polygon (MCP), widely used in activity space research and segregation studies (Palmer, 2013), depicted overgeneralised area estimates of activity spaces and interaction spaces. Although it is the easiest to compute and interpret, MCP activity/interaction spaces included areas across the study area where staff and students did not register any activity location, as suggested by Patterson & Farber, (2015). Therefore, MCP results were not realistic and accurate for analysing activity/interaction space characteristics. Furthermore, it might not be a suitable method to analyse activity space based segregation with GPS data, because it does not capture the accuracy that GPS data provides.

Buffer and grid provided significantly reduced activity/interaction space area estimates. Results showed that the extent of the activity/interaction space has an impact on their characteristics. The smaller the activity/interaction space, the more detailed and realistic results were obtained. Since GPS data provides high precision and space-time granularity (Wong & Shaw, 2011), methods such as buffer and grid capture such data granularity and represent more realistic and accurate activity/interaction space areas that can be more suitable to analyse segregation.

The buffer and grid methods required a more complex computation and analysis of results and both imply different methodological approaches that were considered when calculating and analysing the results.

The buffer widths for the analysis should be chosen considering the maximum accuracy error recorded by GPS network signals and the spatial settings of the study area (Lee et al., 2016; Vich et al., 2017; Zenk et al., 2018). The buffer width for the analysis was chosen considering the spatial context of Tartu city area, however the maximum accuracy error for the collection of data was not considered. Since the data for the study was collected individually through a mobile phone application, each smartphone has a different GPS system, therefore the accuracy of data collection is different for each respondent. Consequently, activity spaces accuracy may vary for each respondent.

The grid method provided more detailed results that revealed additional patterns within the activity/ interaction spaces in a consistent geographical scale of analysis, confirming Hong et al., (2014) findings that the grid approach does not eliminate all the methodological errors, but it could reveal important patterns that would not be possible to find using conventional measures. However, before selecting the spatial resolution of the grid, the study purpose and the spatial context was analysed as suggested by Hillier et al., (1993). Therefore, diverse scenarios with larger and smaller spatial resolutions of the grid were tested. Each scenario provided a totally different perception of spatial behaviour and activity space sizes, corroborating the grid methodological flaw across literature: the modifiable areal unit (Hong et al., 2014; O'Sullivan & Wong, 2007; Wong, 2009).

Grid method detailed and accurate results are more suitable with GPS data. However, methods that provide more detailed and accurate results require higher spatial and temporal accuracy of the stops based on GPS data. The combination of accurate measure activity spaces and the granularity that GPS data provides could contribute to future research and decision makers because it could depict interesting and more realistic insights about segregation. For instance, incorporating points of interest across the study area could allow decision makers and urban planners to identify which are the places where people are more segregated and take actions to implement actions that contribute to the integration of those spaces in the cities.

Despite the methodological differences and results that each of the analysed methods provided, some consistent results about activity spaces and interaction spaces were identified. Students registered larger activity spaces than staff, this confirms findings in Vich et al., (2017) research, and contradicts previous research that suggest that mature adults, normally associated to staff members, due to the higher income and family responsibilities tend to have larger activity spaces than younger adults, normally associated with students (Buliung & Kanaroglou, 2006; Ta et al., 2016). However, as suggested by Kwan, (1999), the size of activity spaces was not enough to provide a complete in-depth overview of the segregation phenomenon across Tartu city.

During weekdays interaction spaces showed to be larger than on weekends. Additionally, the number of interaction spaces during the weekdays is higher compared to weekends. Therefore, activity patterns appear to change during the weekend. This could be explained based on Susilo & Kitamura, (2005) and Buliung et al., (2008) findings: during weekdays fixed activity locations for eating, sleeping, working or going to school are carried out, thus spatial behaviour tends to be repetitive and activities appear to be performed in a wider geographical areas. Whereas, during weekends, activity locations are more variable so spatial behavior tends to be more random and less repetitive, and activities appear to be performed over smaller areas.

Interaction spaces revealed that student's spatial behaviour appeared to have a greater variation than staff both during weekdays and weekends, this corroborates Susilo & Kitamura, (2005) findings about temporal variation of individual's activity spaces: for workers and students on weekdays, the spread of activity locations are more stable than on weekends. However, as workers tend to have more rigid activity schedules they will have less variable activity spaces. Masso et al., (2019) also found that younger adults are more spatially active compared to mature adults.

The most visited districts by staff and students during the weekdays and weekends are Kesklinn, Maarjamõisa and Ülejõe. Maarjamõisa as a working/ study district emerges during weekdays. This confirms that a university environment could be considered a potential interaction space because it brings together local and international students and academics in one geographic location (Rose-redwood & Rose-redwood, 2013). Additionally, including some aspects of the urban context of the study area would have helped to understand with more detail why these districts are more visited. For instance, housing amenities and pricing, university buildings, student facilities, shopping malls, restaurants and other places that staff and students might be visiting.

Consistent results were found about the extent of activity spaces, spatial differences of the use of space between staff and students, and characteristics of spatial behaviour within interaction spaces. However, considering that spatial segregation is a complex spatial process, multidimensional and specific phenomenon (Kaplan & Holloway, 2001; Massey & Denton, 1988), the indicators used for the analysis are not sufficient to confirm and state segregation between staff and students. To deepen the analysis to determine segregation additional indicators widely used in segregation literature should be incorporated (i.e. time spent in locations, unique locations, transportation mode, place of residence, place of work) as suggested by Kwan, (1999). Including an activity domain approach would be complementary to understand the semantic meaning of the visit to activity locations (Lee et al., 2016). The temporal dimension of the study could also be more specific, focusing on day by day analysis, or certain hours of the day considering important aspects of the studied groups and the context of the study. For instance, in the case of students and staff, it would be interesting to analyse pick hours during the week or night hours during the weekends.

With the spatial accuracy that GPS data provides, there is the opportunity to study possible interaction spaces that could provide insights about segregation from a different perspective. It is not a common approach used in segregation literature because segregation studies have mostly focused on how to measure the spatial separation of groups. However, if one analyses segregation from a different perspective, based for instance on Wong's, (2009) approach: the nature of spatial segregation is the lack of interaction among population groups provided by the geographical settings and the locations of people, segregation focus would not only be on how spatially separated two groups are from each other, but also which are the spaces in a district/city/region that could bring together individuals from different social groups.

Exploring the social, economic, political and urban context, as well as the spatial settings of these interaction spaces might provide insights of which are the factors that influence whether individuals from different social groups visit certain places across a city, or not. For instance, this study has demonstrated that interaction spaces where staff and students may have higher chances of interaction across Tartu city are within the university environments, and therefore segregation within these environments might be reduced. This suggest that the factors shaping segregation between staff and students might be also related to other socioeconomic characteristics of the respondents, for instance ethnicity and gender; urban context: tolerance to immigrants from locals; or, the spatial settings of Tartu city like transportation systems.

Besides studying the spatial separation, segregation discipline perhaps also needs to identify these potential interaction places which could be very relevant for the newer segregation studies that try to tackle how to break the vicious circle of segregation (van Ham et al., 2018).

The limitations of the study must be considered. The results were affected by the composition of the provided dataset, described in section 2.2, and the computational complications that arose to calculate each method. Due to the higher number of students than staff members, preliminary analysis stages demonstrated biased results in which students, as a majority in the dataset, had significantly higher values than staff for all the analysed indicators. Therefore, to avoid this methodological flaw 50 students and 50 staff members were randomly selected by their user ID.

In addition to this, not all the respondents were represented with the same number of days in the research period, therefore the activity spaces may not be equally accurate for all respondents.

Another limitation is related to the design and representation of the maps for each method, results with higher level of detail were more complicated to represent so that the maps were clear and readable. Initially, it was planned to represent grid results with a bivariate mapping approach in ArcGIS, however the differences between staff and students were not big enough for the tool to display it. This suggests that with a larger sample bivariate mapping would have been another way to display the results.

This study highlights the importance of the relation between the spatial accuracy and composition of the data, the study purpose, spatial scale and spatial context of the study area to select the most suitable method that reflects GPS data accuracy to analyse segregation. Additionally, besides spatial separation, this thesis provided an approach for studying the interaction spaces between social groups, that might be of interest when investigating the vicious circles of segregation.

Summary

Comparing activity-space based segregation methods: a study with GPS data

Eliana Ortiz G.

Spatial segregation is a multidimensional and complex process that has been identified as a critical issue since the beginning of segregation studies because it affects different domains of the society and people's life. Activity space is a multidimensional (spatial, temporal and cognitive) concept that reflects an individual experience of segregation with other population groups across daily life activities (Park and Kwan, 2018).

The precision, space-time granularity and availability of new datasets, like global positioning data (GPS) has allowed the development of more accurate measures of activity spaces to evaluate segregation levels (Wong and Shaw, 2011; Cornwell and Cagney, 2017). The range of possibilities is from activity space methods widely used in segregation literature because of its easiness to compute and interpret (i.e. standard deviation ellipse (SDE), minimum convex hull polygon (MCP)), to methods that require advanced methodologies for computation and interpretation (i.e. daily path areas (DPA), kernel densities).

Thus, the aim of the thesis is to compare activity space methods for studying segregation and interaction with GPS data. Groups under comparison are staff and students, and the activity space indicators are calculated for weekdays and weekends. To achieve the aim of the research, following research questions are stated:

1) How does the activity space of staff and students vary across different methods?

2) How do the characteristics of interaction space of staff and students vary across different methods?

The author of the thesis compared three activity space methods of different complexity: minimum convex polygon, buffer, inspired by the daily path area method, and finally the grid method.

Data used for the research has been provided by the Department of Geography consisting of members of University of Tartu. GPS data has been obtained through MobilityLog smartphone application. The dataset consists of 50 staff members and 50 students. The month covered by the data is March for the years 2015, 2017 and 2018. The data consisted of pre – calculated stops based

on aggregated sequential GPS points and the social characteristic of the respondents: academic role. The study area is Tartu city including a 5 km buffer from Tartu city administrative borders.

Each method, according to its complexity, depicted different results. Students registered larger activity spaces than staff. However, buffer and grid methods provided more detailed and realistic insights of the activity spaces compared to MCP. MCP produces overgeneralised results which seems to be unsuitable for GPS data. Buffer and grid based calculations are more suitable for GPS data. With the spatially precise data, there is the opportunity to measure the overlap of activity spaces (i.e. interaction spaces). Interaction spaces showed that student's spatial behaviour is more variable than staff both during weekdays and weekends. The most visited districts by staff and students during the weekdays and weekends are Kesklinn, Maarjamõisa and Ülejõe. Maarjamõisa as a working/ study district emerges during weekdays.

Considering that spatial segregation is a complex spatial process, multidimensional and specific phenomenon (Massey & Denton, 1988; Kaplan & Holloway, 2001), the indicators used did not provide sufficient evidence to determine segregation between staff members and students. However, the existence of interaction spaces suggests that segregation might be reduced between staff and students, especially within the university environment.

This study highlights the importance of the relation between the spatial accuracy and composition of the data, the study purpose, spatial scale and spatial context of the study area to select the most suitable method that reflects GPS data accuracy to analyse segregation. Additionally, besides spatial separation, this thesis provided an approach for studying the interaction spaces between social groups, that might be of interest when investigating the vicious circles of segregation.

Kokkuvõte

Tegevusruumipõhiste segregatsioonimeetodite võrdlus GPS andmetel

Eliana Ortiz G.

Ruumiline segregatsioon on keerukas ja mitmemõõtmeline protsess, mida on segregatsiooniuuringute algusest peale käsitletud kriitilise probleemina, kuna see puudutab ühiskonna ja inimeste elu mitmesuguseid valdkondi. Tegevusruum on mitmemõõtmeline (ruumiline, ajaline ja kognitiivne) mõiste, mis peegeldab indiviidi segregatsiooni kogemust igapäevategevuste lõikes (Park and Kwan, 2018).

Uute andmete, näiteks GPS-andmete ajalis-ruumiline täpsus võimaldab välja arendada täpsemaid meetodeid tegevusruumi ja tegevusruumipõhise segregatsiooni hindamiseks (Wong and Shaw, 2011; Cornwell and Cagney, 2017). Selleks on mitmesuguseid võimalusi alustades segregatsiooni uuringutes enam levinumatest meetoditest, mida on võrdlemisi lihtne arvutada ja tõlgendada (nt standardhälbe ellipsid, minimaalne kumer hulknurk), lõpetades meetoditega, mis on töömahukamad ja tõlgendamise mõttes keerulisemad (nt igapäeva trajektoori pindala (*daily path area*), kerneli tihedus).

Käesoleva magistritöö eesmärk on võrrelda erinevaid tegevusruumi meetodeid segregatsiooni ja interaktsiooni uurimiseks GPS andmete abil. Käesoleva töö autor võrdleb kolme tegevusruumi meetodit: minimaalne kumer hulknurk, puhver, ja võre. Vaadeldavad grupid on Tartu Ülikooli töötajad ja üliõpilased. Tegevusruumi indikaatorid arvutatakse tööpäevade ja nädalavahetuse lõikes. Eesmärgi saavutamiseks on püstitatud järgmised uurimisküsimused.

1) Kuidas varieerub töötajate ja üliõpilaste tegevusruum erinevate meetodite lõikes?

2) Kuidas varieeruvad töötajate ja üliõpilaste interaktsiooniruumi indikaatorid erinevate meetodite lõikes?

Töös kasutatav andmestik Tartu Ülikooli töötajate ja üliõpilaste kohta on saadud Geograafia osakonnast. GPS-andmed on kogutud MobilityLog nutitelefoni rakenduse abil ja järjestikku paiknevad GPS punktid on eelnevalt agregeeritud peatusteks. Töös kasutatav andmestik koosneb 50 töötaja ja 50 üliõpilase 2015, 2017 ja 2018 aasta märtsikuu peatustest. Uurimisala on Tartu linn ja seda ümbritsev 5 km laiune piirkond.

Sõltuvalt oma keerukusest andis iga meetod erisuguseid tulemusi. Üliõpilaste tegevusruumid olid ulatuslikumad kui töötajate omad. Puhvri ja võre meetodid andsid detailsema ja realistlikuma arusaama tegevusruumide ulatusest võrreldes minimaalse kumera hulknurga meetodiga. Viimane andis liialt üldistatud pildi tegevusruumide ulatusest, mis tundub olevat GPS-andmete jaoks ebasobiv. Puhvri- ja võrepõhised arvutused on GPS-andmete jaoks sobivamad. Ruumiliselt täpsete andmetega on võimalik mõõta ka tegevusruumide kattumist, nn interaktsiooniruume. Interaktsiooniruumide põhised arvutused näitasid, et üliõpilaste ruumikasutus on töötajatega võrreldes muutlikum nii tööpäevadel kui ka nädalavahetusel. Kesklinn, Maarjamõisa ja Ülejõe olid nii töötajate kui ka üliõpilaste poolt kõige enam külastatud linnaosad tööpäevadel ja nädalavahetusel. Maarjamõisa linnaosa, kui töö- ja õppetööga seotud piirkond eristus selgelt just tööpäevadel.

Arvestades ruumilise segregatsiooni keerukat olemust ja mitmemõõtmelisust (Massey & Denton, 1988; Kaplan & Holloway, 2001), ei võimaldanud töös kasutatud tegevusruumi indikaatorid kindlaks teha töötajate ja üliõpilaste vahelist segregatsiooni. Interaktsiooniruumide olemasolu viitab sellele, et kahe grupi vaheline eraldatus võib olla väiksem ja seda eriti ülikooli keskkonnas.

Käesolev uurimistöö toob esile, kui oluline on segregatsiooni analüüsimiseks sobiva meetodi valimisel arvestada andmete ruumilise täpsuse ja koosseisuga, uurimistöö eesmärgiga, uuritava ala ruumilise ulatuse ja kontekstiga, et tulemused oleksid samaaegselt mõistetavad ja peegeldaksid GPS-andmete eeliseid (täpsust). Lisaks võimaldab magistritöös välja pakutud interaktsiooniruumide analüüs ruumilise eraldatuse kõrval anda huvitavat teavet kahe grupi vahelistest kokkupuutepunktidest, mis on segregatsiooni nõiaringi uurimisel oluline.

Acknowledgements

Especial thanks to the Department of Geography from University of Tartu, for the support and the data provided. My deepest gratitude to my supervisors Siiri Silm PhD and Veronika Mooses MSc for the patience and guidance through this research. Especial thanks to Veronika Mooses MSc for the extra time invested in helping and supporting me in every stage of the master thesis. Finally, this would have not be possible without the support and love from my parents, family and friends.

References

Aggarwal, C. (2016). Spatial Outlier Detection. In *Outlier Analysis* (pp. 345–368). Springer, Cham. https://doi.org/10.1007/978-3-319-47578-3

Ahas, R., Aasa, A., Silm, S., & Aunap, R. (2007). Mobile Positioning in Space–Time Behaviour Studies: Social Positioning Method Experiments in Estonia. *Cartography and Geographic Information Science*, *34*(4), 259–273. https://doi.org/10.1559/152304007782382918

Ahas, R., Poom, A., Aasa, A., & Silm, S. (2017). Experimental study for planning Tartu campuses based on mobile phone tracking. In M. Schewenius, P. Keränen, & R. al Rafaw (Eds.), *Dreams and Seeds: The role of campuses in sustainable urban development.* (pp. 83–85).

Barber, C., Dobkin, D., & Huhdanpaa, H. (1996). The Quickhull Algorithm for Convex Hulls. *ACM Transactions on Mathematical Software*, *22*(4), 469–483. https://doi.org/10.1145/235815.235821

Bender, K., Donohue, S., & Heywood, J. (2005). Job satisfaction and gender segregation. *Oxford Economic Papers*, *57*(3), 479–496. https://doi.org/10.1093/oep/gpi015

Blair, S., & Lichter, D. (1991). Measuring the Division of Household Labor. Gender Segregation of Housework Among American Couples. *Journal of Family Issues*, *12*(1), 91–113. https://doi.org/https://doi.org/10.1177/019251391012001007

Blau, F., & Hendricks, W. (1979). Occupational Segregation by Sex : Trends and Prospects Author. *The Journal of Human Resources*, *14*(2), 197–210. https://doi.org/10.2307/145642

Borjas, G. (1998). To Guetto or Not to Guetto: Ethnicity and Residential Segragation. *Journal of Urban Economics*, *44*(2), 228–253. https://sites.hks.harvard.edu/fs/gborjas/publications/journal/JUE1998.pdf

Brimicombe, A. (2007). Ethnicity, religion, and residential segregation in London: Evidence from a computational typology of minority communities. *Environment and Planning B: Planning and Design*, *34*(5), 884–904. https://doi.org/10.1068/b3309

Buliung, R., & Kanaroglou, P. (2006). A GIS toolkit for exploring geographies of household activity / travel behavior. *Journal of Transport Geography 14*, *14*, 35–51. https://doi.org/10.1016/j.jtrangeo.2004.10.008

Buliung, R., Roorda, M., & Remmel, T. (2008). Exploring spatial variety in patterns of activity-travel behaviour : initial results from the Toronto. *Transportation*, *35*, 697–722. https://doi.org/10.1007/s11116-008-9178-4

Chan, D., Helfrich, C., Hursh, N., Rogers, E., & Gopal, S. (2014). Measuring community integration using Geographic Information Systems (GIS) and participatory mapping for people who were once homeless. *Health & Place*, *27*, 92–101. https://doi.org/10.1016/j.healthplace.2013.12.011

Cornwell, E., & Cagney, K. (2017). Aging in Activity Space : Results From Smartphone-Based GPS-Tracking of Urban Seniors. *Journals of Gerontology: Social Sciences, 72*(5), 864–875. https://doi.org/10.1093/geronb/gbx063

Day, A. (2018). *Interquartile Range (IQR)*. https://www.tutor2u.net/geography/reference/interquartile-range

Duncan, O., & Duncan, B. (1995). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, *20*(2), 210–217. https://doi.org/181.199.43.87

Dwyer, R. (2010). Poverty , Prosperity , and Place : The Shape of Class Segregation in the Age of Extremes. *Social Problems*, *57*(1), 114–137. https://doi.org/10.1525/sp.2010.57.1.114.SP5701

Ellis, M., Wright, R., & Parks, V. (2004). Work together, live apart? Geographies of racial and ethnic segregation at home and at work. *Annals of the Association of American Geographers*, *94*(3), 620–637. https://doi.org/10.1111/j.1467-8306.2004.00417.x

Field, E., Levinson, M., Pande, R., & Visaria, S. (2008). Segregation , Rent Control , and Riots : The Economics of Religious Conflict in an Indian City. *American Economic Review: Papers & Proceedings 2008, 98*(2), 505–510. https://doi.org/http://www.aeaweb.org/articles.php?doi=10.1257/aer.98.2.505

Grusky, D. (2001). *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Westview Press.

Hanson, S., & Pratt, G. (1991). Job Search and the Occupational Segregation of Women. *Annals of the Association of American Geographers*, *81*(2), 229–253. https://doi.org/10.1111/j.1467-8306.1991.tb01688.x

Harrison, J., & Lloyd, S. (2013). New Jobs , New Workers , and New Inequalities : Explaining Employers ' Roles in Occupational Segregation by Nativity and Race. *Social Problems, 60*(3), 281–301. https://doi.org/10.1525/sp.2013.60.3.281.This

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Planning and Design, 20*, 29–66. https://doi.org/10.1068/b200029

Hirsch, J., Winters, M., Clarke, P., & McKay, H. (2014). Generating GPS activity spaces that shed light upon the mobility habits of older adults: A descriptive analysis. *International Journal of Health Geographics*, *13*(51), 1–14. https://doi.org/10.1186/1476-072X-13-51

Hong, S., O'Sullivan, D., & Sadahiro, Y. (2014). Implementing spatial segregation measures in R. *PLoS ONE*, *9*(11), 1–18. https://doi.org/10.1371/journal.pone.0113767

Horton, F., & Reynolds, D. (1971). Effects of Urban Spatial Structure on Individual Behavior. *Economic Geography*, *47*(1), 36–48. https://doi.org/10.2307/143224

Jahn, J., Schmid, C., & Schragt, C. (1947). The Measurement of Ecological Segregation. *American Sociological Review*, *12*(3), 293–303. https://doi.org/193.40.12.10

Järv, O., Ahas, R., & Witlox, F. (2014). Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*, *38*, 122–135. https://doi.org/10.1016/j.trc.2013.11.003

Järv, O., Müürisepp, K., Ahas, R., Derudder, B., & Witlox, F. (2015). Ethnic differences in activity spaces as a characteristic of segregation : A study based on mobile phone usage in Tallinn , Estonia. *Urban Studies*, *52*(14), 2680–2698. https://doi.org/10.1177/0042098014550459

Järv, O., Tenkanen, H., & Toivonen, T. (2017). Enhancing spatial accuracy of mobile phone data using multi-temporal dasymetric interpolation. *International Journal of Geographical Information Science*, *31*(8), 1630–1651. https://doi.org/10.1080/13658816.2017.1287369

Jones, M., & Pebley, A. (2014). Redefining Neighborhoods Using Common Destinations : Social Characteristics of Activity Spaces and Home Census Tracts Compared. *Demography*, *51*, 727–752. https://doi.org/10.1007/s13524-014-0283-z

Kaplan, D., & Holloway, S. (2001). Scaling ethnic segregation : causal processes and contingent outcomes in Chinese residential patterns. *GeoJournal*, *53*, 59–70. https://doi.org/https://doi.org/10.1023/A:1015822117915

Kaplan, D., & Woodhouse, K. (2004). Research in Ethnic Segregation I : Causal Factors. *Urban Geography*, *25*(6), 579–585. https://doi.org/10.2747/0272-3638.25.6.579

Karadimitriou, S. (2019). *Checking normality for parametric tests in R*. https://www.sheffield.ac.uk/polopoly_fs/1.579191!/file/stcp-karadimitriou-normalR.pdf

Kestens, Y., Lebel, A., Chaix, B., Clary, C., Daniel, M., Pampalon, R., Theriault, M., & Subramanian, S. (2012). Association between activity space exposure to food establishments and individual risk of overweight. *PLoS ONE*, *7*(8), 1–13. https://doi.org/10.1371/journal.pone.0041418

Kestens, Y., Lebel, A., Daniel, M., Thériault, M., & Pampalon, R. (2010). Using experienced activity spaces to measure foodscape exposure. *Health and Place, 16*, 1094–1103. https://doi.org/10.1016/j.healthplace.2010.06.016

Kivisto, P. (2016). Introduction : Transatlantic comparisons : analyzing immigrant integration in six countries. *Ethnic and Racial Studies*, *1–4*. https://doi.org/10.1080/01419870.2016.1203451

Kwan, M. (1999). Gender and individual access to urban opportunities: A study using space-time measures. *Professional Geographer*, *51*(2), 210–227. https://doi.org/https://doi.org/10.1111/0033-0124.00158

Kwan, M. (2013). Beyond Space (As We Knew It): Toward Temporally Integrated Geographies of Segregation, Health, and Accessibility. *Annals of the Association of American Geographers*, *103*(5), 1078–1086. https://doi.org/10.1080/00045608.2013.792177

Lee, N., Voss, C., Frazer, A., Hirsch, J., Mckay, H., & Winters, M. (2016). Does activity space size in fl uence physical activity levels of adolescents ? — A GPS study of an urban environment. *Preventive Medicine Reports Journal*, *3*, 75–78. https://doi.org/10.1016/j.pmedr.2015.12.002

Li, F., & Wang, D. (2017). Measuring urban segregation based on individuals ' daily activity patterns : A multidimensional approach. *Environment and Planning A*, *49*(2), 467–486. https://doi.org/10.1177/0308518X16673213

Lund Research Ltd. (2018). *Mann-Whitney U Test using SPSS Statistics*. https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php

Massey, D., & Denton, N. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67*(2), 281–315. https://doi.org/130.212.93.143

Masso, A., Silm, S., & Ahas, R. (2019). Generational differences in spatial mobility : A study with mobile phone data. *Popul Space Place*, *25*(e), 1–15. https://doi.org/10.1002/psp.2210

Miller, H. (2004). Activities in Space and Time. *Handbook of Transport 5: Transport Geography and Spatial Systems*.

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*(3), 205–219. https://doi.org/https://onlinelibrary.wiley.com/doi/full/10.1002/psp.2210

Musterd, S. (2005). Social and Ethnic Segreation in Europe: Levels, Causes and Effects. *Journal of Urban Affairs*, *27*(3), 1–16. https://doi.org/https://doi.org/10.1111/j.0735-2166.2005.00239.x

Nelson, J. (2010). Religious segregation and teacher education in Northern Ireland. *Research Papers in Education*, *25*(1), 1–20. https://doi.org/10.1080/02671520802298464

O'Sullivan, D., & Wong, D. (2007). A Surface-Based Approach to Measuring Spatial Segregation. *Geographical Analysis*, *39*, 147–168. https://doi.org/https://doi.org/10.1111/j.1538-4632.2007.00699.x

Oakes, J., & Kaufman, J. (2017). *Methods in Social Epidemiology*. Jonh Wiley & Sons, Inc.

Owens, A., Reardon, S., & Jencks, C. (2016). Income Segregation Between Schools and School Districts. *American Educational Research Journal*, *53*(4), 1159–1197. https://doi.org/10.3102/0002831216652722

Palmer, J. (2013). *Activity-Space Segregation : Understanding Social Divisions in Space and Time* [Princeton University]. http://arks.princeton.edu/ark:/88435/dsp01k643b130h

Park, Y., & Kwan, M. (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

Patterson, Z., & Farber, S. (2015). Potential Path Areas and Activity Spaces in Application : A Review. *Transport Reviews*, *35*(6), 679–700. https://doi.org/10.1080/01441647.2015.1042944

Raanan, M., & Shoval, N. (2014). Mental maps compared to actual spatial behavior using GPS data : A new method for investigating segregation in cities. *Journal of Cities*, *36*, 28–40. https://doi.org/10.1016/j.cities.2013.09.003

Rai, R., Balmer, M., Rieser, M., Vaze, V., Schönfelder, S., & Axhausen, K. (2007). Capturing Human Activity Spaces: New Geometries. *Transportation Research Record: Journal of the Transportation Research Board*, *2021*(1), 70–80. https://doi.org/10.3141/2021-09

Reardon, S., & Bischoff, K. (2011). Income Inequality and Income Segregation. *American Journal of Sociology*, *116*(4), 1092–1153. https://doi.org/193.40.12.10

Reardon, S., & O'Sullivan, D. (2004). Measures of spatial segregation. In *Sociological Methodology* (pp. 121–162). https://doi.org/https://doi.org/10.1111/j.0081-1750.2004.00150.x

Ren, F. (2016). *Activity space*. Activity Space. https://doi.org/10.1093/OBO/9780199874002-0137

Rose-redwood, C., & Rose-redwood, R. (2013). Self-Segregation or Global Mixing ?: Social Interactions and the International Student Experience. *Journal of College Student Development*, *54*(4), 413–429. https://doi.org/doi.org/10.1353/csd.2013.0062

Rumberger, R., & Willms, J. (1992). The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools. *Educational Evaluation and Policy Analysis*, *14*(4), 377–396. https://doi.org/10.3102/01623737014004377

Sadahiro, Y. (2019). Statistical analysis of spatial segregation of points Yukio. *Computers, Environment and Urban Systems*, *76*, 123–138. https://doi.org/10.1016/j.compenvurbsys.2019.04.008

Schnore, L. (1965). Social Class Segregation Among Nonwhites in Metropolitan Centers. *Demography*, *2*, 126–133. https://doi.org/193.40.12.10

Schönfelder, S., & Axhausen, K. (2002). *Measuring the size and structure of human activity spaces The longitudinal perspective* (Issue Arbeitsbericht Verkehrs-und Raumplanung, 135).

Schönfelder, S., & Axhausen, K. (2003). Activity spaces : measures of social exclusion ? *Transport Policy*, *10*(4), 273–286. https://doi.org/10.1016/j.tranpol.2003.07.002

Shdema, I., Haj-yahya, N., & Schnell, I. (2018). The social space of Arab residents of mixed Israeli cities cities. *Geografiska Annaler: Series B, Human Geography*, 1–18. https://doi.org/10.1080/04353684.2018.1428496

Shen, Y. (2019). Segregation through space: A scope of the flow-based spatial interaction model. *Journal of Transport Geography*, *76*(February), 10–23. https://doi.org/10.1016/j.jtrangeo.2019.02.007

Sherman, J., Spencer, J., Preisser, J., Gesler, W., & Arcury, T. (2005). A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics*, *4*(24), 1–21. https://doi.org/10.1186/1476-072X-4-24

Shoval, N., & Isaacson, M. (2007). Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers*, *97*(2), 282–297. https://doi.org/10.1111/j.1467-8306.2007.00536.x

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*(3), 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 segregation in activity spaces using mobile phone data. *Journal of Ethnic and Migration Studies*, *44*(11), 1797–1817. https://doi.org/10.1080/1369183X.2017.1400425

Smith, A. (2001). Religious Segregation and the Emergence of Integrated Schools in Northern Ireland. *Oxford Review of Education*, *27*(4), 559–575. https://doi.org/10.1080/03054980120086248

Statistics Estonia. (2019). *Statistical database RV0240: Population by sex, age and place of residence after the 2017 administrative reform, 1 January 2019*. andmebaas.stat.ee

Susilo, Y., & Kitamura, R. (2005). Analysis of Day-to-Day Variability in an Individual ' s Action Space Exploration of 6-Week Mobidrive Travel Diary Data. *Journal of the Transportation Research Board*, *1902*, 124–133. https://doi.org/https://doi.org/10.1177/0361198105190200115

Ta, N., Kwan, M., & Chai, Y. (2016). Urban form , car ownership and activity space in inner suburbs : A comparison between Beijing (China) and Chicago (United States). *Urban Studies Journal*, *53*(9), 1784–1802. https://doi.org/10.1177/0042098015581123

Tan, Y., Kwan, M., & Chai, Y. (2017). Examining the impacts of ethnicity on space-time behavior : Evidence from the City of Xining , China. *Cities*, *64*, 26–36. https://doi.org/10.1016/j.cities.2017.02.003

Tao, S., He, S., Kwan, M., & Luo, S. (2020). Does low income translate into lower mobility ? An investigation of activity space in Hong Kong between 2002 and 2011. *Journal of Transport Geography*, *82*, 1–12. https://doi.org/10.1016/j.jtrangeo.2019.102583

Toomet, O., Silm, S., Saluveer, E., Ahas, R., & Tammaru, T. (2015). Where Do Ethno-Linguistic Groups Meet ? How Copresence during Free-Time Is Related to Copresence at Home and at Work. *PLoS ONE*, *10*(5), 1–16. https://doi.org/10.1371/journal.pone.0126093

University of Tartu. (2019). *Number of students in UT*. www.ut.ee/oppijate-arv/

van Ham, M., & Tammaru, T. (2016). New perspectives on ethnic segregation over time and space . A domains approach. *Urban Geography*, *37*(7), 953–962. https://doi.org/10.1080/02723638.2016.1142152

van Ham, M., Tammaru, T., & Janssen, H. (2018). *A multi-level model of vicious circles of socio-economic segregation*. https://doi.org/10.1787/9789264300385-en

van Kempen, R., & Ozuekren, A. (1998). Ethnic Segregation in Cities : New Forms and Explanations in a Dynamic World. *Urban Studies*, *35*(10), 1631–1656. https://doi.org/https://doi.org/10.1080/0042098984088

Vich, G., Marquet, O., & Miralles-Guasch, C. (2017). Suburban commuting and activity spaces: using smartphone tracking data to understand the spatial extent of travel behaviour. *The Geographical Journal*, 1–14. https://doi.org/10.1111/geoj.12220

Wang, D., & Li, F. (2016). Daily activity space and exposure : A comparative study of Hong Kong ' s public and private housing residents ' segregation in daily life. *Journal of Cities*, *59*, 148–155. https://doi.org/10.1016/j.cities.2015.09.010

Wang, D., Li, F., Chai, Y., Tong, K., Kong, H., & Chai, Y. (2012). Activity Spaces and Sociospatial Segregation in Beijing. *Journal of Urban Geography*, *33*(2), 256–277. https://doi.org/10.2747/0272-3638.33.2.256

Weber, J. (2003). Individual accessibility and distance from major employment centers : An examination using space-time measures. *Journal of Geographical Systems*, *5*(1), 51–70. https://doi.org/https://doi.org/10.1007/s101090300103

Wessel, T. (2000). Social polarisation and socioeconomic segregation in a welfare state: The case of Oslo. *Urban Studies*, *37*(11), 1947–1967. https://doi.org/10.1080/713707228

Wissink, B., Schwanen, T., & Kempen, R. (2016). Beyond residential segregation : Introduction. *Journal of Cities*, *59*, 126–130. https://doi.org/10.1016/j.cities.2016.08.010

Wong, D. (2002). Modeling local segregation: A spatial interaction approach. *Geographical and Environmental Modelling*, *6*(1), 81–97. https://doi.org/10.1080/13615930220127305

Wong, D. (2009). The Modifiable Areal Unit Problem (MAUP). In S. Fotheringham & P. Rogerson (Eds.), *Spatial Analysis* (pp. 105–124). SAGE. https://books.google.ee/books?hl=en&lr=&id=phEgXfbCU_YC&oi=fnd&pg=PA105&dq=wong+2009+the+modifiable+areal+unit+problem+sage&ots=JSfyD42vQi&sig=h5vozr-k-zG_VLguDrlC_n-Con4&redir_esc=y#v=onepage&q=wong 2009 the modifiable areal unit problem sage&f=false

Wong, D., & Shaw, S. (2011). Measuring segregation : an activity space approach. *Journal of Geographical Systemsgr Syst*, *13*(2), 127–145. https://doi.org/10.1007/s10109-010-0112-x

Wong, S. (2018). The limitations of using activity space measurements for representing the mobilities of individuals with visual impairment : A mixed methods case study in the San Francisco Bay Area. *Journal of Transport Geography*, *66*(July 2017), 300–308. https://doi.org/10.1016/j.jtrangeo.2017.12.004

Wunderlich, F. (2008). Symphonies of Urban Places : Urban Rhythms as Traces of Time in Space A Study of ‘ Urban Rhythms .’ *KOHT Ja PAIK / PLACE and LOCATION Studies in Environmental Aesthetics and Semiotics VI*, 91–111. http://www.eki.ee/km/place/pdf/kp6_07_wunderlich.pdf

Xu, W., & Leffler, A. (1992). Gender and Race Effects on Occupational Prestige, Segregation, and Earnings. *Gender and Socierty*, *6*(3), 376–392. https://doi.org/193.40.12.10

Yao, J., Wong, D., Bailey, N., & Minton, J. (2018). Spatial Segregation Measures: A Methodological Review. *Tijdschrift Voor Economische En Sociale Geografie*, *00*(00). https://doi.org/10.1111/tesg.12305

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

Zenk, S., Matthews, S., Kraft, A., & Jones, K. (2018). How many days of global positioning system (GPS) monitoring do you need to measure activity space environments in health research? *Health and Place*, *51*, 52–60. https://doi.org/10.1016/j.healthplace.2018.02.004

Zenk, S., Schulz, A., Matthews, S., Odoms-Young, A., Wilbur, J., Wegrzyn, L., Gibbs, K., Braunschweig, C., & Stokes, C. (2011). Activity Space Environment and Dietary and Physical Activity Behaviors: A Pilot Study. *Health Place*, *17*(5), 1150–1161. https://doi.org/10.1016/j.healthplace.2011.05.001.Activity

Zhang, R., Stanley, K., Fuller, D., & Bell, S. (2020). Differentiating Population Spatial Behavior Using Representative Features of Geospatial Mobility (ReFGeM). *Transactions on Spatial Algorithms and Systems*, *6*(1), 1–25. https://doi.org/https://doi.org/10.1145/3362063

Non-exclusive licence to reproduce thesis and make thesis public

I, Eliana Solange Ortiz Gavela

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

master thesis

"Comparing activity-space based segregation methods:

a study with GPS data"

supervised by Siiri Silm, PhD and Veronika Mooses, MSc.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Eliana Solange Ortiz Gavela

25.06.2020.