Examination of residential and running environments of urban and rural runners in the Netherlands

by

Shiyuan Zhang

Student number: 6698611

E-mail: s.zhang9@students.uu.nl

At Utrecht University, 28-02-2021

in partial fulfillment of the requirements for the degree of

Master of Geographical Information Management and Applications

Supervisors: Prof. Dick Ettema

Dr. Zhiyong Wang

GIMA responsible professor: Prof. Stan Geertman

Abstract

In the course of urban growth, living environment and lifestyle of residents in urban and rural areas are undergoing gradual changes (e.g., increase in dependency on automobiles and loss in public open spaces to exercise). This transformation places the importance of protecting or improving the environments that encourage people to participate physical activities. While currently, little is known about the environments of a specific physical activity – running – which is conveniently accessible and considerably benefits health.

Thus, this research aims to examine the differences between urban and rural runners regarding their residential and running environments. It will also explore the differences between residential and running environments of the runners (urban and rural). Furthermore, it will conduct further investigation on the running environments of those urban and rural runners running different distance, duration, at weekends/on weekdays, inside/outside their neighbourhoods and with different running frequency.

To achieve the goal, the following processes will be carried out. The first step is to collect routes taken by runners living in the Netherlands via a GPS mobile tracking application, Endomondo, and then to extract features (e.g., geolocations of routes, running duration and start time) from the GPS data. Subsequently, defining runner's urban/rural status and measuring the environmental variables (e.g., street connectivity and residential building density) of residential areas and surroundings of GPS tracks with the support of geographical and census data. Finally, carrying on t-test to compare the objectively measured environmental variables of running environments between urban and rural runners and the sub-groups (i.e., weekdays/weekends, inside/outside neighbourhoods, different running distance, duration, and frequency) and compare the environmental variables of the runners (urban and rural)' residential and running environments.

The results showed that urban runners were exposed in running and living environments with relatively heterogeneously distributed land uses, more streets, addresses and coverage of residential buildings, yet less green and blue areas than their rural counterparts. Both urban and rural runners preferred running in places with more vegetations while running outside their neighbourhoods. Besides, when urban runners run at weekends, they choose more heterogeneously distributed land uses and streets. In addition, the longer tracks (long distance or duration) and the tracks of more frequently practiced runners were taken in environments with more vegetation and less addresses. Moreover, urban and rural runners chose more environments with more blue spaces for running.

This research found that the running environments of urban and rural runners had distinct characteristics. Their choices of running environments were restricted by their residency. It therefore suggests policy to promote running acknowledging these differences between running environments in urban and rural areas.

Acknowledgements

I would first like to thank my supervisors, Prof. Dick Ettema and Dr. Zhiyong Wang for their continued support and advice throughout the research, especially their guidance for finishing touch. Besides, this research would never go on stream without their consideration. I will never forget your kindness!

I would also like to thank Drs. Maarten Zeylmans Van Emmichohven for his guidance on using virtual computer while I was having difficulties with collecting BBG data; Zhengdong Yuan for help with fixing abrupt database malfunction; Lisanne O for listing me a bunch of popular mobile running applications in the Netherlands.

I would like to dedicate this piece of work to my parents, Mary X, Stanley Y and Laalie J. You always help me fight depression and back me up when I feel I am baffled by the world. I made my way by keeping in mind your words, ‘a stitch in time saves nine’. Your generous support and love keep me safe and warm, and I cannot be more grateful.

Abbreviations

ATMs – Active travel modes

ANOVA – Analysis of variance

BAG – Basisregistratie Adressen en Gebouwen

BBG – Bestand Bodemgebruik

BRK – Basisregistratie Kadaster

BRT – Basisregistratie Topografie

CBS – Centraal Bureau van Statistiek

ETM+ – Enhanced Thematic Mapper Plus

FUAs – Functional Urban Area

GIS – Geographic Information System

GNSS – Global Navigation Satellite System

GPS – Global Positioning System

HTML – Hyper Text Markup Language

IDE - Integrated Development Environment

JSON – Java Script Object Notation

LGN – Land Use Database of the Netherlands

LVPA – Light to vigorous physical activities

MVPA – Moderate to vigorous physical activities

NDVI – Normalized Difference Vegetation Index

OECD – Organisation for Economic Co-operation and Development

OLI – Operational Land Imager

PA – Physical activities

PC6 – 6-digit postcode area

SQL – Structed Query Language

TOP10NL – Key Register Topography at 1:10,000

Contents

1. Introduction

1.1. Research context

Urban growth demands densification and spatial expansion, simultaneously, putting urban open spaces such as public parks in urban and rural areas under increased pressure (Claassens et al., 2020). These public open spaces are tuned to multifarious and balanced environments regarding urban development and public health by city planners and scientists. One of the focuses nowadays is to explore which environments motivate the performance of physical activities that contributes significant benefits to people's health (Maddison et al., 2009; Cohen et al., 2007; Deelen et al., 2019). A specific physical activity, running, has drawn considerable concerns according to its increasing popularities in the Netherlands (Hover et al., 2015). In the meanwhile, technological advancement (e.g., GPS tracking system and GIS) provides the viability of applying, storing, processing, and sharing a diverse range of data for analysing different activities.

1.1.1. Urban growth and physical activity

The term 'urban growth' has appeared in the scientific literature for over half a century (Ioannides and Rossi-Hansberg, 2010). Ever since environmental issues such as global warming became a hot topic, urban growth has been implicated in social and environmental consequences including increasing traffic jams, the dependency of cars, water and air pollution and destruction of the countryside. The Netherland, like many other European countries, has been aware of the issues brought by city expansion and released several national-wise spatial planning policies to avoid sacrificing open spaces (e.g., parks) since a few decades ago. While with densification always being proposed in the policies, it remains unclear to which extent these policies steer urban development to the desired degree (e.g., what are the impacts of increasing density on urban neighbourhoods and whether will the numbers of public places be diminished) (Claassens et al., 2020). Discussing the implications of environmental or spatial planning policies is beyond the scope of this research, therefore, the intention here is to acknowledge the influences of unbridled urban growth and enshrine the importance of developing proper and well-defined strategies to preserve or upgrade the existing open spaces in the Netherlands.

Preserving open spaces raises residents' quality of life and helps them maintain a healthy life. Concerning the complicated relationships between people's health and their environmental exposures due to multiple pathways (European Environment Agency, n.d.), this research will only concentrate on the characteristics of environments motivate physical activities that benefit people's physical and mental health. Many studies reported that physical activities (e.g., cycling and walking) that help to reduce a chance of catching diseases (e.g., obesity, diabetes, and cancer) are more likely to happen in supportive outdoor spaces (e.g., presence of green and recreational space and absence of occupied roads) (Maddison et al., 2009; Gladwell et al., 2013). In this regard, higher priority should be placed on these environments for promoting physical activities during the urban growth.

1.1.2. Running in the Netherlands

Amongst various kinds of physical activities, running has grown in popularity as it is a readily accessible form of exercise and a straightforward way to get benefits such as improving cardiovascular health, training endurance and alleviating depression. While back in the 1940s, running was seen as a strenuous and dismal sport among the Dutch. The health benefits of running were only recognized by the national government in the Netherlands till 1968. With government intervention, running as a sport to promote a healthy lifestyle, raised to its first wave (i.e., a sudden and sharp increase) (Hover et al., 2015). As a result, in the 1990s, about 10% Dutch (6 – 79 years old) participated in a running event at least once a year (Hover et al., 2015). The second wave came after the first wave in the late 1990s, which was primarily driven by the pursuit of weight loss and fitness, growing running opportunities (e.g., commercial running events) and improvement of professionalism. In 2006 and 2012, 10 and 13 of Dutch population between the age of 6 – 79 years run 12 times a year, among which approximately 4.7 million people and 4.8 million people joined a sports club (Scheerder et al., 2015). The third waved started from 2012, as recreational runners – runners not urging to remain attached to a formal sports club but concerning about running advice and news of athletics – were certified as a part of athletics community. In particular, many people are engaged in running as recreational runners. They start running individually or in casual groups in various environments, and favour greater and wider usage of some open spaces (e.g., public parks and natural forest) (Deelen et al., 2019). The open spaces with the presence of greenery are associated with positive health impacts and therefore are preferable to encourage participants to keep running (Pretty et al., 2007; Gladwell et al., 2013). In this respect, understanding the characteristics of environment encouraging running will potentially motivate more people living in the Netherlands to get a start.

1.1.3. A movement to GIS and GPS data-based analysis

A geographical information system is a software, framework, or computer system for storing, managing, analysing, and displaying geospatial data from a range of data sources. Data sets stored and used in a GIS comprised of spatial information (e.g., global coordinates) and non-spatial attributes (e.g., population density for a city). Differing from paper maps, GIS maps allow users to interact with data (e.g., changing mapping direction and zooming in or out on the map). Governments in many counties have implemented open data policy to spur public sectors (ESRI, n.d.a.). Thus, since the middle of last century, GIS technology has been implemented in different fields (e.g., urban planning and educational reform).

Global Positioning System is a satellite system capable of sending receivers their accurate locations on the surface of the Earth. It originated from satellite navigation experiments conducted by United States Stated Navy to track nuclear missiles in the mid-1960s (NASA, 2017). Through the years, the United States Space Force has consistently upgraded positioning capabilities and made GPS service available not only for the military but also for civil and commercial users all over the world.

Before GPS being widely applied in geographical science, studies (e.g., human travel and physical behaviour and animal migration patterns) rely substantially on data collected by direct manual observations of locations in a fixed period (Romanillos, 2016; Sileryte, 2015). The main sources employed including consultation with residents (Lopex and Hynez, 2006), in-depth qualitative interview with participants (Barnfield, 2020), historic journey data recording months or years of activities (Romanillos et al., 2016; Cervero and Kockelman, 1997), systematic observations of geographical areas using accelerometers (Cohen, 2006) and telephone survey of residents living in a neighbourhood (Lopez, 2004). Research using these sources are limited to small-scale studies as some data are not accessible for privacy reasons. In addition, because of the lack of data on geographical locations, these data potentially bring the occurrence of bias (Deelen et al., 2019).

In the recent decades, geographical studies started to consider GPS is an appropriate tool for retrieving coordinate data and storing those data in GIS database (Lange and Gilbert, 1999). The reason behind this transformation is that comparing with human perception, GPS shows a couple of practical advantages. First, GPS can obtain a wide variety of spatial data without any cost, which causes an explosion of data-driven research (Romanillos, 2016). Besides, duration analysis, GPS's capability facilitates these geographical studies to objectively measuring characteristics of environments where the subjects sit, move, or live, etc and addressing the lack of specificity regarding the geolocation of the subjects. In this way, they can scale up in terms of the amount of data applied and the geographical extent and draw a more accurate and unbiased conclusion (Romanillos, 2016; Duncan et al, 2009; Stopher et al., 2008).

1.2. Problem statement and research objective

This research will focus on a specific physical activity – running – based on its growing popularity and easy accessibility. In the context of running, running behaviours can be affected by the presence of green areas, an encounter with dogs, surfaces for running, road safety (e.g., verbal and physical harassment) and the presence of a bunch of fellow exercisers (Ettema, 2016; Collinson, 2008; Deelen et al., 2017). Thus, to promote running entails an understanding of the influences of the surrounding environment. While taking this into account, it is also worthy to note the different environmental characteristics in urban and rural areas (James et al, 2014). However, nowadays little is known about the difference among residential and running environments of runners living in urban and rural areas at present, letting alone capturing running behaviours through GPS tracking system and objectively examining the environments.

Thus, this research aims to examine the differences of environments (running and residential) between urban and rural runners and the differences between runners (urban and rural)' running and residential environments by applying GIS-based measures on GPS data. Within the urban and rural groups, it will carry out stratified analyses regarding runners running different distance, duration, at weekends/on weekdays (Rodriguze et al., 2012), inside/outside their neighbourhoods (Jones et al., 2009) and with different frequency. Expectedly, it will assist urbanists and government in garnering insights of the environmental characteristics encouraging running in urban and rural areas. With this in mind, when a spatial shift is needed due to urban growth, they will make more informed decisions. These decisions can make growth and transformation of urban areas an opportunity about long-term economic viability, environmental compatibility, and sustainable development (Hassan and Lee, 2015).

1.3. Research scope

On the account of the limitation of GPS in indoor condition, the accessibility of open geographical data, and the requirement of a supportive sports climate, this research focuses on public open space in the Netherlands. To develop a method that objectively examines the open spaces, only geographical data analysable in GIS are employed to describe their environmental characteristics. The research subjects are all runners living in the Netherlands using Endomondo and uploading their information publicly in the year of 2015.

1.4. Research questions

The general research question is:

What are the differences of running and residential environments of runners living in urban and rural areas in the Netherlands based on analysing data collected by mobile GPS tracking application?

The sub-research questions are as following:

1) How to determine whether a runner's residency is in urban or rural areas using GPS data?

2) Which environmental factors should be involved regarding their reflections of environmental characteristics of running routes?

3) Which methods should be applied to measure residential and running environments?

4) How to compare the differences of the measured environmental variables of runners living in urban and rural areas and subgroups (i.e., at weekends/on weekdays, inside/outside their neighbourhoods and different running distance, duration, and frequency)?

5) What are the differences between residential and living environments of urban and rural runners?

1.5. Research framework and method description

This section introduces the framework of this research, and provides a brief description of data, tools, and methodologies applied in this research.

1.5.1. Research framework

Figure 1.2 represents the framework of this research by providing an overview of how the research proceeds. The research begins with a literature review, moving on to acquiring and processing data and analysing results with a set of statistical models, and finishing with a discussion and conclusion.

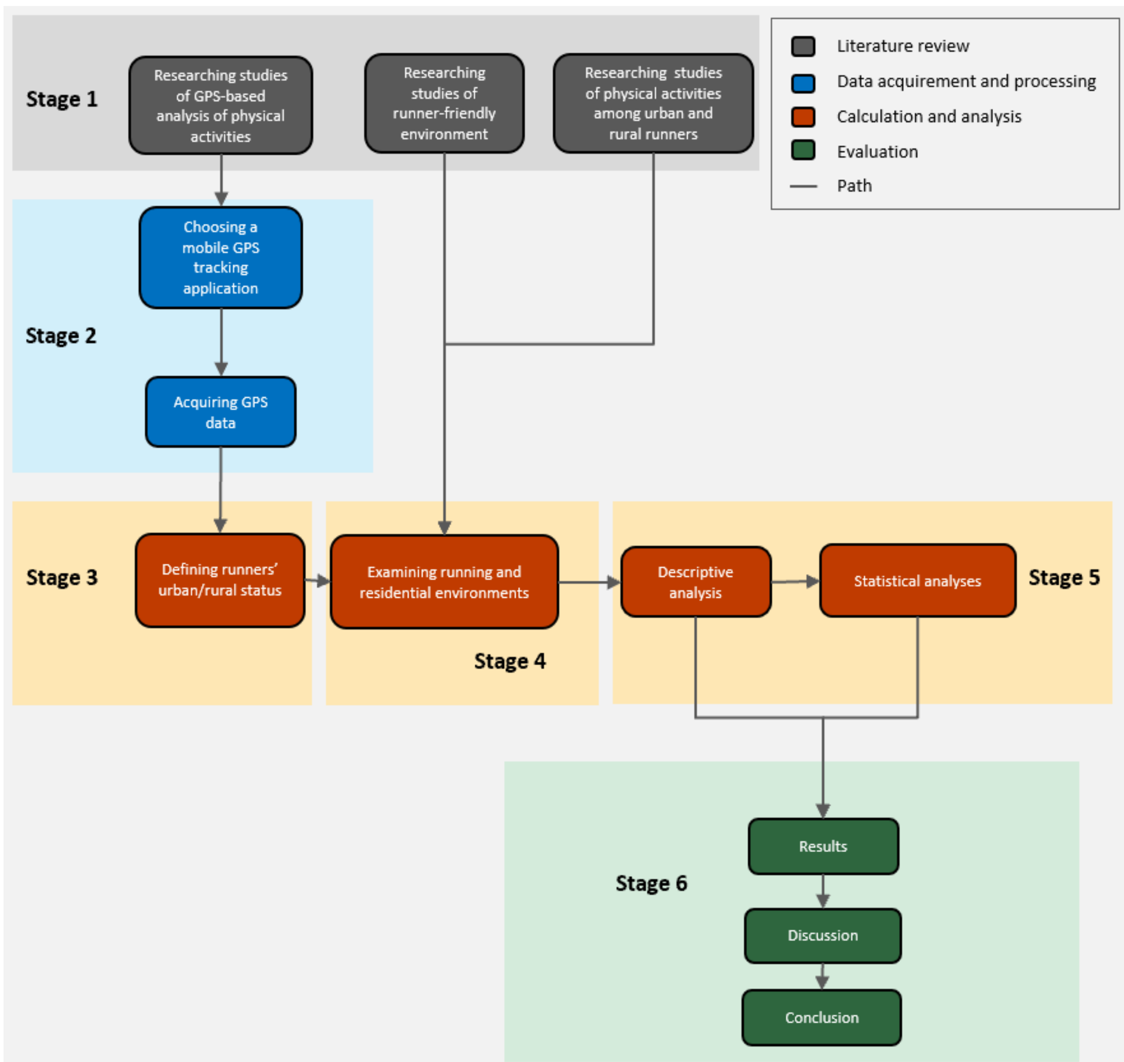


Figure 1.2: Research framework

1.5.2. Methodology description

This section presents a brief description of the methodology part (Fig. 1.2, Stage 1 – 5) in the research. These are:

Stage 1: Literature research starts at the beginning of the processes. Once the main research question is formulated, this step summarizes, analyses, and synthesizes other sources relevant to the research problem being investigated. It aims to reveal the weaknesses and inconsistencies of prior work and notify existing mythology. Literature research is also augmented throughout the research and covers the topic of GPS-based analysis of physical activities, runner-friendly environments, and physical activities among urban and rural residents.

Stage 2: A GPS tracking application is chosen based on consideration of financial investment, compatible operating system, accessible data, and its popularity in the Netherlands. Requirement of the data acquisition process is to collect open data which contains information of GPS tracked running within the study area of this research (Section 1.3). A large sample is desirable in terms of giving more reliable results. The collected data will then be stored in a GIS database as a table with unique identifications of each GPS track.

Stage 3: This process determines whether a runner living in an(a) urban/rural area. It employs the GPS data to identify the home address of each runner and then detect whether a runner's home sits in a city/countryside. Thus, the process will be carried out in a Geographic Information System with the aid of an open-source programming language. To avoid breaching General Data Protection and Regulation, the home address is computed at 6-digital postcode level and the residency is defined as urban/rural if the 6-digital polygon is inside/outside an(a) urban/rural area.

Stage 4: Environmental variables of running and residential environments were selected based on their reflection of environmental characteristics (Sallis and Owen, 1997) and availability of data. From literature review, expected environmental variables are land use mix, the level of greenness, street connectivity, residential building density, urbanization index, and presence of water body. To objectively measure these environmental variables, the measurement is performed in a Geographic Information System which is assisted by an open-source programming language. As their representations of environmental characteristics are influenced by the geographical extents (James et al., 2014), various buffers will be applied in the measurement.

Step 5: Descriptive statistic accompanied by a type of statistical analysis are conducted to compare the values of measured environmental variables of environments (running and residential) between urban and rural runners. The same descriptive and statistical analyses are used to compare the environmental variables between runners (urban and rural)' running and residential environments. Besides, the urban and rural runners are then divided into subgroups (i.e., tracks taken at weekdays/on weekends, inside/outside neighbourhoods and taken by urban and with different distance, duration, and frequency). The differences between the subgroups are also detected by applying the same analyses as the main group.

2. Theoretical underpinnings

To gain broad theoretical insights on devising research method (Chapter 3), this chapter places this research in a wide context. This chapter will review earlier principal studies in three fields: physical activities among urban and rural residents (Section 2.1), a friendly environment for running (Section 2.2) and GPS-based analysis of physical activities (Section 2.3) (Fig. 2.1).

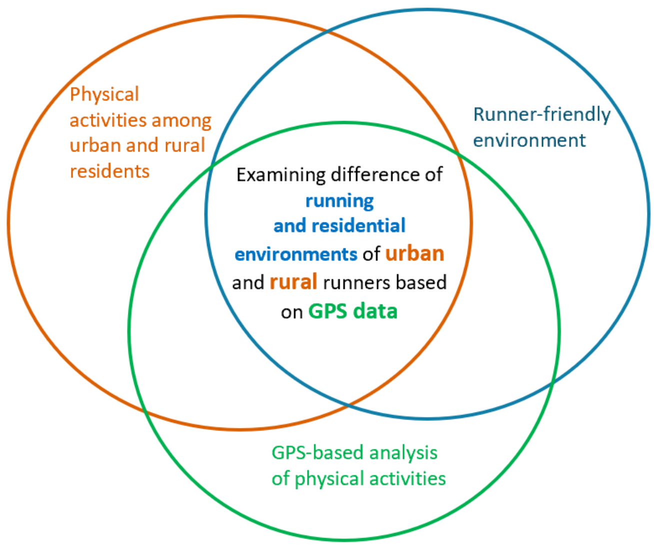


Figure 2.1: Placing this research in a wide context of previous studies. (Figure structure modified from Sileryte, 2015).

2.1. Physical activities among urban and rural residents

The concepts of urban and rural emphasize a dichotomy between characteristics of urban and rural areas. Generally, an urban area is a populated place with infrastructures of built environments; a rural area is comparatively less populous with sparse network and lower accessibility, yet more green space and closer proximity to nature (National Geographic, n.d.). In terms of lives of residents (e.g., people's socioeconomic status and working pattern), rural dwellers habitually live in larger households and are engaged in fewer types of occupational activities than urban residents (European Foundation for the Improvement of Living and Working Conditions, 2014).

The discrepancies generate substantial differences in types and levels of physical activities (e.g., active travel such as running and cycling) conducted by urban and rural residents in different age groups. Rural children spend more time on outdoor playing in their gardens or open spaces in near neighbourhoods, while urban children have higher attendance in sports clubs and more exercise equipment at homes (Loucaides et al., 2004). Consequently, rural children have higher flexibility, strength level, endurance level and lower chance to get obesity than urban children (Karkera et al., 2014; Tsimeas et al., 2005); and those urban children who often participate in sports clubs or extracurricular sports after school have better physical fitness development (e.g., upper-extremity muscle strength) than their peers not involved in any training (Golle et al 2014). Amongst adolescents, those living in rural areas spend less time being sedentary and more time on active leisure than their urban counterparts who often use computers or play video games (Regis et al., 2016). When it comes to young adults, the adults living in the countryside are more physically active and less inclined to consume excessive food when feeling overburdened comparing with adults living in the cities (Suliburska et al., 2012). For older adults and retired residents, however, their PA levels are more associated with socioeconomic factors (e.g., employment status and educational achievement), whereas their places of residents still play an important role (Mashili et al., 2018).

With the differences in mind, it is also important to note that the boundaries between urban and rural areas are being blurred by urban sprawling developments (Brody, 2013). This gradual change of living environments (e.g., segregating different land uses, reducing air quality) (Nabielek et al., 2014) will potentially affect physical activities among urban and rural residents (Ewing et al., 2014). In urban areas, the impacts include increasing people's dependency on automobiles by transforming the public green spaces (e.g., public parks) into highways and buildings. Further sprawling development requires expanding away from urban cores and out in the rural areas over highways or adjacent lands at the suburbs (EverythingConnects, 2013) (Fig. 2.2 and 2.3 illustrate the positive associations between the proportion of pollution and degrees of urbanization that is presented as a higher percentage of urban and suburban areas and a lower percentage of rural areas in European countries. While negative associations between the proportion of green areas and degrees of urbanization, respectively.). The encroachment on the countryside compacts the land-use patterns, which decreases the tendency of residents to be more active, and finally enhances 'rural' obesity rate (Ewing et al., 2014).

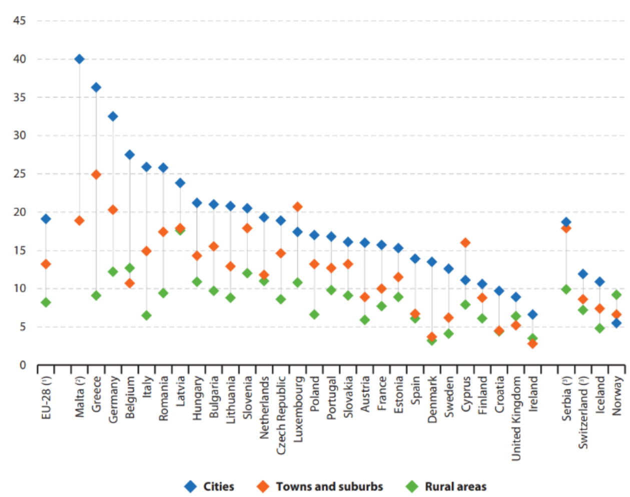


Figure 2.2: Proportion of people who perceive that they live in an area with problems related to pollution, grime, or other environmental problems, by the degree of urbanisation, 2014 (Kotzeva, 2016).

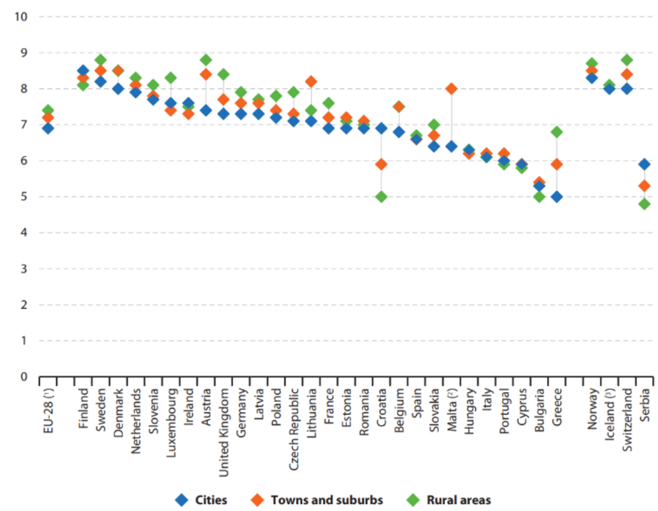


Figure 2.3: Proportion of people who perceive that they live close to recreational and green areas, by the degree of urbanisation, 2013 (Kotzeva, 2016).

Nonetheless, some researchers reported a positive association between urban growth and physical activities. According to these researchers, people living in urban areas have closer proximity to sports facilities and sports clubs, which leads more opportunities to engage in various physical activities (Liu and Walker, 2013; García et al, 2011). It is a widely acknowledged fact (O'Reilly et al., 2015); however, all previously mentioned research focused on different study areas with specific definitions of urban and rural areas. For instance, García et al's research is centred on several areas in Spain with distinct population density and Tsimeas et al (2004) defined urban areas as any towns in Greece with more than 10, 000 inhabitants. Regardless of the results, most studies investigating the performance or bouts of PA among urban and rural residents exclude factors related to the spatial context (e.g., land use density). According to those studies, the factors are not associated with the amount of time which people allocate for exercising (García et al, 2011), or the information from demographical data directly reflecting the motivations of the participants (Liu and Walker, 2013; Hoekman et al., 2017). While consequently, their results might lack an understanding of not only where people perform physical activities, but also what types of sports infrastructures are available and which are keys for reconstructing the current infrastructures or building new facilities in the future (O'Reilly et al., 2015). Thus, the influence of urbanization on physical activities examined in a certain area cannot be directly extrapolated or compared to the effect analysed in other places of the same country or a different country regarding their spatial and functional heterogeneity (Nabielek et al., 2014).

In conclusion, physical activities among urban and rural residents are affected by their socioeconomic discrepancies and geographical differences in cities and countryside. In this case, understanding the differences which have been changing during urban sprawling development is of great importance to promote physical activities.

2.2. Runner-friendly environment

Comparing with other popular physical activities such as cycling and walking, running is a typical recreational activity regarding speed, intensity, frequency, spatial constraints, and sensory experience (Ettema, 2016; Qviström, 2020). In general, running is encouraged by the routes with close connectedness to green areas such as parks and blue areas such as catchments rather than built environment (Deelen et al., 2017; Bodin and Hartig, 2003). The green spaces, especially, have restorative power and deliver aesthetic enjoyment to the runners. This affiliation for nature as explained by biophilia hypothesis also signifies the importance of the presence of green spaces to human health (e.g., beneficial emotional changes) (Gladwell et al., 2013; Bodin and Hartig, 2003). Therefore, it stands as an essential determiner of a runner-friendly environment for all runners.

In the recent decades, there are increasing recreational runners who takes up running individually or enjoy 'light' support from a casual community instead of being firmly attached to a sports club or pulled into a discernible sub-culture of running (Hitchings and Latham, 2017). Hitchings and Latham (2017) argued the overplayed role of sociality in the recreational running. In their regards, practitioners who promote the social aspects of running (e.g., organizing running events and forming running clubs) should instead ensuring the rudimentary physical infrastructures (e.g., parks and connected streets).

Ettema (2016) analysed the results from a survey of 1,581 Dutch novice runners and revealed that runners perceived the roads as more attractive and restorative when they encountered with a dog or ran on a comfortable running surface (e.g., unpaved, or paved without cars); while the perceived attractiveness and restorativeness decrease when runners anticipated hazards, experienced verbal and physical harassments and interacted with other traffic (e.g., pedestrians, cyclists and cars). Later, Deelen et al (2019) extended the research through carrying on their analysis based on survey data of 2,477 runners and dividing the runners into experienced and less experienced groups. Their outcomes were following the previous results, while all these perceived environmental characteristics exerted a stronger influence on the less experienced runners. The importance of a smooth and clean running surface with no boggy swelling and road safety (e.g., runners do not suffer accidental injuries and verbal and physical harassment) were emphasized and explained by Collinson (2008). A minor difference was that Collinson (2008) mentioned: "dogs constitute a persistent nuisance for runners in general" (49). Besides, Collinson (2008) also stated being accompanied can keep runners motivated as "assist establishing and maintaining running synchronization" (54), while to achieve running-together, the accompanied runners have to be sociable and consistently concentrate on activities.

This section discusses the findings of some previous studies to explore the determining factors of a runner-friendly environment. To summarize the discussion, running as an aerobic exercise and a recreational activity can be influenced by the physical environments. Generally, a runner-friendly environment demands the proximity to green and blue spaces, availability of physical infrastructures (e.g., parks and connected streets), a comfortable surface, road safety (e.g., avoidance of verbal and physical harassment) and the presence of a bunch of fellow exercisers.

2.3. GPS-based analysis of physical activities

GPS technology (e.g., calculate geolocation through collecting signals from satellites) spurred the development of many applications (e.g., Google Earth Application and Google Maps) (Dyal Singh College, n.d.) as the real-time geographical data are retrieved when signals of the objects are transmitted by satellites and collected by receivers after embedding their devices (e.g., phones and cars) with GPS receiver chips (Smelser and Baltes, 2001).

The emergence of these applications caused an explosion of data-driven studies of physical activities (Romanillos, 2016) due to two main advantages of GPS data. First, comparing with the previous research relying entirely on manually collected sources (e.g., historic journal data and survey), the studies scale up in terms of the amount of data applied and geographical extent (Romanillos, 2016; Duncan et al, 2009; Stopher et al., 2008). Apart from this, during analysis, the manually collected sources show the characteristics of the environments based on subjects' perceptions. They could have generated uncertainties in what subjects perceive to be a certain kind of environments due to a lack of geographical information, potentially leaving room for bias (Romanillos, 2016; Jankowska et al., 2015). By contrast, the way of measuring environmental characteristics (e.g., calculating the density of greenery within a buffer around a GPS track) is considered more objective through applying GPS data in analysing processes (Deelen, 2019; Duncan et al., 2009).

Back in 1997, Cervero and Kockelman (1997) made a start of using US census data (i.e., GPS data from US Census Bureau). They investigated if built environment influenced the travel physical behaviours (i.e., travel behaviours of physical movement) along three dimensions – density, diversity and design. Among these, density related to the compactness of neighbourhoods; diversity referred to the distribution of grocery shops, retail shops or the like in neighbourhoods; design represented the presence of green areas and the convenience of the cyclist and pedestrians to reach the surrounding neighbourhoods. They found neighbourhood density, land-use diversity, and pedestrian-oriented designs were significantly associated with the behaviours of residents.

Following Cervero and Kockelman, Troped et al (2009) were one of the most pioneer researchers using results created by subjects wearing accelerometer and GPS unit. They tested associations between built environment variables (i.e., land use mix, residential population, housing unit density, vegetation index and intersection density) and moderate-to-vigorous physical activities (e.g., cycling and running) ground on GPS data obtained from accelerometers. Their results showed that population and housing unit density were positively associated with moderate to vigorous physical activities land use mix had a similar effect; however, vegetation index surprisingly displayed an inverse relationship.

Apart from Troped et al's research, many studies in the recent decade are dedicated to exploring travel and physical activities in different environments by using GPS data. Jones et al (2009) compared the effects of a variety of environments (e.g., buildings, roads, pavements, gardens, parks and farmland) on MVPA bouts between boys and girls and stratified their analysis into urban and rural groups. Their results showed that urban children are more likely to exercise in gardens and street environments, whereas rural children, especially the girls, are more active by conducting their activities in farmland and grassland.

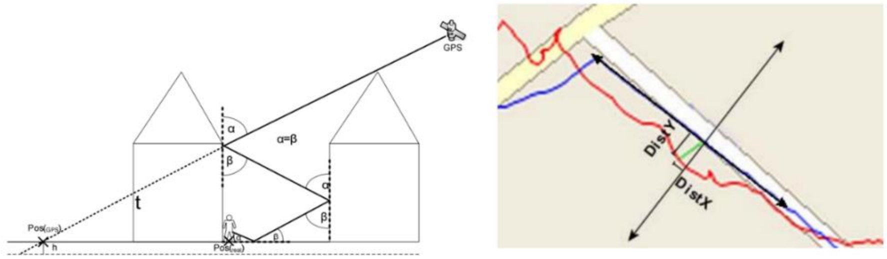
Rodríguez et al (2011) considered five environmental characteristics (i.e., population density, road length and presence of parks, schools, and food outlets) to examine associations between MVPA and surrounding environments along the GPS tracked roads. They also stratified their analysis, but instead of urban and suburban, the subgroups are weekdays and weekends. They found out that people improve their performance MVPA in roads close to parks, schools, and neighbourhoods with high population density during weekdays.

Tamura et al (2009) investigated the association between population, street density and land use mix with GPS-tracked MVPA and LVPA. They concluded a negative association between MVPA with street density and land use mix and a negative relationship between population density and LAPA. Comparatively, Dessing et al (2016) also included characteristics of land use mix in their study along with residential density and presence of water body and parks and traffic to examine which characteristics affect pupils' choice of daily commuting roads measured by a GPS device. The results displayed a preference of green areas, the presence of water and streets with fewer crossings and fewer transport areas.

The research discussed above considers physical activities as a range of activities occurring during leisure time and transportation. There are also several GPS-based studies focusing only on one or a couple of specific physical activities. Yun (2019) examined the environmental factors (i.e., walkability, walking amenities, urbanization, land use mix-diversity and accessibility and bicycle lanes) that are associated with walking behaviours of old adults. He found out the relationship between urbanization and accessibility with walking are different for subjects living in America who prefer areas with low density and in Hong Kong who disregard density. Sileryte (2015) built an urban network model on the base of associations of greenness, land use mix, network centrality and residential density with running behaviours of all runners in the Netherlands. There are quite a several studies on this topic; however, it is impractical to describe all research in the text. Thus, the reviewed literature using GPS data to analyse physical activities and the characteristics of environments that impact these activities alongside the characteristics of a runner-friendly environment (Section 2.2) are listed in Table 2.1. The inconsistency (i.e., different influences of a same environmental variable from multiple reports) in these research outcomes could be due to that these studies took different types of environmental characteristics and/or physical activities into account.

In addition to exploring the effect of surrounding environments on physical activities, GPS data are also applied to explore physical activities in diverse manners. Benson et al (2015) assessed the validity and reliability of an application stored in the iPhone in terms of its capability to monitor distance and the intensity of physical activities. Allahbakhshi et al (2020) examined GPS's role in detecting types of physical activities in real-life. Miller et al (2015) tested if light rail transit encourages new types of physical activities in a neighbourhood whose locations are reported by GPS devices. Noticeably, GPS has become a means to address the lack of specificity regarding where the activities occur, how likely will the activities happen and what activities can happen or can be generated in the future.

However, the error of GPS data should also be acknowledged, particularly the data collected from mobile tracking GPS apps (Romanillos, 2016). These errors are primarily raised by weak signals. As GPS demands uninterrupted signals from four satellites at minimum, any interference (e.g., obstruction due to tree canopy or heavy foliage) will degrade the accuracy and quality of signals. In a case that subjects wearing GPS devices travel through urban areas, any objects like tunnel, railway or premises of restaurants can hide the geolocation data of subjects. Besides, GPS devices need to balance the accuracy of signal and succession of using satellites. This means that in urban areas in particular, the signal can be influenced by the interaction between built areas and satellite constellation. This effect is often called urban canyon (Modsching et al., 2006). Moreover, when there are less than four satellites, many GPS receivers cannot provide proper estimates of geolocations (Sileryte, 2015). Urban canyon, blockage of sights of satellites and less than four satellites can potentially cause dropouts, creating a gap ranging from few seconds to minutes (Duncan et al., 2009). The common strategies to correct these errors are: 1) building a 3D model to predict the reflection of a satellite which causes an increase of time values received (Fig. 2.4 (a)); 2) using street maps as a base (Fig. 2.4 (b)) (Modsching et al., 2006).



(a). Predicting reflection in a 3-D model

(b). Correcting errors along a street

Figure 2.4: Common methods to rectify GPS errors (Modsching et al., 2006)

This section includes the development of GPS, advantages and examples of using GPS to examine physical activities, GPS errors found during analysis and commonly used strategies for correction. It provides insights for this research regarding which characteristics of surrounding environments of GPS-tracked running can be objectively measured.

2.4. Summary

This chapter is a literature review that put this research into a ground context. It helps reach decisions on which environmental characteristics can reflect the running environment of urban and rural runners. The results are displayed in table 2.1, on the base of each environmental characteristic. The environmental characteristics examined in previous studies include the presence of green areas, the presence of water body, residential density, land use mix, street connectivity/density, proximity to other facilities, the surface of the road, road safety, fellow exerciser, and urbanization. However, characteristics (e.g., presence of green area and surface of the road) of a runner-friendly environment were not measured objectively based on GPS in those earlier literature work (Hitchings and Latham, 2017; Deelen et al, 2019; Collinson, 2008). Besides, for those studies applying GPS-based analysis to explore physical activities (Rodríguez et al, 2011; Tamura et al, 2009; Jones et al., 2009), and those studies analysing associations between physical activities and residential status (urban/rural) (Loucaides et al., 2004; Karkera et al., 2014; Tsimeas et al., 2005; Golle et al 2014), none of them focused specifically on running. Thus, this study will fill these gaps by examining running and residential environments of urban and rural runner. Considering this research is ground on GPS data, the suitable environment characteristics for objective measurements are the presence of green areas, the presence of water body, land use mix, street connectivity and urbanization.

Table 2.1: Environmental characteristics analysed in GPS and non-GPS based studies. (+) indicates positive effects on physical activities, (-) means negative effects on physical activities, while if no direct or consistent effects present, no sign will be appended.

Environmental characteristics	**Non-GPS-based analysis (e.g., analysis using survey, interview, and diary data)**	**GPS-based analysis**
Presence of green area	Deelen et al., 2019 (+); Deelen et al, 2017(+); Bodin and Hartig, 2003 (+); Gladwell et al., 2013(+); Ettema, 2016(+); Hitchings and Latham, 2017(+)	Troped et al, 2009 (-); Lee and Kwan(+), 2019; Dessing et al. (+), 2016; Sileryte, 2015; Rodríguez et al, 2011 (+)
Presence of water body	Deelen et al, 2017 (+); Bodin and Hartig, 2003 (+)	Dessing et al., 2016 (+)
Residential density	Cervero and Kockelman, 1997 (+)	Rodríguez et al., 2011(+); Dessing et al., 2016; Tamura et al., 2019 (-); Sileryte, 2015; Badland et al., 2010 ; Lee and Kwan, 2019 ; Hurvitz and Moudon, 2012; Cervero and Kockelman, 1997 (+)
Land use mix	-	Troped et al, 2009(+); Jones et al., 2009 (+); Dessing et al., 2016; Tamura et al., 2019(-); Sileryte, 2015; Badland et al., 2010; Boruff et al., 2012; Yun, 2019; Cervero and Kockelman, 1997
Street connectivity/density	Hitchings and Latham, 2017(+)	Troped et al, 2009(+); Dessing et al., 2016(-); Tamura et al., 2019(-); Sileryte, 2015 ; Badland et al, 2010 ; Hurvitz and Moudon, 2012 ;
Proximity to other facilities (e.g., food outlet and supermarket)	-	Harrison et al., 2014 ; Rodríguez et al., 2011(-) ; Reitzel et al., 2013 (+)
Surface of the road	Ettema, 2016(+); Deelen et al, 2019(+); Collinson, 2008(+)	-
Road safety (e.g., avoidance of verbal and harassment)	Ettema, 2016(+); Deelen et al, 2019(+); Collinson, 2008(+)	-
Fellow exerciser	Collinson, 2008(+), Hitchings and Latham, 2017	-
Urbanization (e.g., population density, numbers of household and built and natural environment in rural/urban areas)	Loucaides et al., 2004; Karkera et al., 2014; Tsimeas et al., 2005; Golle et al 2014; Regis et al., 2016; Suliburska et al., 2012; Mashili et al., 2018; Ewing et al., 2014(-); Liu and Walker, 2013(+); García et al, 2011(+); Hoekman et al., 2017(-)	Troped et al, 2009(+); Jones et al., 2009(-); Tamura et al., 2019 (-); Klompmaker et al., 2019; Yun, 2019

3. Methods

Given findings from the literature review, this chapter elaborates on the data, tools, and methods of objectively estimating and analysing environmental variances among urban and rural runners. Characteristics of preferred environments of urban/rural runners running different distance, duration, at weekends/on weekdays, inside/outside their neighbourhoods and with different running frequency are also considered in the analysis. Thereby, this research intends to develop a method that effectively compares running environment favoured by runners according to their urban/rural status, running extent, time, duration, distance and frequency.

3.1. Tools

There are many tools that could be used for this research. These software products range from commercial software such as ESRI's well-developed ArcGIS and StataCorp's powerful Stata to open-source offerings such as commonly used QGIS and user-friendly Python. This research chooses the open-source software where possible as the commercial software require a licence that will restrict the methodology to be repeatable and extendable for future studies. All tools applied in this research are described below.

Postgres database system with its external extension PostGIS: Postgres is an open-source relational database management system which allows efficient querying and data management (Galpern, 2018). PostGIS is implemented as a Postgres external extension that adds supports for geographical objects. With the extension, the database is widely used by geographical studies.

Python: Python is an open-source programming language. It holds a collection of packages (e.g., psycopg2) to connect with Postgres and PostGIS database and process geospatial data, which makes measurement effective. In addition, Python is widely implemented for eliminating GPS errors and performing a calculation on GPS data (Kwan et al., 2019).

QGIS: *QGIS is a* Desktop GIS that helps to edit and visualise data. It provides a range of plugin (e.g., DBconnection and GDAL) for interaction with database and raster/vector data transformation (Gengeç et al., 2009; Sileryte, 2015). Its efficiency of data processing (e.g., data access, filtering) and capabilities of rendering high-quality images support studies using large datasets.

RStudio: RStudio is an integrated development environment for a programming language R***.*** R is commonly used for statistical computing and data analysis. It is a free software environment with multiple built-in packages for GPS data manipulation and statistical computing. Some of these packages (e.g., dbConnect) allow written-up algorithms to interact with Postgres and PostGIS database (Bivand et al., 2008; Viana et al, 2019).

3.2. Data collection and description

This section includes the details of all data needed for analysing running environments among urban and rural runners in this research. It starts with collecting GPS data. Then, it presents the basics of supplementary geographical and census data (e.g., data of environmental factors, functional urban areas). After, it ends up with processing data as inputs for subsequent analysis.

3.2.1. Running tracks

To make this research replicable and extendable without the need for significant financial investment, mobile GPS tracking application than a dedicated GPS device is chosen as a more suitable option. Endomondo, Strava and Runtastic are three of the most popular mobile tracking applications in Europe and chosen in a variety of studies analysing physical activities. Their compatible operating system, methods for accessing activities data and downloadable information are compared in Table 3.1. As shown in the table, Endomondo is installable in more operating systems, provides accessible data and more downloadable information. Considering these aspects, Endomondo is chosen for this research.

Table 3.1: Comparison between Strava, Endomondo and Runtastic.

Application	Compatible operating system	Data access	Downloadable information
Endomondo	IOS, Android, Blackberry, Windows Phone	Through reading html for data uploaded and set as public	Workout id, author id, workout type, workout duration, workout distance, maximum speed, average speed, minimum altitude, maximum altitude, total accent, total descent, start time, heart rate*, weather type, hydration, geographical location (Endomondo, 2020)
Strava	Android, IOS, Web bower	Paid access	Workout type, duration, activity counts during a workout, time, geographical location with a distinction between subscribed and non-subscribed users (Strava, 2020)
Runtastic	Android, IOS, Windows phone	Paid access	Distance, temperature, goal and motivation (e.g. yearly running goal, weight goal), training plan information (e.g. start date, training plan, associated fitness activities) and shoe information (brand, model, size, colour) (Runtastic, n. d.)

*: Available for premium users.

Endomondo tracks information locally on individual device and users can upload their tracks automatically via a connection with Garmin Connect, Polar Flow and Jabra Sport Life app (Endomondo, n. d. a) or manually by drawing their workout route on Endomondo map and filling in other workout information (e.g., duration). If users make their workouts publicly available, it will be viewed on “www.endomondo.com/workouts/ + workout ID”.

Every publicly available workout is also downloadable as a JSON object containing a range of attributes depending on the uploaded information (Table 3.1.). Some of this information (e.g., distance, duration, maximum speed, and maximum altitude) are generated directly through the creation of GPS trajectories, while other information (e.g., hydration and calories) are personal information uploaded by an individual user. Date and time are attained as the local time displayed on users' mobile devices. Endomondo will automatically record weather data through its cooperation with AccuWeather.com, but which require users to active data connection while conducting their activities (Endomondo, n. d. c.). For perineum users, Endomondo has a tool named heart rate zones to monitor the intensity of workouts. For a thorough heart rate analysis that brings information about heart and running performance (Fig. 3.1), users need to wear a compatible heart wear monitor (Endomondo, n. d. b.).

Figure 3.1: Heart rate analysis by *heart rate zones* (Endomondo, n. d. a.).

In this research, the downloaded JSON data range from February to December 2015. The original intention was to collect all data for the year of 2015, however, the data in January were not retrieved successfully. The reason behind this might be that the athletic apparel maker, Under Armour, acquired Endomondo in February 2015, which made Endomondo more popular in a large number of people using Under Amour. For each retrieved workout, it comprises overall information such as the whole workout duration, starting and ending time, and weather, and also detailed information for each count in the workout such as the geolocation and time of a specific GPS point at an interval of 7 to 15 seconds. The data are then stored in a database as a table, in which every workout has unique identification and information (e.g., total duration, distance, descent, a maximum speed of the whole workout, and geometry as a line from assembling discrete GPS points). Then, the workout types other than 0 which indicates running (Isoteemu, 2014)) are filtered out. The remaining data are shown in Fig. 3. 2, from which the numbers of the running tracks vary from month to month.

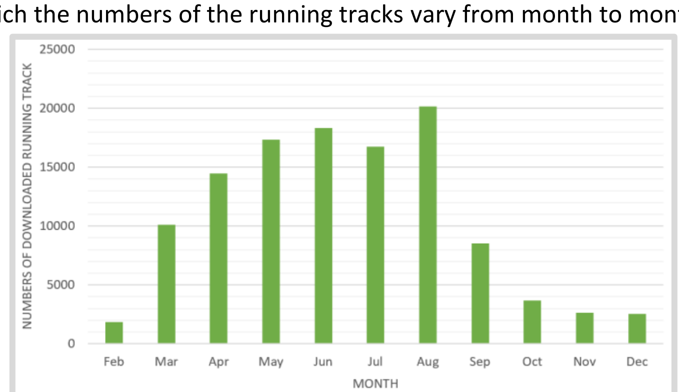


Figure 3.2: Downloaded running tracks per month in 2015

Due to the extensive amount of time required for downloading data, this research only uses data from one mobile GPS tracking application, in which case the runners who do not take their phones with them while running or runners use other applications are not accounted. Besides, among all Endomondo users, those who prefer not showing their routes and do not know how to upload their information may not be represented. Thus, the data sample might not be representable in terms of all runners in the Netherlands.

Another issue of the data is incorrect GPS data. Hundreds of tracks are recorded as less than 10 metres or taken for less than 1 minute. The abrupt losses of the signal are most likely caused by that the runners turn off their GPS setting too soon after they start working out or other applications on their mobile devices interrupted the Endomondo's GPS connection, unexpectedly or purposely. While after cleaning out these data, other GPS errors emerge. Some examples of jumping points, which are viewed as elongated tracks, are marked in Fig. 3.3 in red.

Figure 3.3: Running tracks after eliminating tracks less than 10 metres or taken for less than 1 minute. A couple of *jumping points* are marked in red. The blue layer of European countries is derived from Countries 2020 from Eurostat (Eurostat, n.d.) for visual representation of GPS data inside and outside the Netherlands.

There are several reasons for the defective GPS points. First, satellite reception is lost (Duncan et al., 2009), or when there are only less than four satellites available and they produce a rough estimate varying drastically from the real locations (Sileryte, 2015). Besides, as discussed in section 2.3, any blockage like buildings or tree canopy, and foggy weather will degrade the quality and accuracy of GPS data, or in a worse case, cease GPS data to be generated (Modsching et al., 2006; Duncan et al., 2009). These two are the most common cases and happen frequently when it comes to using smartphone devices which allows users to obtain global positioning data on a phone via mobile networks (Merry et al., 2019). It is also important to address that normally recreational GPS implanted on a smartphone brings less accurate spatial positioning data than a professional GPS (ESRI, n.d.a.).

Preferably, all these error data should be corrected for an unbiased sample. However, rectifying these errors require intricate GIS techniques and some other information from the users (e.g., why there were sudden dropouts after they only start running), which are beyond the scope of this research. Thus, the aforementioned errors will be removed.

3.2.2. Geographical and census data

Decisions of what data to collect for measuring environmental variables are based on the discussion from the literature review. This research favours open data as which provide an opportunity for more future studies to follow and extend the method applied in this research. Besides, as this research has conveyed its intention to objectively analyse the characteristics of running and residential environments of urban and rural runners, it uses geographical data which can be processed and visualized by open-source software or open programming languages. Apart from the geographical data for environmental characteristics, this section also includes data census delineating urban areas, neighbourhoods and 6-digit postcode areas in the Netherlands.

Basic information (e.g., source and publication year) of data obtained are presented in Table 3.2. More detailed description (e.g., content) are shown in the paragraphs below.

NDVI image: The satellite-derived Normalized Difference Vegetation Index image has a spatial resolution of 30m. NDVI values range from −1 to 1, in which negative values correspond to water, values near 0 correspond to barren areas of rocks, sands, and snow, and positive values correspond to the presence of greenery (values increase as the quantity greenery detected increase).

Bestand Bodemgebruik: The land-use dataset contains digital geometry represents land use (e.g., business park, buildings, recreation parks and indoor and outdoor water) in the Netherlands (PDOK, 2019).

Basisregistratie Adressen en Gebouwen: The data consists of all addressable objects (e.g., buildings) and their addresses in the Netherlands (PDOK, n. d. a.).

Key Register Topography at 1:10,000: The file is a digital object-oriented topographic base map as a part of BRT, which contains various elements (e.g., roads) in the Netherlands (PDOK, n. d. b.).

Land Use Database of the Netherlands: The LGN is based on satellite imagery and additional data (e.g., BAG) in the Netherlands, which divides land use up into thirty-nine classes with a resolution of 5m (Wageningen University & Research, n.d.).

Functional urban areas: An FUA comprises densely inhabited city based on population density and travel-to-work flows and surrounding areas (commuting zone) with labour market highly integrated with the city (OECD, n.d.).

6-digit postcode areas: Postcode area on average contains 15 to 20 addresses (PC6 - 6 - digit postal NL, n.d.).

Kerncijfers wijken en buurten: Datasets holds the geometry of all municipalities, districts and neighbourhoods in the Netherlands. The boundaries of districts and neighbourhoods are drawn from information reported by the municipalities to Statistics Netherlands and the municipal boundaries are derived from Basisregistratie Kadaster (PDOK, 2020).

Table 3.2: Basic information of supplementary geographical data and census data.

Data		**Publication year**	**Source**
Geographical data	NDVI image derived from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager	2015	Google Earth Engine cloud computing platform
	Bestand Bodemgebruik	2015	Statistics Netherlands
	Basisregistratie Adressen en Gebouwen	2015	Statistics Netherlands
	Key Register Topography (BRT) at 1:10,000	2015	Statistics Netherlands
	Land Use Database of the Netherlands	2015	Wageningen University & Research
Census data	Functional urban areas	2015	Organisation for Economic Co-operation and Development
	6-digit postcode area	2015	ArcGIS Hub
	Kerncijfers wijken en buurten	2015	Statistics Netherlands

3.3. Data processing

This section details on how this research divides the runners into urban and rural runners and subdivides the main groups according to runners' running distance, time, frequency and extent.

3.3.1. Dividing groups regarding running distance, duration and frequency

A runner's running frequency is calculated by summing up the counts of "workout_id" (i.e., unique index of each GPS track) of the runner. On account of those runners with one record probably were not willing to stick with Edomondo for the entire year, here only considers runners taking multiple tracks. The recommended level of beginners is twice a week (NHS, n.d.). With this in mind, this research classifies data into runners practicing 24 times a year, between 24 to 48 times and more than 48 time a year. The three groups of runners are shown in Table 3.3.

Runners' running duration were recorded alongside "workout_id" of each GPS track. From most runner-friendly websites and mainstream media, the recommendation of beginners is to start running up to 30 minutes. Hence, this research further divides urban and rural runner into groups running less 30 minutes and above 30 minutes (Table 3.3).

Running distance (length) of a GPS track if computed by using *ST_length* function from PostgreSQL. research uses 0 – 3000m for novice runner as 3000m is commonly considered as middle-distance. A long-distance track is usually viewed as a track more than 3000m and an extremely long track is referred to a track more than 5000m. Here sets 3 levels of running distance (Table 3.3).

Table 3.3: Classifications of GPS data regarding running distance, duration, and frequency.

Running duration (min)	Running frequency	Running distance (m)
<= 30	< = 24	< = 3000
> 30	24 – 48	3000 – 5000
	> 48	> 5000

3.3.2. Defining urban/rural status

Following James et al (2014), this research will define runners’ urban/rural status based on their home location. Due to a lack of demographic data on whether the runners living in cities or countryside, the collected data (i.e., GPS tracks, functional urban areas and 6-digit postcode areas) are employed to define the runners urban/rural status. First, GPS data and 6-digit postcode areas are used to seek out the runner’s home. Then, the functional urban areas are added to divide the postcode areas into urban and rural. A runner will be determined as urban/rural if one’s home is sitting in an urban/rural area.

There are two parts in the progress of detecting a runner’s home at a 6-digital postcode level. Part 1 calculates the most referred pc6 polygons containing residential buildings and intersecting with the starting point (the first point) of each track. In the case that the runners who have multiple courses, the most frequently referred pc6(s) are/is chosen. While if there is more than one pc6 have the same frequency, a candidate will be randomly placed. This execution is explained in pseudocode 1:

```
for each runner:
        starting_points = first points of all GPS tracks
        select pc6 intersected with starting_points in all pc6s and starting_points
        pc6 = pc6 intersecting with GPS tracks
        if length(pc6) > 1:
                select pc6 with highest frequency
                pc6_1 = a most frequently called pc6
                if length(pc6_1) > 1:
                    select pc6_1 randomly
                     pc6_2 = a random pc6_1
                     return pc6_2
                 return pc6_1
        return pc6
```

Pseudocode 1: Part 1 for detecting at a 6-digital postcode level.

Part 2 considers both starting and ending points and requires the starting and ending points (the first and last points) locating in the same pc6 polygon where stands residential buildings. Then repeat the steps in part 1 when there are multiple routes taken by a runner. This execution is explained in pseudocode 2:

```
for each runner:
        starting_points = first points of all GPS tracks
        select pc6 intersected with starting_points in (all pc6s and starting_points)
        pc6_strs = pc6 selected
        return pc6_strs

for each runner:
        ending_points = last points of all GPS tracks
        select pc6 intersected with ending_points in (all pc6s and ending_points)
        pc6_ends = pc6 selected
        return pc6_ends

pc6_str_end = an empty list
for pc6_str in pc6_strs:
        if pc6_str is in pc6_ends:
                pc6_str_end.append(pc6_str)
                if length(pc_str_end) > 1:
                select pc6_str_end with highest frequency
                pc6_1 = a most frequently called pc6
                if length(pc6_1) > 1:
                    select pc6_1 randomly
                    pc6_2 = a random pc6_1
                    return pc6_2
                return pc6_1
                if else:
                return pc6_str
```

Pseudocode 2: Part 2 for detecting at a 6-digital postcode level.

The reasoning behind the execution of part 1 is that for most tracks, the starting points jammed in a couple of seamlessly connected 6-digit postcode polygons. As a runner who intends to record their route is expected to switch on GPS connection with Endomondo when starting on the journey, a 6-digit postcode polygon where most starting points sit could potentially be the home area of the runner (Fig. 3.4 (a)). However, it is not the same case for the ending points. Most ending points disperse over a large extent, which diminishes the chance that they can be spotted in the same polygon (Fig. 3.4 (a)). Another scenario (part 2) is that the runner starts running from home and take back home while disabling GPS connection. If it happens, the polygon contains the starting and ending points would plausibly processes the runner's home address (Fig. 3.4 (b)) (practically, *intersection* is used in the scripts as there could be points on the edge of the polygon). To increase the reliability of the home-detection method, only the runners taking multiple tracks (i.e., running frequency higher more than 1, Section 3.3.1.) are considered when performing part 1 and part 2. After completing the performance, the results of part 1 and part 2 are compared and the matched results are determined as the home area of the runners.

Figure 3.4: All tracks of two runners. (a) shows scenario one in which several starting points of runner one clustering in a couple of adjacent pc6 polygons. (b) demonstrates scenario two in which runner two starts and ends running in a same pc6 polygon A.

Resultantly, the method detected home areas for a total of 3799 runners. The next progression is to divide these areas into urban/rural areas. The polygon will be defined as urban/rural if its centre locates inside/outside the FUAs (Eurostat, 2018) (Fig. 3.5). In this step, the *ST_Intersect* (spatial function from PostgreSQL) instead of *ST_Within* (spatial function from PostgreSQL) is utilized and the centroid points p6c polygons rather than the whole polygons are utilized for intersection in case that the edges of some pc6 polygons extend out the FUAs while most of the polygon sits inside the FUAs (Fig. 3.6) and edges of some pc6 polygons intersect with the FUAs while most of the polygon is outside the FUAs (Fig. 3.6). Finally, the runner's urban/rural status is defined as urban/rural if they live in a pc6 whose centre is inside/outside the FUAs.

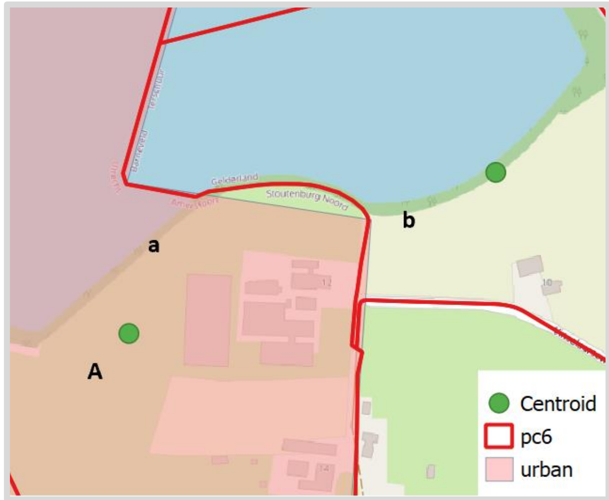


Figure 3.6: Example of edges of pc6 polygons and edges of FUAs not overlapping. Pc6 polygons a and b both intersects with urban area A, but regarding locations of their centroids, only pc6 polygon a is inside A.

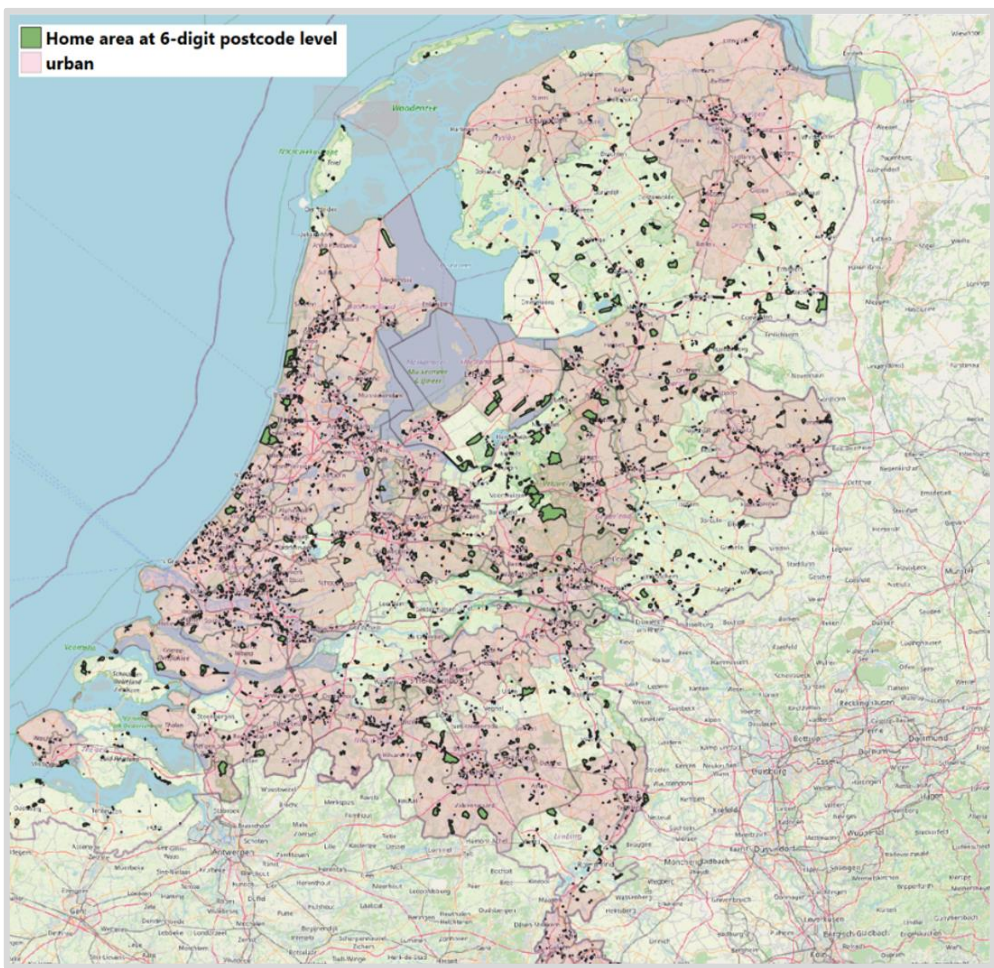


Figure 3.5: Detected home areas at 6-digit postcode level inside/outside FUAs.

3.3.3. Identifying tracks inside/outside neighbourhood

This section will elaborate on the processes to determine which tracks are inside or outside the neighbourhood where lives an urban/rural runner. Accounting the lack of information on which neighbourhood the runner lives, the method is rooted on home defined at the 6-digit postal code level in the preceding section.

The first step is to find neighbourhoods for the runners. Namely, in which neighbourhoods the runners' home determined at pc6 level locate. The neighbourhoods are extracted from District and Neighbourhood Map released by CBS in 2015, each of which holds several 6-digit postal code areas. During the computational process, the *ST_Intersect* function from PostgreSQL is utilized for calculating which neighbourhoods intersect with the pc6 polygons containing runners' home addresses. It happens that the edges of pc6 polygons intersect with the outlines of multiple neighbourhoods (Fig. 3.6). Therefore, the neighbourhood of a runner is defined as the neighbourhood which intersects the centroid of the runner's home pc6 (explained in pseudocode 3).

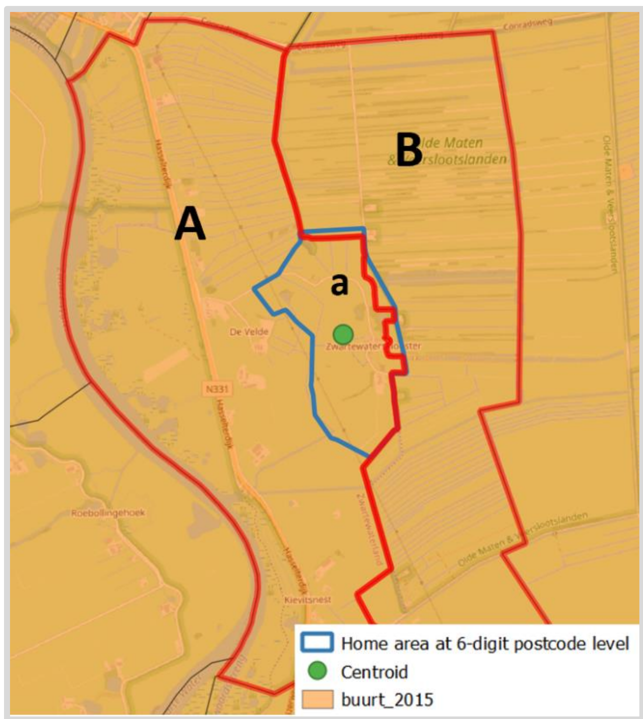


Figure 3.9: Examples of 6-digit postal code areas intersecting with multiple neighbourhoods. Centroid of pc6 polygon a is situated in neighbourhood A (outlined in red), while some of its edges intersect with the outline of neighbourhood B (outlined in red).

```
for each runner:
        select neighbourhood intersected with the runner's home pc6
        neighb = neighbourhood selected
        return neighb
```

Pseudocode 3: Identify the neighbourhood of each runner.

Afterwards, the neighbourhoods intersected with all tracks of each runner are computed also by using *ST_Intersect* function from PostgreSQL. If a track taken by a runner extends to other neighbourhoods than the residential neighbourhood of the runner, this track is defined as a track outside neighbourhood (explained in pseudocode 4). In the other case (a track does not extend outside the runner's residential neighbourhood), the track is defined as a track inside neighbourhood. An example of all routes taken by a runner inside/outside the runner's neighbourhood is shown in Figure 3.10.

```
for each runner:
        select neighbourhood intersected with the runner's home pc6
        neighb = neighbourhood selected
        return neighb

for each track of urban/rural runner:
        neighbourhoods_al = all neighbourhoods intersected with the track
        track_n  = the neighbourhood of a runner who took the track
        n_neighbourhoods_al = the number of neighbourhoods_al
        if track_n is in neighbourhoods_al and n_neighbourhoods_al > 0:
                if n_neighbourhoods_al = 1:
                        return 'Track_inside_neighbourhood'
                if n_neighbourhoods_al > 1:
                        return 'Track_outside_neighbourhood'
        else:
                return 'No_track_al_all'
```

Pseudocode 4: Defining all tracks of a runner taken inside/outside the runner's neighbourhood.

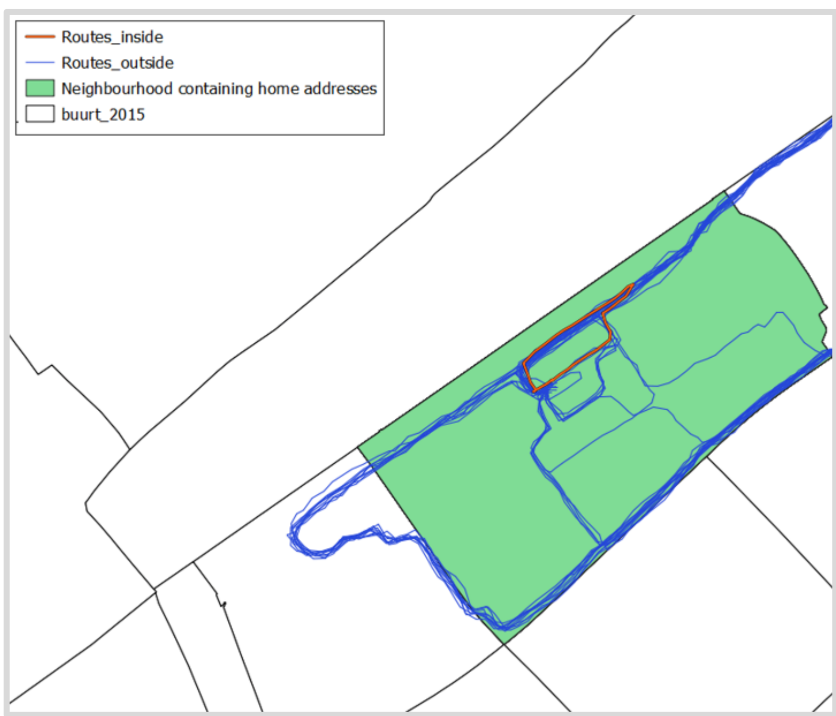


Figure 10: Routes taken a runner inside/outside neighbourhood containing the runner's home address.

3.3.4. Measuring environmental variables

According to previous findings of literature review (Chapter 2), six objective environmental variables need to be measured: the level of greenness (i.e., presence of green areas), residential building density, land use mix, street connectivity, levels of urbanization and the proportion of blue space.

These environmental variables are contextual variables; therefore, their representations of the relevant spatial contexts are influenced by the geographical extents in which they are measured (James et al., 2014). There are multiple types of buffers used by researchers to delineate the geographical context. Circular buffers (e.g., radial, and Euclidean buffers) are created by drawing a circle along a straight line at some distance between two points (e.g., 500m from home) (James et al., 2014; Mavoa et al., 2019). This type of buffer captures the areas where physical activities happen, thus, it is popularly used to investigate the effects of neighbourhood environments (Lee and Kwan, 2019). For describing dynamic exposures (e.g., running environments), a commonly applied buffer line-based network buffer which defines a distance to limit the line traced along with a street network (Oliver et al., 2007). It can provide a more accurate delineation of the spatial context in terms of specific physical activity types (e.g., cycling and running), but it requires to find the appropriate buffer sizes for the activities. Another approach is to measure the dynamic environmental exposures along the GPS trajectories. As found by Lee and Kwan (2019), the buffer sizes affected the measured values of environmental variables, and larger buffer size (> 150m) were less accurate than smaller buffers in terms of the representation of the real-world environment.

In research, since running is regarded as a recreational activity, running distance, speed and spatial context vary from one participant to another. It adds a burden on estimating and selecting an appropriate buffer for the network-based measurements that are particularly sensitive to the buffer size. To avoid making an arbitrary decision and only considering the influences of the surrounding environment of each GPS track and home address, the environmental variables are calculated with a set of 25 (Dessing et al., 2016; Farrell et al., 2015; Badland et al., 2010), 50 (Tamura et al., 2019; Rodríguez et al., 2012; Farrell et al., 2015; Oliver et al., 2007; Feldt and Schlecht., 2016; Tamura et al, 2019; James et al., 2014) and 100m (Badland et al., 2010) buffers along the GPS tracks and within 1,000 radius (Sileryte, 2015; Madsen et al., 2014; Berke et al., 2007) of centroid of each pc6.

As discussed in Section 1.2, the primary goal of measuring these environmental variables is to detect their differences among urban and rural runners. In this regard, it is necessary to address here that the focus is not to compare the urban and rural running environments in the Netherlands, but the environments to which the urban and rural runners are exposed to while running. The difference lies in the fact that urban runners might travel from the vibrant city cores to the countryside for a peaceful run and runners living in rural areas would like to extend their journey to sightsee in city. Description and explanation of the definition and measurement of the individual environmental variable will be detailed in the following sub-sections.

3.3.4.1. Level of greenness

NDVI calculates the difference between reflectance in near-infrared and red wavelengths from the objects' surface divided by their sum (Tucker, 1979), which ranges from -1 to 1:

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

It is an extensively used remote sensing indicator to predict the quantity of greenery (Tamura et al., 2019; Viana et al., 2019; Helbich, 2019). The reason for it is that plants, especially healthy plants with high chlorophyll content and well-formed cell structure, actively absorb red light and reflect a large amount of near-infrared (Fig. 3.6). Comparatively, non-vegetated features such as barren areas of rocks, sands, and snow have zero values due to low reflectance in both red and near-infrared bands (Fig 3.6).

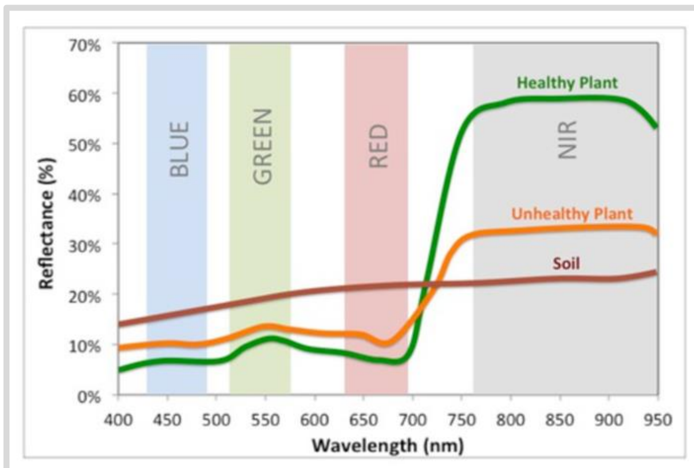


Figure 3.6: Spectral reflectance of healthy plant, unhealthy plant, and soil in visible and NIR wavelengths (Warren, 2011).

In terms of this research, the level of the greenness of areas enveloping each GPS track and home address are calculated as the average of Normalized Difference Vegetation Index of all cells from an NDVI satellite image derived from Landsat 7 Enhanced Thematic Mapper Plus and Landsat 8 Operational Land Imager with spatial resolution of 30m × 30m within a buffered zone. The calculated values of the presence of green areas are expected to vary between -1 and 1, where a high value indicates a higher coverage of greenery within the buffered zones.

3.3.4.2. Residential building density

Residential building density is the intensity of buildings with a residential function to a certain extent. Generally, urban runners travel around blocks with higher residential density than runners living in rural areas (Jones et al., 2009). While as stated earlier, it is also possible that they run out their living neighbourhoods to places with less compacted housing and more lively countryside scene. With this in mind, this method accounts both the residential building density along each GPS track by those runners (Fig. 3.8) and the density of dwellings in the runner's living neighbourhood. The first step of the measurement is to extract the residential buildings among all types of buildings in BAG published on CBS in 2015. For any residential buildings intersected with a buffer, the floor areas of the buildings are computed and accumulated as the total area of residential buildings. Then the total area is divided by the area of the buffer. When the total area of residential buildings is larger than the area of the buffer, the value of the ratio will be set as 1 which is the maximum value of measured residential building density. The density will be 0 (i.e., minimum value) if there is no residence in the buffer area.

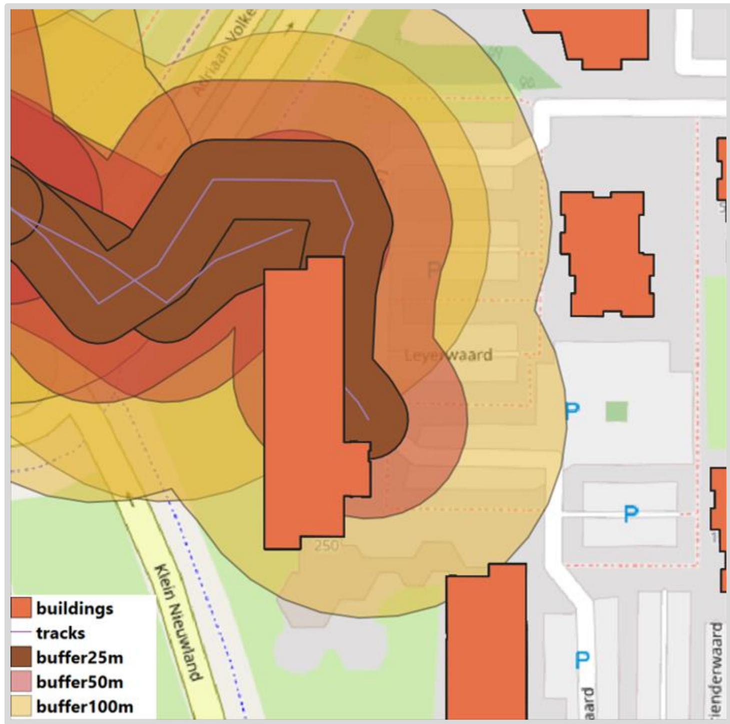


Figure 3.8: Residential buildings extract from BAG 2015 intersecting with 25, 50 and 100m buffers of a GPS track.

3.3.4.3. Urbanization index

Urbanization index is measured as the degree of urbanization in a region. Residents in urbanized areas have relatively restricted open spaces than peopling living in less compacted communities. In this regard, the degree of urbanization is chosen to examine whether running environments are different among urban and rural runners (Lee et al., 2009).

Degree of urbanization can be measured based on the density of addresses (Klompmaker et al, 2019; Deelen et al., 2017). In Klompmaker et al's (2019) study, highly urbanized areas are defined as blocks with more than 8,500 units per km2 and non-urbanized areas hold less than 8,500 units per km2 and more than 900 units per km2. While Deelen et al (2017) categorised urban density into three classes: rural areas with less than 500 addresses per km2, barely to moderately urbanized areas with 500 to 1500 addresses per km2, and strongly urbanized areas with more than 1500 addresses per km2. Following their studies, this research employs urbanization index to indicate the degree at which running environment of runners living in urban/rural areas are urbanized. In this sense, urbanization index is a diagnostic factor that reflects the influence of urbanization on runners' preference of routes and investigates whether their choices are restricted by habitats. The measurement is similar to residential building density, address density is calculated by firstly selecting all addressable units in BAG from CBS in 2015, then summing up the number of units intersected with a buffer and finally dividing the total number of intersected units by the area of the buffer. If a calculated value equals zero, it (minimum) indicates no residence in the buffered zone.

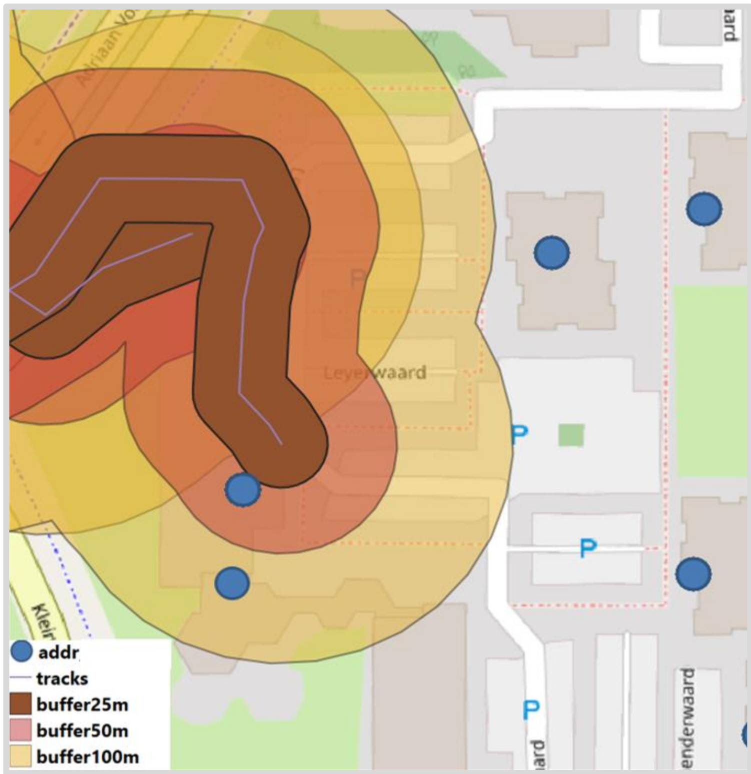


Figure 3.9: Residential units extract from BAG 2015 intersecting with 25, 50 and 100m buffers of a GPS track. Addresses are shown as blue dots.

3.3.4.4. Land use mix

Land use mix describes the land-use heterogeneity in a delimited environment. Many previous studies revealed that neighbourhoods with diverse land use increase residents' propensity to walk (Tamura et al., 2019; Oliver et al., 2007; Turrell et al, 2013). Being inspired by the prior work, this research examines land use diversity in runners' residential and running environments.

The land parcel information was obtained from Dutch land-use dataset released in 2015 that contains digital geometries represent 37 classes of land use. Following Turrell et al (2012), they are grouped into five categories: recreational, residential, industrial, commercial and the other (Table 3.4).

The five categories of land use are then used in calculation of land use diversity which is based on an entropy equation created by Leslie et al (2007), where all parameters are calculated within a buffered zone (e.g., GPS track based buffers, Fig. 3.10):

$$- \frac{\sum_k (p_k \, ln p_k)}{\ln N}$$

k – land use category

p – proportion of land area

N – the number of land-use groups

The result ranges from 0 to 1, with 0 representing homogeneity (i.e., the whole area is dominated by a single type of land use) of land use within a buffer, and 1 representing all groups of land use classes are evenly distributed within the buffer.

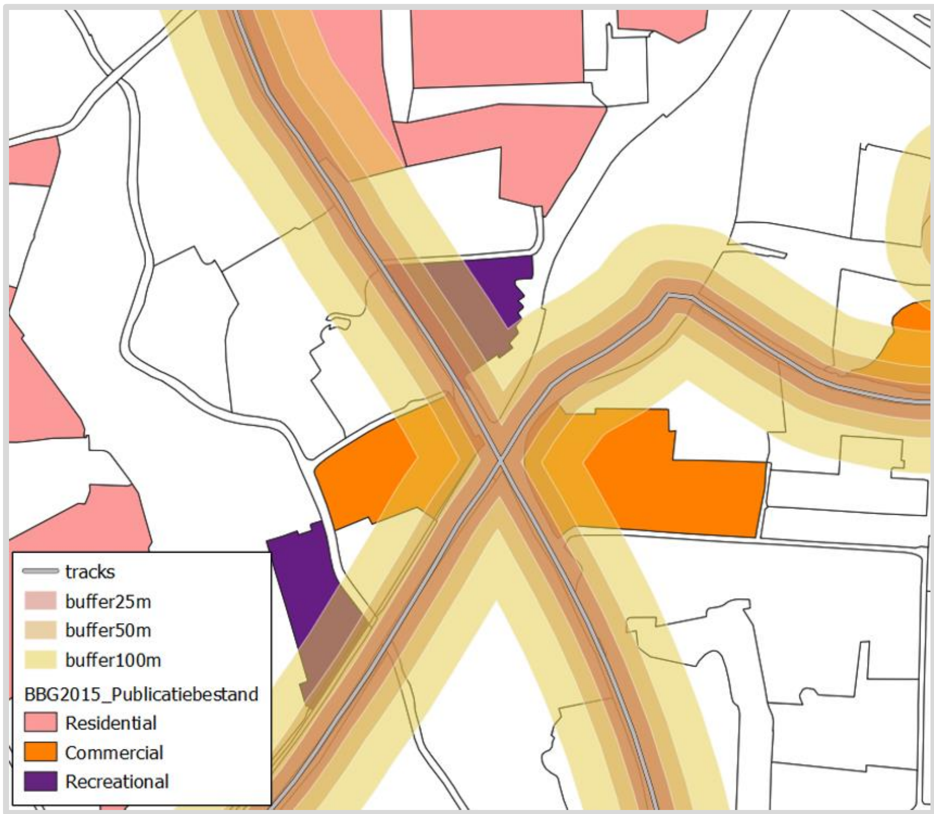


Figure 3.10: 3 categories (residential, commercial and recreationalof grouped from 37 land uses in BBG intersecting with 25, 50 and 100m buffers of multiple paths taken by a runner.

Table 3.4: 37 land use classes and groups.

Group	Land use classes
Recreational	Areas for public facilities
	Area for social cultural facilities social and cultural
	Business park
	Park and public garden parks
	Sport terrain
	Allotment
	Day recreation area recreational areas (daytime)
	Recreational area Recreational areas (campsites etc ...)
	Land for greenhouse horticulture greenhouses
	Other agricultural lands
	Forest
	Open dry natural terrain or open natural areas (dry)
	Open wet natural terrain or open natural areas (wet)
	Recreational backwater
	Wadden Sea, Ems, Dollard
	Eastern Scheldt Eastern Scheldt
	Western Scheldt Western Scheldt
	North Sea North Sea
Residential	Residential land-dwelling areas
Commercial	Terrain for retail and hospitality industry (e.g., cafes, restaurants)
	Retail area
Industrial	Landfill dump
	Wrecks storage depot wrecks storage
	Cemetery
	Mineral extraction site mining
	Building site building area
	Semi paved other terrain other build-up areas
The other	Road traffic area
	Airport area
	Enclosed estuary enclosed water (former sea)
	Rhine and Maas Rhine & Maas
	Randee enclosed water in the centre of the Netherlands
	Water storage reservoir
	Inland water for mineral extraction backwater for mining
	Fluid and / or sludge field sludge area
	Other inland water and backwater
	Abroad foreign

3.3.4.5. Street connectivity

Street connectivity refers to the direct accessibility from one street to the others, and can be represented by the density of their connections (i.e., intersections) in a particular spatial context (Mercedy et al., 2011). Several studies of physical activities (e.g., cycling, walking and running) (Koohsari et al., 2014) have showed that in a well-connected street network, streets have scarce cul-de-sacs and plenty of intersections of multiple streets (Mercedy et al., 2011)

Following previous studies (Badland et al., 2010; Yun, 2019), this research analyses the effect of street connectivity on runners’ running in three situations: crossings intersected with more than 4 streets (i.e., >= 4-way crossings), crossings intersected with 3 streets (i.e., >= 3-way crossings) and a cul-de-sac (i.e., one-way crossing). The types of crossings are derived from the numbers of edges of nodes along the streets from Top10NL (e.g., a node connecting three edges indicates a three-way crossing). Then the counts of crossings within a buffer are summed up (Fig. 3.11).

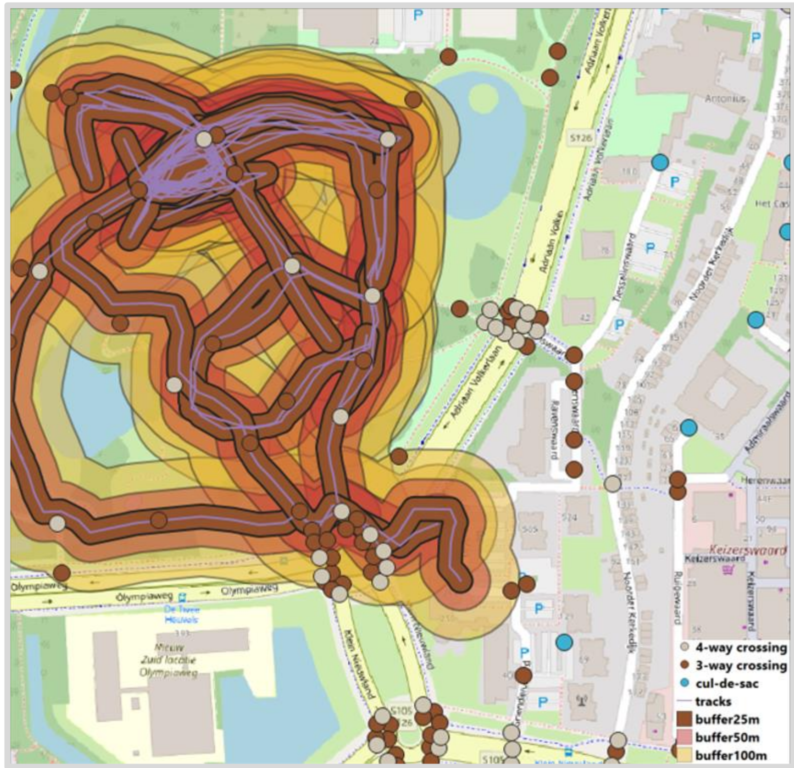


Figure 3.11: 3 types of crossings intersecting with 25, 50 and 100m buffers of multiple paths taken by a runner.

3.3.4.6. Presence of water body

The presence of water is the percentage of visible surface water (e.g., rivers, canals and lakes) in an environment (Dessing et al., 2016). Earlier studies have showed that both cyclists (Krenn et al., 2014) and pedestrians (Dessing et al., 2016) prefer routes which have a higher percentage of visible surface water, which inspires this research to consider the effect of water body on running.

Water bodies are delineated by all 5 × 5m pixels attributed to salt water or fresh water in LGN released in the year of 2015 (Hazeu et al., 2014). First, the intersected areas of these pixels are calculated, then areas within each GPS track and neighbourhood are added up. The calculation is identical to the measurement of the presence of greenery (e.g., GPS track based buffers, Fig., 3.12). When the total area of pixels is larger than the area of the buffer, the value of the ratio will be set as 1 which is the maximum value of the measured presence of water body. The density will be 0 (minimum value) if there is no fresh or saltwater in the buffer area.

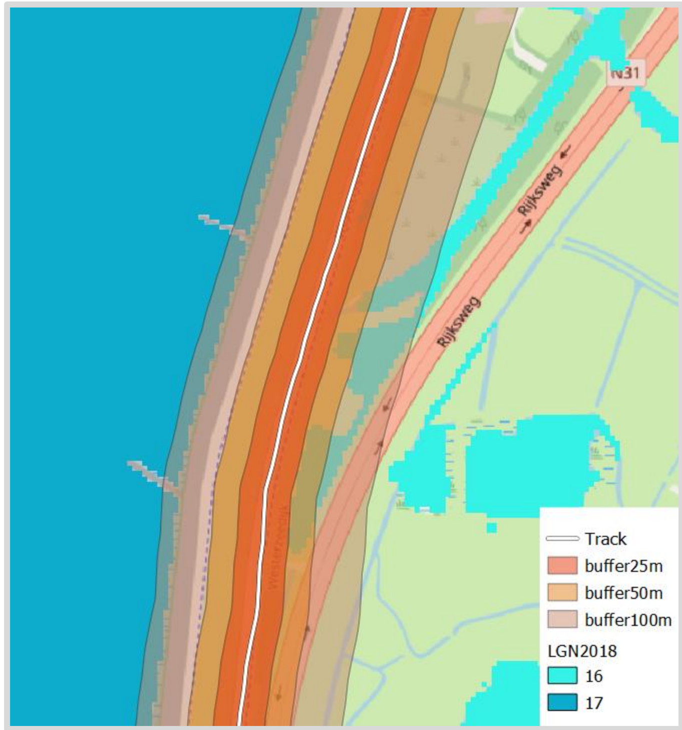


Figure 3.9: Surface water represented 5 × 5 pixels intersecting with 25, 50 and 100m buffers of a GPS track. In LGN, class 16 corresponds to salt water and 17 demonstrates fresh water.

3.4. Statistical analysis

The first objective of analysing the measured environmental variables is to investigate the differences of environments (running and residential) between urban and rural runners and the differences between runners (urban and rural)' running and residential environments. Descriptive data were summarised as mean values with standard deviations at an individual runner level. Differences between urban and rural runners are observed utilizing independent t-test, with the significance level set at $p < 0.05$ and 0.01 (Jones et al., 2009; Tsimeas et al., 2005; Hoekman et al., 2017). Before carrying on independent t-test, the normality of the distribution of each objectively measured environmental variables is tested. From visualization of the distributions, no environmental variables are found normally distributed or following a specific type of non-normal distribution. Besides, as some measured environmental variables (e.g., levels of greenness) assume 0 and negative values, the variables cannot undergo logarithmic transformations. However, as the data sample is quite large (about 30,000 tracks of 3,000 runners), the non-normal distribution can be neglected.

Previous studies have shown the variety in running environments among runners with different experience level (Deelen et al., 2019), journeying on weekdays/weekends (Rodriguze et al., 2012) and roaming inside/outside their neighbourhoods (jones et al., 2009). Therefore, this research will stratify results according to runners running time, spatial extent, distance, duration, and frequency. Considering the distribution of numbers of tracks collected over a year (Fig. 3.2), data from April to August will be utilized for stratified analysis. Thereby, it will increase the robustness of the methods as frequency of uploading information in theses months are consistent and higher than other months (Section 3.2.1). Urban and rural runners are sub-divided into weekend/weekends runners based on their starting time and duration and runners cruising in/out neighbourhoods (i.e., the neighbourhoods where the runners dwell) (Section 3.1.3.3). The two main groups are also classified into subgroups base on their running duration (<= 30 min and > 30 min) running distance (<= 3,000m, 3,000 – 5,000m and > 5,000m) and frequency counts (<= 24, 24 – 48, > 48 in which 10 and 20 times in 5 months equals 24 and 48 times a year) (Section 3.1.3.1).

Descriptive data of the first two stratified analysis (i.e., tracks taken at weekdays/ on weekends and tracks inside/outside neighbourhoods) are summed up as means and standard deviations at individual track level as there are multiple routes inside and outside neighbourhoods or at weekday and on weekends. The same presentation is applied to the subgroups formed by different distances and durations because a < 3,000m and a > 5,000m tracks or a <30mins and >30mins tracks can be taken by a same runner. For the groups created by different running frequencies, the means and standard deviations are summarized at the base of each runner. It is due to the fact that the running frequency of each runner is defined by the numbers of tracks taken by the runner in a year. The Independent t-test with significance level set at $p < 0.05$ and 0.01 will be applied to compare running environments among the first two subgroups. Separate pairwise t-test with same significance level will be utilized to identify which pairwise comparison between groups with different distances and frequencies is statistically different. Thus, to determine which group is significantly different from others. All analyses are performed in RStudio.

4. Results

This chapter describes the results of applying t-test (Section 3.4) in several analyses. The analyses are: comparison of environments (running and residential) between urban and rural runners (Section 4.1) , comparison between runners (urban and rural)' running and residential environments (Section 4.1) and comparison of running environments between the subgroups created by different classifications (e.g., running on weekdays/at weekends) (Section 4.2 – 4.6). As the values of urbanization index are relatively small, the mean values of urbanization index (of running and residential environments) times 1,000 are shown in comparison between urban and rural runners and also between the subgroups.

4.1. Urban and rural runners

Among all runners, there were 25,750 tracks of 2,853 runners living in the cities, and 8,140 tracks of 946 runners living in the countryside (Table 4.1). Rural runners were exposed to running environments with higher mean values of levels of greenness ($p < 0.01$), while lower main values of residential building density ($p < 0.01$), urbanization index ($p < 0.01$), land use mix ($p < 0.01$), and presence of surface water ($p < 0.01$). Relatively, urban runners ran across places with higher averages of the numbers of dead ends ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.01$) than rural inhabitants ($p < 0.01$).

There are no major variances among the groups when their running environments are measured within 25, 50, 100m buffers (Table 4.1, 4.2 and 4.3), therefore, only the data calculated when applying a 25m buffer are displayed in the section of subgroups (Section 4.2 – 4.6) for a concise presentation. Statistics calculated with 50 and 100m buffer are presented in Appendix A1 – 10.

Table 4.1: Descriptive statistics of running environments of urban and rural runners (buffer = 25m).

Environmental variables	**Urban** (n^x = 2853) (n^y = 25750) Mean (SD)	**Rural** (n^x = 946) (n^y = 8140) Mean (SD)
Level of greenness 25m	0.503 (0.093)**	0.546 (0.078)**
Residential building density 25m	0.379 (0.339)**	0.228 (0.268)**
Urbanization index 25m	1.100 (1.190)**	0.614 (0.597)**
Land use mix 25m	0.510 (0.178)**	0.486 (0.183)**
Count of 1-way crossings 25m	2.660 (2.380)**	2.160 (2.020)**
Count of 3-ways crossings 25m	29.20 (20.60)**	23.10 (18.60)**
Count of 4-way crossings 25m	14.60 (14.70)**	9.690 (11.60)**
Proportion of water body 25m	0.262 (0.186)**	0.218 (0.176)**

**: Statistically significant difference between urban and rural runners at $p < 0.01$ of t-test.
n^x: Number of runners.
n^y: Number of tracks.

Table 4.2: Descriptive statistics of running environments of urban and rural runners (buffer = 50 m).

Environmental variables	**Urban** (n = 946) Mean (SD)	**Rural** (n = 2853) Mean (SD)
Level of greenness 50m	0.500 (0.092)**	0.546 (0.076)**
Residential building density 50m	0.380 (0.343)**	0.216 (0.251)**
Urbanization index 50m	1.130 (1.360)**	0.626 (0.748)**
Land use mix 50m	0.505 (0.174)**	0.465 (0.182)**
Count of 1-way crossings 50m	5.950 (4.990)**	4.770 (4.310)**
Count of 3-ways crossings 50m	39.90 (28.90)**	30.50 (25.30)**
Count of 4-way crossings 50m	18.80 (19.40)**	11.90 (14.10)**
Proportion of water body 50m	0.278 (0.197)**	0.246 (0.185)**

**: Statistically significant difference between urban and rural runners at p < 0.0 of t-test.
n: Number of runners.

Table 4.3: Descriptive statistics of running environments of urban and rural runners (buffer = 100 m).

Environmental variables	**Urban** (n = 946) Mean (SD)	**Rural** (n = 2853) Mean (SD)
Level of greenness 100m	0.497 (0.093)**	0.542 (0.076)**
Residential building density 100m	0.382 (0.345)**	0.216 (0.259)**
Urbanization index 100m	1.270 (1.240)**	0.662 (0.583)**
Land use mix 100m	0.504 (0.170)**	0.450 (0.183)**
Count of 1-way crossings 100m	12.20 (9.660)**	9.630 (8.320)**
Count of 3-ways crossings 100m	62.80 (45.20)**	46.40 (38.60)**
Count of 4-way crossings 100m	24.59 (24.10)**	15.00 (16.70)**
Proportion of water body 100m	0.289 (0.202)**	0.259 (0.196)**

**: Statistically significant difference between urban and rural runners at p < 0.01 of t-test.
n: Number of runners.

Table 4.4 shows that rural runners were living in environments with higher means of level of greenness ($p < 0.01$), lower means of residential building density ($p < 0.01$), urbanization index ($p < 0.01$), land use mix ($p < 0.01$), less one, three and four-way crossings ($p < 0.01$) and more less proportion of water body ($p < 0.01$).

Table 4.4: Descriptive statistics of residential environments of urban and rural runners (buffer = 1000m).

Environmental variables	Urban (n^x = 2853) (n^y = 25750) Mean (SD)	Rural (n^x = 946) (n^y = 8140) Mean (SD)
Level of greenness 1000m	0.472 (0.101)**	0.547 (0.088)**
Residential building density 1000m	0.238 (0.296)**	0.187 (0.294)**
Urbanization index 1000m	1.640 (1.600)**	0.675 (0.599)**
Land use mix 100m	0.609 (0.205)**	0.498 (0.248)**
Count of 1-way crossings 1000m	55.40 (32.30)**	41.70 (28.80)**
Count of 3-ways crossings 1000m	251.0 (138.0)**	165.0 (122.0)**
Count of 4-way crossings 1000m	83.00 (62.60)**	41.40 (37.50)**
Proportion of water body 1000m	0.060 (0.078)**	0.050 (0.091)**

**: Statistically significant difference between urban and rural runners at $p < 0.01$ of t-test.
n^x: Number of runners.
n^y: Number of tracks.

From table 4.5, both urban runners ran in environments that had higher means of proportion of water bodies ($p < 0.01$) and residential building density ($p < 0.01$) and lower means of urbanization index ($p < 0.01$) and one ($p < 0.01$), three ($p < 0.01$) and four-way ($p < 0.01$) crossings than those processed by their residential environments. Besides, for urban runners, the mean values of level of greenness ($p < 0.01$), urbanization index ($p < 0.01$) and land use mix ($p < 0.01$) of running environments were comparatively higher than those of their living environments.

Table 4.6 and 4.7 show similar results regarding differences between runners (urban and rural)' residential environments and running environments. The only dissimilarities were that for rural runners the means of land use mix of the two environments were significantly different ($p < 0.01$) with 50 and 100m buffers, while the averages of residential building density were not significant ($p > 0.05$) with these buffers.

Table 4.5: Descriptive statistics of environmental variables of residential (1,000m) and running environments (25m).

Environmental variables		Residential buffer 1,000m	GPS-based buffer 25m
Urban	Level of greenness	0.472 (0.101)**	0.503 (0.093)**
	Residential building density	0.238 (0.296)**	0.379 (0.339)**
	Urbanization index	1.640 (1.600)**	1.100 (1.190)**
	Land use mix	0.609 (0.205)**	0.510 (0.178)**
	Count of 1-way crossings	55.40 (32.30)**	2.660 (2.380)**
	Count of 3-ways crossings	251.0 (138.0)**	29.20 (20.60)**
	Count of 4-way crossings	83.00 (62.60)**	14.60 (14.70)**
	Proportion of water body	0.060 (0.078)**	0.262 (0.186)**
Rural	Level of greenness	0.547 (0.088)	0.546 (0.078)
	Residential building density	0.187 (0.294)**	0.228 (0.268)**
	Urbanization index	0.675 (0.599)	0.614 (0.597)
	Land use mix	0.498 (0.248)	0.486 (0.183)
	Count of 1-way crossings	41.70 (28.80)**	2.160 (2.020)**
	Count of 3-ways crossings	165.0 (122.0)**	23.10 (18.60)**
	Count of 4-way crossings	41.40 (37.50)**	9.690 (11.60)**
	Proportion of water body	0.050 (0.091)**	0.218 (0.176)**

**: Statistically significant difference between residential (1,000m) and running environments (25m) at $p < 0.01$ of t-test.
n^a: Number of tracks of urban runners.
n^b: Number of tracks of rural runners.

Table 4.6: Descriptive statistics of environmental variables of residential (1,000m) and running environments (50m).

Environmental variables		Residential buffer 1,000m	GPS-based buffer 50m
Urban	Level of greenness	0.472 (0.101)**	0.500 (0.092)**
	Residential building density	0.238 (0.296)**	0.380 (0.343)**
	Urbanization index	1.640 (1.600)**	1.130 (1.360)**
	Land use mix	0.609 (0.205)**	0.505 (0.174)**
	Count of 1-way crossings	55.40 (32.30)**	5.950 (4.990)**
	Count of 3-ways crossings	251.0 (138.0)**	39.90 (28.90)**
	Count of 4-way crossings	83.00 (62.60)**	18.80 (19.40)**
	Proportion of water body	0.060 (0.078)**	0.278 (0.197)**
Rural	Level of greenness	0.547 (0.088)	0.546 (0.076)
	Residential building density	0.187 (0.294)	0.216 (0.251)
	Urbanization index	0.675 (0.599)	0.626 (0.748)
	Land use mix	0.498 (0.248)**	0.465 (0.182)**
	Count of 1-way crossings	41.70 (28.80)**	4.770 (4.310)**
	Count of 3-ways crossings	165.0 (122.0)**	30.50 (25.30)**
	Count of 4-way crossings	41.40 (37.50)**	11.90 (14.10)**
	Proportion of water body	0.050 (0.091)**	0.246 (0.185)**

**: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.01$ of t-test.
n^a: Number of tracks of urban runners.
n^b: Number of tracks of rural runners.

Table 4.7: Descriptive statistics of environmental variables of residential (1,000m) and running environments (100m).

Environmental variables		Residential buffer 1,000m	GPS-based buffer 100m
Urban	Level of greenness	0.472 (0.101)**	0.497 (0.093)**
	Residential building density	0.238 (0.296)**	0.382 (0.345)**
	Urbanization index	1.640 (1.600)**	1.270 (1.240)**
	Land use mix	0.609 (0.205)**	0.504 (0.170)**
	Count of 1-way crossings	55.40 (32.30)**	12.20 (9.660)**
	Count of 3-ways crossings	251.0 (138.0)**	62.80 (45.20)**
	Count of 4-way crossings	83.00 (62.60)**	24.59 (24.10)**
	Proportion of water body	0.060 (0.078)**	0.289 (0.202)**
Rural	Level of greenness	0.547 (0.088)	0.542 (0.076)
	Residential building density	0.187 (0.294)	0.216 (0.259)
	Urbanization index	0.675 (0.599)	0.662 (0.583)
	Land use mix	0.498 (0.248)**	0.450 (0.183)**
	Count of 1-way crossings	41.70 (28.80)**	9.630 (8.320)**
	Count of 3-ways crossings	165.0 (122.0)**	46.40 (38.60)**
	Count of 4-way crossings	41.40 (37.50)**	15.00 (16.70)**
	Proportion of water body	0.050 (0.091)**	0.259 (0.196)**

**: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

4.2. Runners running on weekdays and at weekends

The differences between running environments of runners running on weekends and at weekdays are shown in Table 4.8. More routes were taken on weekdays than at weekends. For urban runners mean values of urbanization index ($p < 0.05$), land use mix ($p < 0.05$) and more one ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.05$) were relatively higher on weekends than on weekdays. While for rural runners, no significant difference between weekdays and weekends were shown ($p > 0.05$).

Table 4.8: Descriptive statistics of surroundings (<=25m) of routes taken at weekdays and on weekends (buffer = 25m).

Environmental variables		Weekdays (n^a = 15746) (n^b = 4983) Mean (±SD)	Weekends (n^a = 6773) (n^b = 2091) Mean (±SD)
Urban	Level of greenness 25m	0.502 (0.095)	0.508 (0.095)
	Residential building density 25m	0.380 (0.354)	0.372 (0.355)
	Urbanization index 25m	1.110 (1.220)*	1.030 (1.210)*
	Land use mix 25m	0.512 (0.181)*	0.522 (0.185)*
	Count of 1-way crossings 25m	2.620 (2.440)**	2.840 (2.920)**
	Count of 3-ways crossings 25m	28.90 (21.40)**	31.10 (23.20)**
	Count of 4-way crossings 25m	14.70 (16.20)*	15.80 (16.80)*
	Proportion of water body 25m	0.263 (0.196)	0.267 (0.199)
Rural	Level of greenness 25m	0.547 (0.079)	0.550 (0.077)
	Residential building density 25m	0.225 (0.276)	0.217 (0.288)
	Urbanization index 25m	0.609 (0.575)	0.627 (1.040)
	Land use mix 25m	0.496 (0.186)	0.479 (0.194)
	Count of 1-way crossings 25m	2.090 (2.030)	2.280 (2.370)
	Count of 3-ways crossings 25m	22.50 (18.50)	24.10 (21.50)
	Count of 4-way crossings 25m	9.520 (11.60)	10.10 (13.50)
	Proportion of water body 25m	0.223 (0.184)	0.217 (0.184)

*: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.05$ of t-test.

**: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

4.3. Tracks inside and outside neighbourhoods

Table 4.9 presents that more urban and rural runners roamed outside their neighbourhoods. For both urban and rural runners, mean values of level of greenness of tracks ($p < 0.01$), land use mix ($p < 0.01$), urbanization index ($p < 0.01$), the number of cul-de-sac ($p < 0.01$), three-way crossings ($p < 0.01$) and four-way crossings ($p < 0.01$) and mean of proportion of water bodies ($p < 0.01$) was higher of surroundings of tracks outside neighbourhoods than that of the tacks inside their neighbourhoods. Besides, the means of residential building density within 25m around tracks also increased significantly when urban runners running outside their neighbourhoods.

Table 4.9: Descriptive statistics of surroundings of tracks taken inside and outside neighbourhoods (buffer = 25m).

Environmental variables		**Inside neighbourhood** (n^a = 1469) (n^b = 494) Mean (±SD)	**Outside neighbourhood** (n^a = 20857) (n^b = 6033) Mean (±SD)
Urban	Level of greenness 25m	0.464 (0.119)**	0.596 (0.092)**
	Residential building density 25m	0.344 (0.338)**	0.389 (0.345)**
	Urbanization index 25m	2.140 (3.040)**	1.060 (1.140)**
	Land use mix 25m	0.316 (0.235)**	0.531 (0.168)**
	Count of 1-way crossings 25m	1.000 (1.620)**	2.840 (2.630)**
	Count of 3-ways crossings 25m	9.890 (11.60)**	30.90 (21.70)**
	Count of 4-way crossings 25m	4.090 (6.370)**	15.60 (15.60)**
	Proportion of water body 25m	0.166 (0.212)**	0.271 (0.189)**
Rural	Level of greenness 25m	0.519 (0.114)**	0.550 (0.075)**
	Residential building density 25m	0.202 (0.256)	0.236 (0.270)
	Urbanization index 25m	1.130 (1.190)**	0.581 (0.500)**
	Land use mix 25m	0.303 (0.214)**	0.516 (0.169)**
	Count of 1-way crossings 25m	0.894 (0.146)**	2.260 (2.080)**
	Count of 3-ways crossings 25m	7.750 (8.950)**	24.60 (19.10)**
	Count of 4-way crossings 25m	2.170 (3.380)**	10.50 (12.20)**
	Proportion of water body 25m	0.149 (0.206)**	0.232 (0.176)**

**: Statistically significant difference between tracks taken inside/outside neighbourhood at $p < 0.01$ of t-test.
n^a: Number of tracks of urban runners.
n^b: Number of tracks of rural runners.

4.4. Tracks classified by running duration

From Table 4.10, there were less <= 30 minutes tracks than >30 mins tracks taken by urban and rural runners. For both urban and rural runners taking runs for longer than 30 minutes, they were exposed in running environments with higher averages of level of greenness ($p < 0.01$), land use mix ($p < 0.01$), density of one ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.01$) and proportion of water bodies ($p < 0.01$) and lower means of urbanization index ($p < 0.01$).

Table 4.10: Descriptive statistics of surroundings of <=30mins and >30mins tracks (buffer = 25m).

Environmental variables		**<= 30 minutes** (n[a] = 8540) (n[b] = 2638) Mean (±SD)	**> 30 minutes** (n[a] = 13982) (n[b] = 4436) Mean (±SD)
Urban	Level of greenness 25m	0.487 (0.106)**	0.512 (0.092)**
	Residential building density 25m	0.371 (0.356)	0.382 (0.354)
	Urbanization index 25m	1.280 (1.700)**	0.910 (1.170)**
	Land use mix 25m	0.342 (0.228)**	0.522 (0.168)**
	Count of 1-way crossings 25m	2.360 (3.280)**	6.330 (5.440)**
	Count of 3-ways crossings 25m	13.70 (15.70)**	42.30 (30.30)**
	Count of 4-way crossings 25m	5.420 (8.410)**	20.00 (20.50)**
	Proportion of water body 25m	0.195 (0.218)**	0.286 (0.200)**
Rural	Level of greenness 25m	0.535 (0.088)**	0.554 (0.072)**
	Residential building density 25m	0.223 (0.285)	0.233 (0.286)
	Urbanization index 25m	0.736 (0.744)**	0.549 (0.592)**
	Land use mix 25m	0.457 (0.211)**	0.514 (0.180)**
	Count of 1-way crossings 25m	1.420 (1.554)**	2.700 (2.550)**
	Count of 3-ways crossings 25m	15.00 (12.40)**	27.90 (22.10)**
	Count of 4-way crossings 25m	6.090 (7.260)**	11.90 (14.00)**
	Proportion of water body 25m	0.199 (0.184)**	0.232 (0.183)**

**: Statistically significant difference between runners running less and more than 30 mins at $p < 0.01$ of t-test.

n[a]: Number of tracks of urban runners.

n[b]: Number of tracks of rural runners.

4.5. Tracks classified by running distance

Table 4.11 presents the environmental variables of surroundings of tracks at different distances. A large number of urban or rural runners ran more than 5000m and comparatively, only a smart part of whole run less than 3000m. For urban runners, as the running distance increased, mean values of level of greenness and land use mix increased. Runners taking routes between 3000 and 5000m ran across places with significantly higher averages of one ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.01$). Runners running less than 3000m chose environments with comparatively higher means of residential building density ($p < 0.01$), urbanization index ($p < 0.01$) and proportion of water bodies ($p < 0.01$).

For rural runners, those who running more than 5000m also were exposed to environments a with a considerably higher means of level of greenness ($p < 0.01$) and land use mix ($p < 0.05$) than others and those running less than 3000m came cross places with a higher average urbanization index ($p < 0.01$). While regarding the runners taking 3000-5000m routes, they also ran into environments with higher mean values of proportion of water bodies ($p < 0.01$), residential building density ($p < 0.01$) and numbers of one ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.01$).

Table 4.11: Descriptive statistics of surroundings of tracks with different lengths (buffer = 25m).

Environmental variables		**<= 3000m** (n^a = 4383) (n^b = 1382) Mean (±SD)	**3000 – 5000m** (n^a = 7550) (n^b = 2439) Mean (±SD)	**> 5000m** (n^a = 10589) (n^b = 3253) Mean (±SD)
Urban	Level of greenness 25m	0.482 (0.103)**	0.505 (0.092)**	0.517 (0.098)**
	Residential building density 25m	0.394 (0.368)**	0.376 (0.351)	0.382 (0.411)
	Urbanization index 25m	1.448 (2.011)**	1.062 (1.274)**	0.841 (1.085)**
	Land use mix 25m	0.488 (0.187)**	0.515 (0.176)**	0.552 (0.192)**
	Count of 1-way crossings 25m	8.720 (8.160)**	14.90 (13.50)**	3.707 (3.985)**
	Count of 3-ways crossings 25m	45.10 (35.90)**	75.90 (59.00)**	43.02 (34.75)**
	Count of 4-way crossings 25m	17.20 (18.20)**	30.30 (32.40)**	22.89 (26.42)**
	Proportion of water body 25m	0.300 (0.211)**	0.275 (0.228)**	0.297 (0.215)**
Rural	Level of greenness 25m	0.534 (0.089)**	0.554(0.074)**	0.560 (0.075)**
	Residential building density 25m	0.210 (0.269)	0.199 (0.253)**	0.208 (0.320)
	Urbanization index 25m	0.919 (1.214)**	0.594 (0.650)**	0.432 (0.470)**
	Land use mix 25m	0.445 (0.202)	0.463 (0.185)	0.530 (0.188)*
	Count of 1-way crossings 25m	6.580 (6.370)**	11.70 (10.20)**	3.032 (3.381)**
	Count of 3-ways crossings 25m	31.60 (26.90)**	55.50 (45.60)**	34.41 (28.32)**
	Count of 4-way crossings 25m	9.960 (11.30)**	18.20 (20.10)**	14.85 (17.68)**
	Proportion of water body 25m	0.244 (0.198)**	0.271 (0.201)**	0.245 (0.210)**

*: Statistically significant difference between runners running at different levels of distances at $p < 0.05$ of pairwise t-test.

**: Statistically significant difference between runners running at different levels of distances at $p < 0.01$ of pairwise t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

4.6. Tracks classified by running frequency

Table 4.12 shows that there were approximately same number of runners in each group. For urban runners, mean values of urbanization index, density of three and four-way crossings were higher of groups at higher running frequency levels. The averages of level of greenness ($p < 0.05$) and proportion of water bodies ($p < 0.01$) of runners running more than 48 times a year were significantly higher. Urban runners participated running less than 24 time a year has comparatively lower means of land use mix ($p < 0.05$) and urbanization index ($p < 0.01$).

Among rural runners, those running more than 48 times a year came across places with considerably higher mean values of urbanization index ($p < 0.01$), land use mix ($p < 0.01$), one ($p < 0.01$), three ($p < 0.01$) and four-way crossings ($p < 0.01$). A higher mean value of proportion of water bodies was seen when the level of running frequency increase. Runners with frequency between 24 and 48 times a year were exposed in environments with relatively higher mean of level of greenness ($p < 0.05$), but lower means of residential building density ($p < 0.05$) and urbanization index ($p < 0.01$).

Table 4.12: Descriptive statistics of running environments of runners running at different frequencies (buffer = 25m).

Environmental variables		**<= 24** (n[a] = 7983) (n[b] = 2634) Mean (±SD)	**24 - 48** (n[a] = 7410) (n[b] = 2136) Mean (±SD)	**> 48** (n[a] = 7174) (n[b] = 2304) Mean (±SD)
Urban	Level of greenness 25m	0.507 (0.104)	0.505 (0.109)	0.511 (0.104)*
	Residential building density 25m	0.375 (0.406)	0.378 (0.410)	0.385 (0.413)
	Urbanization index 25m	1.077 (1.473)**	1.018 (1.400)	1.000 (1.285)
	Land use mix 25m	0.513 (0.214)*	0.522 (0.210)	0.521 (0.215)
	Count of 1-way crossings 25m	2.621 (3.219)**	2.825 (3.241)	2.851 (3.524)
	Count of 3-ways crossings 25m	28.58 (25.29)**	31.92 (28.45)**	33.92 (32.26)**
	Count of 4-way crossings 25m	14.23 (18.13)**	16.70 (21.93)**	17.62 (22.55)**
	Proportion of water body 25m	0.267 (0.224)	0.263 (0.222)	0.288 (0.216)**
Rural	Level of greenness 25m	0.546 (0.094)	0.560 (0.080)*	0.544 (0.087)
	Residential building density 25m	0.220 (0.323)	0.199 (0.316)*	0.234 (0.337)
	Urbanization index 25m	0.614 (0.833)	0.488 (0.612)**	0.636 (0.768)
	Land use mix 25m	0.483 (0.224)	0.494 (0.211)	0.529 (0.183)**
	Count of 1-way crossings 25m	2.093 (2.792)	2.058 (2.463)	2.582 (3.065)**
	Count of 3-ways crossings 25m	21.73 (22.33)	22.82 (21.25)	29.41 (25.78)**
	Count of 4-way crossings 25m	9.963 (14.25)	9.452 (11.98)	12.30 (15.34)**
	Proportion of water body 25m	0.221 (0.211)*	0.206 (0.218)*	0.241 (0.214)**

*: Statistically significant difference between runners running at different levels of frequencies at $p < 0.05$ of pairwise t-test.

**: Statistically significant difference between runners running at different levels of frequencies at $p < 0.01$ of pairwise t-test.

n[a]: Number of urban runners.

n[b]: Number of rural runners.

5. Discussion

This research examined running and residential environments of urban and rural runners. It showed that urban and rural runners are exposed to different running and residential environments, but not in a dichotomous manner. There were more runner living and running in urban areas than rural areas. For both urban and rural runners, they ran in environments with more blue spaces and less streets than their residential areas. Among them, urban runners preferred to run in environments also with more vegetation and lower address density. Thus, this research showed that both urban and runners (especially urban runners) would select more natural environments while running. Running routes and residential settings of urban runners were found to be linked with less greenery, more coverage of residential buildings, higher address density, a higher degree of mixed land uses and more cul-de-sacs and well-connected streets, which are common features comprising a typical environment in more urbanized areas (Ewing et al., 2003; Lee et al., 2009). Those features lead to distinct urban lifestyle (i.e., reduced level of fitness among adolescents) (Regis et al; 2016; Deelen et al., 2017). While the case of rural runners was the opposite in terms of the characteristics of living and running environments. In this sense, this research reported that even urban and rural runners chose more natural environments for running, their routes were still constrained by their residency. Loucaides et al (2004) have revealed that pupils living in the countryside were more engaged in outdoor physical activities than pupils dwelling in cities because of more available open spaces. Hence, when considering running as a recreational activity (Ettema, 2016), this research also showed that rural runners to a large degree were running in environments with more green vegetation and less compacted residential buildings, mixed land uses and streets. Besides, proportion of water body, which was regarded as an environmental characteristic motivating running (Deelen et al., 2017), were more involved in routes taken by urban runners than rural runners. It indicated urban runners were more likely to run in environments with water than rural runners.

This research also found that urban runners participated running more on weekdays than at weekends. At weekends, they chose running environments with land use distribution with higher heterogeneity and density of addresses and more streets with one, three or four-way crossings. It happened probably because when urban runners' leisure time was more abundant, their running trails were less restricted by their living environments. To some degree, the finding extends Rodriguze et al (2012)'s the conclusion of the relationship among physical activity, physical environments and time. More specifically, Rodriguze et al (2012) claimed that moderate to vigorous physical activities (e.g., cycling and running) were positively associated with parks, high population density and weekdays and negatively associated with roads and outlets, while this finding provided the spatial context, a complex environment, where running as a physical activity occurred on weekdays. In other words, it proved the coexistence of a variety of characteristics of urban and rural runners' running environments at weekends or on weekdays, despite those characteristics exert different effects on running.

Another outcome is that more rural and urban runners travelled outside their neighbourhoods. When urban and rural runners toured outside their neighbourhoods, they came across more vegetation, streets, water body and land use with higher heterogeneity. It could indicate that urban and rural runners preferred green and blue spaces while running outside their neighbourhoods, but in the meantime, they ran through more streets and places with mixed land uses. For urban runners, another reason is that those who travelled outside their neighbourhoods were living in environments containing less greenery but more residential units and streets. Urban runners also ran into places with more coverage of residential buildings, which was probably because there were more residential buildings in urbanized environments.

Apart from that, this research divided urban and rural runners into different sub-groups according to their running durations, distances, and frequencies. More green and blue spaces, streets, less address, and land uses with higher heterogeneity were in the surroundings of > 30 mins tracks. Similarly, surrounding environments of longer tracks (> 5,000m) and tracks of runners who trained more frequently contained more vegetation and less addresses. It demonstrated that the long tracks (>5,000m or > 30 mins) and the tracks taken by more frequently practiced runners were exposed in running environments with more natural and less built-in features. The result is partially in line with the study of James et al (2014) and Tamura et al (2019) where they found that individuals who spent more time on PA tended to seek green spaces.

Moreover, this research observed variances of all running groups when their running environments were measured at different geographical scales (Appendix A1 – 12). The variances also differed from one environmental variable to another and from urban runners to rural dwellers. For all runners (urban and rural), more streets crossings, addresses and water bodies around the tracks were accounted while the geographical extent expanded. For rural runners, the coverage of the residential building decreased when buffer size increased, which could be due to the typical phenomenon that residential properties generally do not extend more than 50m from the tracks upon where they locate (Oliver et al., 2007). While for runners living in the cities, if a runner toured in neighbourhoods with compacted housing, more areas of residences were accounted within a larger spatial extent. Apart from this, according to Oliver et al (2007), a buffer less than 50m would exclude less residential areas than other types of land uses (e.g., recreational, and industrial) on account of the size of residential units. Thus, the land use mix will expectedly increase when the buffer size increase. However, this research found the opposite result. An explanation from the visual representation (Fig. 3.10) is that normally the land uses (i.e., residential, industrial, recreational, industrial and others) involving buildings and ground fields exceeded much more than 100m from the GPS tracks to their borders at the opposing side. Therefore, despite more types of land uses were likely to be included at larger buffer sizes, more areas were predominantly occupied by a particular land use (e.g., residential lands) (Fig. 3.10). As a result, the land uses were less evenly distributed while increasing size of the buffers. The level of the greenness of a track which was measured as NDVI does not change noticeably at 25, 50, and 100m. A potential reason is that for tracks of rural runners in forests, parks or fields their surrounding environments were almost totally covered by greenery (e.g., trees), and for tracks were taken by urban runners, trees and shrubs were scattered in the residential sites where their tracks go through. In either case, the measured NDVI was not substantially influenced by geographical extent.

6. Conclusion, limitation, and future work

In the Netherlands, due to the supportive sports climates, a growing number of people has recognized running as a conveniently accessible form of exercise that contributes to good health. This research found that residential and running environments between urban and rural runners were still distinctive (e.g., rural runners were exposed in places with more greenery and less residential units), despite urban growth had demanded a progressive transformation of the landscapes in the suburbs (Nabielek et al., 2014). When urban runners had more leisure time at weekends, they travelled into places with more heterogeneously distributed land uses and streets. In addition, both urban and rural runners preferred running in places with more vegetations while running outside their neighbourhoods. Moreover, the longer tracks (long distance or duration) and the tracks of more frequently practiced runners were taken in environments with more vegetation and less built-in places. Regarding the differences between runners (urban and rural)' running and residential environments, both urban and rural runners chose environments with more blue spaces for running. Among them, urban runners also ran through places with more presence of greenery and less density of addresses.

To the author's knowledge, this research developed an GIS-based method to examine the residential and running environment of runners with urban or rural status (Section 3.1.2). The method is different from earlier works that distinguish the difference between urban and rural runners by involving their prior knowledge of the study areas (James et al., 2014) or residential density (Klompmaker et al, 2019; Lee et al., 2009) and how people perceive their environment (Hoekman, 2017). This method starts from defining runners urban/rural status and neighbourhood by detecting their home at pc6 level. Then, if the runner lives in a pc6 where inside/outside the functional urban areas released by OECD, the runner is regarded as an urban/rural runner (Section 3.1.3.). Environmental characteristics of a track taken by an urban/rural runner are mapped as environmental variables which are determined from reviewing previous studies based on people's perception and GIS measurements (Chapter 2). In this sense, environmental variables examined in this research to a large degree reflect the runners' observations of their surroundings and these variables can also be mapped in a GIS. Methods of measuring environmental variables are devised in GIS by following prior work (section 3.3.4.). After the measurement, the urban and rural runners are further divided into sub-groups by different classifications (i.e., running different distance, durations, at weekdays/on weekends and inside/outside neighbourhood, and with different frequency). Significances of the effect of running distance and frequency on their running environment were analysed by using separate pairwise t-test; differences of runners living in urban and rural areas, running different durations, at weekdays/on weekends and inside/outside neighbourhood were detected using independent t-test (Tsimeas et al., 2005; Jones et al., 2009).

The research also has some limitations. First, the analysis of this research rests entirely on data gathered from a mobile GPS tracking application. This is to say, no demographic or socioeconomic data enriched in contextual information (e.g., runners real urban/rural status, age, and social class) for understanding the propensity for running in a specific environment substantiate analysing the results (Romanillos et al., 2016). The method (section 3.3.2.) in this research is developed to overcome the lack of knowledge regarding the runners' home addresses. However, people could drive to a favoured environment to start and end running whenever they record their track. In this case, the defined urban/rural runners are not in accordance with the runners' true residency, and its threat to the validity of the method will depend on the numbers of runners being defined falsely. Besides, only one mobile GPS application, Endomondo, was selected to collect the GPS tracks and the process is under the condition that runners made their information available for the public. The reason is to make this research repeatable and extensible without the requirement of substantial financial investment by obtaining data from professional GPS and other mobile application (e.g., Strava) and a privacy breach. Whereas it also makes this research simply leads to self-selective samples. As a result, the data acquired might not be representative for running behaviours of all runners across the Netherlands and potentially causes some uncertainties in results. For example, this result showed there are more urban runners in the Netherlands, but it is possibly because less rural runners are using Endomondo. Apart from this, the results are yielded from data in the year of 2015. Hence, taking urban development in the Netherlands into account (Nabielek et al., 2014), the results of differences of running environments of runners (urban and rural) cannot certainly predict the running environments of the runners afterwards (same for differences between runners (urban and rural)' running and residential environments). A plausible solution is to this problem is to use both demographic data and GPS data (Romanillos et al., 2016).

Secondly, usually a smartphone GPS, in this case, Edomondo, is less accurate than a dedicated GPS device (ESRI, n.d.). Their deviation from the true position due to blockages like buildings or tree canopy is wildly acknowledged in urban dynamic analyses (Merry et al., 2019). While because of the extensive time and intricate GIS techniques are required to correct GPS errors, the anomalies (e.g., jumps of GPS coordinates and erroneous GPS coordinates) due to incorrect GPS positioning and user errors (Sileryte, 2015) are only ruled out rather than being rectified. Considering the size of the data, its effect on results is marginal. To improve from this research, a method to overcome the limitation is to retrieve data from multiple mobile GPS applications.

Thirdly, NDVI satellite image was utilized to calculate the levels of greenness in running environments (i.e., with a 25m/50m/100m around a running track) to which runners are exposed. While it does not represent runners' perception of greenery while they are running as there is no extract information of what type of green spaces are of urban and rural runners' interests. One solution to address this problem is to use street view maps or land cover maps to measure the presence of green spaces (e.g., parks and grasslands) (Jones et al., 2009; Rodriguze et al., 2012). Urbanization index was calculated by the density of addresses in the running environment. Comparing with other studies using complex index accounting population density, the average block size in urban areas (Lee et al., 2009) and runners' socioeconomic statuses at the neighbourhood level (Klompmaker et al., 2019), the absence of rigorous consideration for calculating urbanization index in this research is noticed. Besides, because of the time limit, other environmental variables (e.g., weather, temperature, and air quality) which are revealed as situational barriers of running (Wang et al., 2021) are not accounted in this research. Thus, an improvement of this research could include more environmental variables and devising algorithm specific for urban sprawling in the Netherland (Ewing et al., 2003).

This research suggests policies encouraging running acknowledging the differences between environments (running and residential environments) between urban and rural runners. Besides, for both urban and rural runners, their running environments were constrained by their residency. Furthermore, preserving or upgrading green spaces in urban and rural areas is important to encourage more people to get a start.

For future studies intending to extend the research, one could enrich the background of runners by collecting and employing demographical data, which can furnish background information and be set as a base (e.g., age and income) to subdivide the runner groups. Besides, it is more desirable to use professional GPS devices or collect and pool data from multiple mobile GPS applications. Apart from that, including more environmental variables (e.g., weather, temperature, and air quality) are recommended. Moreover, as the Netherlands has a quite distinct physical and social environment (e.g., limited space, high population density, robust demographic and economic growth and considerable spatial and functional heterogeneity in suburb areas) (Nabielek et al., 2014), developing an algorithm to calculate urbanization index especially for the Netherlands would be practical to investigate the effect of urban growth on runners’ running environments.

Reference

Allahbakhshi, H., Conrow, L., Naimi, B., & Weibel, R. (2020). Using Accelerometer and GPS Data for Real-Life Physical Activity Type Detection. Sensors, 20(3), 588.

Badland, H. M., Duncan, M. J., Oliver, M., Duncan, J. S., & Mavoa, S. (2010). Examining commute routes: applications of GIS and GPS technology. Environmental health and preventive medicine, 15(5), 327-330.

Benson, A. C., Bruce, L., & Gordon, B. A. (2015). Reliability and validity of a GPS-enabled iPhoneTM “app” to measure physical activity. Journal of sports sciences, 33(14), 1421-1428.

Berke, E. M., Koepsell, T. D., Moudon, A. V., Hoskins, R. E., & Larson, E. B. (2007). Association of the built environment with physical activity and obesity in older persons. American journal of public health, 97(3), 486-492.

Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V., & Pebesma, E. J. (2008). Applied spatial data analysis with R (Vol. 747248717, pp. 237-268). New York: Springer.

Bodin, M., & Hartig, T. (2003). Does the outdoor environment matter for psychological restoration gained through running?. Psychology of sport and exercise, 4(2), 141-153.

Boruff, B. J., Nathan, A., & Nijënstein, S. (2012). Using GPS technology to (re)-examine operational definitions of ‘neighbourhood’in place-based health research. International journal of health geographics, 11(1), 22.

Brody, S. (2013). The characteristics, causes, and consequences of sprawling development patterns in the United States. Nature Education Knowledge, 4(5), 2.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. Transportation research. Part D, Transport and environment, 2(3), 199-219.

Claassens, J., Koomen, E., & Rouwendal, J. (2020). Urban density and spatial planning: The unforeseen impacts of Dutch devolution. PloS one, 15(10), e0240738.

Coetzee, S., Odijk, M., Van Loenen, B., Storm, J., & Stoter, J. (2018). Stakeholder analysis of the governance framework of a national SDI dataset–whose needs are met in the buildings and address register of the Netherlands?. International Journal of Digital Earth.

Cohen, D. A., McKenzie, T. L., Sehgal, A., Williamson, S., Golinelli, D., & Lurie, N. (2007). Contribution of public parks to physical activity. American journal of public health, 97(3), 509-514.

Collinson, J. A. (2008). Running the routes together: Corunning and knowledge in action. Journal of contemporary ethnography, 37(1), 38-61.

Dessing, D., de Vries, S. I., Hegeman, G., Verhagen, E., Van Mechelen, W., & Pierik, F. H. (2016). Children’s route choice during active transportation to school: difference between shortest and actual route. International journal of behavioral nutrition and physical activity, 13(1), 48.

Deelen, I., Jansen, M., Dogterom, N. J., Kamphuis, C. B., & Ettema, D. (2017). Do objective neighbourhood characteristics relate to residents' preferences for certain sports locations? A cross-sectional study using a discrete choice modelling approach. BMC public health, 17(1), 943.

Deelen, I., Janssen, M., Vos, S., Kamphuis, C. B., & Ettema, D. (2019). Attractive running environments for all? A cross-sectional study on physical environmental characteristics and runners' motives and attitudes, in relation to the experience of the running environment. BMC public health, 19(1), 366.

Ding, Y., & Lu, H. (2017). The interactions between online shopping and personal activity travel behavior: an analysis with a GPS-based activity travel diary. Transportation, 44(2), 311-324.

Duncan, M. J., Badland, H. M., & Mummery, W. K. (2009). Applying GPS to enhance understanding of transport-related physical activity. Journal of Science and Medicine in Sport, 12(5), 549-556.

Dyal Singh College. University of Delhi. (n.d.). Advantages and disadvantages of GPS. Retrieved from: http://dsc.du.ac.in/wp-content/uploads/2020/04/BAProg_Sem4_RemoteSensing_3_NamrataHandique.pdf
Endomondo. (2020, November 12) Retrieved from: endomondo.com

Endonmondo. (n. d. a.). Connected Service. Retrieved from: https://support.endomondo.com/hc/en-us/sections/206181227-Connected-Services

Endomondo. (n. d. b.). Heart Rate Zones (Premium feature). Retrieved from: https://support.endomondo.com/hc/en-us/articles/201861343-Heart-Rate-Zones-Premium-feature-

Endomondo. (n. d. c.). Weather. Retrieved from: https://support.endomondo.com/hc/en-us/articles/202228438-Weather

Ettema, D. (2016). Runnable Cities: How Does the Running Environment Influence Perceived Attractiveness, Restorativeness, and Running Frequency?. Environment and Behavior, 48(9), 1127-1147.

ESRI. (n.d.a). Recreational Versus Professional GPS: What's the Difference? Retrieved from: https://www.esri.com/news/arcuser/0104/rec-gps.html.
ESRI. (n.d.b.). GIS & OPEN DATA. Retrieved from: https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/industries/government/State/GL_Guide_Esri_Gis_R4.pdf.

European Environment Agency. (n.d.). Environment, health, life expectancy and social inequalities are linked. Chapter 5: environment, health and quality of life.

European Foundation for the Improvement of Living and Working Conditions. (2014). Quality of life in urban and rural Europe.

Eurostat. (2018, December 11). Territorial typologies manual - cities, commuting zones and functional urban areas. Retrieved from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Territorial_typologies_manual_-_cities,_commuting_zones_and_functional_urban_areas#:~:text=As%20a%20final%20step%2C%20the,as%20a%20functional%20urban%20area.

Eurostat. (n.d.) COUNTRIES. Retrieved from: https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/countries.

EverythingConnects. (2013, November 20). Urban Sprawl. Retrieved from: https://www.everythingconnects.org/urban-sprawl.html#:~:text=Although%20some%20would%20argue%20that,agricultural%20capacity%2C%20increased%20car%20dependency%2C.

Ewing, R., Meakins, G., Hamidi, S., & Nelson, A. C. (2014). Relationship between urban sprawl and physical activity, obesity, and morbidity–Update and refinement. Health & place, 26, 118-126.

Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., & Raudenbush, S. (2003). Relationship between urban sprawl and physical activity, obesity, and morbidity. American journal of health promotion, 18(1), 47-57.

Farrell, W. J., Cavellin, L. D., Weichenthal, S., Goldberg, M., & Hatzopoulou, M. (2015). Capturing the urban canyon effect on particle number concentrations across a large road network using spatial analysis tools. Building and environment, 92, 328-334.

Galpern, P., Ladle, A., Uribe, F. A., Sandalack, B., & Doyle-Baker, P. (2018). Assessing urban connectivity using volunteered mobile phone GPS locations. Applied Geography, 93, 37-46.

García, J., Lera-López, F., & Suárez, M. J. (2011). Estimation of a structural model of the determinants of the time spent on physical activity and sport: Evidence for Spain. Journal of Sports Economics, 12(5), 515-537.

Gengeç, N. E., & Tarı, E. (2009). Practical Performance Evaluation of GPS Trajectory Rasterization Methods. Journal of Spatial Information Science.

Gil, J., & Read, S. (2014). Patterns of sustainable mobility and the structure of modality in the Randstad city-region. A| Z ITU Journal of the Faculty of Architecture, 11(2), 231-254.

Gladwell, V. F., Brown, D. K., Wood, C., Sandercock, G. R., & Barton, J. L. (2013). The great outdoors: how a green exercise environment can benefit all. Extreme physiology & medicine, 2(1), 1-7.

Golle, K., Granacher, U., Hoffmann, M., Wick, D., & Muehlbauer, T. (2014). Effect of living area and sports club participation on physical fitness in children: a 4 year longitudinal study. BMC Public Health, 14(1), 499.

Harrison, F., Burgoine, T., Corder, K., van Sluijs, E. M., & Jones, A. (2014). How well do modelled routes to school record the environments children are exposed to?: a cross-sectional comparison of GIS-modelled and GPS-measured routes to school. International journal of health geographics, 13(1), 5.

Hassan, A. M., & Lee, H. (2015). Toward the sustainable development of urban areas: An overview of global trends in trials and policies. Land Use Policy, 48, 199-212.

Hazeu, G. W., Schuiling, C., Dorland, G. J., Roerink, H.S.D. Naeff & R.A. Smidt. (2014). Landelijk Grondgebruiksbestand Nederland versie 7 (LGN7). Vervaardiging, nauwkeurigheid en gebruik. Report.

Helbich, M. (2019). Spatiotemporal contextual uncertainties in green space exposure measures: exploring a time series of the normalized difference vegetation indices. International journal of environmental research and public health, 16(5), 852.

Hitchings, R., & Latham, A. (2017). How 'social'is recreational running? Findings from a qualitative study in London and implications for public health promotion. Health & place, 46, 337-343.

Hoekman, R., Breedveld, K., & Kraaykamp, G. (2017). Sport participation and the social and physical environment: explaining differences between urban and rural areas in the Netherlands. Leisure studies, 36(3), 357-370.

Hover, P., van der Werff, H., & Breedveld, K. (2015). The Netherlands. In Running across Europe (pp. 187-207). Palgrave Macmillan, London.

Hurvitz, P. M., & Moudon, A. V. (2012). Home versus nonhome neighborhood: quantifying differences in exposure to the built environment. American journal of preventive medicine, 42(4), 411-417.

Ioannides, Y. M., & Rossi-Hansberg, E. (2010). Urban growth. In Economic Growth (pp. 264-269). Palgrave Macmillan, London.

Isoteemu (2014) Sports-Tracker-Liberator [Source code]. Retrieved from: https://github.com/isoteemu/sports-tracker-liberator/blob/master/endomondo/workout.py.
James, P., Berrigan, D., Hart, J. E., Hipp, J. A., Hoehner, C. M., Kerr, J., ... & Laden, F. (2014). Effects of buffer size and shape on associations between the built environment and energy balance. Health & place, 27, 162-170.

Jankowska, M. M., Schipperijn, J., & Kerr, J. (2015). A framework for using GPS data in physical activity and sedentary behavior studies. Exercise and sport sciences reviews, 43(1), 48.

Jones, A. P., Coombes, E. G., Griffin, S. J., & van Sluijs, E. M. (2009). Environmental supportiveness for physical activity in English schoolchildren: a study using Global Positioning Systems. International Journal of Behavioral Nutrition and Physical Activity, 6(1), 42.

Karkera, A., Swaminathan, N., Pais, S. M., Vishal, K., & Rai, S. (2014). Physical fitness and activity levels among urban school children and their rural counterparts. The Indian Journal of Pediatrics, 81(4), 356-361.

Klompmaker, J. O., Hoek, G., Bloemsma, L. D., Wijga, A. H., van den Brink, C., Brunekreef, B., ... & Janssen, N. A. (2019). Associations of combined exposures to surrounding green, air pollution and traffic noise on mental health. Environment international, 129, 525-537.

Koohsari, M. J., Sugiyama, T., Lamb, K. E., Villanueva, K., & Owen, N. (2014). Street connectivity and walking for transport: role of neighborhood destinations. Preventive medicine, 66, 118-122.

Kotzeva, M. M., & Brandmüller, T. (Eds.). (2016). Urban Europe: statistics on cities, towns and suburbs. Publications office of the European Union.

Krenn, P. J., Titze, S., Oja, P., Jones, A., & Ogilvie, D. (2011). Use of global positioning systems to study physical activity and the environment: a systematic review. American journal of preventive medicine, 41(5), 508-515.

Kwan, M. P., Wang, J., Tyburski, M., Epstein, D. H., Kowalczyk, W. J., & Preston, K. L. (2019). Uncertainties in the geographic context of health behaviors: A study of substance users' exposure to psychosocial stress using GPS data. International Journal of Geographical Information Science, 33(6), 1176-1195.

Lange, A. F., & Gilbert, C. (1999). Using GPS for GIS data capture. Geographical Information Systems, 1, 467-476.

Lee, K., & Kwan, M. P. (2019). The Effects of GPS-Based Buffer Size on the Association between Travel Modes and Environmental Contexts. ISPRS International Journal of Geo-Information, 8(11), 514.

Lee, I. M., Ewing, R., & Sesso, H. D. (2009). The built environment and physical activity levels: the Harvard Alumni Health Study. American journal of preventive medicine, 37(4), 293-298.

Leslie, E., Coffee, N., Frank, L., Owen, N., Bauman, A., & Hugo, G. (2007). Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes. Health & place, 13(1), 111-122.

Liu, H., & Walker, G. J. (2015). The effects of urbanization, motivation, and constraint on Chinese people's leisure-time physical activity. Leisure Sciences, 37(5), 458-478.

Loenen, B. (2006). Developing geographic information infrastructures: The role of information policies. IOS Press.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2001). Geographic information systems and science. John Wiley & Sons.

Lopez, R. (2004). Urban sprawl and risk for being overweight or obese. American journal of public health, 94(9), 1574-1579.

Lopez, R. P., & Hynes, H. P. (2006). Obesity, physical activity, and the urban environment: public health research needs. Environmental Health, 5(1), 1-10.

Loucaides, C. A., Chedzoy, S. M., & Bennett, N. (2004). Differences in physical activity levels between urban and rural school children in Cyprus. Health education research, 19(2), 138-147.

Maddison, R., Vander Hoorn, S., Jiang, Y., Mhurchu, C. N., Exeter, D., Dorey, E., ... & Turley, M. (2009). The environment and physical activity: The influence of psychosocial, perceived and built environmental factors. International Journal of Behavioral Nutrition and Physical Activity, 6(1), 1-10.

Madsen, T., Schipperijn, J., Christiansen, L. B., Nielsen, T. S., & Troelsen, J. (2014). Developing suitable buffers to capture transport cycling behavior. Frontiers in public health, 2, 61.
Mashili, F. L., Mbatia, J., Nana, A., Saguti, G., & Kagaruki, G. B. (2018). Physical activity and associated socioeconomic determinants in rural and urban Tanzania: results from the 2012 WHO-STEPS Survey.

Mavoa, S., Bagheri, N., Koohsari, M. J., Kaczynski, A. T., Lamb, K. E., Oka, K., O'sullivan, D., and Witten, K. (2019). How do neighbourhood definitions influence the associations between built environment and physical activity? International journal of environmental research and public health, 16(9):1501

Mecredy, G., Pickett, W., & Janssen, I. (2011). Street connectivity is negatively associated with physical activity in Canadian youth. International journal of environmental research and public health, 8(8), 3333-3350.

Merry, K., & Bettinger, P. (2019). Smartphone GPS accuracy study in an urban environment. PloS one, 14(7), e0219890.

Miller, H. J., Tribby, C. P., Brown, B. B., Smith, K. R., Werner, C. M., Wolf, J., ... & Oliveira, M. G. S. (2015). Public transit generates new physical activity: evidence from individual GPS and accelerometer data before and after light rail construction in a neighborhood of Salt Lake City, Utah, USA. Health & place, 36, 8-17.

Modsching, M., Kramer, R., & ten Hagen, K. (2006, March). Field trial on GPS Accuracy in a medium size city: The influence of built-up. In 3rd workshop on positioning, navigation and communication (Vol. 2006, pp. 209-218).

Nabielek, K., Kronberger-Nabielek, P., & Hamers, D. (2014). The rural-urban fringe in the Netherlands: recent developments and future challenges. Spool, 1(1), 101-120.

National Geographic. (n.d.). Urban area. Retrieved from: https://www.nationalgeographic.org/encyclopedia/urban-area/.

NASA. (2012, October 27). Global Positioning System History. Retrieved from: https://www.nasa.gov/directorates/heo/scan/communications/policy/GPS_History.html.

National Research Council (U.S.). Committee on Physical Activity, Land Use, Transportation Research Board, & Institute of Medicine. (2005). Does the Built Environment Influence Physical Activity?: Examining the Evidence--Special Report 282 (Vol. 282). Transportation Research Board.

NHS. (n.d.). Running for beginners. Retrieved from: https://www.nhs.uk/live-well/exercise/running-tips-for-beginners/#:~:text=Regular%20running%20for%20beginners%20means,for%20the%20next%203%20weeks.

OECD. (n. d.). Functional urban areas by country. Retrieved from: https://www.oecd.org/regional/regional-statistics/functional-urban-areas.htm.

Oliver, L. N., Schuurman, N., & Hall, A. W. (2007). Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. International journal of health geographics, 6(1), 1-11.

O'Reilly, N., Berger, I. E., Hernandez, T., Parent, M. M., & Séguin, B. (2015). Urban sportscapes: An environmental deterministic perspective on the management of youth sport participation. Sport Management Review, 18(2), 291-307.

PC6 - 6 - digit postal NL. https://www.element.nl/maps-data/postal-maps-and-administrative-boundaries-netherlands/6-digit-postal-maps/.

PDOK. (2019, April 24). The Land Use 2015 file from CBS now with PDOK. Retrieved from: https://www.pdok.nl/-/bestand-bodemgebruik-2015-van-cbs-nu-bij-pdok.

PDOK. (2020, December 18). Statistics Netherlands Districts and neighbourhoods updated. Retrieved from: https://www.pdok.nl/-/cbs-wijken-en-buurten-geactualiseerd.

PDOK. (n. d. a). Dataset: Key Register of Addresses and Buildings (BAG). Retrieved from: https://www.pdok.nl/introductie/-/article/basisregistratie-adressen-en-gebouwen-ba-1.

PDOK. (n. d. b.). Dataset: Key Register Topography (BRT) TOPNL. Retrieved from: https://www.pdok.nl/introductie/-/article/basisregistratie-topografie-brt-topnl.

Pretty, J., Peacock, J., Hine, R., Sellens, M., South, N., & Griffin, M. (2007). Green exercise in the UK countryside: Effects on health and psychological well-being, and implications for policy and planning. Journal of environmental planning and management, 50(2), 211-231.

Qviström, M., Fridell, L., & Kärrholm, M. (2020). Differentiating the time-geography of recreational running. Mobilities, 1-13.

Regis, M. F., Oliveira, L. M. F. T. D., Santos, A. R. M. D., Leonidio, A. D. C. R., Diniz, P. R. B., & Freitas, C. M. S. M. D. (2016). Estilos de vida urbano versus rural em adolescentes: associações entre meio-ambiente, níveis de atividade física e comportamento sedentário. Einstein (São Paulo), 14(4), 461-467.

Reitzel, L. R., Regan, S. D., Nguyen, N., Cromley, E. K., Strong, L. L., Wetter, D. W., & McNeill, L. H. (2014). Density and proximity of fast food restaurants and body mass index among African Americans. American Journal of Public Health, 104(1), 110-116.

Rodríguez, D. A., Cho, G. H., Evenson, K. R., Conway, T. L., Cohen, D., Ghosh-Dastidar, B., ... & Lytle, L. A. (2012). Out and about: association of the built environment with physical activity behaviors of adolescent females. Health & place, 18(1), 55-62.

Romanillos, G., Zaltz Austwick, M., Ettema, D., & De Kruijf, J. (2016). Big data and cycling. Transport Reviews, 36(1), 114-133.

Runtastic. (n. d.). Privacy policy. Retrieved from: https://www.runtastic.com/en/privacy-policy#headline_5.

Glanz, K., Rimer, B. K., & Viswanath, K. (Eds.). (2008). Health behavior and health education: theory, research, and practice. John Wiley & Sons.

Sallis, J. F., & Owen, N. (1997). Ecological models. Health behavior and health education: Theory, research, and practice, 2, 403-424.

Scheerder, J., Breedveld, K., & Borgers, J. (Eds.). (2015). Running across Europe: the rise and size of one of the largest sport markets. Springer.

Sileryte, R. (2015). Analysis of Urban Space Networks for Recreational Purposes based on Mobile Sports Tracking Application Data.

Smelser, N. J., & Baltes, P. B. (Eds.). (2001). International encyclopedia of the social & behavioral sciences (Vol. 11). Amsterdam: Elsevier.

Stopher, P., FitzGerald, C., & Zhang, J. (2008). Search for a global positioning system device to measure person travel. Transportation Research Part C: Emerging Technologies, 16(3), 350-369.

Strava. (2020, November 12). Retrieved from: strava.com

Suliburska, J., Bogdanski, P., Pupek-Musialik, D., Glod-Nawrocka, M., Krauss, H., & Piatek, J. (2012). Analysis of lifestyle of young adults in the rural and urban areas. Annals of Agricultural and Environmental Medicine, 19(1).

Tamura, K., Wilson, J. S., Goldfeld, K., Puett, R. C., Klenosky, D. B., Harper, W. A., & Troped, P. J. (2019). Accelerometer and GPS data to analyze built environments and physical activity. Research quarterly for exercise and sport, 90(3), 395-402.

Troped, P. J., Wilson, J. S., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). The built environment and location-based physical activity. American journal of preventive medicine, 38(4), 429-438.

Tsimeas, P. D., Tsiokanos, A. L., Koutedakis, Y., Tsigilis, N., & Kellis, S. (2005). Does living in urban or rural settings affect aspects of physical fitness in children? An allometric approach. British journal of sports medicine, 39(9), 671-674.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote sensing of Environment, 8(2), 127-150.

Turrell, G., Haynes, M., Wilson, L. A., & Giles-Corti, B. (2013). Can the built environment reduce health inequalities? A study of neighbourhood socioeconomic disadvantage and walking for transport. Health & place, 19, 89-98.

Vazquez-Prokopec, G. M., Stoddard, S. T., Paz-Soldan, V., Morrison, A. C., Elder, J. P., Kochel, T. J., ... & Kitron, U. (2009). Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. International journal of health geographics, 8(1), 68.

Viana, C. M., Oliveira, S., Oliveira, S. C., & Rocha, J. (2019). Land use/land cover change detection and urban sprawl analysis. In Spatial modeling in GIS and R for earth and environmental sciences (pp. 621-651). Elsevier.

Wageningen University & Research. (n.d.). The Land Use Database of the Netherlands (LGN database). Retrieved from: https://www.wur.nl/en/Research-Results/Research-Institutes/Environmental-Research/Facilities-Tools/Kaarten-en-GIS-bestanden/Land-use-database-of-the-Netherlands.htm.

Wang, S., Scheider, S., Sporrel, K., Deutekom, M., Timmer, J., & Kröse, B. (2020). What Are Good Situations for Running? A Machine Learning Study Using Mobile and Geographical Data. Frontiers in Public Health, 8.

Wu, X., Dong, J., & Lin, Z. (2014). Cost analysis of plug-in hybrid electric vehicles using GPS-based longitudinal travel data. Energy Policy, 68, 206-217.

Yun, H. Y. (2019). Environmental factors associated with older adult's walking behaviors: A systematic review of quantitative studies. Sustainability, 11(12), 3253.

Appendix A

Table A1: Descriptive statistics of surroundings of routes taken at weekdays and on weekends (buffer = 50m).

Environmental variables		Weekdays (n^a = 15746) (n^b = 4983) Mean (±SD)	Weekends (n^a = 6773) (n^b = 2091) Mean (±SD)
Urban	Level of greenness 50m	0.500 (0.094)	0.505 (0.095)
	Residential building density 50m	0.381 (0.355)	0.370 (0.363)
	Urbanization index 50m	1.180 (1.210)	1.110 (1.220)
	Land use mix 50m	0.506 (0.178)	0.515 (0.181)
	Count of 1-way crossings 50m	5.900 (5.180)*	6.310 (5.880)*
	Count of 3-ways crossings 50m	39.80 (30.90)**	43.10 (33.60)**
	Count of 4-way crossings 50m	18.90 (21.10)*	20.50 (22.40)*
	Proportion of water body 50m	0.279 (0.207)	0.285 (0.208)
Rural	Level of greenness 50m	0.546 (0.079)	0.551 (0.077)
	Residential building density 50m	0.213 (0.265)	0.206 (0.277)
	Urbanization index 50m	0.630 (0.590)	0.639 (0.640)
	Land use mix 50m	0.473 (0.186)	0.456 (0.192)
	Count of 1-way crossings 50m	4.640 (4.310)	4.860 (0.191)
	Count of 3-ways crossings 50m	29.70 (24.80)	31.60 (29.20)
	Count of 4-way crossings 50m	11.70 (14.10)	12.50 (16.50)
	Proportion of water body 50m	0.248 (0.192)	0.249 (0.192)

*: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.05$ of t-test.

**: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

Table A2: Descriptive statistics of surroundings of routes taken at weekdays and on weekends (buffer = 50m).

Environmental variables		**Weekdays** (n^a = 15746) (n^b = 4983) Mean (±SD)	**Weekends** (n^a = 6773) (n^b = 2091) Mean (±SD)
Urban	Level of greenness 100m	0.496 (0.095)	0.501 (0.096)
	Residential building density 100m	0.388 (0.355)*	0.366 (0.353)*
	Urbanization index 100m	1.280 (1.250)	1.210 (1.270)
	Land use mix 100m	0.506 (0.173)	0.511 (0.178)
	Count of 1-way crossings 100m	12.20 (10.10)*	12.90 (11.30)*
	Count of 3-ways crossings 100m	62.60 (48.60)**	67.50 (53.40)**
	Count of 4-way crossings 100m	24.60 (26.10)*	26.70 (28.20)*
	Proportion of water body 100m	0.289 (0.217)	0.295 (0.212)
Rural	Level of greenness 100m	0.546 (0.081)	0.551 (0.078)
	Residential building density 100m	0.204 (0.256)	0.191 (0.260)
	Urbanization index 100m	0.662 (0.593)	0.653 (0.701)
	Land use mix 100m	0.458 (0.186)	0.441 (0.191)
	Count of 1-way crossings 100m	9.410 (8.340)	9.900 (9.380)
	Count of 3-ways crossings 100m	45.30 (37.90)	47.80 (44.10)
	Count of 4-way crossings 100m	14.80 (16.70)	15.50 (19.30)
	Proportion of water body 100m	0.261 (0.203)	0.264 (0.199)

*: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.05$ of t-test.

**: Statistically significant difference between tracks taken at weekdays/on weekends at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

Table A3: Descriptive statistics of surroundings of tracks taken inside and outside neighbourhoods (buffer = 50m).

	Environmental variables	**Inside neighbourhood** (n^a = 1469) (n^b = 494) Mean (±SD)	**Outside neighbourhood** (n^a = 20857) (n^b = 6033) Mean (±SD)
Urban	Level of greenness 50m	0.464 (0.115)**	0.503 (0.092)**
	Residential building density 50m	0.341 (0.346)**	0.389 (0.347)**
	Urbanization index 50m	2.010 (2.290)**	1.140 (1.140)**
	Land use mix 50m	0.342 (0.228)**	0.522 (0.168)**
	Count of 1-way crossings 50m	2.360 (3.280)**	6.330 (5.440)**
	Count of 3-ways crossings 50m	13.70 (15.70)**	42.30 (30.30)**
	Count of 4-way crossings 50m	5.420 (8.410)**	20.00 (20.50)**
	Proportion of water body 50m	0.195 (0.218)**	0.286 (0.200)**
Rural	Level of greenness 50m	0.520 (0.111)**	0.550 (0.076)**
	Residential building density 50m	0.183 (0.238)*	0.223 (0.263)*
	Urbanization index 50m	1.100 (1.060)**	0.612 (0.514)**
	Land use mix 50m	0.320 (0.208)**	0.489 (0.172)**
	Count of 1-way crossings 50m	1.910 (2.710)**	5.050 (4.480)**
	Count of 3-ways crossings 50m	10.70 (11.90)**	32.50 (25.70)**
	Count of 4-way crossings 50m	2.860 (4.410)**	248.0 (222.0)**
	Proportion of water body 50m	0.183 (0.227)**	0.261 (0.179)**

*: Statistically significant difference between tracks taken inside/outside neighbourhoods at $p < 0.05$ of t-test.

**: Statistically significant difference between tracks taken inside/outside neighbourhoods at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

Table A4: Descriptive statistics of surroundings of tracks taken inside and outside neighbourhoods (buffer = 100m).

Environmental variables		Inside neighbourhood (n^a = 1469) (n^b = 494) Mean (±SD)	Outside neighbourhood (n^a = 20857) (n^b = 6033) Mean (±SD)
Urban	Level of greenness 100m	0.464 (0.11)**	0.499 (0.093)**
	Residential building density 100m	0.375 (0.369)	0.387 (0.347)
	Urbanization index 100m	1.960 (2.030)**	1.240 (1.210)**
	Land use mix 100m	0.385 (0.217)**	0.517 (0.167)**
	Count of 1-way crossings 100m	4.950 (5.930)**	13.00 (10.50)**
	Count of 3-ways crossings 100m	23.40 (24.40)**	66.40 (47.90)**
	Count of 4-way crossings 100m	8.020 (11.10)**	26.00 (25.60)**
	Proportion of water body 100m	0.237 (0.209)**	0.294 (0.209)**
Rural	Level of greenness 100m	0.522 (0.106)**	0.549 (0.077)**
	Residential building density 100m	0.203 (0.273)	0.209 (0.246)
	Urbanization index 100m	1.060 (0.962)**	0.639 (0.533)**
	Land use mix 100m	0.366 (0.209)**	0.468 (0.175)**
	Count of 1-way crossings 100m	3.960 (5.230)**	10.20 (8.620)**
	Count of 3-ways crossings 100m	18.10 (20.00)**	49.40 (39.30)**
	Count of 4-way crossings 100m	4.300 (5.830)**	16.20 (17.40)**
	Proportion of water body 100m	0.202 (0.241)**	0.272 (0.180)**

**: Statistically significant difference between tracks taken inside/outside neighbourhoods at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

Table A5: Descriptive statistics of surroundings of <=30mins and >30mins tracks (buffer = 50m).

Environmental variables		**<= 30 minutes** (n^a = 8540) (n^b = 2638) Mean (±SD)	**> 30 minutes** (n^a = 13982) (n^b = 4436) Mean (±SD)
Urban	Level of greenness 50m	0.485 (0.104)**	0.509 (0.091)**
	Residential building density 50m	0.378 (0.361)	0.380 (0.353)
	Urbanization index 50m	1.360 (1.600)**	0.910 (1.170)**
	Land use mix 50m	0.479 (0.197)**	0.522 (0.177)**
	Count of 1-way crossings 50m	4.110 (4.250)**	7.250 (6.790)**
	Count of 3-ways crossings 50m	27.80 (21.70)**	48.80 (37.10)**
	Count of 4-way crossings 50m	12.80 (14.10)**	23.50 (25.60)**
	Proportion of water body 50m	0.259 (0.218)**	0.292 (0.204)**
Rural	Level of greenness 50m	0.535 (0.088)**	0.554 (0.073)**
	Residential building density 50m	0.213 (0.274)	0.219 (0.274)
	Urbanization index 50m	0.761 (0.718)**	0.575 (0.599)**
	Land use mix 50m	0.445 (0.206)**	0.486 (0.183)**
	Count of 1-way crossings 50m	3.210 (3.360)**	5.790 (5.330)**
	Count of 3-ways crossings 50m	20.10 (17.20)**	36.70 (29.70)**
	Count of 4-way crossings 50m	7.610 (9.150)**	14.60 (16.90)**
	Proportion of water body 50m	0.227 (0.193)**	0.260 (0.191)**

**: Statistically significant difference between running takes less than/above 30 minutes at $p < 0.01$ of t-test.
n^a: Number of tracks of urban runners.
n^b: Number of tracks of rural runners.

Table A6: Descriptive statistics of surroundings of <=30mins and >30mins tracks (buffer = 100m).

Environmental variables		**<= 30 minutes** (n^a = 8540) (n^b = 2638) Mean (±SD)	**> 30 minutes** (n^a = 13982) (n^b = 4436) Mean (±SD)
Urban	Level of greenness 100m	0.482 (0.103)**	0.505 (0.092)**
	Residential building density 100m	0.394 (0.368)	0.376 (0.351)
	Urbanization index 100m	1.470 (1.600)**	1.100 (1.220)**
	Land use mix 100m	0.488 (0.187)**	0.515 (0.176)**
	Count of 1-way crossings 100m	8.720 (8.160)**	14.90 (13.50)**
	Count of 3-ways crossings 100m	45.10 (35.90)**	75.90 (59.00)**
	Count of 4-way crossings 100m	17.20 (18.20)**	30.30 (32.40)**
	Proportion of water body 100m	0.300 (0.211)**	0.275 (0.228)**
Rural	Level of greenness 100m	0.534 (0.089)**	0.554(0.074)**
	Residential building density 100m	0.210 (0.269)	0.199 (0.253)
	Urbanization index 100m	0.783 (0.715)**	0.596 (0.571)**
	Land use mix 100m	0.445 (0.202)	0.463 (0.185)
	Count of 1-way crossings 100m	6.580 (6.370)**	11.70 (10.20)**
	Count of 3-ways crossings 100m	31.60 (26.90)**	55.50 (45.60)**
	Count of 4-way crossings 100m	9.960 (11.30)**	18.20 (20.10)**
	Proportion of water body 100m	0.244 (0.198)**	0.271 (0.201)**

**: Statistically significant difference between running takes less than/above 30 minutes at $p < 0.01$ of t-test.
n^a: Number of tracks of urban runners.
n^b: Number of tracks of rural runners.

Table A7: Descriptive statistics of surroundings of tracks with different lengths (buffer = 50m).

Environmental variables		**<= 3000m** (n[a] = 4383) (n[b] = 1382) Mean (±SD)	**3000 – 5000m** (n[a] = 7550) (n[b] = 2439) Mean (±SD)	**> 5000m** (n[a] = 10589) (n[b] = 3253) Mean (±SD)
Urban	Level of greenness 50m	0.488 (0.119)**	0.500 (0.104)**	0.515 (0.098)**
	Residential building density 50m	0.367 (0.404)	0.396 (0.407)**	0.378 (0.402)
	Urbanization index 50m	1.597 (1.804)**	1.162 (1.277)**	0.921 (1.101)**
	Land use mix 50m	0.425 (0.235)**	0.521 (0.200)**	0.538 (0.195)**
	Count of 1-way crossings 50m	2.821 (3.389)**	5.279 (4.870)**	8.339 (8.091)**
	Count of 3-ways crossings 50m	17.66 (15.73)**	35.82 (23.94)**	58.48 (48.70)**
	Count of 4-way crossings 50m	7.542 (9.957)**	16.43 (16.71)**	29.33 (34.45)**
	Proportion of water body 50m	0.244 (0.248)**	0.280 (0.229)**	0.310 (0.227)**
Rural	Level of greenness 50m	0.524 (0.110)**	0.549 (0.009)**	0.561 (0.076)**
	Residential building density 50m	0.195 (0.292)	0.231 (0.224)**	0.197 (0.308)
	Urbanization index 50m	0.926 (1.124)**	0.642 (0.684)**	0.461 (0.501)**
	Land use mix 50m	0.411 (0.230)**	0.491 (0.206)	0.495 (0.191)
	Count of 1-way crossings 50m	2.402 (3.086)**	4.251 (0.206)**	6.633 (7.102)**
	Count of 3-ways crossings 50m	13.67 (12.86)**	26.23 (19.91)**	44.78 (38.59)**
	Count of 4-way crossings 50m	4.586 (6.767)**	9.958 (11.30)**	19.09 (21.56)**
	Proportion of water body 50m	0.201 (0.226)**	0.246 (0.214)**	0.275 (0.214)**

**: Statistically significant difference between tracks with different lengths at $p < 0.01$ of t-test.

n[a]: Number of tracks of urban runners.

n[b]: Number of tracks of rural runners.

Table A8: Descriptive statistics of surroundings of tracks with different lengths (buffer = 100m).

Environmental variables		**<= 3000m** (n[a] = 4383) (n[b] = 1382) Mean (±SD)	**3000 – 5000m** (n[a] = 7550) (n[b] = 2439) Mean (±SD)	**> 5000m** (n[a] = 10589) (n[b] = 3253) Mean (±SD)
Urban	Level of greenness 100m	0.483 (0.116)**	0.487 (0.112)**	0.511 (0.099)**
	Residential building density 100m	0.387 (0.403)	0.398 (0.400)	0.370 (0.393)*
	Urbanization index 100m	1.595 (1.737)**	1.278 (1.334)**	1.025 (1.182)**
	Land use mix 100m	0.443 (0.226)**	0.519 (0.196)**	0.527 (0.197)**
	Count of 1-way crossings 100m	2.821 (3.389)**	10.96 (8.966)**	17.07 (15.78)**
	Count of 3-ways crossings 100m	29.68 (25.17)**	57.15 (38.68)**	90.17 (76.30)**
	Count of 4-way crossings 100m	10.64 (12.72)**	21.84 (21.26)**	37.45 (42.37)**
	Proportion of water body 100m	0.268 (0.235)**	0.285 (0.241)**	0.319 (0.232)**
Rural	Level of greenness 100m	0.525 (0.108)**	0.547 (0.086)**	0.561 (0.078)**
	Residential building density 100m	0.202 (0.298)	0.213 (0.305)	0.180 (0.284)*
	Urbanization index 100m	0.932 (0.980)**	0.677 (0.718)**	0.484 (0.535)**
	Land use mix 100m	0.424 (0.223)**	0.472 (0.209)	0.462 (0.194)
	Count of 1-way crossings 100m	4.888 (5.567)**	8.717 (8.023)**	13.55 (13.75)**
	Count of 3-ways crossings 100m	22.61 (21.24)**	41.10 (32.02)**	66.21 (59.62)**
	Count of 4-way crossings 100m	6.434 (8.569)**	8.032 (497.2)**	14.85 (17.68)**
	Proportion of water body 100m	0.224 (0.231)**	0.261 (0.220)**	0.281 (0.235)**

*: Statistically significant difference between tracks with different lengths at $p < 0.05$ of t-test.

**: Statistically significant difference between tracks with different lengths at $p < 0.01$ of t-test.

n[a]: Number of tracks of urban runners.

n[b]: Number of tracks of rural runners.

Table A9: Descriptive statistics of surroundings of runners running at different frequencies (buffer = 50m).

	Environmental variables	**<= 24** (n^a = 7983) (n^b = 2634) Mean (±SD)	**24 - 48** (n^a = 7410) (n^b = 2136) Mean (±SD)	**> 48** (n^a = 7174) (n^b = 2304) Mean (±SD)
Urban	Level of greenness 50m	0.504 (0.103)	0.545 (0.093)	0.502 (0.108)*
	Residential building density 50m	0.374 (0.401)	0.206 (0.305)	0.377 (0.403)
	Urbanization index 50m	1.151 (1.403)*	1.104 (1.356)	1.090 (1.251)
	Land use mix 50m	0.506 (0.208)	0.461 (5.692)	0.514 (0.208)
	Count of 1-way crossings 50m	5.863 (6.455)**	28.82 (30.09)	6.336 (6.708)*
	Count of 3-ways crossings 50m	39.16 (35.58)**	11.36 (17.08)	43.78 (39.81)**
	Count of 4-way crossings 50m	18.27 (23.84)**	222.9 (278.8)	21.63 (28.52)*
	Proportion of water body 50m	0.284 (0.234)	0.251 (0.217)**	0.277 (0.237)
Rural	Level of greenness 50m	0.545 (0.093)	0.560 (0.081)*	0.545 (0.088)
	Residential building density 50m	0.206 (0.305)	0.196 (0.312)	0.222 (0.327)
	Urbanization index 50m	0.640 (0.810)	0.525 (0.614)**	0.667 (0.767)
	Land use mix 50m	0.461 (5.692)	0.471 (0.212)	0.499 (0.185)**
	Count of 1-way crossings 50m	28.82 (30.09)	4.610 (5.222)	5.806 (6.314)**
	Count of 3-ways crossings 50m	11.36 (17.08)	30.38 (29.47)	38.30 (34.55)**
	Count of 4-way crossings 50m	222.9 (278.8)	11.75 (15.23)	14.95 (18.72)**
	Proportion of water body 50m	0.251 (0.217)**	0.233 (0.232)**	0.267 (0.205)**

*: Statistically significant difference between runners running at different frequencies at $p < 0.05$ of t-test.

*: Statistically significant difference between runners running at different frequencies at $p < 0.01$ of t-test.

n^a: Number of tracks of urban runners.

n^b: Number of tracks of rural runners.

Table A10: Descriptive statistics of surroundings of runners running at different frequencies (buffer = 100m).

	Environmental variables	<= 24 (n^a = 7983) (n^b = 2634) Mean (±SD)	24 - 48 (n^a = 7410) (n^b = 2136) Mean (±SD)	> 48 (n^a = 7174) (n^b = 2304) Mean (±SD)
Urban	Level of greenness 100m	0.500 (0.103)	0.497 (0.107)	0.503 (0.107)
	Residential building density 100m	0.377 (0.394)	0.379 (0.398)	0.392 (0.401)*
	Urbanization index 100m	1.246 (1.407)	1.206 (1.299)	1.208 (1.408)
	Land use mix 100m	0.506 (0.200)	0.510 (0.204)	0.508 (0.212)
	Count of 1-way crossings 100m	12.13 (12.59)**	13.34 (13.65)	13.11 (12.92)
	Count of 3-ways crossings 100m	61.88 (56.34)**	68.23 (62.26)**	72.43 (68.73)**
	Count of 4-way crossings 100m	23.89 (29.90)**	27.96 (34.80)**	29.45 (36.33)**
	Proportion of water body 100m	0.297 (0.235)*	0.284 (0.247)*	0.312 (0.227)**
Rural	Level of greenness 100m	0.545 (0.094)	0.559 (0.083)*	0.547 (0.088)
	Residential building density 100m	0.195 (0.290)	0.184 (0.290)	0.207 (0.302)
	Urbanization index 100m	0.658 (0.734)	0.560 (0.633)**	0.686 (0.770)
	Land use mix 100m	0.448 (0.214)	0.451 (0.290)	0.476 (0.185)**
	Count of 1-way crossings 100m	9.142 (10.84)	9.454 (10.10)	11.94 (12.46)**
	Count of 3-ways crossings 100m	43.81 (45.94)	46.29 (45.30)	57.55 (53.40)**
	Count of 4-way crossings 100m	14.32 (20.21)	14.85 (18.60)	18.56 (22.22)**
	Proportion of water body 100m	0.263 (0.231)*	0.247 (0.247)*	0.277 (0.211)*

*: Statistically significant difference between runners running at different frequencies at $p < 0.05$ of t-test.

*: Statistically significant difference between runners running at different frequencies at $p < 0.01$ of t-test.

na: Number of tracks of urban runners.

nb: Number of tracks of rural runners.