Walkability in Amsterdam

A research investigating the most suitable walkability index for predicting walking behavior in Amsterdam

Name: Maaike Steenbeek
Student number: 6549004

Date: July 8th, 2022
Supervisors: dr. mr. S. Scheider, dr. mr. E. Nyamsuren

Human Geography
Master's thesis

Utrecht University,
Faculty of Geosciences

Image on front page retrieved from: https://www.clinkhostels.com/guide/things-to-do-alone-in-amsterdam/.
Logo Utrecht University retrieved from: https://www.uu.nl/organisatie/huisstijl/downloads/logo.

Contents

Preface

Before you lies my research titled *'Walkability in Amsterdam: a research investigating the most suitable walkability index for predicting walking behavior in Amsterdam.'* This research was conducted to fulfill the graduation requirements of the Master *Human Geography* at the University of Utrecht. I started this research in February 2022 and finished it in July 2022.

I would like to thank my supervisors Mr. Scheider and Mr. Nyamsuren for their excellent supervision and support during these months. You both provided me with practical feedback, tips, and ideas throughout the entire thesis process, which kept me motivated.

A special note of thanks to my family and to my significant other. You always offered me support during these past months. Writing a thesis comes with ups and downs, and you were there for me at all times.

I hope you enjoy reading this research!

Maaike Steenbeek
Driebergen-Rijsenburg, 8 juli 2022

Abstract

This research investigates the most suitable walkability index for predicting walking behavior in Amsterdam. While walkability has already been widely researched, there is no one-size-fits-all model to predict walkability. Variables that are used in North American and Australian Walkability Indices (WIs) may not be directly applicable to a European context. To find the most suitable walkability index for Amsterdam, this study uses a review of related work, a comparison of the three chosen walkability indices, two types of sensitivity analysis, a regression analysis, and a correlation analysis. The regression analysis pointed out that the Urban Walkability Index has the largest share of explained variance out of all three existing Walkability Index models that were compared; the Graz Walkability Index and Frank's Walkability Index were found to be less suitable. In this study, a walkability index is constructed especially for the context of Amsterdam. This Amsterdam Walkability Index turns out to be the most suitable walkability index for predicting walking behavior in Amsterdam, since 30.6% of variance in mean walking distance per postal code area can be ascribed to the index.

Keywords: walkability, walkability indices, walking behavior, built environment.

1. Introduction

Walkability has become an important concept in urban planning and transportation (Liao et al., 2020). The term can be simply defined as a measure of how friendly an area is to walking (Adkins et al., 2017 in Liao, Van den Berg, Van Wesemael, & Arentze 2020; Vale et al., 2016 in Liao et al., 2020). According to Southworth (2005), Moura, Cambra & Gonçalves (2017), and Zakaria and Ujang (2015), walkability is the basis of a sustainable city, and it has environmental, social and health benefits. A walkable city leads to less automobile use, less air pollution, less noise pollution, and it can also contribute to the reduction of greenhouse gas emissions and the per capita rate of resource use (Liao et al., 2020; Southworth, 2005). In addition, walkability can boost social interaction and therefore has the potential to help community building (Whyte, 2012; Southworth, 2005). Walkability is also related to physical activity; people living in more walkable and bikeable neighborhoods with homes in proximity to non-residential destinations are less likely to be overweight or obese than people living in more suburban neighborhoods that rely on motorized transportation (Frank, Andresen & Schmid, in Frank, Sallis, Saelens, Leary, Cain, Conway, & Hess, 2010; Southworth, 2005). Walkability thus forms an essential facilitator for sustainable urban development, with benefits for people and planet (Liao et al., 2020).
The built environment has been known as an influencing factor on an individual's tendency to walk, and thus researchers try to account for the effective environmental factors on walking. More than 200 studies show that the urban environment can influence walking behavior (Ruiz-Padillo, Pasqual, Uriarte, & Cybis, 2018). A number of researchers put efforts to develop walkability indices (WIs) to show the status of walkability for specific zones (Habibian & Hosseinzadeh, 2018). Dobesova and Krivka (2012) argue that walkability indices can be useful in urban planning. A walkability index gives urban planners the opportunity to assess the built environment and to evaluate proposed urban plans by the effect the plan has on possible physical activity of the city's inhabitants (Dobesova & Krivka, 2012).

1.1 Problem statement

It becomes clear that walkability is a relevant and important concept, but there are several problems concerning walkability research. Most walkability indices (WIs) are based on cities in North America and Australia, with less research conducted in European contexts (Stockton, Duke-Williams, Stamatakis, Mindell, Brunner, & Shelton, 2016). European cities typically feature an old town in the center, with small streets, medieval houses, and pedestrian areas (Grasser, Van Dyck, Titze, & Stronegger, 2016). Unlike European cities, urban areas in the USA are characterized by low population density, low land use mix and low connectivity (Grasser et al., 2016). The fundamental differences between the USA and Europe concerning their urban form and transportation infrastructure mean that extrapolation of findings to European cities is not appropriate (Reyer, Fina, Siedentop, & Schlicht, 2014; Stockton et al., 2016). Research in the European context is thus necessary (Grasser, 2014). There are some European studies on WIs, but the variables that are used are rather similar to variables used in North American studies. For example, a WI that was constructed for London used land use mix, street connectivity, and residential dwelling density as variables (Stockton et al., 2016). Rather similar, Grasser, Titze & Stronegger (2017) constructed a WI for the Austrian city of Graz and used four (or more)-way intersection density, land use mix, and residential density. For Stuttgart, the Walk Score, a WI developed by a company for Seattle, was seen as promising in the German context (Reyer et al., 2014). It seems that the variables are used in European context are mostly the same as the variables that are used in North American WI studies. For example, a Canadian walkability study shows that population density, residential density, intersection density and the availability of retail and services were significant variables that correlate with walking in Toronto (Glazier, Weyman, Creatore, Gozdyra, Moineddin, Matheson, Dunn & Booth, 2012). Another study from North America uses net residential density, retail floor area ratio, land use mix, and intersection density as variables in their WI (Frank, Sallis, Saelens, Leary, Cain, Conway, & Hess, 2010). These frequently used variables are also the areas where differences between the USA and Europe become apparent. As stated above, the differences between the two continents when it comes to urban form and transportation infrastructure are fundamental (Reyer et al., 2014; Stockton et al., 2016). More research is thus needed to investigate which variables can best explain walking behavior in a European context.

1.2 Research question

To address these problems, the goal of this study is to find the variables that best estimate walking behavior in Amsterdam, by comparing three existing walkability indices and constructing a walkability index specifically for Amsterdam using the variables that occur in the three existing models. Using regression analysis and walking data from the city of Amsterdam (ODiN (*Onderweg in Nederland*) dataset) as an empirical ground truth, it will be analyzed which model with which variables best predicts the walking behavior in Amsterdam.
The research question that this thesis aims to answer is thus as follows:

What is the most suitable walkability index for predicting walking behavior in Amsterdam?

In addition, to help answer the research question, several sub questions (SQs) are used:
SQ1: What are the most commonly used variables in walkability indices?
SQ2: How can variables in a walkability index be operationalized for the context of Amsterdam in a way that is comparable to the original source?
SQ3: How consistent are the walkability indices?
SQ4: In what ways do the walkability indices differ from each other?
SQ5: Which set of variables most accurately predicts empirical walkability data for Amsterdam?

1.3 Scientific relevance

The built environment has been known as an influencing factor on an individual's tendency to walk, and thus researchers try to account for the effective environmental factors on walking. More than 200 studies show that the urban environment can influence walking behavior (Ruiz-Padillo et al., 2018). A number of researchers put efforts to develop walkability indices (WIs) to show the status of walkability for specific zones (Habibian & Hosseinzadeh, 2018). WI studies often address the criteria of transportation network design, land use diversity and population density (the so-called 3Ds) to capture the built environment (Habibian & Hosseinzadeh, 2018). However, different walkability models use different variables. Some WIs include just the 3Ds, those being density, diversity, and design (Cervero & Kockelman, 1997 in Habibian & Hosseinzadeh, 2018), while other WIs also include distance to transit and destination accessibility (Habibian & Hosseinzadeh, 2018). Such different models may produce significantly different results. After three decades of research, considerable inconsistency remains over how walkability characteristics are measured and defined (Frank, Appleyard, Ulmer, Chapman & Fox, 2021). The results from studies using dissimilar methods are often conflicting, making it difficult to interpret and compare findings (Frank et al., 2021). Such differences in methods can be noticed for example in the use of either topographical or perceived or environmental data, and the geographic aggregation scale (Brownson et al., 2009 in Frank et al., 2021). Comparing WIs on methodology is thus essential because methods can largely influence the outcomes.
In addition, most WIs are based on cities in North America and Australia, with less research conducted in European contexts (Stockton et al., 2016). As explained in section 1.1, the differences between Europe and the USA in urban form and transportation infrastructure mean that extrapolation of findings to European cities is not appropriate (Reyer et al., 2014; Stockton et al., 2016). Grasser (2014) similarly claims that there is a particular need for European walkability studies, because of the large differences between the built environment in the US and in Europe. This thesis thus adds to existing research by comparing the reliability of different models and by investigating which variables best predict actual walking behavior per postal code in Amsterdam.

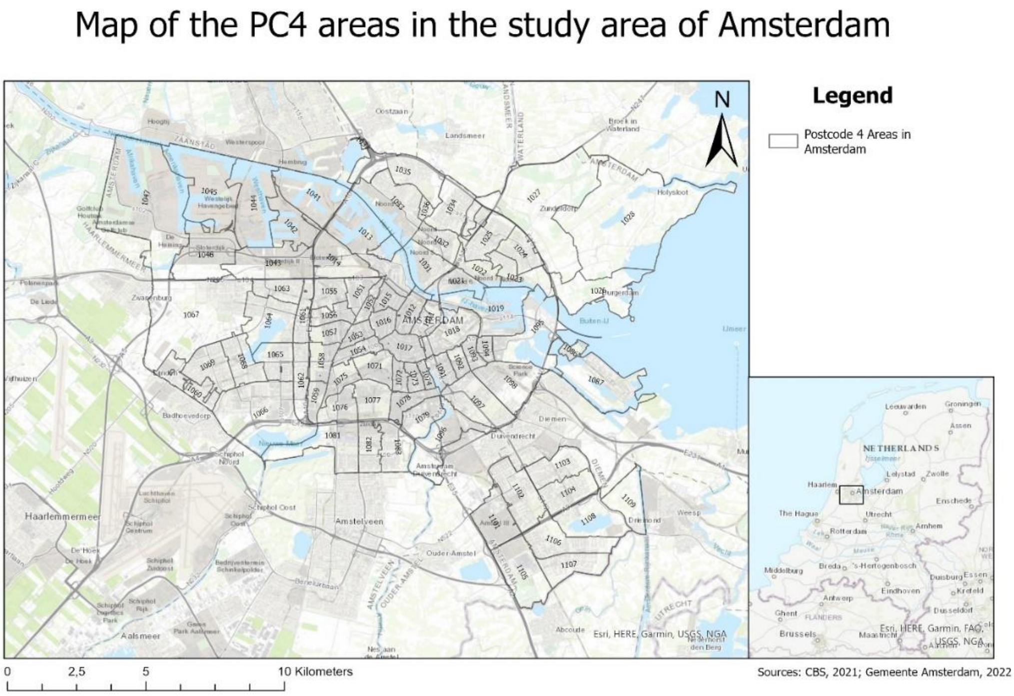
1.4 Societal relevance

Cities are facing serious environmental and public health challenges (Tsiompras & Photis, 2017). Walking is an environmental-friendly and human powered transportation option and it seems to help combat both climate change and the epidemic of obesity (Tsiompras & Photis, 2017). Walkability has environmental, social and health benefits and it forms the basis of a sustainable city (Moura et al., 2017; Southworth, 2005). When a city government knows what factors influence walkability in their city, actions can be undertaken in public space to ameliorate the walkability of the city. However, a body of contradictory results of walkability studies has led to considerable confusion over policy directions for decision-makers, argue Frank et al. (2021).
Next to that, investigating the reliability of WIs provides clarity and can increase the reliability of science for the larger public. Indices, e.g. those used in health models, are used to draw conclusions by city governments. Methods and models can thus have an impact on society, and this thesis aims to investigate the WIs for the context of Amsterdam.
This thesis thus adds to existing research by aiming to provide clarity for decision-makers on what policy could boost walkability in Amsterdam.

1.5 Study area

Amsterdam is selected as the area to study walkability (see figure 1). As a European city, its urban form and transportation infrastructure is fundamentally different from US cities (Reyer et al., 2014; Stockton et al., 2016). The capital city of the Netherlands had 873000 inhabitants in 2021 and a surface area of 219 square kilometers (CBS, 2021; KadastraleKaart, 2022). The city consists of 81 postal codes (PC4), ranging from 1011 to 1109 with some gaps in between. Gilderbloom, Hanka and Lasley (2009) claim that Amsterdam is an earth-friendly city, which is illustrated by the city's lack of petroleum dependency. Nearly half of the trips in Amsterdam are petroleum free (Gilderbloom et al., 2009). Most pedestrian flows can be found in the historic inner city, especially near the central train station (Zonghao, 2019). Amsterdam thus forms an interesting case study area to investigate which factors in the built environment influence walking.

Figure 1: The study area of this research, including the PC4 areas in Amsterdam and the position of the capital city within the Netherlands.



2. Related work

This chapter aims to provide insight into the meaning of the term walkability, its importance, its characteristics and the ways it can be measured. The terms walkability and WI are embedded in the existing body of relevant literature. In addition, various aspects of walkability are discussed, which lays the foundation for identifying the variables of walkability used in WIs. The WIs that are compared in this research are introduced in the concluding section of this chapter and are elaborated upon in later chapters.

2.1 Definitions of walkability

According to Southworth (2005), a clear, theory-based definition of walkability is lacking. Walkability is sometimes simply defined as a measure of how friendly an area is to walking (Adkins et al., 2017 in Liao et al., 2020; Vale et al., 2016 in Liao et al., 2020). Other definitions explicitly include the importance of the built environment and define walkability as *'the extent to which the built environment is walking-friendly'* (Abley, 2005, p. 3). Similarly, Steve (2005, in Zakaria & Ujang, 2015) defines walkability as *'the extent to which walking is readily available as a safe, connected, accessible, and pleasant mode of transport.'* Here, the importance of walking as a mode of transport is stressed. Southworth (2005) focuses more on safety and aesthetics and defines walkability as *'the extent to which the built environment supports and encourages walking by providing for pedestrian comfort and safety, connecting people with varied destinations within a reasonable amount of time and effort, and offering visual interest in journeys throughout the network'* (Southworth, 2005, p. 248). Galanis and Eliou (2010) argue that the definition of walkability is not specific, but that the term can be explained as *'the suitability that the urban road environment offers to pedestrians'* (Galanis & Eliou, 2010, p. 386). This shows a clear focus on roads and road connectivity. It becomes clear that there is not one universal definition, but different authors stress various aspects of walkability. This research uses the following definition: walkability is *'the extent to which the built environment provides factors that are known to support walking.'* This definition is loosely based on that of Abley (2005) and Southworth (2005) and this focus is chosen because this research focuses of walkability variables that can be observed in the built environment (see section 2.5).

2.2 Importance and advantages of a walkable city

As mentioned in the introduction, walkability has become an important concept in the fields of urban planning and transportation (Liao et al., 2020). In addition, the topic has gained increasing interest as a relevant approach of urban revitalization (Conticelli, Tondelli, Papageorgiou & Maimaris, 2018). This is because walkability has several advantages.

First of all, a high walkability yields environmental benefits. Walking is a mode of transport that does not emit pollution, moreover, it is the primary form of active travel which has zero emissions (Conticelli et al., 2018). Walkability has the potential to reduce congestion and car dependency (Conticelli et al., 2018), which helps enhance air quality.

Secondly, walking has health benefits. Areas with a high walkability can facilitate outdoor walking and exercise, and therefore can enhance a healthy lifestyle. Saelens, Sallis and Frank (2003) found that neighborhoods with a high walkability can add 15 to 30 walking minutes per week per resident. While this is a small effect on the individual level, it can have a large positive public health effect (Grasser, 2014; Forsyth, Hearst, Oakes, & Schmitz, 2008). The built environment is shown to be an important determinant of health (Grasser, 2014), and physical activity is a correlate of the built environment (Duncan, Spence and Mummery, 2005 in Grasser, 2014). Therefore, modifying the built environment can have a permanent impact on the entire population living in that area (Grasser, 2014). Thus, by creating a health-promoting built environment, one that is walkable, the healthier choice simultaneously becomes the easier choice (Grasser, 2014).

Thirdly, the economic development benefits of a high walkability (Litman, 2003). Retail and employment centers are affected by the quality of their pedestrian environment, especially in urban

areas and resorts. A shopping mall or office complex may become more economically competitive if walking conditions improve. Walkability improvements can also boost regional economic development by shifting consumer expenditures from vehicles and fuel to other consumer goods that provide regional business activity (Litman, 2003).
Lastly, walkability could in some cases promote social interaction and community building through face-to-face contact (Conticelli et al., 2018; Whyte, 2012). Walkability has major impacts on community livability, argues Litman (2003). Streets are a large portion of public space and that is where people interact with their community. More attractive, safe, and walkable streets increase community livability (Litman, 2003). Moreover, walkability has recently been emphasized as an important element that enhances social capital (Leyden, 2003 in Jun & Hur, 2015). However, Sallis et al. (2009) found no association between walkability and social cohesion. The effect of walkability on social interaction and cohesion is therefore not clear.
A walkable neighborhood thus is important, because it is beneficial for the environment, for the health of inhabitants, and for economic development in the area (Liao et al., 2020; Southworth, 2005; Moura et al., 2017; Litman, 2003).

2.3 Types of walking

In walkability research, a distinction is made between transport-related walking and recreational walking (Giao, Kamphuis, Helbich, & Ettema, 2020). Wunderlich (2008) calls this respectively purposive walking and discursive walking. Transport-related walking, or purposive walking, means walking to reach a destination for a specific purpose (Mirzaei, Kheyroddin, Behzadfar, & Mignot, 2018). It is essential and destination-oriented, and it entails a rapid pace (Hsieh & Chuang, 2021). Purposive walking thus is associated with road networks, but not strongly linked to the aesthetics of the urban environment (Careri, 2017 in Hsieh & Chuang, 2021). Recreational walking on the other hand entails walking for relaxation (Mirzaei et al., 2018). It can be a spontaneous walk without a specific destination, and it is therefore sometimes called *urban roaming* (Wunderlich, 2008). Discursive walking therefore differs from walking to a destination for recreational activity (walking for recreation) because a discursive walk can be treated as the trip purpose of itself (Hsieh & Chuang, 2021). This view, that traveling for its own sake has intrinsic utility, is opposed to traditional travel demand theory, which claims travel is merely a means to reach a destination or to participate in an activity (Mirzaei et al., 2018). In this research, the focus is mainly on walking as a mode of transportation.

2.4 Factors of walking behavior

Walking as a mode of transportation, purposive walking, is affected by the physical environment (Saelens, 2003). The physical environment consists of seven D’s that can influence transport behavior: density, diversity, design, destination accessibility, distance to transit, demand management, and demographics (Ewing & Cervero, 2010). The three WIs that are discussed in the methodology chapter, mostly use a mix of some of these seven D’s to construct their models. Density, the first D, is measured as the variable of interest per unit of area. Examples of variables of interest are population, and dwelling units (Ewing & Cervero, 2010). Dense neighborhoods encourage active transport, since daily needs are likely to be reached in a short distance (Saelens et al., 2003). The second D is diversity and this concerns the different land use types in a certain area and the extent to which they occur. It is argued that neighborhoods with mixed land use allow people to reach their daily needs within shorter distances, thereby encouraging active transport modes like walking (Saelens et al., 2003). Design, being the third D, includes street network characteristics within an area (Ewing & Cervero, 2010). The availability of convenient walking paths or routes is an important factor in encouraging people to use this active way of transportation (Maibach, Steg, & Anable, 2009). Destination accessibility is the fourth D and measures the ease of access to trip attractions. Neighborhoods with access to daily needs such as work, food, and leisure within walking distance can encourage walking as a form of transportation (Maibach et al., 2009). The fifth D is distance to transit and this is often measured as an average of the shortest street routes from residences or workplaces to the nearest rail station or bus stop. Demand management, which forms the sixth D, is involved with parking supply and cost. This included in some studies. The seventh and last D is demographics, which can entail socio-economic status (SES). Though demographics is not part of the environment, the factor is sometimes used in travel studies (Ewing & Cervero, 2010).

Out of these factors concerned with walking behavior, Saelens et al. (2003) argue that the main elements influencing walkability are residential density, land use mix, and street connectivity. This corresponds with Park, Deakin, and Lee (2014), who claim that the most frequently used urban form variables used to measure walkability are population and housing density, entropy-based land use mix, and variables representing street patterns like intersection density. These factors are used as variables in several WIs, for example in Frank's WI, which is one of the most commonly used methods in walkability studies (Liao et al., 2020). Frank, Engelke, & Schmid (2003) and Frumkin, Frank and Jackson (2004) claim that next to street connectivity, proximity to destinations is another fundamental aspect of walkability. Proximity is determined by density and land use mix. Density measures the quantity of people, household or jobs distributed over a certain area unit, and land use mix is a measure of how many types of land use (like offices, housing, retail, entertainment, services and so on) are located in a given area (Grasser et al., 2017). Connectivity calculates the directness of the routes between destinations, based on the design of the street network (Frank et al., 2003; Frumkin, et al., 2004).

In some studies, like in the Neighborhood Quality of Life Study (Frank et al., 2006) and in the Physical Activity in Localities and Community Environments (Leslie, Coffee, Frank, Owen, Bauman, & Hugo, 2007), the neighborhood's socio-economic status (SES) was also studied as a potential moderator of the associations between walkability and physical activity (Van Dyck, Cardon, Deforche, Sallis, Owen, & De Bourdeaudhuij, 2010). SES is defined as a measure of one's combined economic and social status, and tends to be positively associated with better health (Baker, 2014). The term encapsules education, work, and economic resources (Ross & Mirowski, 2008). A low neighborhood SES means there is inequality in contextual or macro-level resources (Ross & Mirowski, 2008). Disadvantaged places for example have few resources and its residents have few opportunities (Ross & Mirowski, 2008). The influence of neighborhood SES on walkability is still under debate (Van Dyck et al., 2010). McNeill, Kreuter and Subramanian (2006) showed that neighborhood SES is positively related to physical activity and walkability; a higher neighborhood SES means a higher walkability in that neighborhood. This was explained by the fact that a low neighborhood SES is also related to other factors that discourage physical activity such as poorer safety and aesthetic characteristics (Zhu & Lee, 2008 in Van Dyck et al., 2010). Van Lenthe et al. (2005, in Van Dyck et al., 2010) and Ross (2000, in Van Dyck et al., 2010) on the other hand found that a higher walkability was related to lower neighborhood SES. Since there is no clear association between neighborhood SES and walkability, and since it is not a component of any of the WIs, this factor is not taken into account in this research.

Next to the topographic aspects discussed so far, there are certain perceptual factors that can influence walkability. Safety and aesthetics are important aspects of walkability according to Grasser (2014). Measuring this however is challenging, because of the limited availability of geodata, resource-intensive GIS processes, and by the challenges of measuring perceptual concepts using GIS (Grasser, 2014). Since this study focuses solely on topographic ways of measuring walkability, the perceptual aspects of it are not discussed in this research. A point of criticism on studies that focus on perceptual evaluations of the built environment is that they have a limited applicability to planners and decisionmakers (Jun & Hur, 2015). Since this research aims to give recommendations for practical implementations, solely topographic ways of measuring walkability are taken into account.

2.5 Topographic and perceptual ways of measuring walkability

Walking is both associated with person-level characteristics and with environmental characteristics (Sallis, Owen & Fisher, 2015). Many urban design theories implicitly assume that environmental or physical features will make people want to walk (Forsyth, 2015), while the health field on the other hand focuses more on personal characteristics that affect walking, like income, cultural values and individual preferences (Forsyth, 2015). Measuring walkability can thus be done either through the assessment of the physical environment (topographic), which is the current predominant method, or through the gathering of personal perceptions of a location (perceptual) (Glanz, 2011). Grasser (2014) respectively calls this observational measures and perceived measures. Similarly, Bartzokas-Tsiompras and Photis (2020) claim that there are two main categories of walkability indicators: those on macro-level and those on micro-level. Macro-level indicators, or topographical measures, use readily available GIS datasets (Grasser, 2014). Such features include block length and number of intersections; aspects that are visible and tangible in the built environment. This is the type of indicators that are studied in this research. Proper measurement of the built environment requires a thorough understanding of methodological issues, access to appropriate geospatial data sources, the use of GIS software, and the technical ability to process geospatial data (Frank et al., 2021). These requirements will be dealt with in section 3.2. Micro-level walkability tools on the other hand are perceptual, and feature for example the presence of street amenities, perceived safety, and the aesthetics and conditions of the buildings in the neighborhood (Alfonzo, Boarnet, Day, Mcmillan, & Anderson, 2008; Forsyth, 2015; Grasser, 2014).

2.6 Walkability indices

As mentioned in the introduction, several researchers have selected some of the factors that possibly influence walking and have constructed models, called walkability indexes (WIs). These WIs are models that aim to show the status of walkability in certain areas (Habibian & Hosseinzadeh, 2018). In this research, three WIs are studied and compared (see also section 3.2).

Frank's Walkability Index (Frank et al., 2010) is composed of four topographical variables: net residential density, retail floor area ratio, intersection density, and land use mix. These were chosen based on extensive conceptual and empirical literature data, according to Frank et al. (2010). Ewing and Cervero's (2010) first three Ds out of the seven Ds, namely density, diversity, and design, are clearly present in Frank's WI. The study areas in which this WI was originally tested are King County, Washington, and the state of Maryland.

The Graz Walkability Index (Grasser, Titze & Stonegger, 2017) uses three topographical variables: four-way intersection density, proportion of mixed land use, and household unit density (Grasser et al., 2017). This WI bears a striking resemblance to Frank's WI. Again, the three Ds density, diversity, and design are used. The area for which this WI was constructed is Graz, a mid-sized city in the Southeast of Austria.

Lastly, the Urban Walkability Index (Glazier, Weyman, Creatore, Gozdyra, Moineddin, Matheson, & Booth, 2012) uses four topographical variables: dwelling density, population density, availability of all retail and services within 720 m, and number of street intersections within 720 m. Toronto, which is Canada's largest municipality, was selected as the case study area (Glazier et al., 2012).

2.7 Conclusions

At the end of this chapter, SQ1 can be answered. The question is '*What are the most commonly used variables in walkability indices?'* It becomes clear that the most frequently used variables that can be observed in the built environment that are used to measure walkability, are population and housing density, entropy-based land use mix, and variables representing street patterns like intersection density (Park et al., 2014). These most commonly used variables are similar to the main variables that influence walkability according to Saelens et al. (2003), since those are residential density, land use mix, and street connectivity. These factors are used as variables in several WIs, for example in Frank's WI, which is one of the most commonly used methods in walkability studies (Liao et al., 2020).

3. Methodology

This chapter explains all choices surrounding methodology. First, the research method will be explained. After that, the WI choices are explained, which includes a detailed model comparison. Subsequently, the operationalization is dealt with to ensure transparency, followed by an explanation of the dependent variable and the walking data of Amsterdam. Then, an explanation of the way the sensitivity analysis is performed is given. The chapter concludes by answering SQ2 and SQ3.

The research question that this study aims to answer, is as follows:

What is the most suitable walkability index for predicting walking behavior in Amsterdam?

To answer the research question, a quantitative research approach is used, with Geographical Information Systems (GIS) and SPSS as the most important toolboxes. This is because a GIS enables its users to analyze data both spatially and non-spatially and has the capacity of integrating large amounts of data (Mantri, 2008). GIS can be used to assess the built environment (Grasser et al., 2017). SPSS is a software program that is ideally suited for analyzing large amounts of data and for presenting statistical results (De Vocht, 2019). These two tools are both needed for comparing the different models, which will be explained in further detail below.
Several steps must be followed to answer the research question. After the overview of related works in chapter 2, this chapter focuses on the three chosen WIs; Frank's WI, the Graz WI and the Urban WI. The variables and formulas that these WIs use are examined. To be able to see which model best predicts walking behavior in Amsterdam, data on the actual amount and duration of walking trips per postal code in Amsterdam is used to establish a ground truth. For this, the ODiN dataset is used (see section 3.3. By performing several analyses using GIS, a walkability score per postal code (PC4) in Amsterdam is produced. This is done for each walkability index. Two types of sensitivity analysis are performed as well to see how robust the WIs are. To be able to see which model best estimates the walkability, a multiple linear regression analysis is performed, followed by a correlation analysis. The methods behind these steps will be dealt with in this chapter, which closes off by answering SQ2 and SQ3.

3.1 Explanation of WI choice

First of all, Frank's Walkability Index (Frank et al., 2010) was chosen because this is the most commonly used method in walkability studies (Liao et al., 2020). Secondly, the Graz Walkability Index (Grasser et al., 2017) was selected because this is another very commonly used index, according to Liao et al. (2020), and specifically deals with the European context. Lastly, the Urban Walkability Index (Glazier et al., 2012) was selected because this is a data-driven approach (Liao et al., 2020) instead of a theory-driven approach like the other two selected WIs. In theory-driven approaches, the selection, operationalization and weighing of environmental factors are fully based on a conceptual definition of walkability (Liao et al., 2020). Frank's Walkability Index is the most common measure in this approach in the American context. Such theory-based measures are validated by comparing computed scores to observed walking frequencies (Manaugh & El-Geneidy, 2011). This validation however merely provides a test of face-validity and it does not provide convincing empirical evidence that the selecting and weights of the factors are accurate, according to Liao et al. (2020). A data-driven approach on the other hand, derives a measure from regressing walking behavior on physical factors of the local environment. This approach is for example used by Glazier et al. (2012) for the construction of their Urban WI. A list of candidate variables was put together, and from this list the variables for which suitable data sources were readily available. The authors then performed a factor analysis to identify factors in the built environment that statistically uncorrelated with one another. This resulted in dwelling density, population density, street connectivity, and retail and service availability as the variables that create the Urban WI (Glazier et al., 2012). This WI, together with the other two WIs, are discussed in more detail in the next section.

Overall, it can be stated that theory-driven approaches result in measures that potentially are better generalizable, since they are not fitted on specific regional characteristics like data-driven measures (Liao et al., 2020). On the other hand, data-driven approaches consider a wide range of potentially relevant variables. In both approaches, the bias-variance tradeoff can come into play. The theory-driven approaches have a higher bias, meaning an underfit is a possible risk (Data Science Partners, 2021). Data-driven approaches on the other hand have more variance, which can lead to an overfit (Data Science Partners, 2021). Both of these extremes should be avoided; a right balance between bias and variance results in the smallest margin of error (Data Science Partners, 2021). It is therefore interesting to compare the WIs also with regards to the way they were constructed.

3.2 Reconstructing indices

This section will provide a deeper insight into the three WI models. Attention will be paid to different parameters: variables, spatial units of the analysis, allocated weights, and buffers.

3.2.1 Variables and formulas

The WI models all use a different set of variables and certain formulas to compute their walkability indices. This subsection gives an overview of how these are applied in the models.

The Frank Walkability Index by Frank et al. (2010) uses four variables: net residential density, retail floor area ratio, intersection density, and land use mix. Net residential density is operationalized by Frank et al. (2010) as the ratio of residential units to the land area devoted to residential use per block group. This operationalization runs into several problems for the Dutch context. First of all, block groups are not a common spatial unit in the Netherlands, and secondly, the Graz WI used residential density or household density instead of net residential density, because it seemed more fitting in a European context. This alteration by Grasser et al. (2017) is also applied here, hence residential density is operationalized as the number of dwellings divided by the land area of the PC4 area. Retail floor area ratio (FAR) is the next variable and this is computed by dividing the retail building square footage by the retail land square footage. The idea behind this is that a low ratio indicates a retail development likely to have substantial parking, while a high ratio indicates smaller setbacks (plain walls) and less surface parking (Frank et al., 2010). In a European context, this is operationalized in a different way than originally proposed by Frank et al. (2010). In a German study about walkability in Stuttgart, no data could be found on the actual floor space, so the blueprint size of the buildings was used as the nearest approximation available instead of using FAR (Reyer et al., 2014). This generalization has also been accepted by Dobesova and Krivka (2012). Therefore, the FAR in this research is operationalized as the area ratio of commercial buildings per PC4 area.
The third variable, intersection density, measures the connectivity of the street network, represented by the ratio between the number of intersections with three or more legs to the land area of the block group in acres. A higher density means a more direct path between destinations (Frank et al., 2010). The fourth and last variable is land use mix, also called the entropy score. This variable indicates the degree to which a diversity of land use types are present in a block group. Five land use types are used by Frank et al. (2010): residential, retail, entertainment, and institutional. In this research, all land uses from the BBG that were present in the study area were taken into account, to ensure a richer analysis and to make sure that no land uses were left out because they did not fall into one particular category. Values are normalized between 0 and 1, where 0 means single use and 1 means a completely even distribution of area across the uses (Frank et al., 2010). The formula that is used here is as follows:

$$LUM = -1\left(\sum_{i=1}^{n} pi * \ln(pi)\right) \div \ln(n)$$

Here, *LUM* stands for Land Use Mix, pi is the proportion of the neighborhood covered by the land use type i, against the summed area for the total land use categories of interest, and *n* is the number of land use categories of interest.
The scores of the variables are normalized using a Z-score. The Z-score standardizes the values and gives an idea of the relative position of a value within a distribution (De Vocht, 2019). The formula for calculating the Z-score is as follows (De Vocht, 2019):

$$Z = \frac{x - \mu}{\sigma}$$

Here, χ is the raw score, μ is the mean value and σ is the standard deviation.
The final Frank's Walkability Index thus is the sum of the z-scores of the four variables:

$$Walkability = (2 \times Z - intersection\ density) + (Z - net\ residential\ density) \\ + (Z - retail\ floor\ area\ ratio) + (Z - land\ use\ mix)$$

The Graz walkability Index (Grasser et al., 2017) uses three variables: four-way intersection density, proportion of mixed land use, and household unit density (Grasser, Titze, & Stronegger, 2016).
The first variable is four-way intersection density. It was assumed that using four-way intersections, instead of the more commonly used three-way intersections, may enable a better differentiation between walkable and less walkable areas in a European city. This variable is operationalized as the number of at least four-way intersections, divided by the total area of the neighborhood (Grasser et al., 2017).
Proportion of land use mix is the second variable. This is used instead of the land use Shannon Entropy. The land use 'Mainly residential' in Graz is defined as residential, but buildings with commercial, social, religious and cultural services are also included (Grasser et al., 2017). Areas with this land use type are considered mixed use and very walkable. If the area of one neighborhood has only one of these land use types, the Shannon Entropy would be zero although the actual land use mix might be high. To overcome this issue, proportion of mixed land use is used instead by Grasser et al. (2017). The proportion of mixed land use is thus the sum of land use categories 'Mainly residential' and 'Mixed use' within the neighborhood, divided by the total area of the neighborhood. For the context of Amsterdam however, these categories were already separated in the land use dataset (e.g. special categories for retail, public services, residential areas, etc.). Hence, it seemed more fitting to use the actual land use Shannon Entropy index, instead of the alteration made by Grasser et al. (2017).
The third and last variable is the household unit density. This is operationalized as the apportionment of number of household units, which is calculated by dividing the area of the statistical sector within the neighborhood by the total area of the statistical sector, multiplied by the number of household units. The density is thus the sum of the number of apportioned household units within the neighborhood, divided by the total area of the neighborhood.
As noted above, the Graz WI uses almost the same variables as the Frank WI, however, retail floor area ratio (FAR) is omitted. From a conceptual viewpoint, it is argued that there seems to be no need to include FAR in the walkability index (Grasser, 2014). FAR does not take underground parking into account, and only one study has found a positive association between physical activity and FAR. Additionally, it is questionable whether FAR is applicable to the European context (Grasser, 2014). A study conducted in Belgium that used Frank's Walkability Index (Frank et al., 2010), omitted FAR because of 'lack of relevance for a Belgium context and because no GIS data were available' (Van Dyck et al., 2010, p. S76).

The formula used for the Graz Walkability Index is as follows:

$$Walkability = (Z - four - way\ intersection\ density) + (Z - proportion\ of\ land\ use\ mix) \\ + (Z - household\ unit\ density)$$

Lastly, the Urban Walkability Index (Glazier, Weyman, Creatore, Gozdyra, Moineddin, Matheson, & Booth, 2012) uses four variables: dwelling density, population density, availability of all retail and services within 720 m, and number of street intersections within 720 m. Dwelling density and population density are the sum of numerator values (population or dwelling) in all tracts whose polygon area intersected with the 720 meter Euclidean buffer of a given tract centroid, divided by the summed area of all intersecting tracts. The third variable, all retail and services within 720 m, consists of grocery stores and fruit and vegetable stands, convenience and variety stores, bank branches, restaurants and cafes, and other miscellaneous retail and services. This variable is calculated by performing a network analysis, which generates variables describing the availability of selected resources. Availability is measured as the count of locations of a resource type within the 720 m buffer of each tract centroid.
The fourth and last variable is the number of street intersections within a 10 minute (720 m) network buffer. This is operationalized as the count of 3-way or greater intersections within a 720 m network buffer of the tract centroid.

$$
\begin{aligned}
Walkability = \ & \iota 1(Population\ density) + \iota 2(Dwelling\ density) \\
& + \iota 3(Availability\ of\ all\ retail\ and\ services) + \iota 4(Street\ connectivity)
\end{aligned}
$$

Here, ι is the value of factor loading. The Urban Walkability Index uses the following factor loadings for calculating the index on census tract level: 90 for population density, 94 for dwelling density, 77 for availability of all retail and services, and 70 for street connectivity. Therefore, these exact factor loadings will also be used in this research, to closely reproduce the walkability index.

3.2.2 Weights

Frank's Walkability Index assigns a weight of 2 to the Z-score of intersection density. This is based on prior evidence of the strong influence that street connectivity has on non-motorized travel choice (Frank et al., 2010). The other three variables, Z-score net residential density, Z-score retail floor area ratio, and Z-score land use mix, have a weight of 1.

The Graz Walkability Index uses equal weights for all variables; Z-score intersection density, Z-score net residential density, and Z-score land use mix all have a weight of 1. Connectivity was not found to be a strong correlate of health-related outcomes in the uncontrolled bivariate analyses, which is why each component was given the same weight in the Graz WI (Grasser et al., 2017).

Lastly, the Urban Walkability Index by Glazier et al. (2012) uses equal weights for all its four variables. Population density, dwelling density, availability of all retail and services, and street connectivity all have a value of 1. Factor analysis was used to identify variables that were statistically associated with each other and had validity as a measure of walkability.

3.2.3 Units and buffers

Frank's Walkability Index uses block group level data as the most appropriate geographical scale (Frank et al., 2010). Block groups are statistical divisions of census tracts, and are generally defined to contain between 600 and 3000 people (United States Census Bureau, 2022). For each block group, a walkability index is calculated. This unit was chosen because it takes advantage of available socio-economic data like median household income and education (Frank et al., 2010).

The Graz Walkability Index uses the neighborhood area as the unit of measurement. Variables are measured per neighborhood, e.g. the household density is the sum of household units in the neighborhood divided by the total area of the neighborhood. In addition, the residential address of each individual respondent in their study is also used for their analysis (Grasser et al., 2017). Around each address, a 1000 m circular buffer and a 1500 m street network buffer is created. Though walkable distances generally range somewhere between 0.25 and 0.50 miles (Schlossberg & Brown, 2004), which is approximately 400 and 800 m, Grasser et al. (2017) argue that inhabitants of European cities may also walk to destinations that are up to 1.5 km from their home. Therefore, Grasser et al. (2017) chose a buffer size of 1500 m for their European context. However, in this study, the address of the respondents in the ODiN dataset is not provided. The analyses therefore cannot be performed per address, but ODiN data per PC4 area is available. Therefore, the Graz WI is calculated per PC4 area.

The Urban Walkability Index (Glazier et al., 2012) uses census tracts as the main unit of analysis. This is done because data from surveys used for validation were not available at smaller geographical levels, nor at individual level (Glazier et al., 2012). Census tracts are small, relatively stable geographic areas that are located in census metropolitan areas and in census agglomerations with a core population of 50,000 or more inhabitants in the previous census (Statistics Canada, 2018). The closest approximation of this in a Dutch context is the postal code area. Glazier et al. (2012) use a geometric center point (centroid) as a representative 'sample point' for each of the census tracts in their study area, which is Toronto, Canada. For the availability variable and for the connectivity variable, a network buffer of 720 meters around the centroids was made, which is equivalent to 10 minutes of walking time (Glazier et al., 2012). Both of these variables were calculated as the count of locations of a certain resource type (retail and services, and intersections) within a 720 meter network buffer of the tract centroid, based on walkable streets and paths (Glazier et al., 2012). (Glazier et al., 2012). The two density variables (population and residential) were calculated using 720 meter Euclidian buffers.

3.3 Operationalization

Where the previous section described each WI in detail, this paragraph focuses on the calculations and analyses that were performed and how the original variables were operationalized for the context of Amsterdam.

3.3.1 Frank's WI

Frank's Walkability Index (Frank et al., 2010) uses four variables: net residential density, retail floor area ratio, intersection density, and land use mix. The net residential density, which for this research was simplified to residential density (see section 3.1.1) was calculated as follows. First, the dataset (see table 1) was clipped to PC4 Amsterdam. Then, to be able to join the addresses (PC6) to the PC4 file, a new field was calculated with the formula ''postcode'' [:-2]. This removed the letters from the postal code. Then, the address file was intersected with the PC4 Amsterdam file. Subsequently, the *collect events* tool was performed, followed by a *spatial join* with target feature PC4 Amsterdam. After that, to calculate the household density, a new field was made and the following calculation was done: (Join count / (shape area/10000)). Join count is the number of addresses per PC4. The shape area is the shape area of the PC4, and this is divided by 10000 because the shape area was originally given in squared meters but had to be transformed to ha, since that is the unit it is usually measured in. The final step was to add a new field and calculate the Z-score for the residential density per PC4.
Retail Floor Area Ratio (FAR) is the next variable. Frank et al. (2010) calculate retail FAR by dividing the retail building square footage by the retail land square footage. As previously mentioned, the FAR is operationalized in a different way in a European context; the FAR in this research was operationalized as the mean area of commercial buildings per PC4 (see section 3.1.1). To obtain the FAR, the *buildings* dataset (table 1) was first of all clipped to the PC4 Amsterdam extent. Next, all buildings containing one or more retail functions were selected. Then, this selection of the clipped dataset was intersected with the PC4 Amsterdam dataset. Then, the *dissolve* tool was used to obtain the mean area of retail buildings per PC4. Finally, a new field was made where the Z-score was calculated.
Three (or more) way intersection density was calculated as follows: first, the file containing all bicycle paths and footpaths (table 1) was clipped to the PC4 extent of Amsterdam. Then, the bicycle paths were removed. POI line segments were also removed, because these short line segments indicated the

entrance to a building. Next, the *unsplit line* tool was used, in order to be able to retrieve the number of ways per intersection. Subsequently, the *feature vertices to points* tool was used, followed by the *collect events* tool, and after that, all intersections with 2 or less ways were removed. Then, a *spatial join* (one to one) was performed, where the join count indicates the intersections per PC4 area for Amsterdam. The last step was calculating the Z-score of the intersections per PC4, which was done by creating a new field and calculating the Z-score.
The Shannon Entropy land use mix is the last variable in Frank's WI. To calculate the Shannon entropy, the CBS file of BBG (see table 1) was used, a file that contains land use data for the entire Netherlands. This was clipped to the PC4 extent of Amsterdam. Then, the *intersect* tool was used to intersect the PC4 Amsterdam file to the BBG land use file. After that, the *dissolve* tool was executed, by dissolving the PC4 Amsterdam file on land use and postal code, and selecting the land use field as statistic field with the function of Count. This resulted in the total amount of land uses per PC4 area. Then, a new field was made and the formula as mentioned in 3.3.1. was used: -1 * ((area covered by land use type i) / total area of postcode area) + ln (area covered by land use type I / total area of PC4)) / ln (total number of land use types per PC4). Then, *summarize statistic*s was done, to find the mean Shannon entropy per PC4 Amsterdam. Then, another new field was made, where the Z-score was calculated using the previously mentioned formula.

To calculate the final Frank et al. (2010) Walkability Index, the different variable datasets were all joined to the PC4 dataset. There, a new field was added. In that field, the Z-scores were added up and the Z-score of intersection density was multiplied by two, using the formula by Frank et al. (2010):

$$Walkability = (2 \times Z - intersection\, density) + (Z - net\, residential\, density) \\ + (Z - retail\, floor\, area\, ratio) + (Z - land\, use\, mix)$$

This resulted in the final Frank et al. (2010) Walkability Index.

3.3.2 The Graz WI

The Graz Walkability Index (Grasser et al., 2017) uses three variables: four-way intersection density, land use mix, and household unit density. Four-way intersection density was calculated in the same way as described above for Frank's WI. However, for the Graz WI, all intersections with less than three (instead of two) ways were removed. This was done because Grasser et al. (2017) use four (or more) way intersections instead of three (or more) way intersections.
The second variable is the proportion of mixed land use. As explained in section 3.1.1, the regular Shannon Entropy land use mix is used instead of the alteration made by Grasser et al. (2017). Shannon Entropy land use mix was thus calculated in the exact same way as described above for Frank's WI.
The third and last variable is household unit density. Due to alterations (see section 3.2), this was calculated in the same way as residential density which is described in the section above.
To construct the final Graz Walkability Index, the different variable datasets were all joined to the PC4 dataset. There, a new field was added. In that field, the Z-scores were added up, using the formula by Grasser et al. (2017):

$$Walkability = (Z - four - way\, intersection\, density) + (Z - proportion\, of\, land\, use\, mix) \\ + (Z - household\, unit\, density)$$

This resulted in the final Graz Walkability Index.

3.3.3 The Urban WI

Lastly, the Urban Walkability Index (Glazier et al., 2012) uses four variables: dwelling density, population density, availability of all retail and services within 720 m, and number of street intersections within 720 m.

To calculate the availability of all retail and services within 720 m, the following steps were taken. First of all, the *feature to point* tool was used to calculate the centroid per PC4 polygon in Amsterdam. Then, the retail and services dataset, and the restaurants and cafes dataset, were both clipped to the extent of PC4 Amsterdam. To be able to calculate the variable of availability of all retail and services within 720 m, the retail and services dataset was merged with the restaurants and cafes dataset. This had to be done because Glazier et al. (2012) include restaurants and cafes in the retail and services, but in the Dutch context, these were in a different dataset. After the merging, a service area network analysis was performed on the footpath network (the same dataset as was used for the Graz WI and Frank WI for calculating the intersection density). A 720 m network buffer around the centroids was made. Subsequently, the *intersect* tool was used to intersect the newly created 720 m network buffer around the PC4 centroids, with the retail and services dataset (which now also contains restaurants and cafes). Then, to link the centroid buffer polygons to the PC4 area where they have their centroid, a *spatial join* was performed for the network buffer polygons with the PC4 areas as their join feature. After that, another *spatial join* was performed to join the retail and services (including restaurants and cafes) to the network buffer polygons. This resulted in the join count providing the number of retail and services (including restaurants and cafes) within the 720 m network buffers around the centroid of each PC4.

The number of street intersections within 720 m is the next variable. This was calculated using the dataset for three-way or more intersections, the one that was also used for the Frank WI. A *spatial join* was performed to join the intersections to the network buffer polygons. The join count then indicates the number of three-way or more intersections within the 720 m network buffers around the centroid of each PC4.

The population density was obtained by making Euclidian buffers of 720 m around the PC4 centroids. Then, a *tabulated intersection* was performed, followed by *summary statistics*. This resulted in a standalone table showing all 81 PC4 centroid buffers with the summed amount of inhabitants in all PC4 areas that their buffer intersected with, and the table also contained the summed area in hectares of all PC4 areas that their buffer intersected with. In order to calculate the variable in the right unit of measurement, the area in hectares also needed to be changed into squared kilometers, since that unit is used by Glazier et al. (2012). After that step, the population density was calculated by dividing the summed amount of inhabitants by the summed area in km2.

For residential density, these same steps were undertaken. Here, BAG address data was used to identify residences (see chapter 4.4).

After collecting these four variables, the Urban WI was calculated using the following formula:

$$Walkability = 90 * (Population\ density) + 94 * (Dwelling\ density) + 77 * (Availability\ of\ all\ retail\ and\ services + 70 * (Street\ connectivity)$$

The original Urban WI by Glazier et al. (2012) does not use a Z-score normalization for the final walkability score. However, to be able to compare scores between the walkability indices, this research did calculate Z-scores for the Urban WI.

3.4 Dependent variable

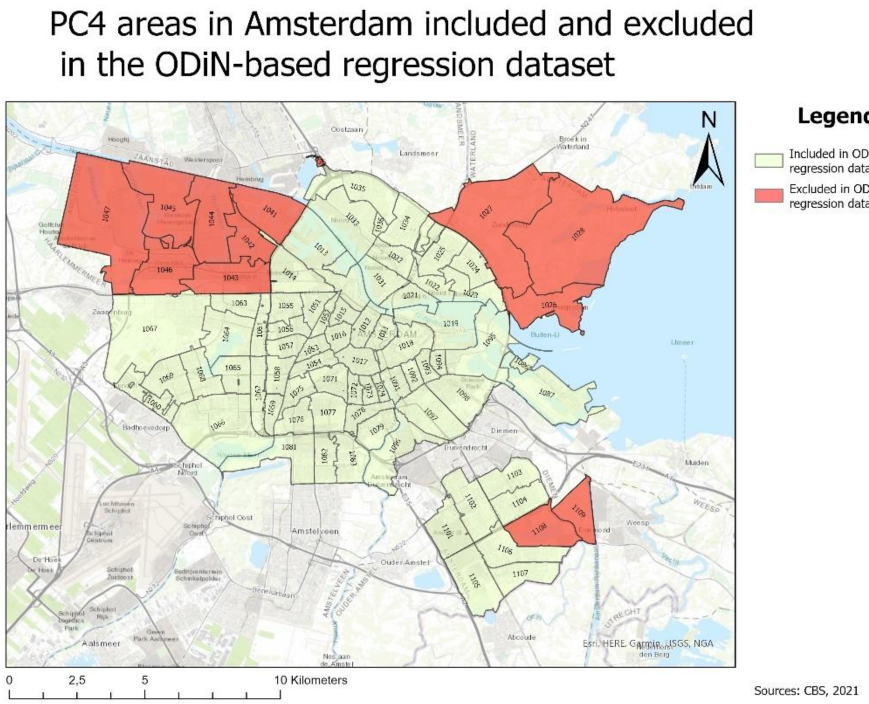
Where the previous sections described the WIs, containing the independent variables, this section focuses on a description of the dependent variable that is used in this research. The dependent variable is the mean walking distance per PC4 area and is operationalized as the mean walking distance per

PC4 area divided by the area in ha of the PC4 area. This dependent variable was also used to measure walkability in the study of Sisson, McClain, and Todor-Locke (2008) and Reyer et al. (2014) also wrote about the association between mean walking distance and the walkability of an area. The idea behind this dependent variable is that in highly walkable areas, the mean walking distance is supposed to be shorter (Reyer et al., 2014). This is because in highly walkable areas, there are many route opportunities for pedestrians so they can make their path as short as possible because of connected infrastructure.

3.5 Walking data of Amsterdam

The Onderweg in Nederland (ODiN) dataset provides one-day trip-diary data of a large nationwide sample of individuals above 6 years old in the Netherlands, in which all days of the week are covered. The residential neighborhood of every individual can be identified based on the postal code of the home address. The dataset also holds information about the distance that is walked by each individual. Using this data, the mean walking distance (Reyer et al., 2014) across individuals per the postal code (PC4) areas is computed. This way, the relevant walking data is gathered per postal code area level of Amsterdam. Then, this number is normalized for PC4 area in ha, by dividing the mean walking distance by the area of the PC4 in ha. This normalization is needed because Amsterdam's PC4 areas have a broad size range: the smallest one is just 8.89 ha (PC4 area 1037), and the largest one is 1367.93 ha (PC4 area 1028). It must be noted here that there was no ODiN walking data for 13 PC4 areas, which is why those areas are left out of the regression analysis (see figure 2). This concerns PC4 area 1026, 1027, 1028, 1037, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1108, and 1109. These areas are situated at the edges of the city of Amsterdam and include the harbor areas in the Northwest of the city (1041, 1042, 1043, 1044, 1045, 1046, 1047), the rural areas in the Northeast of the city (1026, 1027, 1028), a small area around the Noorder IJpolder in the upper North of Amsterdam (1037), and the rural area in the Southeast of the city, east of the Gaasperplas (1108 and 1109). These areas are relatively sparsely populated, and there were no respondents in the ODiN research that lived in these areas and reported any walking.

Figure 2: Map showing the PC4 areas that were included and excluded from the ODiN-based regression dataset.



3.6 Sensitivity analysis

Another method that will be used to determine the most suitable WI for Amsterdam is Sensitivity analysis (SA). This investigates how variations in the model output can be apportioned to different sources of variation (Crosetto & Tarantola, 2001). It is a way to find out how the model depends upon the information fed into it (Crosetto & Tarantola, 2001). In a SA, the various parameters and inputs are varied in order to observe their impacts on the model's results (Goodchild, 2005). For example, a model can be rerun with a parameter value increased by 10% and then reduced by 10% from its original value. If the impact on the results is substantially less than 10%, the parameter is not of critical importance and its accuracy is not a large concern (Goodchild, 2005).

The SA in this research consists of two parts. First of all, the weights of all variables of the WIs are both increased by 10%, one by one. This means that e.g. for the Graz WI, one variable would receive a weight of 1.1 while the other variables would keep their original weight, which is 1. This way, the effect of the changed weight of one variable on the final WI score becomes visible. The difference between that new WI score and the original WI score without changed weights is calculated in percentages and difference maps are produced for each variable.

The second way of checking the sensitivity of a model is changing the spatial units of analysis. In this study, the analyses as described in section 3.1 are rerun, but '*buurten*' (neighborhoods, as defined by CBS) are used as administrative units instead of PC4 areas. For this SA, 479 out of the 481 neighborhoods of Amsterdam are taken into account. The two neighborhoods that cover the lake Buiten-IJ and a part of lake IJmeer are excluded, since they are not covered by the official PC4 areas of Amsterdam either. By assessing walkability on a smaller geographical scale, MAUP (Modifiable Areal Unit Problem) can be noticed and it can be observed whether outliers in certain neighborhoods strongly influence the outcome of the larger PC4 area.

3.7 Regression analysis

A multiple linear regression investigates the correlational linear relationship between a dependent variable and several independent variables (De Vocht, 2019). This analysis is essential in examining the association between walking behavior and the variables that are used in the different walkability indices. Each WI is entered into the regression with the dependent variable being the mean walking distance per PC4 (see section 3.3).

3.8 Correlation analysis

A correlation shows the strength and direction of an association between two variables (De Vocht, 2019). After testing how the WIs relate to actual walking behavior as reported in ODiN, it should also be investigated whether the WIs correlate with each other. Therefore, a correlation is carried out between all three existing WIs and the newly constructed WI. Since the WIs are ratio variables, Pearson's Correlation coefficient is used.

3.9 Data collection

The data needed for the variables was for a large part obtained from sources via the L-drive of Utrecht University. The table below gives a complete overview of all variables used, which index uses which variables and where the data was retrieved from.

Table 1: Overview of data sources that were used for the WI variables, including any specific selections that were applied to the dataset.

Variable	Used by Frank's WI	Used by Graz WI	Used by Urban WI	Data source (see Bibliography for full reference)	Selection
Household density / dwelling density	Yes	Yes	Yes	Overheid (2022).	Use 'Adressen' and make the selection for Amsterdam as 'woongemeente'.
Population density	No	No	Yes	CBS (2020).	
Retail floor area ratio	Yes	No	No	Overheid (2022).	Only include the 'oppervlakte' of addresses with a 'winkelfunctie' of 1 or more.
Intersection density	Yes	Yes	Yes* (number of street intersections within 720 m)	Gemeente Amsterdam (2022).	Leave out the cycle paths and POI's (home entrances)
Land use mix	Yes	Yes	No	CBS (2017).	
Availability of retail and services within 720m	No	No	Yes	Esri Nederland (2021) Esri Nederland (2022)	

3.10 Conclusions

At the end of this chapter, SQ2 and SQ3 can be answered.

SQ2 is as follows: '*How can variables in a walkability index be operationalized for the context of Amsterdam in a way that is comparable to the original source?*' It becomes clear that three variables in the WIs needed to be operationalized in a slightly different way than the original method in order to reproduce them. This was the case for retail FAR, for residential density, and for Shannon Entropy land use mix. Frank's WI calculates retail FAR by dividing the retail building square footage by the retail land square footage (Frank et al., 2010). For the context of Amsterdam, no data could be found on actual floor space, which was also the case for other European WI studies (Reyer et al., 2014; Van Dyck et al., 2010). Instead, the blueprint building size was used to determine the floor area of commercial buildings, an alternative that was accepted by Dobesova and Krivka (2012) and also used by Reyer et al. (2014). This is why in this research the area ratio of commercial buildings per PC4 area was used as an alternative to retail FAR.

Frank et al. (2010) originally use net residential density, but this variable was simplified to residential density. This was done because Grasser et al. (2017) also used residential density or household density instead of net residential density, to make the variable more fitting to European context. Therefore, residential density is operationalized as the number of dwellings divided by the land area of the PC4 area.

The Shannon Entropy land use mix was also operationalized in a slightly different way. Five land use types are used by Frank et al. (2010): residential, retail, entertainment, and institutional. In this research however, all land uses from the BBG that were present in the study area were taken into

account in the entropy calculations. This was done to ensure a richer analysis and to make sure that no land uses were left out because they did not fall into one particular category.
All in all, it was possible to operationalize most variables in the WIs exactly in the intended way for the context of Amsterdam, but retail FAR and Shannon Entropy land use mix required some little alterations.

SQ3 is as follows: '*In what ways do the walkability indices differ from each other?'* This chapter showed that Frank's WI and the Graz WI are rather similar. They are both theory-driven models, meaning that the selection, operationalization and weighing of environmental factors are fully based on a conceptual definition of walkability (Liao et al., 2020). The Graz WI uses three out of the four variables that occur in Frank's WI, including intersection density which was slightly adjusted to European context. Their formulas are a sum of the Z-scores of the variables. Frank's WI uses a weight of 2 for intersection density while the other variables have a weight of one. The Graz WI uses equal weights for all variables.
The Urban WI is rather different from both Frank's WI and the Graz WI. This is a data-driven approach, which means that a measure is derived from regressing walking behavior on physical factors of the local environment (Liao et al., 2020). Furthermore, the Urban WI uses centroids as main units of analysis, from where both network and Euclidean buffers are made. In addition, the Urban WI uses factor loadings instead of weights in the WI formula.
While this chapter dealt with the input into the WIs, the next chapter goes on to investigate how the outputs of the three WIs differ for the context of Amsterdam.

4. Results

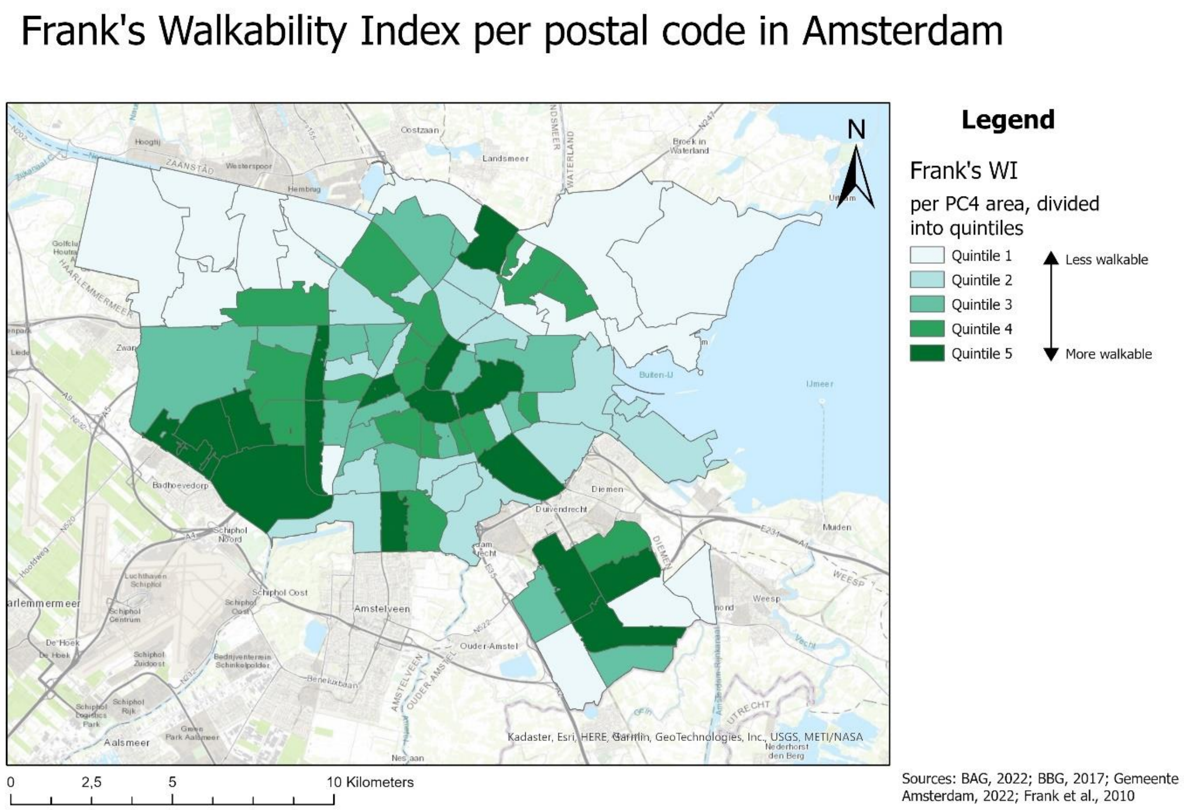
This chapter reports on the most important findings of this study. First of all, the maps that show the walkability in Amsterdam according to the three WIs are presented. Subsequently, the findings of the sensitivity analysis are reported to see how consistent the WIs are. Lastly, the outputs of the regression and correlation analyses are displayed.

4.1 WI maps for Amsterdam

Following the operationalization as described in section 3.3, three maps were produced, one for each WI per PC4 area in Amsterdam. The maps are presented in this section, but possible explanations for spatial patterns are given in the next chapter. All maps that are displayed in this report can be found in full-size in the Appendix.

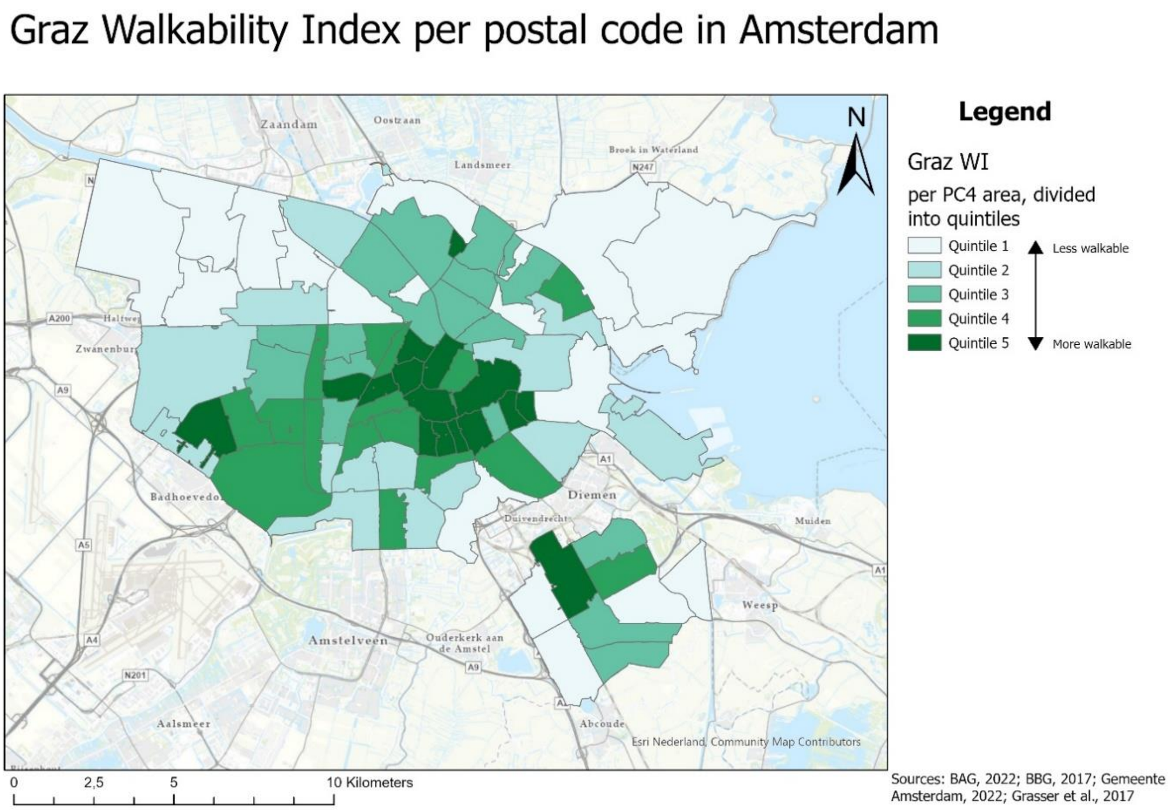
The map below (figure 3) shows Frank's WI per PC4 area in Amsterdam. A darker shade of green indicates a higher walkability in that area. There are large differences between parts of the city regarding the walkability. It can be seen that PC4 areas in the inner city, in the Southwest and in the Southeast of the city have a high walkability.

Figure 3: Map showing Frank's Walkability Index per PC4 area in Amsterdam.



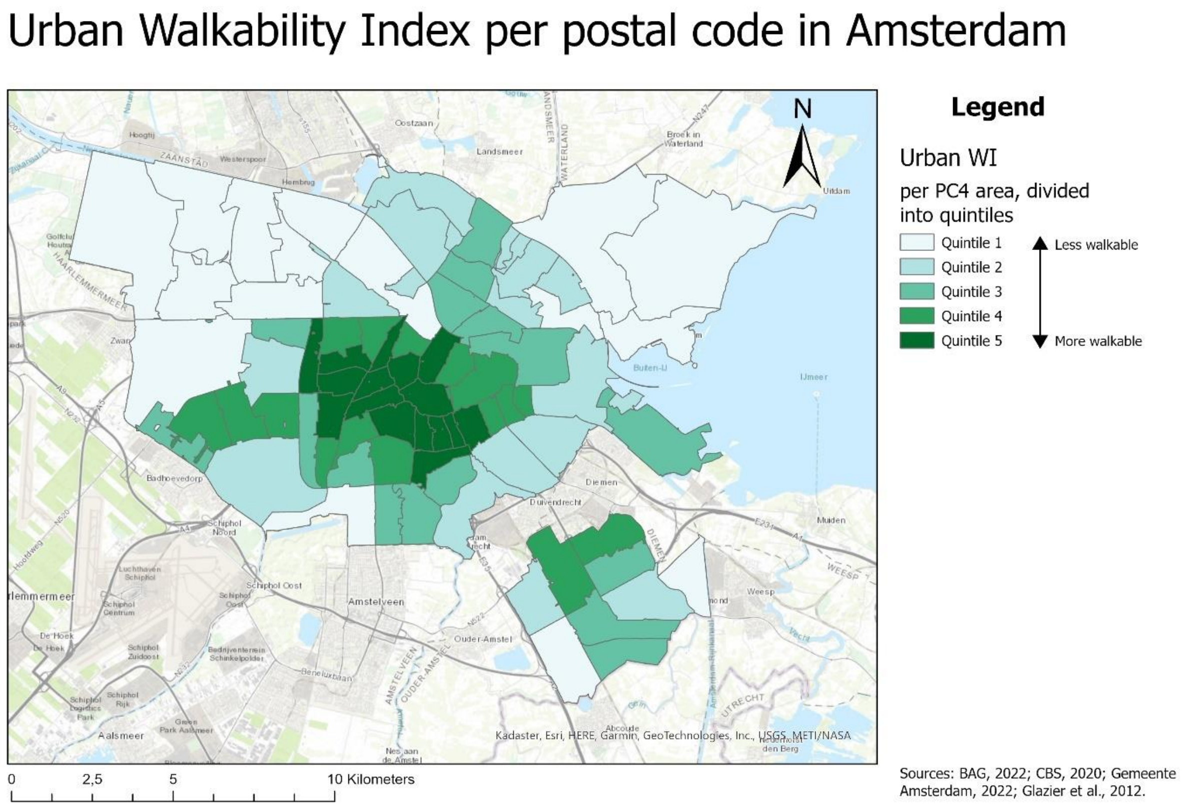
The map below (figure 4) displays the Graz WI per PC4 area in Amsterdam. Again, a darker shade of green indicates a higher walkability in that area. There are large differences between parts of the city regarding the walkability. It can be observed that especially the PC4 areas that are situated in and around the inner city have a high walkability, with some other high-walkable neighborhoods in the West and South of the city.

Figure 4: Map showing the Graz Walkability Index per PC4 area in Amsterdam.



The map below (figure 5) displays the Urban WI per PC4 area in Amsterdam. Again, a darker shade of green indicates a higher walkability in that area. A spatial pattern of walkability can be observed: the closer to the inner city of Amsterdam, the higher the walkability (as measured by the Urban WI). There are some exceptions in the South and West, but nonetheless, the Urban WI shows a rather straightforward pattern. This is elaborated upon in section 5.3.

Figure 5: Map showing the Urban Walkability Index per PC4 area in Amsterdam.



4.2 Sensitivity analysis

As described in section 3.6, the SA consists of two parts. The first SA was performed by increasing the weights of all variables of the WIs by 10%. The second part of the SA was done by using '*buurten'* (neighborhoods, as defined by CBS) as administrative units instead of PC4 areas.

4.2.1 Sensitivity analysis type 1

First of all, every individual weight in each WI was increased by 10% while the other variables kept their original weight. The difference in percentages between the regular WI score and the WI score that was obtained with the weight alteration of one of the variables, was visualized in maps (see below) and in an overview table (table 2). This gives insight into the sensitivity of the WI and the importance of different variables per PC4 area; the larger the change in the WI score in a PC4 area, the more dependent it is on changes in that variable. In the case of a 10% increase, if the impact on results is substantially less than 10%, the parameter is not of critical importance according to Goodchild (2005).

Table 2: Overview of the mean difference and standard deviation in percentage in the WI Z-score of the WI with a 10% increase of each individual variable one by one, compared to the regular WI Z-score.

	Mean	Standard deviation
Frank's WI		
• **Intersection density**	19.82	182.28
• **Residential density**	3.71	68.86
• **Shannon Entropy land use mix**	-52.23	460.34
• **Retail FAR**	38.69	330.90
Graz WI		
• **Intersection density**	0.10	0.53
• **Residential density**	0.06	0.36
• **Shannon Entropy land use mix**	-0.05	0.38
Urban WI		
• **Population density**	-0.008	0.30
• **Residential density**	0.006	0.05
• **Retail and service availability**	0.001	0.04
• **Intersection density**	0.0006	0.36

4.2.1.1 Frank's WI

The SA on Frank's WI is displayed in the four maps below (see figure 6, 7, 8, and 9).
Figure 6 below shows that the weight increase of 10% on intersection density (3 or more legs) resulted in quite some changes. PC4 area 1058 saw its Frank's WI Z-score surge with an increase of over 1500%. Another area that turns out to be sensitive to intersection density weight increase is PC4 area 1051, which saw a decrease in WI Z-score of over 250%. The mean percentage change is 19.82 (see table 2), which indicates that this variable is rather easily influenced by weight changes.

Figure 6: The impact of 10% weight increase of intersection density, measured in percentage change of Frank's WI Z-score.

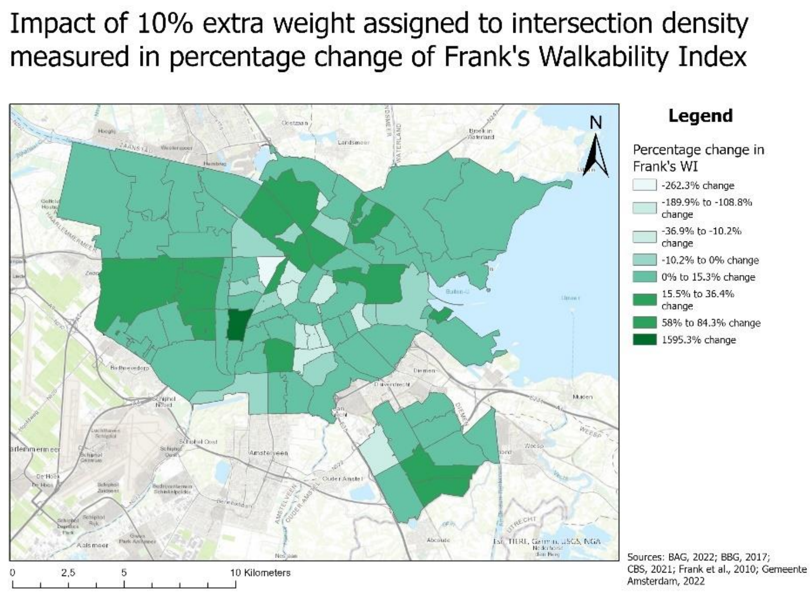


Figure 7 below shows that the impact of a 10% increase in the weight of residential density did not cause many changes. Area 1051 shows a drastic increase in WI Z-score of over 200%, while PC4 area 1058 on the other hand saw a decline in WI Z-score of over -380%. This indicates a high sensitivity of both of these areas to changes in the weight of residential density. The mean percentage change however is 3.71 (see table 2), so this variable is generally robust.

Figure 7: The impact of 10% weight increase of residential density, measured in percentage change of Frank's WI Z-score.

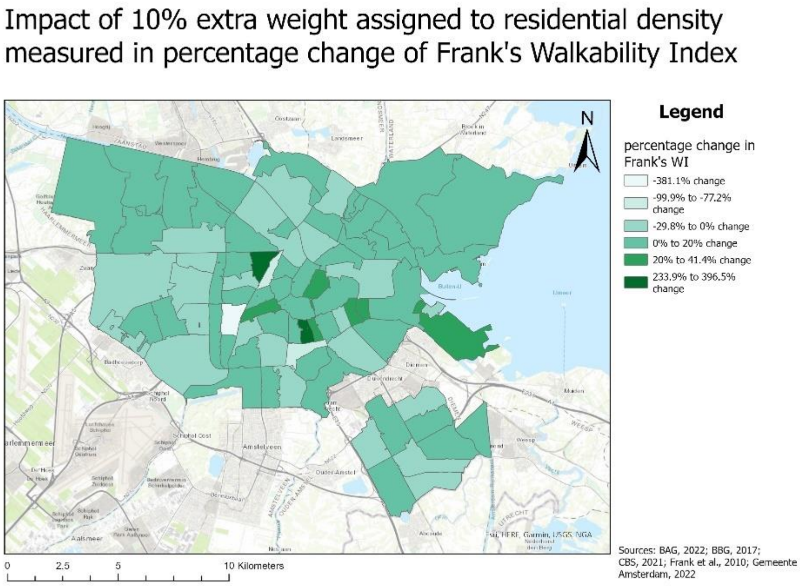


Figure 8 below shows the impact of a 10% increase in the weight of Shannon Entropy land use mix. Generally spoken, this had a rather large, negative effect on Frank’s WI Z-score across Amsterdam. Area 1051 displays a Frank’s WI Z-score that has increased with over 135% compared to the regular Frank’s WI Z-score for that area. PC4 area 1058 shows the largest decrease in WI Z-score of over -4000%. Both of these areas show a high sensitivity to changes in the weight of Shannon Entropy land use mix. The mean percentage change is -52.23 (see table 2)), which also shows that this variable is not robust.

Figure 8: The impact of 10% weight increase of Shannon Entropy land use mix, measured in percentage change of Frank's WI Z-score.

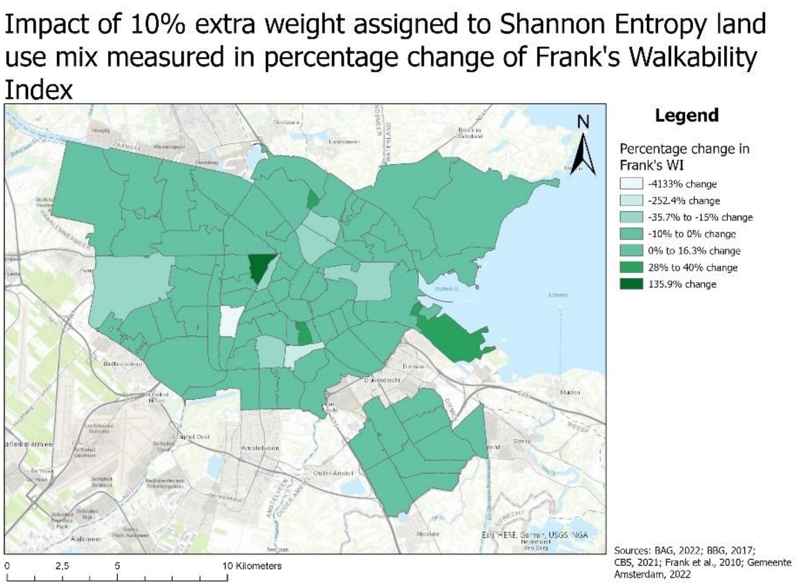
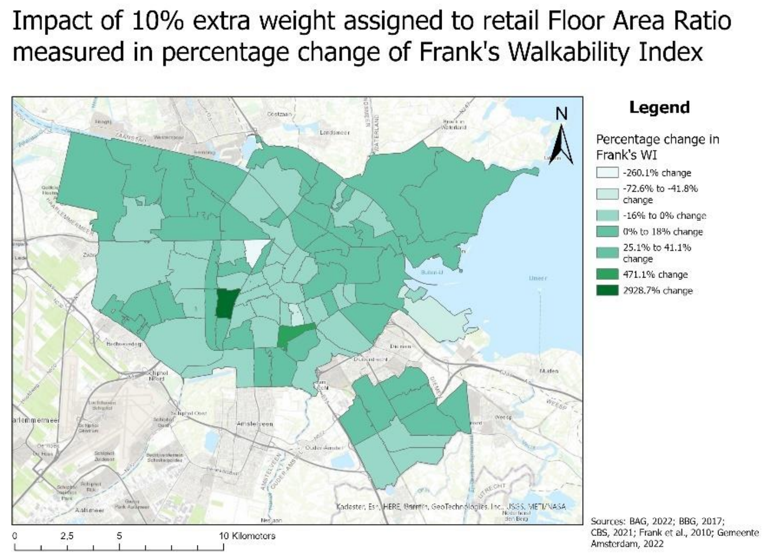


Figure 9 below shows the impact of a 10% increase in the weight of retail FAR. Both PC4 area 1058 and PC4 area 1051 are outliers once again. The former area saw its Frank's WI Z-score increase with over 2900%, while the latter saw a decrease in Frank's WI Z-score of over -260% compared to the regular Frank's WI Z-score. Again, both of these areas show a high sensitivity to changes in the weight of retail FAR. The mean percentage change is 38.69 (see table 2), which also shows that this variable is not robust.

Figure 9: The impact of 10% weight increase of retail FAR, measured in percentage change of Frank's WI Z-score.



4.2.1.2 The Graz WI

The SA on the Graz WI is displayed in the four maps below (figure 10, 11, 12, and 13).

Figure 10 below shows the impact of a 10% increase in the weight of intersection density (4 or more legs). Generally spoken, this weight increase has very little effect on the Graz WI Z-score across Amsterdam. PC4 area 1064 shows the largest increase in Graz WI Z-score of 4.54%. Area 1092 on the other hand displays a Graz WI Z-score that has decreased with -0.89% compared to the regular Graz WI Z-score for that area. Both of these areas show little sensitivity to changes in the weight of intersection density. The mean percentage change across all PC4 areas of Amsterdam is 0.10 (see table 2), which supports that this variable is highly robust.

Figure 10: The impact of 10% weight increase of intersection density, measured in percentage change of the Graz WI Z-score.

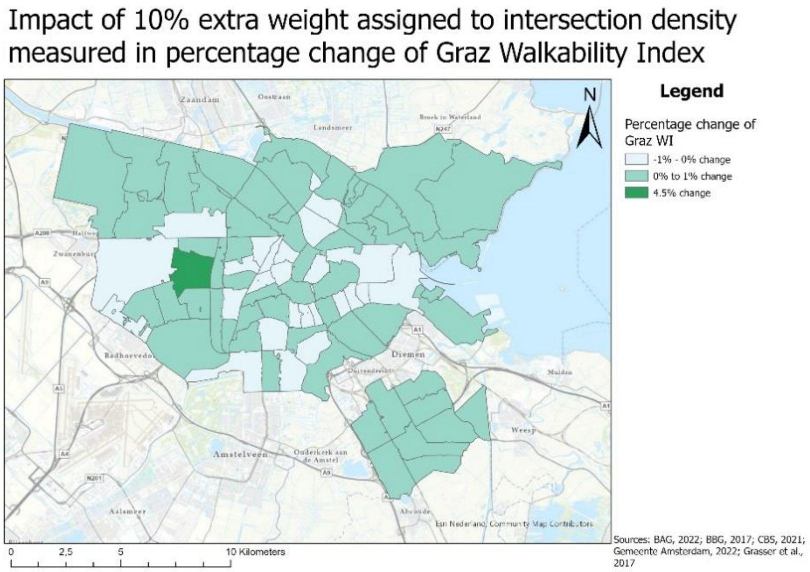


Figure 11 below shows the impact of a 10% increase in the weight of residential density. Generally spoken, this weight increase has very little effect on the Graz WI Z-score across Amsterdam. PC4 area 1056 shows the largest increase in Graz WI Z-score of 2.36%. Area 1064 on the other hand displays a Graz WI Z-score that has decreased with -1.80% compared to the regular Graz WI Z-score for that area. Both of these areas show little sensitivity to changes in the weight of intersection density. The mean percentage change across all PC4 areas of Amsterdam is 0.06 (see table 2), which supports that this variable is highly robust.

Figure 11:The impact of 10% weight increase of residential density, measured in percentage change of the Graz WI Z-score.

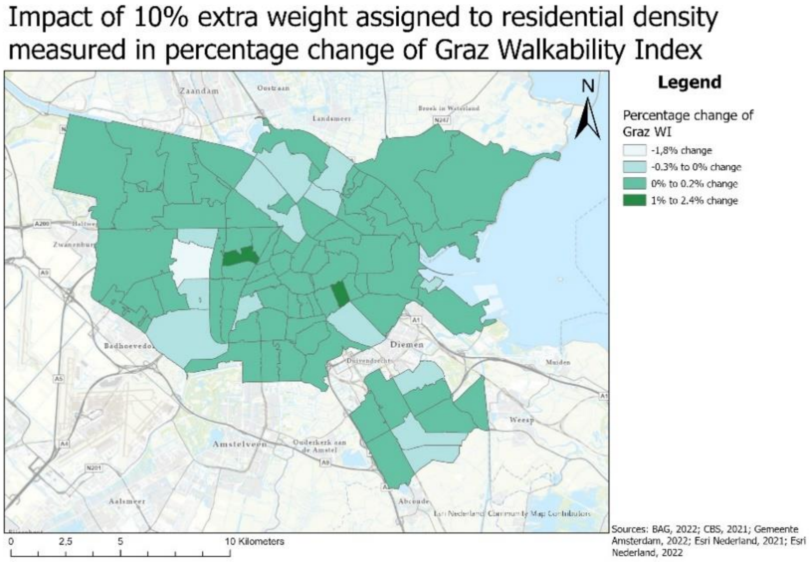
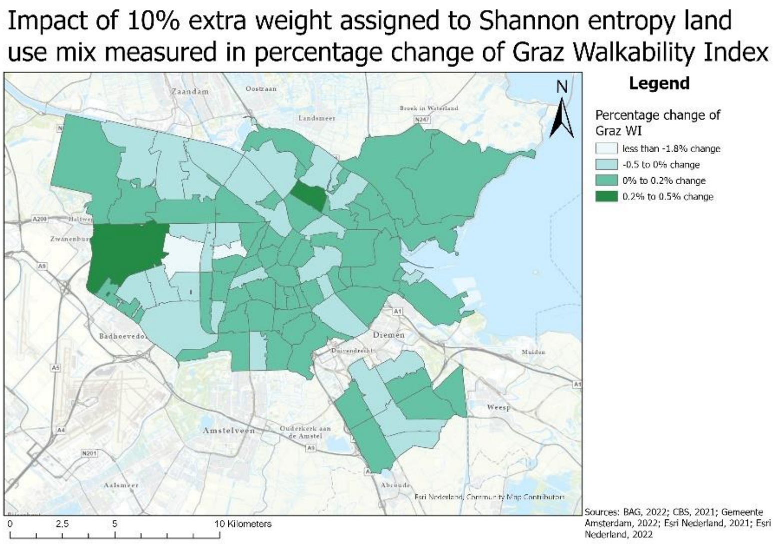


Figure 12 below shows the impact of a 10% increase in the weight of the Shannon Entropy land use mix. Generally spoken, this weight increase has very little effect on the Graz WI Z-score across Amsterdam. PC4 area 1032 shows the largest increase in Graz WI Z-score of 0.40%. Area 1064 is the other extreme and displays a Graz WI Z-score that has decreased with -2.64% compared to the regular Graz WI Z-score for that area. It can therefore be stated that both of these areas show little sensitivity to changes in the weight of the Shannon Entropy land use mix. The mean percentage change across all PC4 areas of Amsterdam is -0.05 (see table 2), which supports that this variable is highly robust.

Figure 12: The impact of 10% weight increase of Shannon Entropy land use mix, measured in percentage change of the Graz WI Z-score.

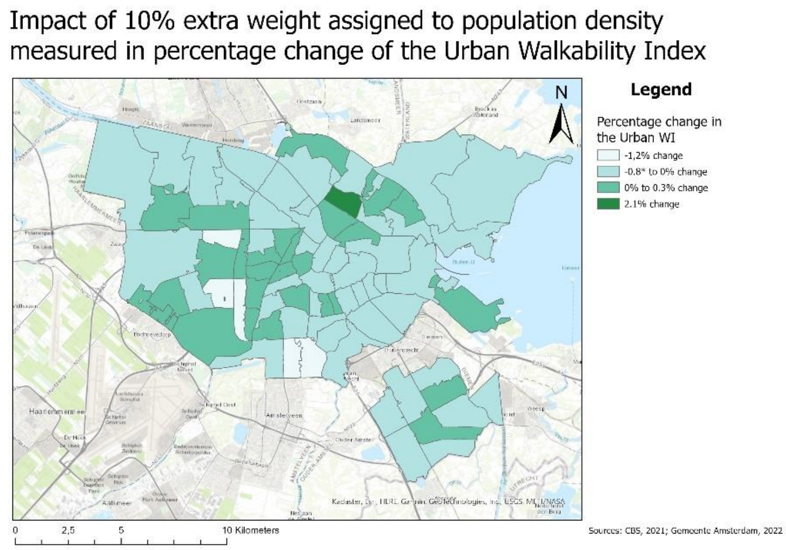


4.2.1.3 The Urban WI

The SA on the Graz WI is displayed in the four maps below (figure 13, 14, 15, and 16).
Figure 13 below shows the impact of a 10% increase in the weight of population density. Generally spoken, this weight increase has little effect on the Urban WI score across Amsterdam. PC4 area 1032

shows the largest increase in the Urban WI Z-score of 2.09%. Area 1083 is the other extreme and displays an Urban WI Z-score that has decreased with -1.24% compared to the regular Urban WI Z-score for that area. Even the extremes show little percentage change, and the mean percentage change across all PC4 areas of Amsterdam is -0.008 (see table 2), which displays that this variable is highly robust.

Figure 13: The impact of 10% weight increase of population density, measured in percentage change of the Urban WI Z-score.



The map below (figure 14) shows the impact of a 10% increase in the weight of residential density. Generally spoken, this weight increase has little effect on the Urban WI Z-score across Amsterdam. PC4 area 1032 shows the largest increase in the Urban WI Z-score of 0.29%. Area 1104 is the other extreme and displays an Urban WI Z-score that has decreased with -0.08% compared to the regular Urban WI Z-score for that area. Even the extremes show very little percentage change, and the mean percentage change across all PC4 areas of Amsterdam is 0.006 (see table 2), which displays that this variable is highly robust.

Figure 14: The impact of 10% weight increase of residential density, measured in percentage change of the Urban WI Z-score.

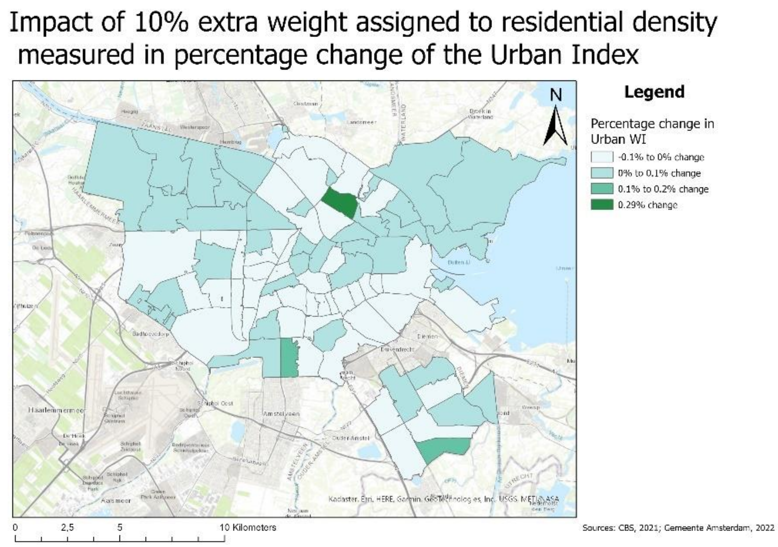
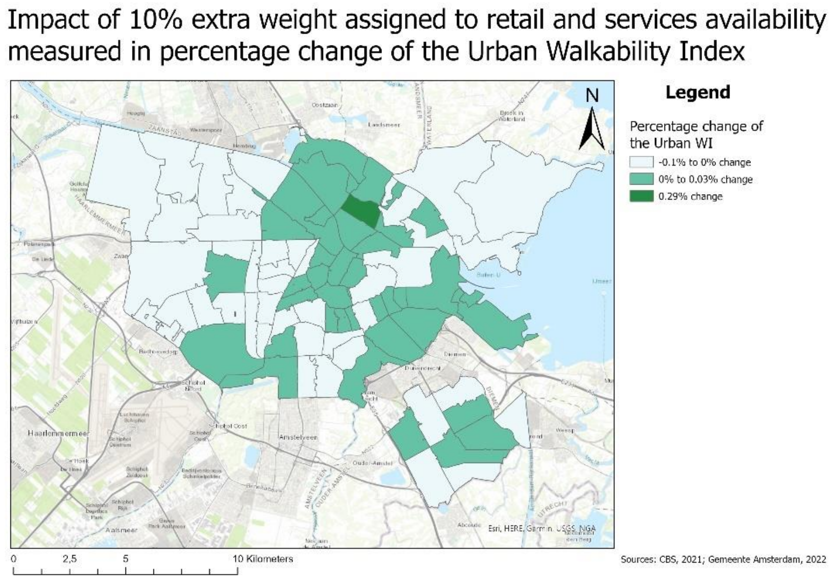


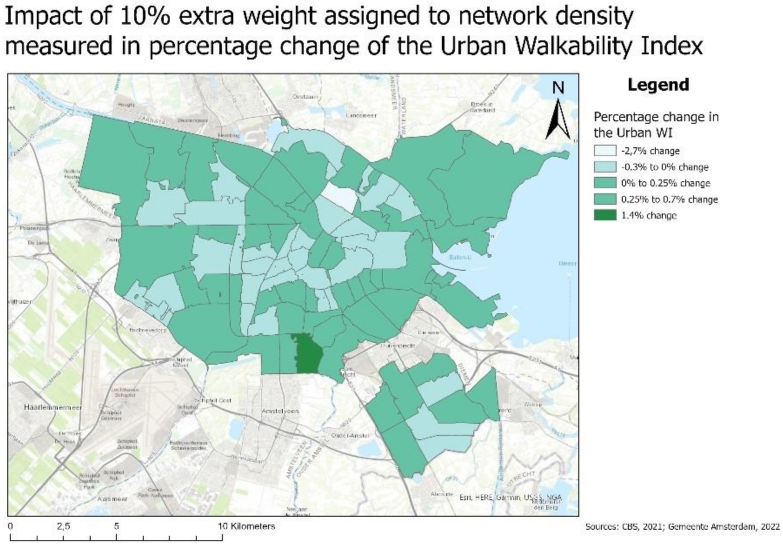
Figure 15 below shows the impact of a 10% increase in the weight of retail and services availability. Generally spoken, this weight increase has very little effect on the Urban WI Z-score across Amsterdam. PC4 area 1032 shows the largest increase in the Urban WI Z-score of 0.29%. Area 1083 is the other extreme and displays an Urban WI Z-score that has decreased with -0.11% compared to the regular Urban WI Z-score for that area. Even both extremes are minimally sensitive to changes in the weight of the retail and services availability. The mean percentage change across all PC4 areas of Amsterdam is small as well; 0.001 (see table 2), which emphasizes that the retail and services availability variable is highly robust.

Figure 15: The impact of 10% weight increase of retail and services availability, measured in percentage change of the Urban WI Z-score.



The map below (figure 16) shows the impact of a 10% increase in the weight of intersection density. Generally spoken, this weight increase has very little effect on the Urban WI Z-score across Amsterdam. PC4 area 1083 shows the largest increase in the Urban WI Z-score, which went up by 1.43%. Area 1032 on the other hand displays an Urban WI Z-score that showed the largest decrease, its Urban WI Z-score went down by –2.72% compared to the regular Urban WI Z-score for that area. Even the extremes show very little percentage change, and the mean percentage change across all PC4 areas of Amsterdam is 0.0006 (see table 2), which displays that this variable is highly robust.

Figure 16: The impact of 10% weight increase of intersection density, measured in percentage change of the Urban WI Z-score.



4.2.2 Sensitivity analysis type 2

This part of the SA was carried out by using '*buurten'* (neighborhoods, as defined by CBS) as administrative units instead of PC4 areas. The steps as described in the operationalization (section 3.3) were exactly repeated, but on neighborhood level instead of on PC4 level to be able to account for differences on a smaller scale that are not visible on PC4 area level.

4.2.2.1 Frank's WI

The four maps below (figure 18, 19, 20, and 21) show the SA performed on the four variables in Frank's WI.

Figure 17 shows the three (or more) way intersection density per neighborhood in Amsterdam. A vague pattern can be detected of a higher intersection density on the outskirts of Amsterdam. The inner city and the neighborhoods in the Northeast on the other hand have a low Z-score.

Figure 17: Map showing three (or more) way intersection density, which is one of the variables in Frank's WI.

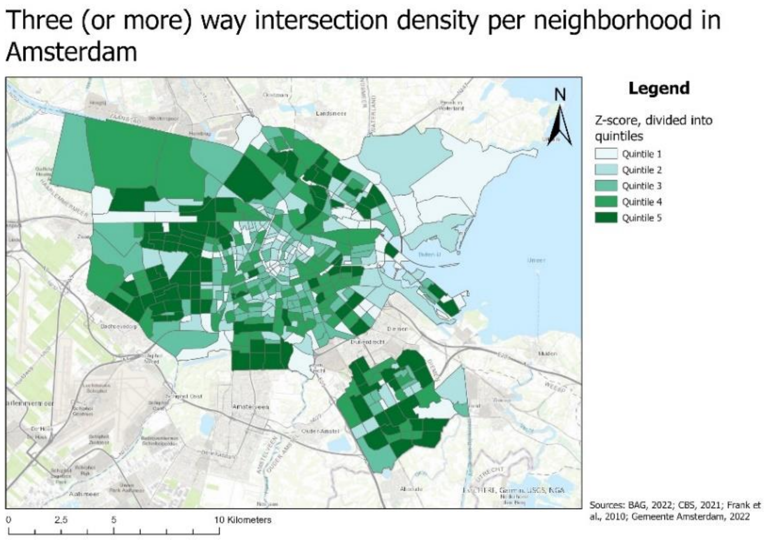


Figure 18 below shows the residential density per neighborhood in Amsterdam. It is clearly visible that the inner city has a high residential density Z-score compared to the rest of Amsterdam.

Figure 18: Map showing the SA performed on residential density, which is one of the variables in Frank's WI.

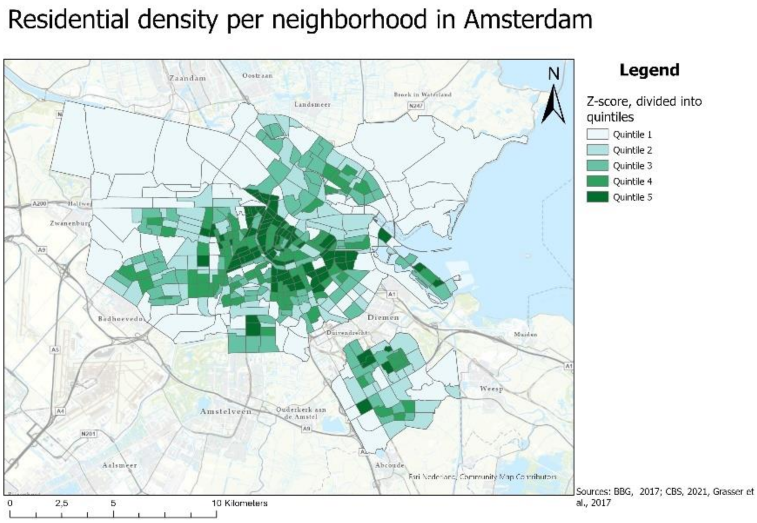


Figure 19 shows the retail FAR per neighborhood in Amsterdam. The pattern seems dispersed and scattered throughout the city, though neighborhoods in the Northeast clearly score low compared to the rest of Amsterdam.

Figure 19: Map showing the SA performed on retail FAR, which is one of the four variables in Frank's WI.

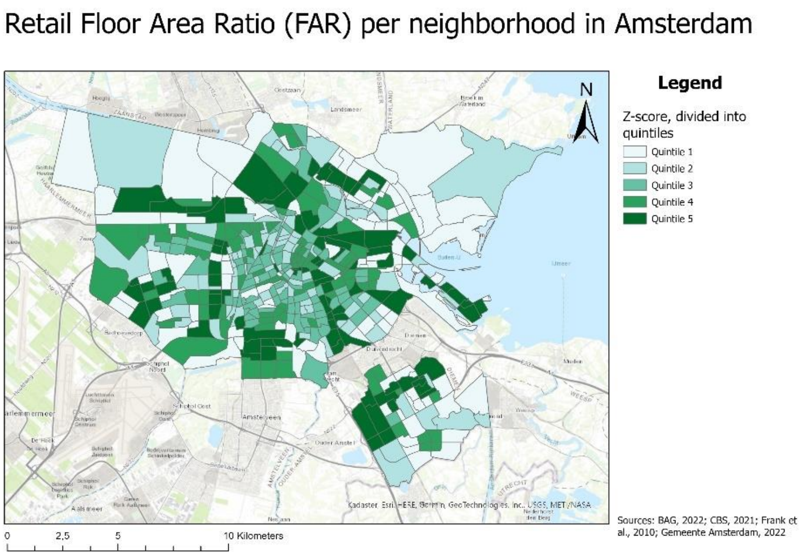
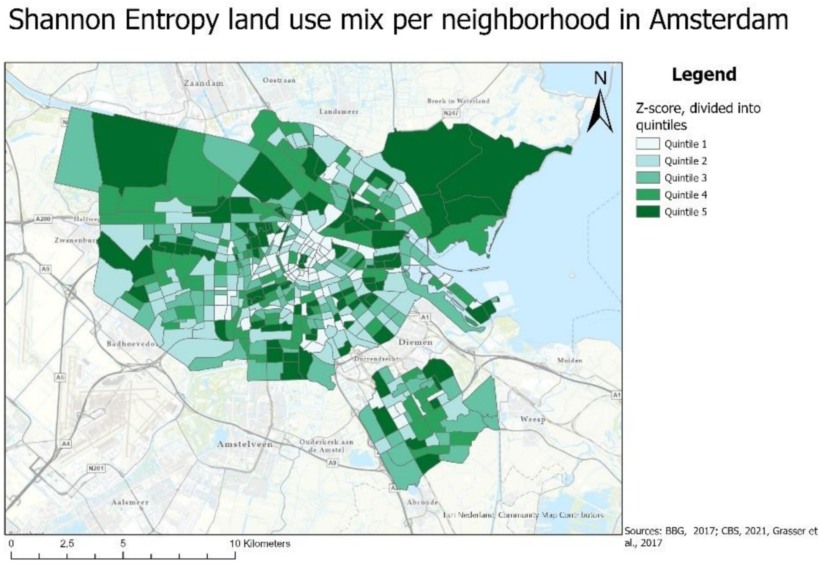


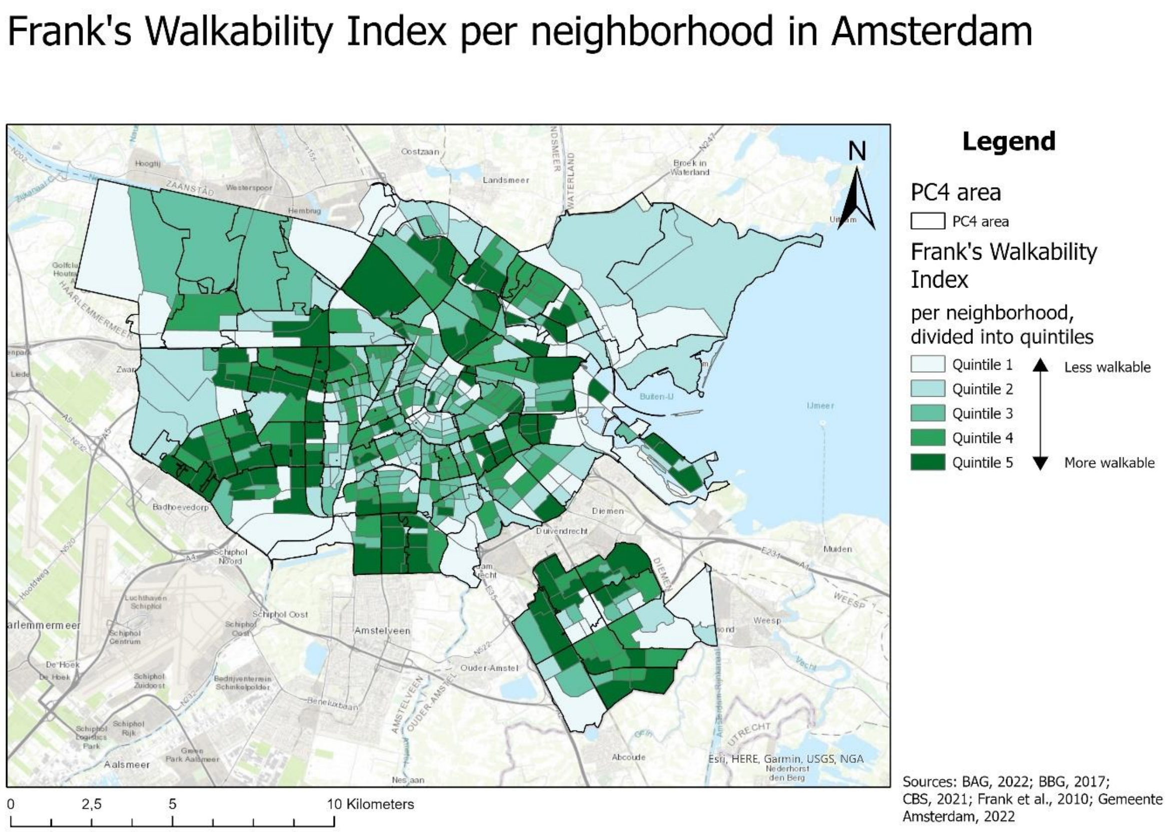
Figure 20 shows the Shannon Entropy land use mix per neighborhood in Amsterdam. It seems that the inner city is generally characterized by a low land use mix Z-score, while the outskirts of the city score highest.

Figure 20: Map showing the SA performed on the Shannon Entropy land use mix, which is one of the four variables in the Frank WI.



Below, in figure 21, the total Frank's Walkability Index per neighborhood in Amsterdam is displayed. A PC4 area outline is placed on top of the neighborhood layer, which gives an indication as of how more detailed the neighborhood level is. The general pattern seems to be that walkability is highest in a circle outside of and around the inner city center. It does offer a lot more detail and nuance compared to Frank's Walkability Index per PC4 area (see figure 3).

Figure 21: Frank's WI score per neighborhood, divided into quintiles.



4.2.2.2 The Graz WI

Figures 22, 23, 24 and 25 show the SA performed on the three variables in the Graz WI.
Below, figure 22 shows the four (or more) way intersection density per neighborhood in Amsterdam. It appears that the outskirts of the city score higher than the inner city. The pattern is very similar to the one that was displayed in figure 17, showing the three (or more) way intersection density.

Figure 22: Map showing the SA performed on intersection density, which is one of the three variables in the Graz WI.

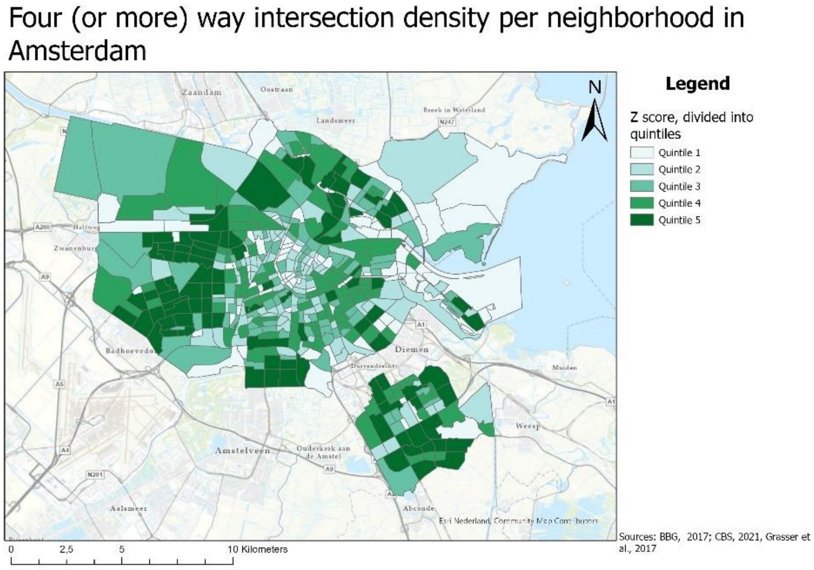
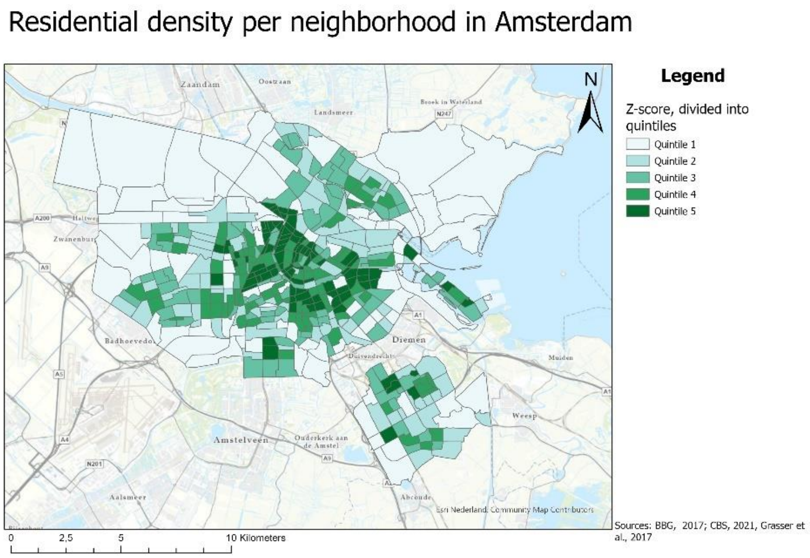


Figure 23 below shows the residential density per neighborhood in Amsterdam. This map is the exact same as the one shown in figure 18, since residential density is a common variable in Frank's WI and the Graz WI. The inner city clearly has a high residential density Z-score compared to the rest of Amsterdam.

Figure 23: Map showing the SA performed on residential density, which is one of the three variables in the Graz WI.



The map below (figure 24) is again the same as was shown for this variable in Frank's WI. The low entropy in the inner city is visible, compared to the higher values on the Northeastern and Northwestern edges of the city.

Figure 24: Map showing the SA performed on the Shannon Entropy land use mix, which is one of the three variables in the Graz WI.

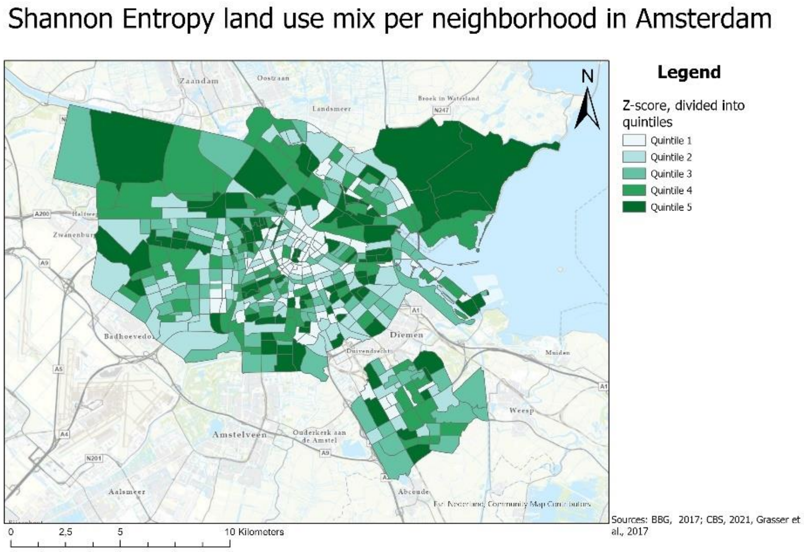
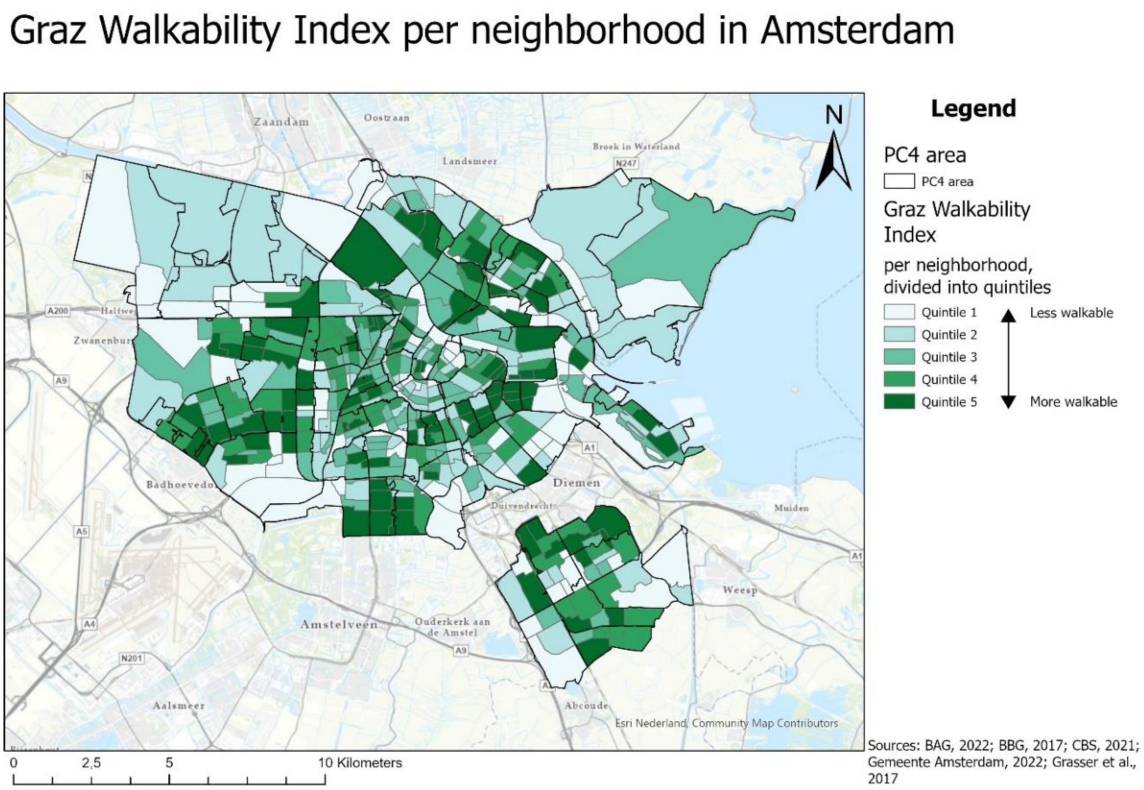


Figure 25 shows the total Graz Walkability Index per neighborhood in Amsterdam. Again, the PC4 area outline is placed on top as a way to compare spatial unit size. The pattern is rather similar to the one that Frank's WI displayed, however the pattern has shifted towards the inner city a bit more.

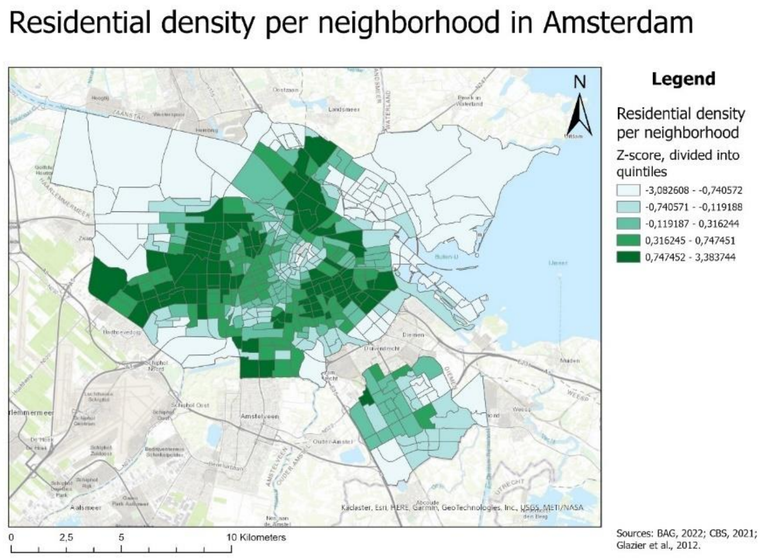
Figure 25: The Graz WI score per neighborhood, divided into quintiles.



4.2.2.3 The Urban WI

Figure 26 and figure 27 both show the same variable per neighborhood; residential density. However, the map in figure 26 was retrieved by following the operationalization as originally proposed by Glazier et al (2012). The Z-score was retrieved from the sum of address (residences) values in all neighborhoods whose polygon area intersected with the 720 meter Euclidean buffer of a given neighborhood centroid, divided by the summed area of all intersecting neighborhoods (Glazier et al., 2012). This aims to estimate the amount of residences within the 720 meter Euclidean buffer zone. However, this method is prone to the modifiable areal unit problem (MAUP). This occurs when arbitrarily defined boundaries are used for the reporting of spatial phenomena (Heywood, Cornelius, & Carver, 2011). Population density is an example of this as it is often reported for areas; it would be wrong to assume the population density is equally spread out through the area.

Figure 26: Map showing the SA performed on the residential density, which is one of the four variables in the Urban WI.



The operationalization made by Glazier et al. (2012) is especially useful when it is not known how the phenomenon is spread out throughout the area, as is the case with population density. For residential density however, the exact geographic location of the units (addresses) is known. Therefore, the residential density can be calculated more precisely using a *spatial join* of the 720 meter Euclidean buffer around the neighborhood centroids and the addresses. This shows the exact amount of addresses in the buffer, and is therefore more precise than the estimation that can be seen in figure 26. The result of this variation on the SA can be seen in figure 27. When comparing figures 26 and 27, the spatial patterns are rather different. Figure 27 shows a much clearer pattern of a high residential density near the inner city center.

Figure 27: Map showing the SA performed on the residential density (in the neighborhood centroid buffer), which is one of the four variables in the Urban WI.

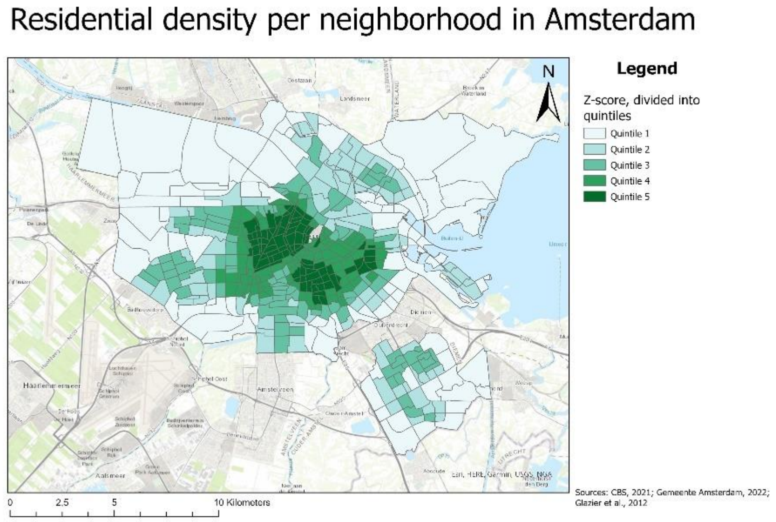


Figure 28 below shows the population density per neighborhood in Amsterdam. A similar pattern to that of residential density is visible, the inner city clearly has the highest population density Z-scores out of all neighborhoods in Amsterdam.

Figure 28: Map showing the SA performed on population density, which is one of the four variables in the Urban WI.

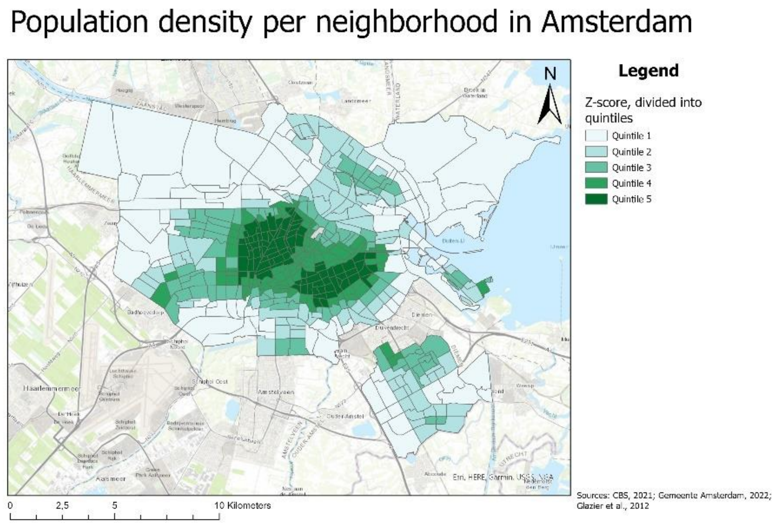


Figure 29 below shows the intersection density per neighborhood in Amsterdam. This shows a rather straightforward pattern of high intersection density Z-scores is found in and just around the inner city center.

Figure 29: Map showing the SA performed on intersection density, which is one of the four variables in the Urban WI.

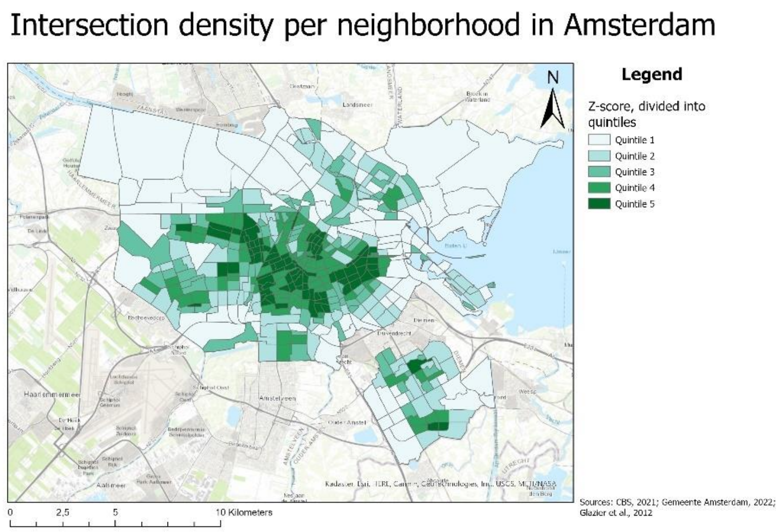
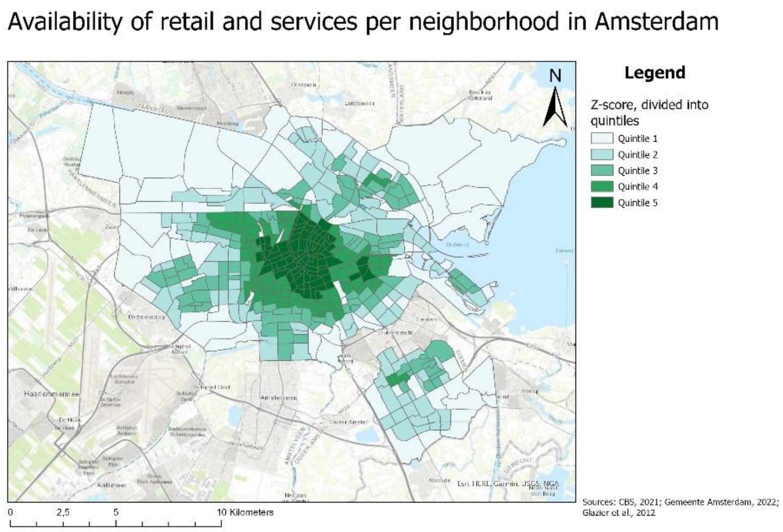


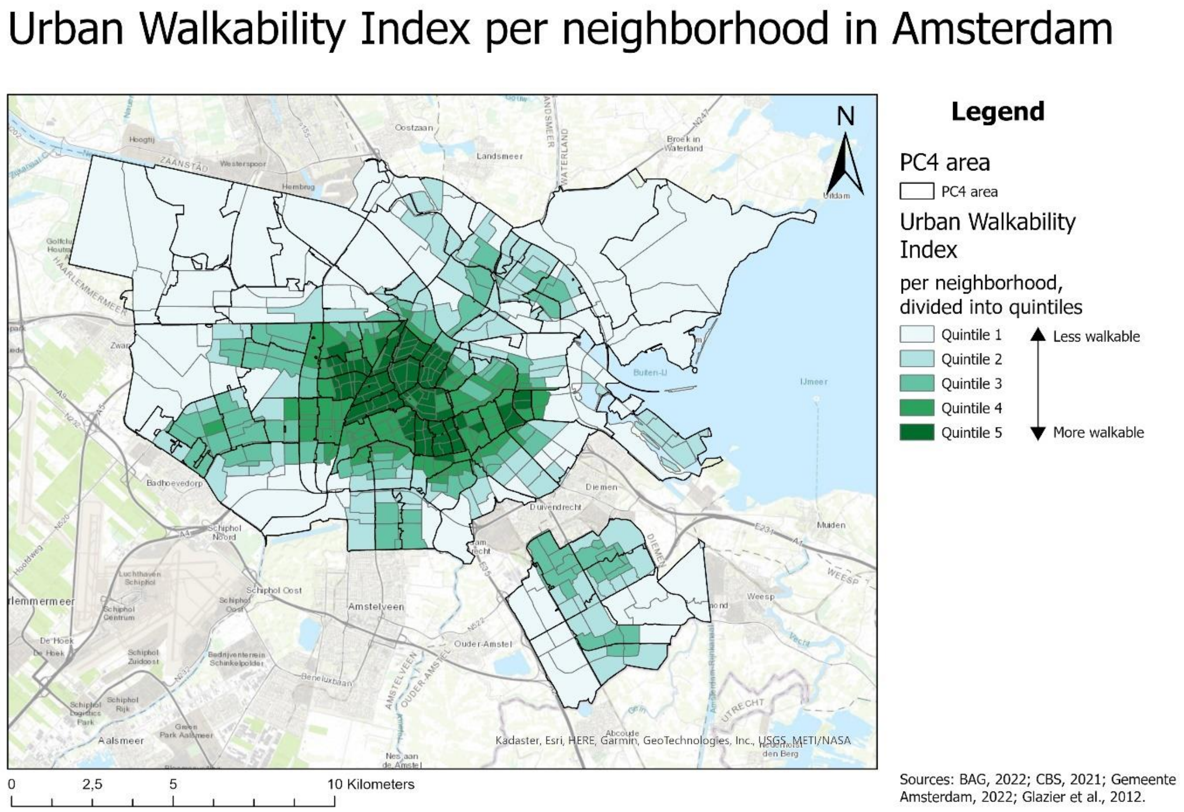
Figure 30 below shows the availability of retail and services per neighborhood in Amsterdam. Again, the spatial clustering of neighborhoods with high Z-scores is found in the inner city. The density of retail and service availability seems to fade away from the inner city towards the outskirts of the city.

Figure 30: Map showing the SA performed on the availability of retail and services, which is one of the four variables in the Urban WI.



The Urban Walkability Index per neighborhood in Amsterdam, as shown in figure 31, displays a clear spatial pattern: there appears to be a concentration of more walkable neighborhoods in the city center. The further the neighborhood is from the city center, the lower the walkability. The PC4 outline again shows the detail that neighborhoods supply compared to the larger PC4 areas.

Figure 31: The Urban WI score per neighborhood, divided into quintiles.



4.3 Regression analysis

To be able to see which model best estimates walking behavior in Amsterdam, a multiple linear regression analysis is performed. As mentioned before, the walking data that functions as a ground truth is retrieved from the ODiN dataset and 13 PC4 areas were left out of the analysis due to no ODiN data (see section 3.5). Thus, instead of the total of 81 PC4 areas that Amsterdam has, the regression analysis was performed with the remaining 68 PC4 areas.

4.3.1 Frank's WI

First of all, the Z-scores of all four variables of Frank's WI (Frank et al., 2010) were selected as the independent variables (3 (or more) way intersection density, FAR, residential density, and Shannon Entropy LUM), with the dependent variable being the mean walking distance per PC4 area in ha.
This resulted in the following tables.

Table 3: Model Summary table of the regression analysis on Frank's WI.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Frank WI	0.540	0.292	0.247	0.10223657414

The model summary above shows that there is a strong correlation between the independent variables and the dependent variable, since R is between 0.5 and 0.7 (De Vocht, 2019). The R Square indicates that 29.2% of the variance of the dependent variable is explained by the four variables in Frank's WI.

The Adjusted R Square is the adjusted share of explained variance for the population, based on sample data, and is therefore a better indicator (De Vocht, 2019). The Adjusted R Square is 24.7%, meaning 24.7% of the variance in mean walking distance is explained by the variables in Frank's WI.

Table 4: ANOVA table of the regression analysis on Frank's WI.

Model	**Sum of Squares**	**Df**	**Mean Square**	**F**	**Sig.**
Regression	0.271	4	0.068	6.481	<0.001
Residual	0.658	63	0.010		
Total	0.929	67			

The ANOVA table above contains a variance analysis, which tests the significance of the entire model (De Vocht, 2019). The row called 'Regression' shows the explained variance, and the 'Residual' row shows the unexplained variance. The multiple linear regression with the mean walking distance in Amsterdam per PC4 area as the dependent variable and the Z-scores of the variables used in Frank's WI model as independent variables, is significant, since $F(4.63) = 6.481$, $p<0.001$.

Table 5: Coefficients table of the regression analysis on Frank's WI.

Model	**Unstandardized B**	**Coefficients Std. Error**	**Standardized Coefficients Beta (b)**	**t**	**Sig.**
(Intercept)	0.145	0.014		10.521	0.001
3 (or more) way intersections	-0.042	0.016	-0.322	-2.686	0.009
FAR	-0.006	0.013	-0.050	-0.458	0.649
Shannon Entropy LUM	0.005	0.018	0.038	0.255	0.799
Residential density	0.038	0.019	0.289	2.024	0.047

The coefficients table above shows the actual multiple linear regression. It becomes clear that only the 3 (or more) way intersections and the residential density are significant predictors of mean walking distance per PC4 area. Three (or more) way intersections is a significant independent variable, since $b = -0{,}322$, $t(67) = -2.686$, $p<0.05$. The b-value is negative, indicating that an increase in 3 (or more) way intersections results in a decrease of the mean walking distance per PC4.
Residential density is another significant independent variable, since $b = 0.289$, $t(67) = 2.024$, $p<0.05$. The b-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. The other two independent variables, FAR and Shannon Entropy Land Use Mix, are not significant since for both, $p>0.05$.

4.3.2 The Graz WI

For the Graz WI, the Z-scores of 4 (or more) way intersection density, residential density and Shannon Entropy Land Use Mix were selected as the independent variables, with again the dependent variable being the mean walking distance per PC4 in ha. The multiple linear regression had the following outputs.

Table 6: Model Summary table of the regression analysis on the Graz WI.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Graz WI	0.544	0.296	0.263	0.10110622900

The model summary above shows that there is a strong correlation between the independent variables and the dependent variable, since R is between 0.5 and 0.7 (De Vocht, 2019). The R Square indicates that 29.6% of the variance of the dependent variable is explained by the three variables in the Graz WI. The Adjusted R Square is the adjusted share of explained variance for the population, based on sample data, and is therefore a better indicator (De Vocht, 2019). The Adjusted R Square is 26.3%, meaning 26.3 % of the variance in mean walking distance is explained by the three variables in the Graz WI.

Table 7: ANOVA table of the regression analysis on the Graz WI.

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	0.275	3	0.092	8.974	<0.001
Residual	0.654	64	0.010		
Total	0.929	67			

The ANOVA table above contains a variance analysis, which tests the significance of the entire model (De Vocht, 2019). The multiple linear regression with the mean walking distance in Amsterdam per PC4 area as the dependent variable and the Z-scores of the variables used in the Graz WI model as independent variables, is significant, since $F(3.64) = 8.974$, $p<0.001$.

Table 8: Coefficients table of the regression analysis on the Graz WI.

Model	Unstandardized B	Coefficients Std. Error	Standardized Coefficients Beta (b)	t	Sig.
(Intercept)	0.144	0.014		10.694	<0.001
4 (or more) way intersections	-0.043	0.015	-0.334	-2.826	0.006
Shannon Entropy LUM	0.006	0.018	0.049	0.330	0.742
Residential density	0.037	0.018	0.287	2.049	0.045

The coefficients table above shows the actual multiple linear regression. It becomes clear that the four (or more) way intersections and the residential density are significant predictors of mean walking distance per PC4 area. Four (or more) way intersections is a significant independent variable, since b = -0,334, $t(67) = -2.826$, $p<0.05$. The b-value is negative, indicating that an increase in four (or more) way intersections results in a decrease of the mean walking distance per PC4.
Residential density is another significant independent variable, since $b = 0.287$, $t(67) = 2.049$, $p<0.05$. The b-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. The other independent variable, Shannon Entropy Land Use Mix, was not found to be significant, since $p>0.05$.

4.3.3 The Urban WI

For the Urban WI, the Z-scores of residential density, population density, intersection availability, and retail and services availability were selected as the independent variables, with again the dependent variable being the mean walking distance per PC4 in ha. The multiple linear regression had the following outputs.

Table 9: Model Summary table of the regression analysis on the Urban WI.

Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
Urban WI	0.582	0.338	0.296	0.09879904988

The model summary above shows that there is a strong correlation between the independent variables and the dependent variable, since R is between 0.5 and 0.7 (De Vocht, 2019). The R Square indicates that 33.8% of the variance of the dependent variable is explained by the four variables in the Urban WI. The Adjusted R Square is the adjusted share of explained variance for the population, based on sample data, and is therefore a better indicator (De Vocht, 2019). The Adjusted R Square is 29.6%, meaning 29.6% of the variance in mean walking distance is explained by the four variables in the Urban WI.

Table 10: ANOVA table of the regression analysis on the Urban WI.

Model	**Sum of Squares**	**Df**	**Mean Square**	**F**	**Sig.**
Regression	0.314	4	0.079	8.055	<0.001
Residual	0.615	63	0.010		
Total	0.929	67			

The ANOVA table above contains a variance analysis, which tests the significance of the entire model (De Vocht, 2019). The multiple linear regression with the mean walking distance in Amsterdam per PC4 area as the dependent variable and the Z-scores of the variables used in the Urban WI model as independent variables, is significant, since $F(4.63) = 8.055$, $p<0.001$.

Table 11: Coefficients table of the regression analysis on the Urban WI.

Model	**Unstandardized B**	**Coefficients Std. Error**	**Standardized Coefficients Beta (b)**	**t**	**Sig.**
(Intercept)	0.142	0.013		10.625	<0.001
Residential density per km2	0.066	0.021	0.506	3.211	0.002
Population density per km2	-0.005	0.017	-0.038	-0.286	0.776
Intersection availability	-0.045	0.014	-0.347	-3.152	0.002
Retail and services availability	-0.029	0.015	-0.265	-2.001	0.050

The coefficients table above shows the actual multiple linear regression. It becomes clear that residential density, intersection availability and retail and services availability are significant predictors of mean walking distance per PC4 area.
Residential density is a significant independent variable, since $b = 0.506$, $t(67) = 3.211$, $p<0.05$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4.
Intersection availability is the second independent variable that is significant, since $b = -0.347$, $t(67) = -3.152$, $p<0.05$. The *b*-value is negative, meaning that an increase in intersections results in a decrease of the mean walking distance per PC4.
Retail and services availability is the third independent variable that is significant, since $b = -0.265$, $t(67) = -2.001$, $p=0.050$. The *b*-value is negative, meaning that an increase in retail and services results in a decrease of the mean walking distance per PC4.
The remaining independent variable, population density, was not found to be significant, since $p>0.05$.

4.4 Construction of a new WI for Amsterdam

With the independent variables in the three WIs, a new WI was constructed that fits Amsterdam as best as possible using the same dependent variable as was used for the other three WIs. This was done by performing regression analyses with the variables that were significant and assessing which combination of significant variables resulted in the highest Adjusted R Square. The Adjusted R Square is the adjusted share of explained variance for the population, based on sample data (De Vocht, 2019). The highest Adjusted R Square was achieved when residential density, 4 (or more) way intersections, and availability of retail and services were selected as independent variables. The first two variables stem from the Graz WI, while the latter comes from the Urban WI. These three variables together form the Amsterdam Walkability Index, which is shown below (figure 32).

Figure 32: The newly constructed Amsterdam Walkability Index per PC4 area in Amsterdam.

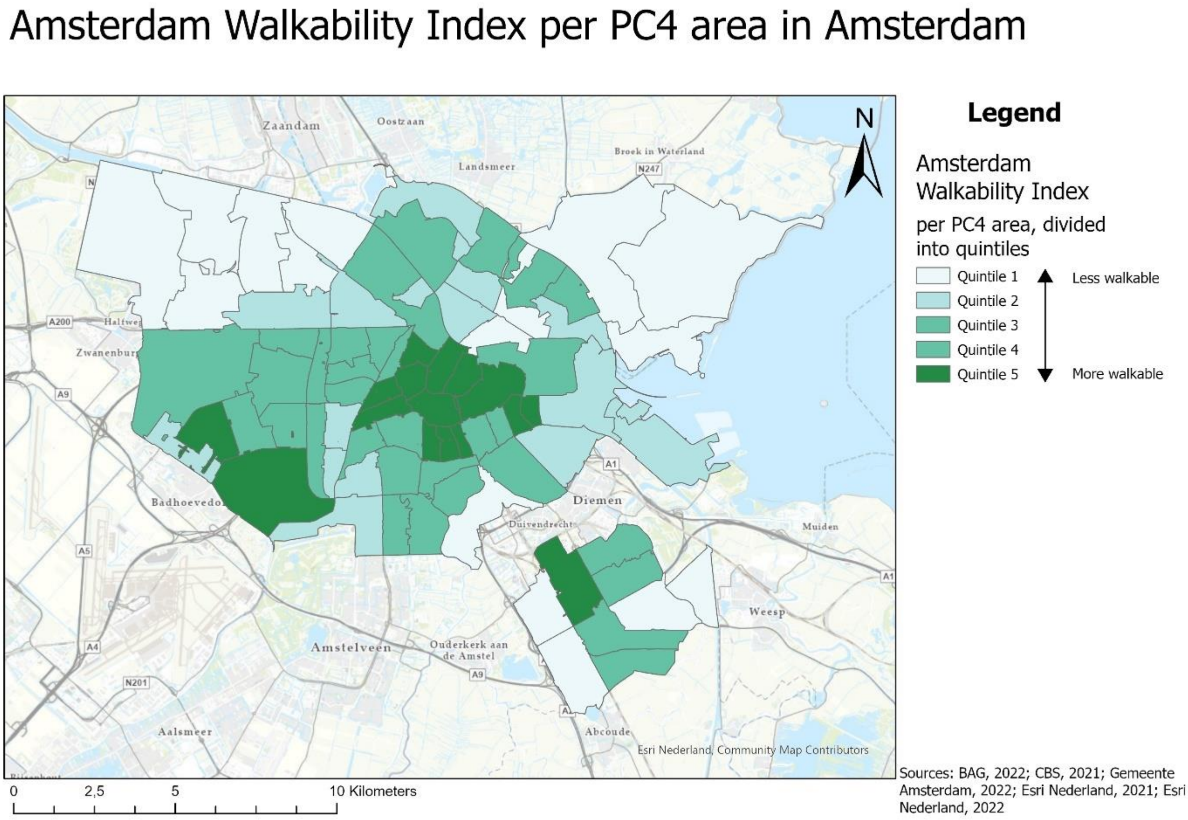


Table 12: Model Summary table of the regression analysis on the Amsterdam WI.

Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
Amsterdam WI	0.581	0.338	0.306	0.09808793367

The model summary above shows that there is a strong correlation between the independent variables and the dependent variable, since R is between 0.5 and 0.7 (De Vocht, 2019). The R Square indicates that 33.8% of the variance of the dependent variable is explained by the three variables in the Amsterdam WI. The Adjusted R Square is the adjusted share of explained variance for the population, based on sample data, and is therefore a better indicator (De Vocht, 2019). The Adjusted R Square is 30.6%, meaning 30.6% of the variance in mean walking distance per PC4 is explained by the three variables in the Amsterdam WI.

Table 13: ANOVA table of the regression analysis on the Amsterdam WI.

Model	**Sum of Squares**	**Df**	**Mean Square**	**F**	**Sig.**
Regression	0.314	3	0.105	10.868	<0.001
Residual	0.616	64	0.010		
Total	0.929	67			

The ANOVA table above contains a variance analysis, which tests the significance of the entire model (De Vocht, 2019). The multiple linear regression with the mean walking distance in Amsterdam per PC4 area as the dependent variable and the Z-scores of the variables used in the Amsterdam WI model as independent variables, is significant, since $F(4.63) = 8.055$, $p<0.001$.

Table 14: Coefficients table of the regression analysis on the Amsterdam WI.

Model	Unstandardized B	Coefficients Std. Error	Standardized Coefficients Beta (b)	t	Sig.
(Intercept)	0.056	0.031		1.791	0.078
Residential density per hectare	0.003	0.001	0.484	3.530	<0.001
Retail and services availability	-0.029	0.015	-0.267	-2.029	0.047
Intersection density 4 (or more) legs	-0.045	0.014	-0.344	-3.163	0.002

The coefficients table above shows the actual multiple linear regression. It becomes clear that residential density, intersection availability and retail and services availability are all significant predictors of mean walking distance per PC4 area.
Residential density is a significant independent variable, since $b = 0.484$, $t(67) = 3.530$, $p<0.001$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4.
Retail and services availability is the second independent variable that is significant, since $b = -0.267$, $t(67) = -2.029$, $p<0.05$. The *b*-value is negative, meaning that an increase in intersections results in a decrease of the mean walking distance per PC4.
Intersection density (intersections with 4 or more legs) is the third independent variable that is significant, since $b = -0.344$, $t(67) = -3.163$, $p<0.05$. The *b*-value is negative, meaning that an increase in retail and services results in a decrease of the mean walking distance per PC4.

4.5 Correlation tests

After testing how the WIs relate to actual walking behavior as reported in ODiN, it is now investigated whether the WIs correlate with each other. A correlation was carried out between all three existing WIs and the newly constructed WI. The results of the correlation are in the table below.

Table 15: Correlation matrix for Frank's WI, Graz WI and Urban WI.

		Frank's WI	Graz WI	Urban WI	Amsterdam WI
Frank's WI	Pearson Correlation	1	0.708	0.269	0.617
	Sig. (2-tailed)		<0.001	0.026	<0.001
	N	68	68	68	68
Graz WI	Pearson Correlation	0.708	1	0.621	0.838

	Sig. (2-tailed)	<0.001		<0.001	<0.001
	N	68	68	68	68
Urban WI	Pearson Correlation	0.269	0.621	1	0.574
	Sig. (2-tailed)	0.026	<0.001		<0.001
	N	68	68	68	68
Amsterdam WI	Pearson Correlation	0.617	0.838	0.574	1
	Sig. (2-tailed)	<0.001	<0.001	<0.001	
	N	68	68	68	68

The correlation in table 15 shows that the Pearson's Correlation coefficient (r) is highest between the Graz WI and the Amsterdam WI, since r = 0.838, *Sig.*<0.001, and *N* = 68. A value of r = 1 would indicate a perfect positive correlation between both WIs. With a value of r = 0.838, there is a strong positive correlation between the Graz WI and the Amsterdam WI. The second highest correlation was found between Frank's WI and the Graz WI, since r = 0.708, *Sig.*<0.001, and *N* = 68. With a value of r = 0.708, there is a rather strong positive correlation between Frank's WI and the Graz WI.

4.6 Conclusion

At the end of this chapter, SQ4 and SQ5 can be answered.
The fourth sub question was *'How consistent are the walkability indices?'* This question can be answered by assessing the sensitivity analyses that were performed. The first SA, which was performed by systematically increasing the weights of variables, showed that the Urban WI is the most consistent out of the three WIs. This is because the mean percentage change in WI score between the WI score that was obtained using a changed weight, and the regular WI score, was the lowest out of all three WIs (see table 2). This shows that the Urban WI is very robust when it comes to weight changes of its variables. The Urban WI was followed by the Graz WI. This index was still found to be relatively robust; the model proved to be a little sensitive to the weight change for intersection density. Frank's WI appeared to be most sensitive to weight changes (see table 2), with mean percentage WI score changes that range from -52.23 for the weight increase on Shannon Entropy land use mix, to 38.69 for the weight increase of retail FAR.
The other SA was performed by reproducing the WIs on neighborhood (CBS *buurt*) level, and this analysis showed that the Urban WI had results on neighborhood that were most consistent compared to the PC4 results out of the three WIs. The neighborhood level Urban WI made the pattern that was already clear in the PC4 level Urban WI even more clear.
After performing two sensitivity analyses, one by systematically increasing the weights of variables, and one by reproducing the WIs on neighborhood level, it can thus be concluded that the Urban WI is the most consistent WI when it comes to both weight changes of variables and changes in spatial units.

SQ5 was: '*Which set of variables most accurately predicts empirical walkability data for Amsterdam?'*
The regression analysis of the independent variables in the WIs and the dependent variable pointed out that out of the existing indices, the Urban WI is the model that fits the dependent variable best. This is because this WI has the highest adjusted R Square of all existing WIs; 29.6% of the variance in mean walking distance is explained by the Urban WI. The newly made Amsterdam WI however is even more accurate, since it accounts for 30.6% of the variance in mean walking distance per PC4. This is the highest explained variance out of all WIs. The Amsterdam WI consists of three variables; four (or more) way intersection density, residential density, and retail and services availability. The combination of these three variables in a WI thus most accurately predicts empirical walking data for Amsterdam.

5. Discussion

This chapter discusses results shown in the previous chapter. The results are explained and embedded into literature where possible. This chapter also reflects on the carried out research and offer possible follow-up research topics. Next to that, several practical recommendations are given for city governments and policymakers aiming to make their city more walkable.

5.1 Discussion of visual and geographic comparison

The three walkability indices as shown in figures 3, 4, and 5 have certain spatial patterns in common. In all three maps, it is clearly visible that the inner city center of Amsterdam is the most walkable area. This is because the inner city has a relatively high population density and a relatively high dwelling density. In addition, roads are often situated on both sides of the canals, with many bridges connecting the streets, resulting in quite some intersections in the inner city. The main shopping streets are also situated in the inner city, which explains why PC4 areas 1012 and 1011 have the highest amount of retail and services out of all PC4 areas in Amsterdam (according to the Graz WI). The presence of many types of shops and services, combined with several parks that the inner city has, accounts for a high land use mix, especially in PC4 area 1012.

Similarly, it can be observed that the harbor area on the Northwest of the map has a poor walkability in all three indices. The harbors have few residents, few retail and services and a low intersection density. Another area that has a poor walkability in all three indices is the somewhat rural area in the Northeast of Amsterdam (PC4 areas 1026, 1027, and 1028). This area contains several small villages like Zunderdam, Durgerdam and Holysloot. The largest part of these areas consists of fields and meadows, with some dairy farms. The character of the built environment here is clearly different from the rest of Amsterdam, as meadows take up more space than the built-up area, in large contrast to other areas in the city where buildings are predominant. The low score is thus explained by a low concentration of retail and services, a low population density and a low land use mix.

At the same time, a quick glance at the three maps is enough to see some differences. This is the case in the Southwest of Amsterdam. This area is Nieuw-Sloten and surroundings; a residential area that was built in the 1990s. It has a mostly square street pattern, but short blocks of buildings, resulting in a high intersection density. This is why PC4 area 1066 (containing Nieuw-Sloten) has the highest intersection density out of all PC4 areas in Amsterdam (according to the Graz WI). Since the three WIs handle intersection density in different ways, the output varies a lot in this area specifically.

5.2 Discussion of sensitivity analysis

As described in section 3.6, the SA consists of two parts. The first SA was performed by increasing the weights of all variables of the WIs by 10%. The second part of the SA was done by using '*buurten*' (neighborhoods, as defined by CBS) as administrative units instead of PC4 areas.

5.2.1 Sensitivity analysis type 1

In the case of a 10% increase, if the impact on results is substantially less than 10%, the parameter is not of critical importance according to Goodchild (2005). In areas where the impact of a 10% weight change is large, it is likely that there is a large homogeneity in variables, e.g. the harbor areas where there are intersections, but hardly any land use mix, retail, or residences.

5.2.1.1 Frank's WI

The weight increase for all variables in Frank's WI; intersection density (3 or more legs), residential density, Shannon Entropy land use mix, and retail FAR, are all especially noticeable in PC4 areas 1051 and 1058.

As can be seen in figure 6, PC4 area 1058 showed the largest percentage increase in WI Z-score caused by a weight increase on intersection density, the largest decrease caused by a weight increase on residential density, the largest increase caused by a weight increase on retail FAR, and the largest decrease caused by a weight increase on Shannon Entropy land use mix. This pattern could possibly be explained by the fact that PC4 area 1058 is an area that contains some residential areas and areas with shops and a large park with footpaths in and around it. This means that this area has many intersections. Therefore, with a higher weight ascribed to intersection density, the total WI Z-score increases. Since the area ratio of commercial buildings is relatively high as well, a higher weight on retail FAR logically results in a large percentage increase of the WI Z-score. 1058 on the other hand has a relatively low residential density and a low Shannon Entropy land use mix, since the park takes up a large amount of space and the residential areas that are there, are not densely populated. This is why this PC4 area saw their WI Z-score plummet when more weight is given to both residential density and Shannon Entropy land use mix.

PC4 area 1051 showed the largest decrease in intersection density, the largest increase in residential density, the largest decrease in retail FAR, and the largest increase in Shannon Entropy land use mix. This can be explained by the fact that PC4 1051 is a largely a residential area, and the area also consists of a business park called *'Centrale markt'* with some wastelands. Next to that, there is a cemetery in the PC4 area. Therefore, the area scores well on both Shannon Entropy land use mix and residential density, so when these variables are given more weight, the WI Z-scores show a large percentage increase compared to the one with regular weights.

However, due to the business park with its long roads without turns, there are relatively few intersections. When intersections are given even more weight, the WI Z-score of this area will logically plunge in comparison to the regular WI score. Since there are hardly any shops and services in the area, the retail FAR is already low and when its weight is increased, the WI Z-score of the area decreases.

Next to these spatial patterns, most variables in Frank's WI have a high (ranging from -52.23% to 38.69%) mean difference in percentage when comparing the WI Z-score of the WI with a 10% variable increase, compared to the regular WI Z-score (see table 2). This shows that the WI as a whole is highly sensitive to weight changes of its variables. Therefore, in this regard, Frank's WI does not seem suitable for the context of Amsterdam.

5.2.1.2 Graz WI

The weight increase for all variables in the Graz WI; intersection density (4 or more legs), residential density, and Shannon Entropy land use mix, are all especially noticeable in PC4 areas 1064 (largest percentage increase for intersection density, largest percentage decreases for residential density and for Shannon Entropy land use mix), 1092 (largest percentage decrease for intersection density), 1056 (largest percentage increase for residential density), and 1032 (largest percentage increase for Shannon Entropy land use mix).

The largest percentage decreases in WI Z-scores caused by a weight increase on Shannon Entropy land use mix and on residential density, are both found in PC4 area 1064. This could possibly be explained by the fact that PC4 area 1064 contains a large part of lake Sloterplas and the remaining parts of the PC4 are mostly residential areas. Because of the relatively monotonous land use, the WI Z-score goes down when Shannon Entropy land use mix gets allocated more weight. While the area does have residential areas, these consist mostly out of single-family terrace houses. This housing type logically holds less residents than an apartment complex. Therefore, residential density is relatively low and the

WI Z-score therefore plummets in this area when residential density is given more weight. Intersection density on the other hand is high in this area, since there are many road connections in and around housing blocks, especially for pedestrians. This explains the surge in WI Z-score after increasing the weight of intersection density.
The most dramatic decrease in WI Z-score after increasing the weight of intersection density was found in PC4 area 1092, since this area mostly consists of long, parallel roads without many connections between them.
After implementing a 10% increase in the weight of residential density, the WI Z-score of PC4 area 1056 saw the largest increase. This can possibly be explained by the presence of many apartment complexes in this postal code, indicating a high residential density because many addresses can be found in a relatively small area. Giving more weight to residential density thus boosts the WI Z-score of this area specifically.
The 10% increase in the weight of Shannon Entropy land use mix caused the largest increase in the WI Z-score of PC4 area 1032. This can be explained by the diverse land use that this area has; it contains a business park, a park, residential areas, and several green strips with trees along pedestrian paths. When this land use mix is given more weight, the WI Z-score of the area logically improves.

Next to these spatial patterns that are present, it must be noted that all variables in the Graz WI have a low (ranging from -0.05% to 0.1%) mean difference in percentage when comparing the WI Z-score of the WI with a 10% variable increase, compared to the regular WI Z-score (see table 2x). This shows that the WI as a whole is highly robust to weight changes of its variables. Therefore, the Graz WI seems suitable for the context of Amsterdam in this regard.

5.2.1.3 Urban WI

The weight increase for all variables in the Urban WI; population density, residential density, retail and services availability, and intersection density (3 or more legs), are all especially noticeable in PC4 areas 1032 (largest percentage increase for residential density, largest percentage increase for population density, largest percentage increase for retail and services availability, and largest decrease for intersection density), 1083 (largest percentage increase for intersection density, largest percentage decrease for population density, largest percentage decrease for retail and services availability), 1104 (largest percentage decrease for residential density), and 1032 (largest percentage increase for intersection density).

As discussed in the previous section, PC4 area 1032 has a diverse land use, including a high residential density, a high population density, and many retail and services. Therefore, when the weight of the variables residential density, population density, and retail and services availability is increased, the WI Z-score of the area improves. Due to the large business park in this area however, the intersection density is relatively low. When that variable is given more weight, the WI Z-score of this area in particular goes down.
PC4 area 1104 sees the largest percentage decrease in WI Z-score as a result of increased weight of residential density. A possible explanation for this decline is that this neighborhood contains a part of the Bijlmer area, but also another residential area with many strips of green along walking paths and many pedestrian passageways between buildings. Therefore, 1104 scores good on intersection density but not on residential density because its area also contains many green areas without addresses but with pedestrian paths and therefore many intersections.

Next to these spatial patterns, it should be noted that all variables in the Urban WI have a very low (ranging from -0.008% to 0.006%) mean difference in percentage when comparing the WI Z-score of the WI with a 10% variable increase, compared to the regular WI Z-score (see table 2x). This shows that the WI as a whole is highly robust to weight changes of its variables. Therefore, the Graz WI seems suitable for the context of Amsterdam in this regard.

5.1.2 Sensitivity analysis type 2

This part of the SA was carried out by using neighborhoods as administrative units instead of PC4 areas. This way, differences in WI Z-score within a larger area become visible. This section gives an overview of the most important changes.

5.1.2.1 Frank's WI

When comparing Frank's WI per neighborhood (figure 22) to the original Frank's WI per PC4 area (figure 3), several differences can be observed. It is striking that especially the Southwest of Amsterdam shows more nuance. Where PC4 area 1066 had a high WI Z-score (quintile 5), part of the area when looking at the neighborhoods now has a low WI Z-score of quintile 1. This indicates large differences within the PC4 area and shows that Frank's WI is rather sensitive to changes in spatial unit of analysis.

5.1.2.1 The Graz WI

Compared to the Graz Walkability Index per PC4 area (figure 4), the Graz Walkability Index per neighborhood (figure 26) shows a more detailed image. The postal code level makes it seem like the Graz WI is especially high in the inner city, but when looking at neighborhood level, this pattern is a lot more scattered. The spatial pattern of a high walkability in the inner city, which was present in figure 4, is not overly clear anymore in figure 26. This indicates large differences within the PC4 area and shows that the Graz WI is also rather sensitive to changes in spatial unit of analysis.

5.1.2.3 The Urban WI

The Urban Walkability Index per PC4 area (figure 5) and the Urban Walkability Index per neighborhood (figure 32) display the same pattern. The highest quintiles, meaning the areas with the highest walkability, are spatially concentrated in the inner city in both maps. Where the pattern is already clear on PC4 level, the neighborhood level analysis gives more detail and shifts the pattern of high walkability even more towards the inner city center. The pattern changes that occur when comparing neighborhood level to PC4 level are very little, and it can be said that the Urban WI is rather robust when it comes to changes in spatial unit of analysis.

5.3 Regression analysis

The regression analysis pointed out which variables were significant in each WI. It also displayed the direction of the association between the independent variable and the dependent variable. This section connects findings to literature and tries to find possible explanations for unexpected results.

5.3.1 Frank's WI

The regression analysis Model Summary on Frank's WI showed that Frank's WI explains 24.7% of the variance in mean walking distance (see table 3). Table 5 displays that only the three (or more) way intersections and the residential density are significant predictors of mean walking distance per PC4 area. Three (or more) way intersections appeared to be a significant independent variable, since b = -0,322, $t(67)$ = -2.686, p<0.05. The b-value is negative, indicating that an increase in 3 (or more) way

intersections results in a decrease of the mean walking distance per PC4. Walkability and mean walking distance are inversely related (the more walkable an area is, the shorter the mean walking distance is) (Reyer et al., 2014), so this finding is in line with Frank et al. (2010). They used intersection density in their model because they found it to be a significant positive predictor of walkability, which is also the case in this research.
Residential density is another significant independent variable, since $b = 0.289$, $t(67) = 2.024$, $p<0.05$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. Given the inverse relationship between walkability and mean walking distance, this output is unexpected and not in line with Frank et al. (2010). An increase in mean walking distance would mean that the area is less walkable (Reyer et al., 2014), since otherwise shorter routes would have been possible. This unexpected association between residential density and mean walking distance could possibly be explained by the fact that car use is discouraged by the municipality of Amsterdam. Parking costs are highest in the inner city center of Amsterdam; it can cost up to €7.50 per hour to park (Gemeente Amsterdam, 2017). This financial downside could spark an incentive for residents to use an active mode of transportation, like walking, instead of using private vehicles. In addition, the inner city center has the second lowest mean car ownership per household of all parts of Amsterdam (Gemeente Amsterdam, 2017). This means there are relatively few cars in the inner city, which is why people have to use other modes of transportation, including walking.
Though the association between residential density and mean walking distance that emerged in the regression analysis was unexpected, other authors was also found this relationship to exist. Duncan, Dansie, Strachan, Munsell, Huang, Moudon, Goldberg, and Buchwald (2010) and Huang, Moudon, Zhou, and Saelens (2019) found high residential density to be associated with more walking in the neighborhood. Liao et al. (2020) therefore argue that high-density areas are more attractive for walking. They suggest that this might be because walking distances are shorter in these areas due to the compactness of the built-up area. It could however also be argued that walking distances in these areas are longer, due to the many opportunities and wide variety of potential destinations that the high-density built-up area has to offer.

The other two independent variables, FAR and Shannon Entropy land use mix, are not significant since for both, $p>0.05$. These independent variables thus do not play a role in determining the walkability in the context of Amsterdam, when mean walking distance per PC4 area is used as the dependent variable.

5.3.2 The Graz WI

The regression analysis Model Summary on the Graz WI showed that the Graz WI explains 26.3% of the variance in mean walking distance (see table 6). Table 8 showed that the four (or more) way intersections and the residential density are significant predictors of mean walking distance per PC4 area. Four (or more) way intersections turned out to be a significant independent variable, since $b = -0,334$, $t(67) = -2.826$, $p<0.05$. The negative *b*-value indicates that an increase in four (or more) way intersections results in a decrease of the mean walking distance per PC4. Walkability and mean walking distance are inversely related (the more walkable an area is, the shorter the mean walking distance is) (Reyer et al., 2014), so this finding corresponds with Grasser et al. (2017), who used intersection density in their WI because they found it to be a significant positive predictor of walkability.
Residential density is another significant independent variable, since $b = 0.287$, $t(67) = 2.049$, $p<0.05$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. Given the inverse relationship between walkability and mean walking distance, this output is unexpected and not in line with Grasser et al. (2017), since an increase in mean walking distance would mean that the area is less walkable (Reyer et al., 2014), since otherwise shorter routes would have been possible. The possible explanation given for this is given in the previous section (5.3.1).

The other independent variable, Shannon Entropy Land Use Mix, was not found to be significant, since $p>0.05$. This independent variable thus does not play a role in determining the walkability in the context of Amsterdam, when mean walking distance per PC4 area is used as the dependent variable.

5.3.3 The Urban WI

The regression analysis Model Summary on the Urban WI showed that the Urban WI accounts for 29.6% of the variance in mean walking distance (see table 9). Table 11 displayed that residential density, intersection availability and retail and services availability are significant predictors of mean walking distance per PC4 area.

Residential density is an independent variable that turned out to be significant, since $b = 0.506$, $t(67) = 3.211$, $p<0.05$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. Again, given the inverse relationship between walkability and mean walking distance, this output is unexpected and not in line with the results of Glazier et al. (2012), who did find residential density to be a significant positive predictor of walkability. The possible explanation given for this is given in the section 5.2.1.

Intersection availability is the second independent variable that was found to be significant, since $b = -0.347$, $t(67) = -3.152$, $p<0.05$. The *b*-value is negative, meaning that an increase in intersections results in a decrease of the mean walking distance per PC4. This was an expected result, since Glazier et al. (2012) also found this to be true, and Reyer et al. (2014) argue that walking distances are shorter when the area is more walkable.

Retail and services availability is the third independent variable that was found to be significant, since $b = -0.265$, $t(67) = -2.001$, $p=0.050$. The negative *b*-value indicates that an increase in retail and services availability results in a decrease of the mean walking distance per PC4. This was also an expected result, given the fact that walking distances are shorter when the area is more walkable (Reyer et al., 2014), and Glazier et al. (2012) found this variable to be a significant positive predictor of walkability as well.

The remaining independent variable, population density, was not found to be significant, since $p>0.05$. This independent variable thus does not play a role in determining the walkability in the context of Amsterdam, when mean walking distance per PC4 area is used as the dependent variable.

5.3.4 The Amsterdam WI

The regression analysis Model Summary on the Amsterdam WI showed that the Urban WI accounts for 30.6% of the variance in mean walking distance (see table 12). Table 14 revealed that residential density, intersection availability and retail and services availability are all significant predictors of mean walking distance per PC4 area.

Residential density is a significant independent variable, since $b = 0.484$, $t(67) = 3.530$, $p<0.001$. The *b*-value is positive, indicating that an increase in residential density results in an increase of the mean walking distance per PC4. Again, given the inverse relationship between walkability and mean walking distance, this output is unexpected and not in line with the results of Frank et al. (2010), Grasser et al. (2017) and Glazier et al. (2012), who did find residential density to be a significant positive predictor of walkability. The possible explanation given for this is given in the section 5.2.1.

Retail and services availability is the second independent variable that is significant, since $b = -0.267$, $t(67) = -2.029$, $p<0.05$. The *b*-value is negative, meaning that an increase in intersections results in a decrease of the mean walking distance per PC4. As discussed in the previous section 5.3.2, this was an expected result as this association was found by Glazier et al. (2012) as well.

Intersection density (intersections with 4 or more legs) is the third independent variable that is significant, since $b = -0.344$, $t(67) = -3.163$, $p<0.05$. The *b*-value is negative, meaning that an increase in retail and services results in a decrease of the mean walking distance per PC4. Since walkability and mean walking distance are inversely related (Reyer et al., 2014), this finding corresponds with Grasser et al. (2017), who used intersection density in their WI because they found it to be a significant positive predictor of walkability.

When comparing the Model Summary tables of all four WIs, it becomes clear that the Urban WI is the best fitting model out of the existing WIs for the context of Amsterdam; 29.6% of the variance in mean walking distance is explained by the Urban WI. The newly constructed Amsterdam WI turns out to be the best fitting model of all WIs for the context of the city, since it accounts for 30.6% of the variance in mean walking distance.

5.4 Discussion of correlation analysis

The correlation matrix displayed in table 15 shows that the Pearson's Correlation coefficient (r) is highest between the Graz WI and the Amsterdam WI, since r = 0.838, *Sig.*<0.001, and *N* = 68. A value of r = 1 would indicate a perfect positive correlation between both WIs. With a value of r = 0.838, there is a strong positive correlation between the Graz WI and the Amsterdam WI. This strong correlation indicates that these indices are quite alike. This is expected, because two out of the three variables in the Amsterdam WI are used in the Graz WI.
The second highest correlation was found between Frank's WI and the Graz WI, since r = 0.708, *Sig.*<0.001, and *N* = 68. With a value of r = 0.708, there is a rather strong positive correlation between Frank's WI and the Graz WI, indicating that those indices are quite alike as well (see section 3.1.1). This is also an expected result, since the Graz WI uses roughly the same variables as Frank's WI, with the exception of retail FAR, which is omitted by the Graz WI.

5.5 Strengths

The research that was carried out has several strengths that can be mentioned.
Firstly, the WIs were thoroughly checked for sensitivity to parameter changes. This was done via the two sensitivity analyses that were performed. This way, both sensitivity to weight changes and sensitivity to spatial unit changes was investigated.
Secondly, this research adds to existing research by filling in a research gap about walkability in Europe. No WI had been made yet for Amsterdam, let alone for any Dutch city. The construction of the Amsterdam WI aims to contribute to WI research in Europe, which is much needed according to Grasser (2014).

5.6 Limitations

The research that was carried out has some important limitations that need to be mentioned.
First of all, the ODIN dataset does not take the walking behavior of tourists into account. Only the walking behavior of a group of statistically representative residents was included in the dataset. Therefore, touristic walks are not included, while Amsterdam welcomed almost 22 million unique tourists in 2019 (Fedorova & Klingen, 2021). Tourists are therefore an important share of the pedestrians in Amsterdam, but this entire group is not represented in this research. This should be kept in mind as this impacts results. For example, tourists may find themselves more likely to walk anyway in order to sightsee, while the local residents have other ways of transportation readily available, like their own bike, their own or a shared car, or public transportation.

A second limitation is that ODiN data is not available for CBS '*buurt'* (neighborhood) level. This hinders a more detailed regression analysis and may cause MAUP and the inaccuracies that come with it. In addition, the ODiN dataset consisted of 1529 cases for the entire city of Amsterdam, but some areas had very little to no cases. This is why 13 PC4 areas were omitted from the regression analysis (see section 3.5). This may raise the question of whether the results of the regression analysis are reliable and generalizable enough. A more detailed and more extensive dataset would offer opportunities for doing a similar research with more precise results.

Thirdly, the Amsterdam Walkability Index was constructed using only the variables that occurred in the three chosen WI models. It could very well be a possibility that more variables outside of those used by the existing WIs play a role in predicting walking behavior in Amsterdam, since the explained variance of this model is 30.6%. This opens up possibilities for further research to finetune the Amsterdam WI by checking if other variables come into play.

A fourth limitation is the buffer sizes of the Urban WI. The WI uses 720m buffers, and it can be questioned whether this is truly an appropriate buffer size for Amsterdam, as PC4 areas in Amsterdam are rather small and the buffer will intersect through several other PC4 areas. Also, MAUP is very much an issue here: the difference between the area of the PC4 within the buffer and the area outside of it, could be big.

5.7 Recommendations

At the end of the discussion, recommendations need to be made. The limitations of this research, as stated above, offer opportunities for follow-up researches to improve the Amsterdam WI. In addition, practical advice to decision-makers is given about ways to ameliorate walkability in Amsterdam.

5.7.1 Follow-up research

This research opens the door for several follow-up studies. First of all, a follow-up research could aim to repeat this research using a more extensive dataset on walking behavior in Amsterdam. This could be obtained by gathering data, or by using a more detailed dataset that might be published in the future. Using a more extensive dataset would ensure a more precise analysis, since in this research, there was missing data in the ODiN dataset for several PC4 areas (figure 2).
A second recommendation is to explore more WIs and more variables. This research used three WIs, but more and other indices could be compared, which might result in finding variables that fit the context of Amsterdam even better. Further research could also follow a data-driven approach, where a measure is derived from regressing walking behavior on a list of physical factors of the local environment (Liao et al., 2020). An example of this is the study of Glazier et al. (2012), where a total of nine variables were selected and from there, the selection narrowed down to four variables that were found to be significant (Glazier et al., 2012).
A third recommendation is to investigate perceptual walkability factors. Reyer et al. (2014) argue that Stuttgart, a German city, has a much less of a 'walking suppressive' urban infrastructure than cities in North America or Australia. Hence, walkability might not be a less powerful factor of personal mobility in the European context but it might be more difficult to detect (Reyer et al., 2014). Since Amsterdam is a Northern-European city as well, this might be applicable to Amsterdam as well and in that case, it may prove to be insightful to look beyond topographical walkability factors.

5.7.2 City government of Amsterdam

As stated in the introduction, a body of contradictory results of walkability studies has led to considerable confusion over policy directions for decision-makers (Frank et al., 2021). This thesis also aimed to provide clarity for decision-makers on what policy could boost walkability in Amsterdam by providing practical recommendations. The first recommendation is to use models that are fitted for the local context. While this may seem obvious, this research showed the large difference in walkability scores between three WIs. The most commonly used WI, which is Frank's WI, turned out to be the least well-fitting out of the compared WIs for the context of Amsterdam. Instead, the WI that was constructed in this research, the Amsterdam WI, was proven to be the most fitting WI model for the dependent variable. While this is currently the only available WI for Amsterdam, future research may improve this WI further by making it more accurate. It is therefore advised for decision-makers to keep up with developments in WIs in order to base policy directions off of the WI that fits the local context best.

Secondly, when decision-makers or urban planners want to ameliorate walkability in a certain existing or in a new construction project, special attention should be given to intersections (with four or more legs). Out of the Amsterdam WI, this variable had the largest positive significant effect on walkability. Constructing pedestrian paths with many intersections makes it easy for people to choose the shortest route to their destination (Reyer et al., 2014). A road network with many intersections can therefore make walking an attractive transport option, and walking for transport is essential for a sustainable city (Liao et al., 2020). Another important predictor of walkability is the availability of retail and services. Mixing residential areas with some shopping facilities and services may make the area more attractive for walking (Glazier et al., 2012), and can thus be used as a strategy to boost neighborhood walkability.

6. Conclusion

This final chapter contains a synergy of the most important findings of this research. The research question is answered and embedded in results and literature.

This study aimed to answer the following research question:

What is the most suitable walkability index for predicting walking behavior in Amsterdam?

To provide an answer to this question, a review of related work, a comparison of the three chosen WIs, two types of sensitivity analysis, a regression analysis, and a correlation analysis were carried out. The results revealed that out of the three existing WIs, the Urban WI proved to be most fitting to the context of Amsterdam with an explained variance of 29.6% (table 9). The results of the two sensitivity analyses furthermore show that the Urban WI is the most consistent out of the existing WIs, when it comes to both weight changes of variables and changes in spatial units.

However, the newly constructed Amsterdam Walkability Index is even more suitable for predicting walking behavior in Amsterdam, since the regression analysis showed that this WI accounts for 30.6% of the variance in mean walking distance per PC4 area (table 12). The Amsterdam WI uses three independent variables that were all found to be significant: intersection density (4 or more legs), residential density, and the availability of retail and services. Intersection density and residential density are two of the most frequently used variables that can be observed in the built environment, that are used to measure walkability (Park et al., 2014). The combination of these three variables in a WI thus most accurately predicts empirical walking data for Amsterdam.

These findings are relevant, since the Amsterdam Walkability Index is the first WI that was developed specifically for a Dutch city. The Amsterdam WI therefore aims to contribute to WI research in Europe, which is much needed according to Grasser (2014). Furthermore, the Amsterdam WI can be used by decision-makers and urban planners that wish to ameliorate walkability in Amsterdam. The newly created index provides a combination of variables that are fitted to the context of Amsterdam, making it a useful tool for assessing the impact that changes in the built environment may have on walking behavior.

In short, this research has shown that the most suitable walkability index for predicting walking behavior in Amsterdam is the Amsterdam WI, and the most suitable existing walkability index is the Urban WI.

Bibliography

Alfonzo, M., Boarnet, M. G., Day, K., Mcmillan, T., & Anderson, C. L. (2008). The relationship of neighbourhood built environment features and adult parents' walking. *Journal of Urban Design*, *13*(1), 29-51.

CBS (2017). Bestand Bodemgebruik. Retrieved from https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/natuur-en-milieu/bestand-bodemgebruik.

CBS (2020). Kerncijfers per postcode. Retrieved from https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode.

CBS (2021). Inwoners per gemeente. Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/regionaal/inwoners.

Conticelli, E., Maimaris, A., Papageorgiou, G., & Tondelli, S. (2018). Planning and designing walkable cities: a smart approach. In: R. Papa, R. Fistola, & C. Gargiulo (eds). *Smart Planning: Sustainability and Mobility in the Age of Change. Green Energy and Technology.* New York City: Springer International Publishing.

Crosetto, M. & Tarantola, S. (2001). Uncertainty and sensitivity analysis: tools for GIS-based model implementation. *International Journal of Geographical Information Science, 15*(5), 415-437.

Data Science Partners (2021, 24 augustus.) Bias-Variance Tradeoff: wat is het? En waarom is het belangrijk? Retrieved from https://pythoncursus.nl/bias-variance-tradeoff/.

De Vocht, A. (2019). *Basishandboek SPSS 25.* Utrecht: Bijleveld Press.

Dobesova, Z., & Krivka, T. (2012). Walkability index in the urban planning: A case study in Olomouc city. In: J. Burian (eds.). *Advances in spatial planning.* Rijeka, Croatia: InTech.

Duncan, G.E., Dansie, E.J., Strachan, E., Munsell, M., Huang, R., Moudon, A.V., Goldberg, J., & Buchwald, D. (2012). Genetic and environmental influences on residential location in the US. *Health Place, 18*(3), 515-519.

Esri Nederland (2021). Winkels (bèta) uit OpenStreetMap. Last accessed on March 30th, 2022, retrieved from https://www.arcgis.com/home/item.html?id=574b374dae2f4764bd93eb36d1093f65.

Esri Nederland (2022). Eten & Drinken (bèta) uit OpenStreetMap. Last accessed on March 30th, 2022, retrieved from https://www.arcgis.com/home/item.html?id=d5183c9c2c4d40beb9c16fcc3eb39a38.

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association, 76*(3), 265-294.

Fedorova, T., & Klingen, J. (2021). Bezoekers in Amsterdam 2019-2020, Bezoekersprognose 2021-2022 [Factsheet]. Retrieved from https://onderzoek.amsterdam.nl/publicatie/derde-raming-bezoekersaantallen-amsterdam-2018-2021.

Frank, L.D., Sallis, J.F., Saelens, B.E., Leary, L., Cain, K., Conway, T.L., & Hess, P.M. (2010). The development of a walkability index: application to the Neighborhood Quality of Life Study. *British journal of sports medicine, 44*(13), 924-933.

Frank, L.D., Appleyard, B.S., Ulmer, J.M., Chapman, J.E., & Fox, E.H. (2021). Comparing walkability methods: creation of street smart walk score and efficacy of a code-based 3D walkability index. *Journal of Transport & Health, 21*(1), 101005.

Frank, L.D., Engelke, P.O., & Schmid, T.L. (2003). *Health and community design: The impact of the built environment on physical activity.* Washington DC: Island Press.

Forsyth, A., Hearst, M., Oakes, J,M., & Schmitz, K.H. (2008). Design and Destinations: Factors Influencing Walking and Total Physical Activity. *Urban Studies, 45*(9), 1973-1996.

Forsyth, A. (2015). What is a Walkable Place? The Walkability Debate in Urban Design. *Urban Design International, 20*(4), 274-292.

Frumkin, H., Frank, L.D., Jackson, R. (2004). *Urban Sprawl and Public Health: Designing, planning, and building for healthy communities.* Washington DC: Island Press.

Galanis, A., & Eliou, N. (2011). Evaluation of the pedestrian infrastructure using walkability indicators. *WSEAS Transactions on Environment and Development, 7*(12), 385-394.

Gemeente Amsterdam (2017). Nota Parkeernormen Auto. Retrieved from https://assets.amsterdam.nl/publish/pages/865234/nota_parkeernormen_auto_1.pdf.

Gemeente Amsterdam (2022). Loop- en fietsnetwerk Amsterdam. Retrieved from https://data.amsterdam.nl/datasets/7hGzsRXqWSGqHw/loop-en-fietsnetwerk-amsterdam/.

Giao, J., Kamphuis, C.B.M., Helbich, M., & Ettema, D. (2020). What is 'neighborhood walkability'? How the built environment differently correlates with walking for different purposes and with walking on weekdays and weekends. *Journal of Transport Geography, 88*(1), 102860.

Gilderbloom, J.I., Hanka, M.J., & Lasley, C.B. (2009). Amsterdam: planning policy for the ideal city? *Local Environment, 14*(6), 473-493.

Glanz, T.A. (2011). *Walkability, social interaction, and neighborhood design* [Master's thesis]. Lincoln: University of Nebraska.

Grasser, G. (2014). *Walkability and Public Health: Development of GIS-based indicators of walkability for surveillance and planning purposes in the city of Graz* [Dissertation]. Graz: Medical University of Graz, Institute for Social Medicine and Epidemiology.

Grasser, G., Titze, S., & Stronegger, W.J. (2016). Are residents of high-walkable areas satisfied with their neighborhood? *Zeitschrift Für Gesundheitswissenschaften, 26*(6), 469-476.

Grasser, G., Van Dyck, D., Titze, S., & Stronegger, W. J. (2017). A European perspective on GIS-based walkability and active modes of transport. *The European Journal of Public Health, 27*(1), 145-151.

Glazier, R.H., Weyman, J.T., Creatore, M.I., Gozdyra, P., Moineddin, R., Matheson, F.I., & Booth, G.L. (2012). Development and validation of an urban walkability index for Toronto, Canada. *Toronto Community Health Profiles Partnership*, 1-21.

Goodchild, M.F. (2005). GIS and modeling overview. In D.J. Maguire, M. Batty, & M.F. Goodchild (eds). *GIS, Spatial Analysis, and Modeling.* Redlands, CA, USA: Esri Press.

Habibian, M., & Hosseinzadeh, A. (2018). Walkability index across trip purposes. *Sustainable Cities and Society, 42*(1), 216-225.

Heywood, I., Cornelius, S., & Carver, S. (2011). *An introduction to geographical information systems* (4th edition). Harlow, UK: Pearson Education Limited.

Hsieh, H. & Chuang, M. (2021). Association of perceived environment walkability with purposive and discursive walking for urban design strategies. *The Journal of Transport and Land Use, 14*(1), 1099-1127.

Huang, R., Moudon, A.V., Zhou, C., & Saelens, B.E. (2019). Higher residential and employment densities are associated with more objectively measured walking in the home neighborhood. *Journal of transport and health, 12*(1), 142-151.

Jun, J.H. & Hur, M. (2015). The relationship between walkability and neighborhood social environment: the importance of physical and perceived walkability. *Applied Geography, 62*(1), 115-124.

Kadastrale Kaart (2022). *Gemeente Amsterdam.* Retrieved from https://kadastralekaart.com/gemeenten/amsterdam-GM0363.

Leslie, E., Coffee, N., Frank, L.D., 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.

Liao, B., Van den Berg, P.E.W., Van Wesemael, P.J.V., & Arentze, T.A. (2020). Empirical analysis of walkability using data from the Netherlands. *Transportation research part D: transport and environment, 85*(1), 102390.

Litman, T.A. (2003). Economic value of walkability. *Transportation Research Record, 1828*(1), 3-11.

Maibach, E., Steg, L., & Anable, J. (2009). Promoting physical activity and reducing climate change: Opportunities to replace short car trips with active transportation. *Preventive medicine, 49*(4), 326-327.

Mantri, A. (2008). *A GIS Based Approach to Measure Walkability of a Neighborhood.* Cincinnati: University of Cincinnati.

McNeill, L.H., Kreuter, M.W., & Subramanian, S.V. (2006). Social environment and physical activity: A review of concepts and evidence. *Social science & medicine, 63*(4), 1011-1022.

Mirzaei, E., Kheyroddin, R., Behzadfar, M., & Mignot, D. (2018). Utilitarian and hedonic walking: examining the impact of the built environment on walking behavior. *European transport research review, 10*(2), 1-14.

Moura, F., Cambra, P., & Gonçalves, A.B. (2017). Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in Lisbon. *Landscape and Urban Planning, 157*(1), 282-296.

Overheid (2022). Dataset Basisregistratie Adressen en Gebouwen (BAG). Last accessed on May 11th, 2022, retrieved from https://data.overheid.nl/dataset/basisregistratie-adressen-en-gebouwen--bag-.

Park, S., Deakin, E., & Lee, J.S. (2014). Perception-based walkability index to test impact of microlevel walkability on sustainable mode choice decisions. *Transportation Research Record, 2464*(1), 126-134.

Ruiz-Padillo, A., Pasqual, F.M., Uriarrte, A.M.L., & Cybis, H.B.B. (2018). Application of multi-criteria decision analysis methods for assessing walkability: A case study in Porto Alegre, Brazil. *Transportation Research Part D, 63*(2018), 855-871.

Saelens, B.E., Sallis, J.F., & Frank, L.D. (2003). Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of behavioral medicine, 25*(2), 129-137.

Sallis, J.F., Owen, N., & Fisher, E. (2015). Ecological models of health behavior. *Health Behavior: Theory, Research, and Practice, 5*(1), 43-64.

Schlossberg, M., & Brown, N. (2004). Comparing transit-oriented development sites by walkability indicators*. Transportation Research Record, 1887*(1), 34-42.

Sisson, S.B., McClain, J.J., & Tudor-Locke, C. (2008). Campus Walkability, Pedometer-Determined Steps, and Moderate-to-Vigorous Physical Activity: A Comparison of 2 University Campuses. *Journal of American College Health, 56*(5), 585-592.

Southworth, M. (2005). Designing the walkable city. *Journal of urban planning and development, 131*(4), 246-257.

Statistics Canada (2018). Census tract: Detailed definition. Last accessed on March 22, 2022, retrieved from https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/ct-sr/def-eng.htm.

Stockton, J.C., Duke-Williams, O., Stamatakis, E., Mindell, J.S., Brunner, E.J., & Shelton, N.J. (2016). Development of a novel walkability index for London, United Kingdom: cross-sectional application to the Whitehall II Study. *BMC public health, 16*(1), 1-12.

Tsiompras, A.B., & Photis, Y.N. (2017). What matters when it comes to ''Walk and the city''? Defining a weighted GIS-based walkability index. *Transportation research procedia, 24*(1), 523-530.

Tsiompras, A.B., & Photis, Y.N. (2020). Does neighborhood walkability affect ethnic diversity in Berlin? Insights from a spatial modeling approach. *European Journal of Geography, 11*(1), 163-187.

United States Census Bureau (2022). Glossary. Last accessed on April 8, 2022, retrieved from https://www.census.gov/glossary/.

Van Dyck, D., Cardon, G., Deforche, B., Sallis, J.F., Owen, N., & De Bourdeaudhuij, I. (2010). Neighborhood SES and walkability are related to physical activity behavior in Belgian adults. *Preventive Medicine, 50*(Supplement), S74-S79.

Wunderlich, F.M. (2008). Walking and rhythmicity: Sensing urban space. *Journal of Urban Design, 13*(1), 125–139.

Zakaria, J., & Ujang, N. (2015). Comfort of Walking in the City Center of Kuala Lumpur. *Procedia – Social and Behavioral Sciences, 170*(1), 642-652.

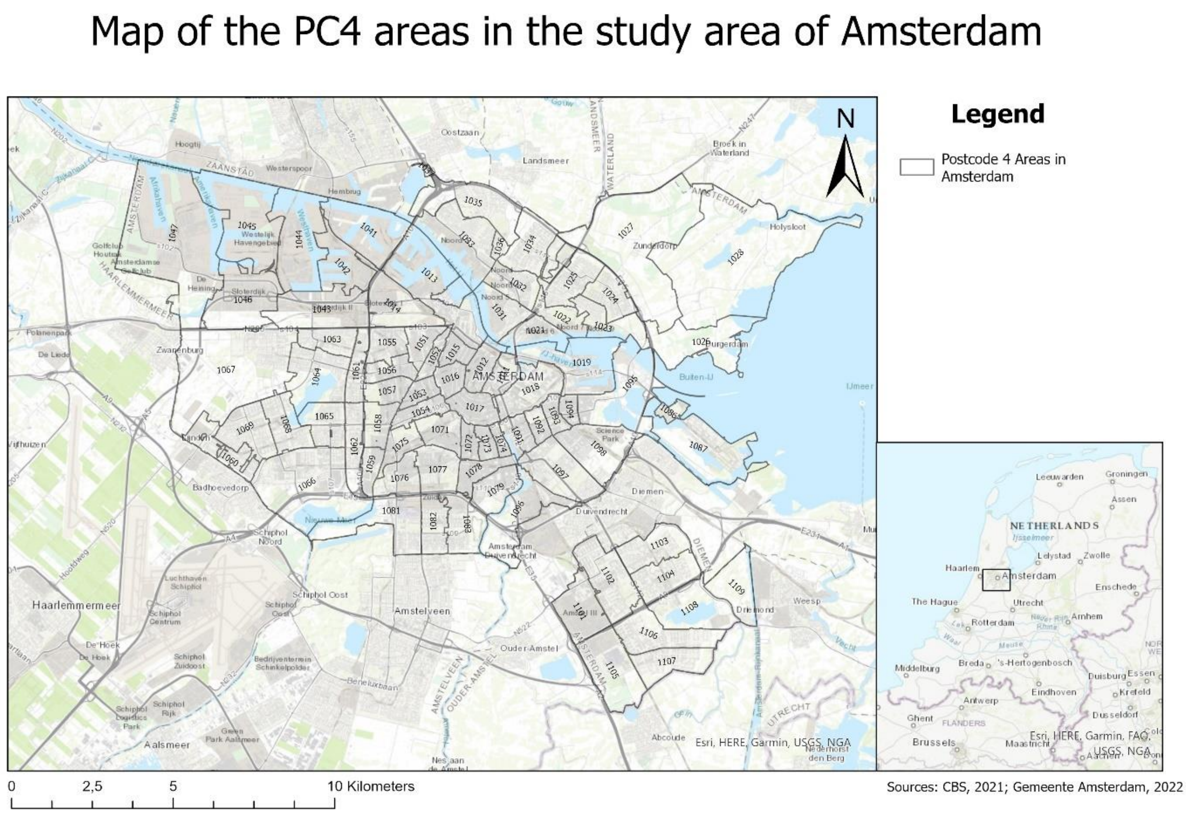
Zonghao, W. (2019). Shifting urbanity: urban experiment on walkability and traffic efficiency [Master thesis]. Retrieved from https://repository.tudelft.nl/islandora/object/uuid:3074ed7e-2fb7-4516-8c8d-964fc98e73e6.

Appendix

This appendix includes full-sized images of all maps used in this research. The titles of the appendices refer to the chapter in which the maps are used.

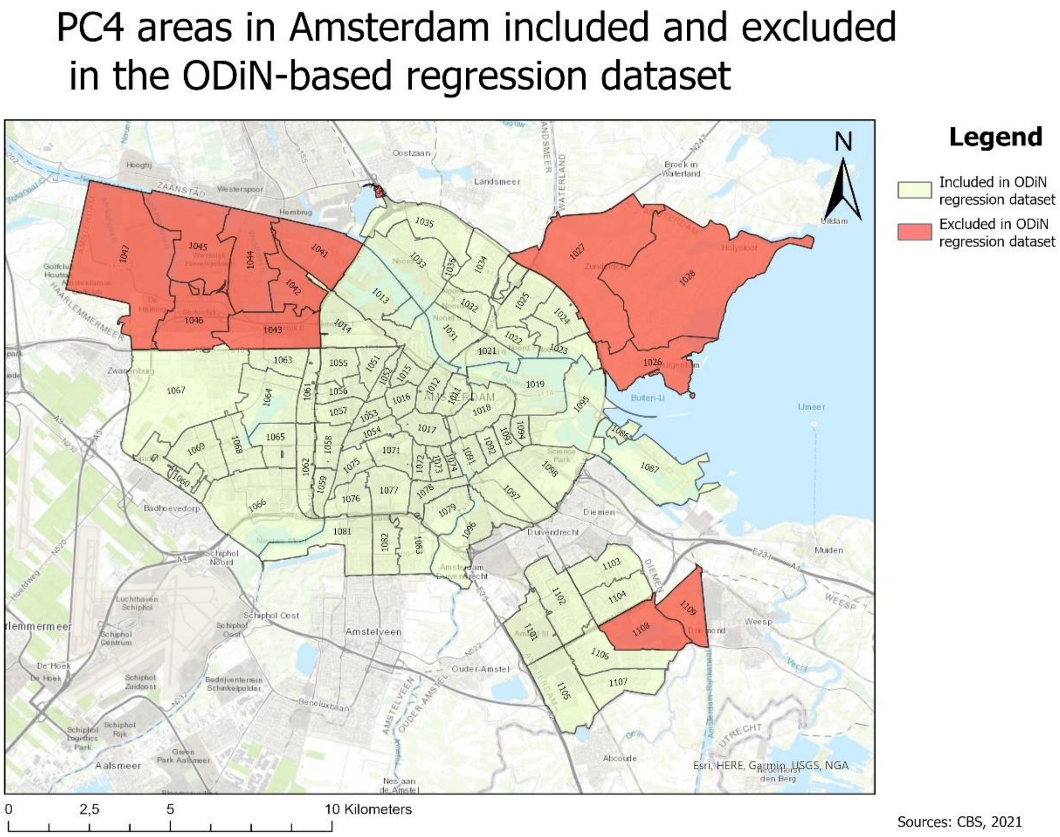
Appendix A: Introduction

Figure 1: The study area of this research, including the PC4 areas in Amsterdam and the position of the capital city within the Netherlands.



Appendix B: Methodology

Figure 2: Map showing the PC4 areas that were included and excluded from the ODiN-based regression dataset.



Appendix C: Results

Figure 3: Map showing Frank's Walkability Index per PC4 area in Amsterdam.

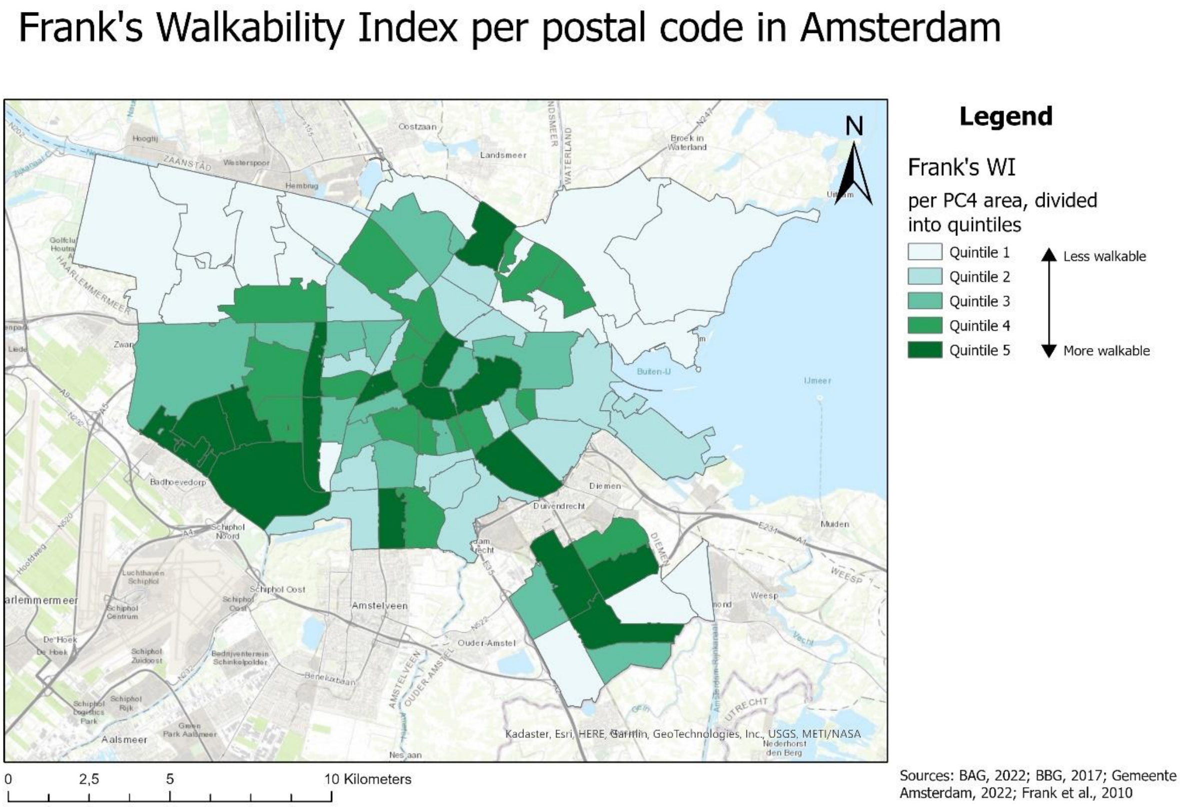


Figure 4: Map showing the Graz Walkability Index per PC4 area in Amsterdam.

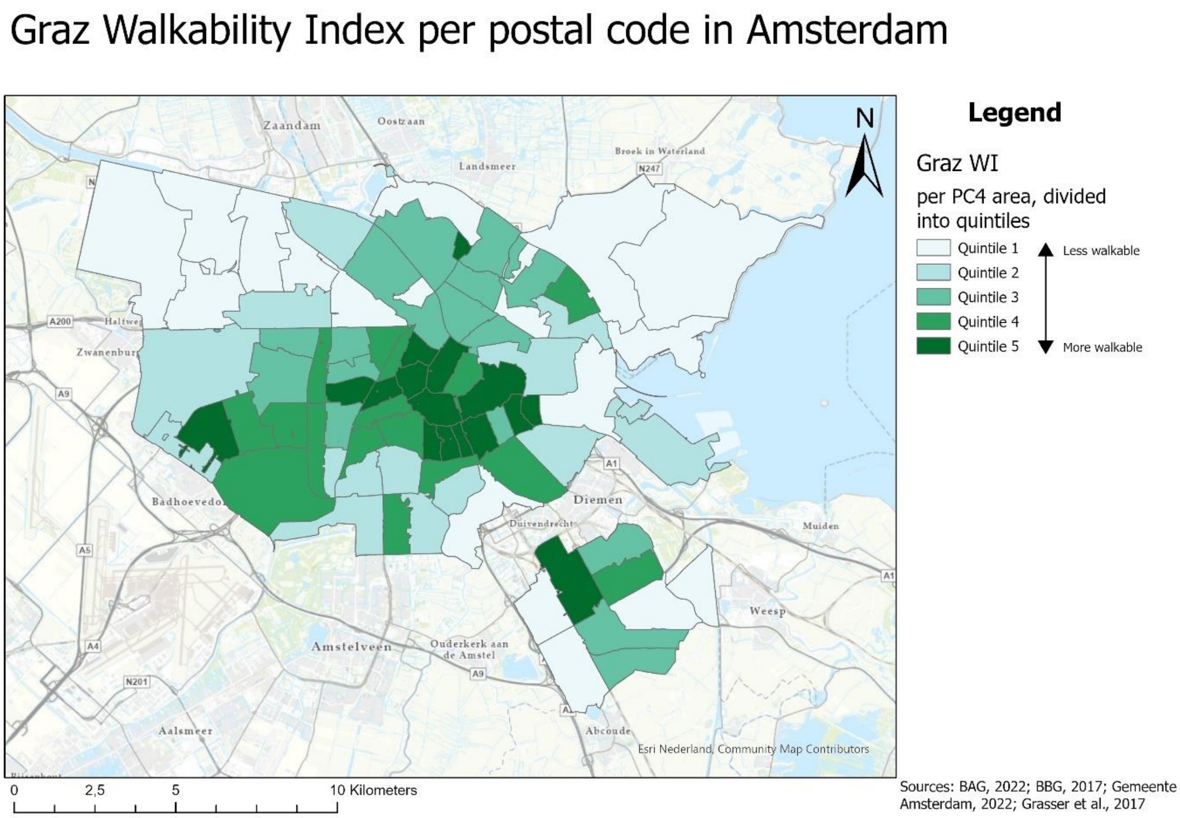
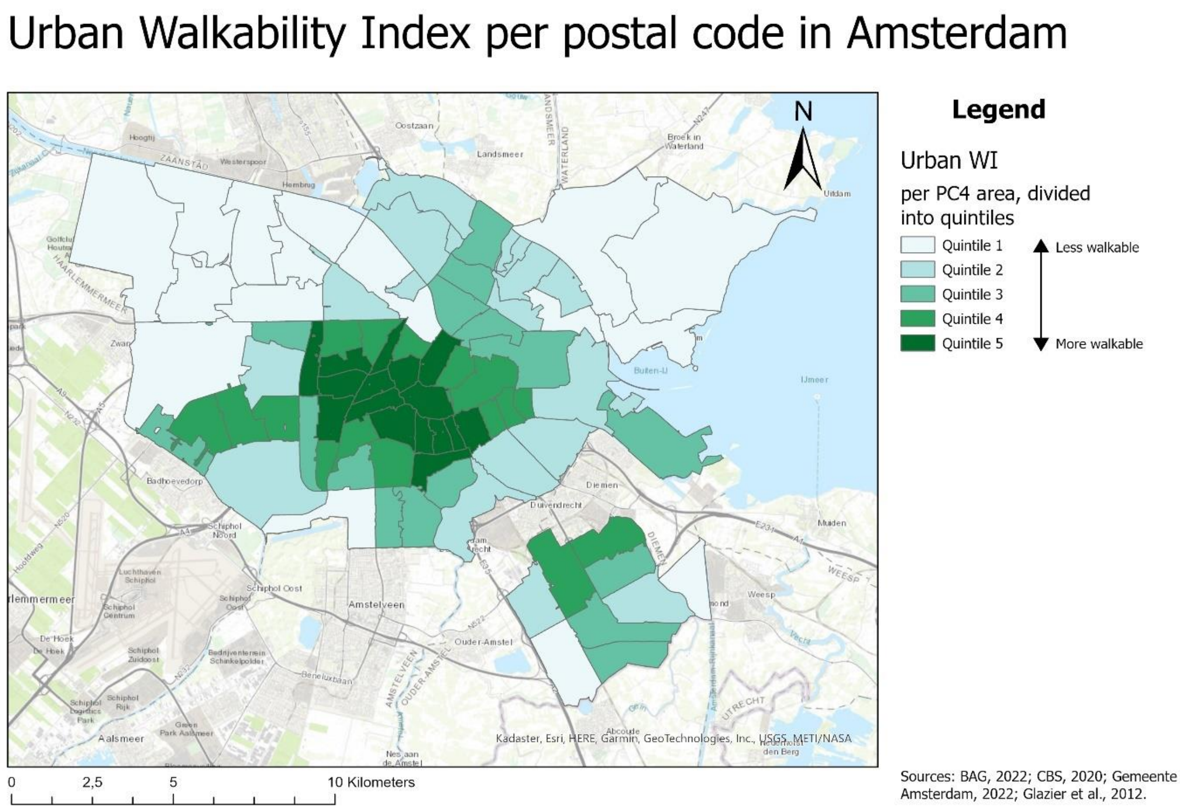


Figure 5: Map showing the Urban Walkability Index per PC4 area in Amsterdam.



Sensitivity analysis type 1

Figure 6: The impact of 10% weight increase of intersection density, measured in percentage change of Frank's WI Z-score.

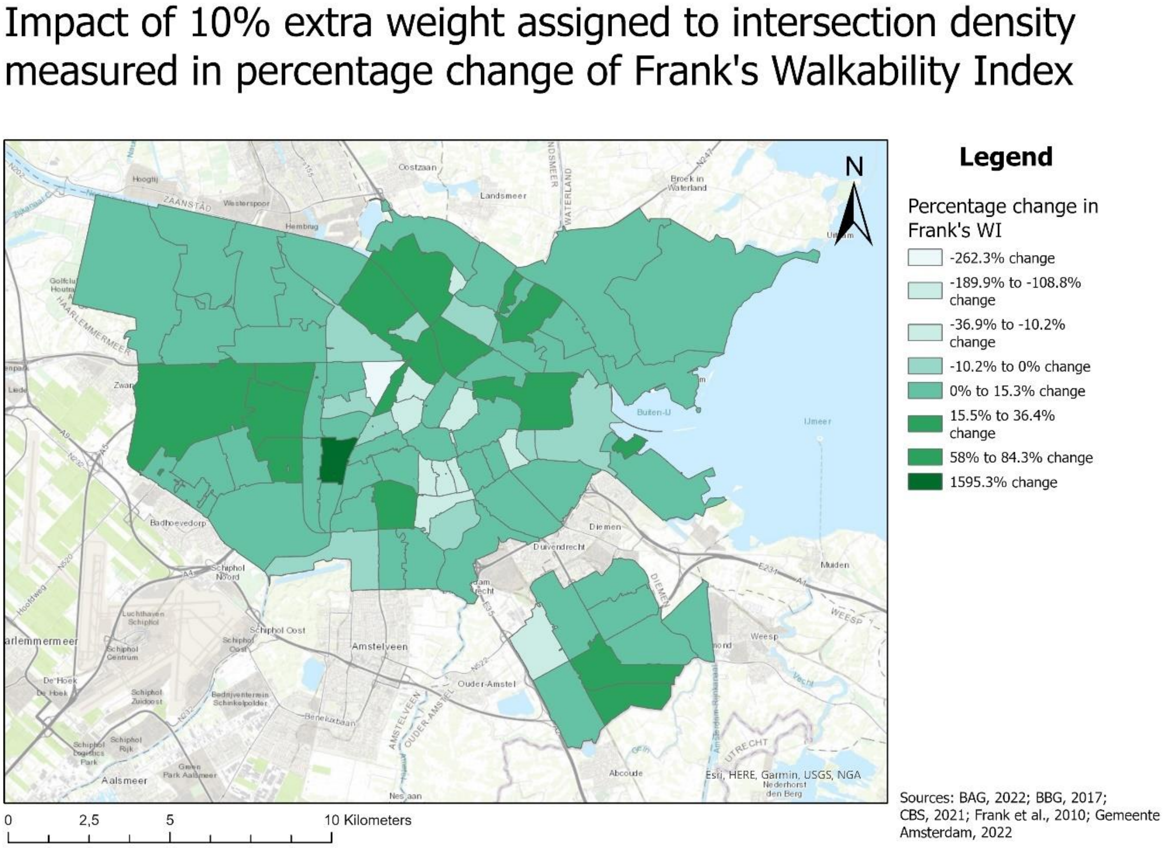


Figure 7: The impact of 10% weight increase of residential density, measured in percentage change of Frank's WI Z-score.

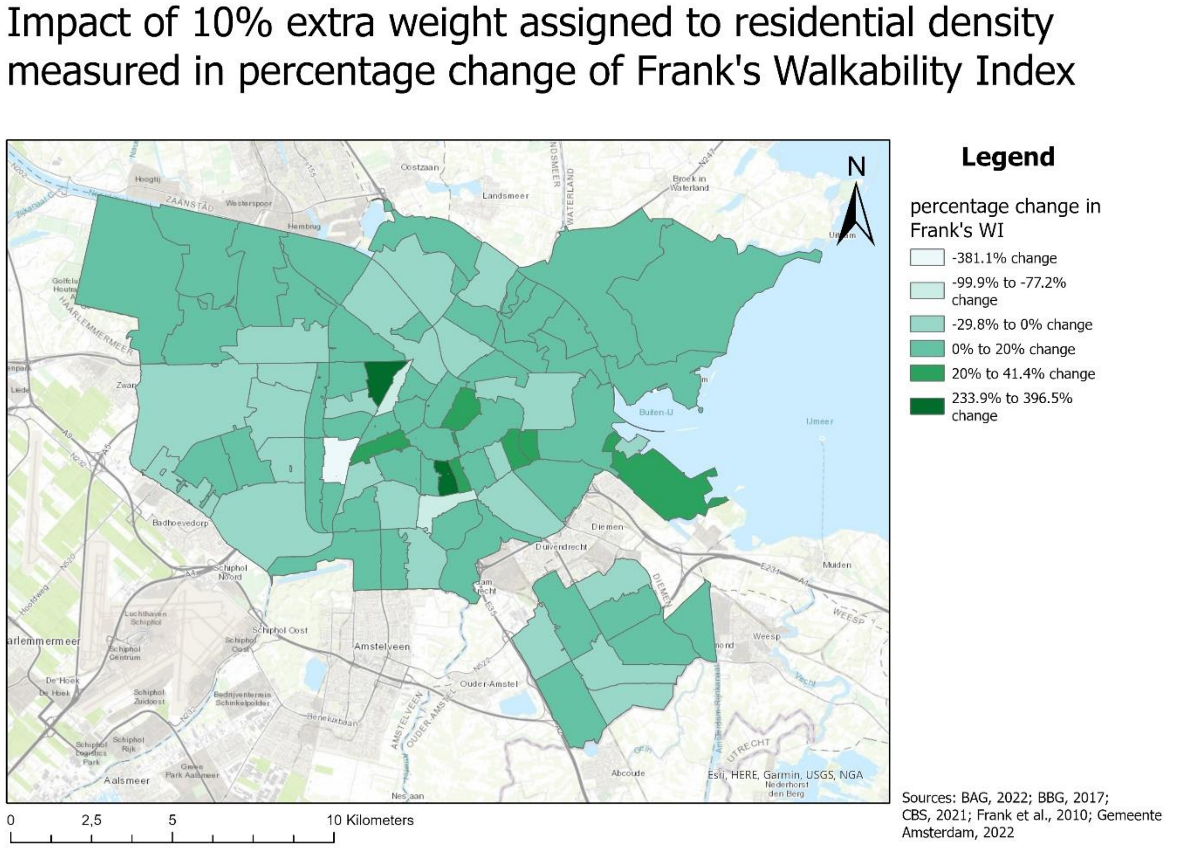


Figure 8: The impact of 10% weight increase of Shannon Entropy land use mix, measured in percentage change of Frank's WI Z-score.

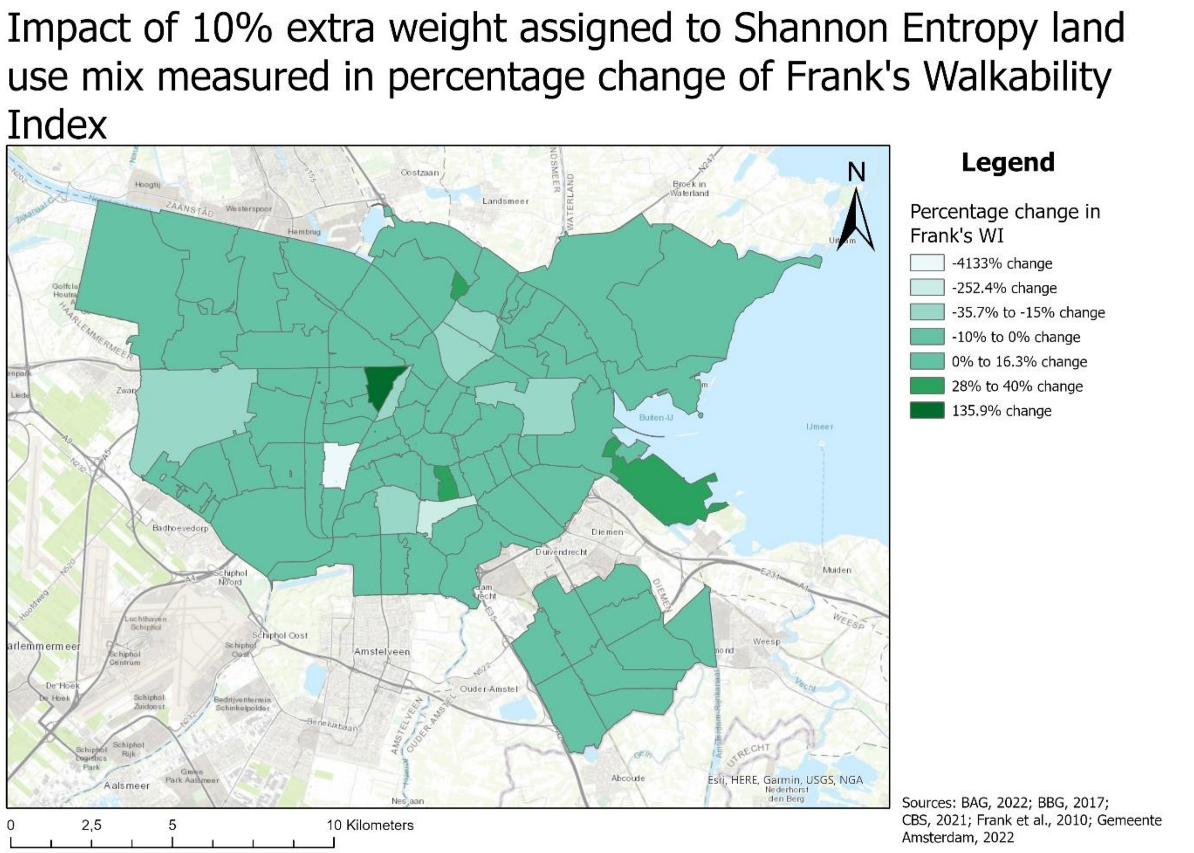


Figure 9: The impact of 10% weight increase of retail FAR, measured in percentage change of Frank's WI Z-score.

Impact of 10% extra weight assigned to retail Floor Area Ratio measured in percentage change of Frank's Walkability Index

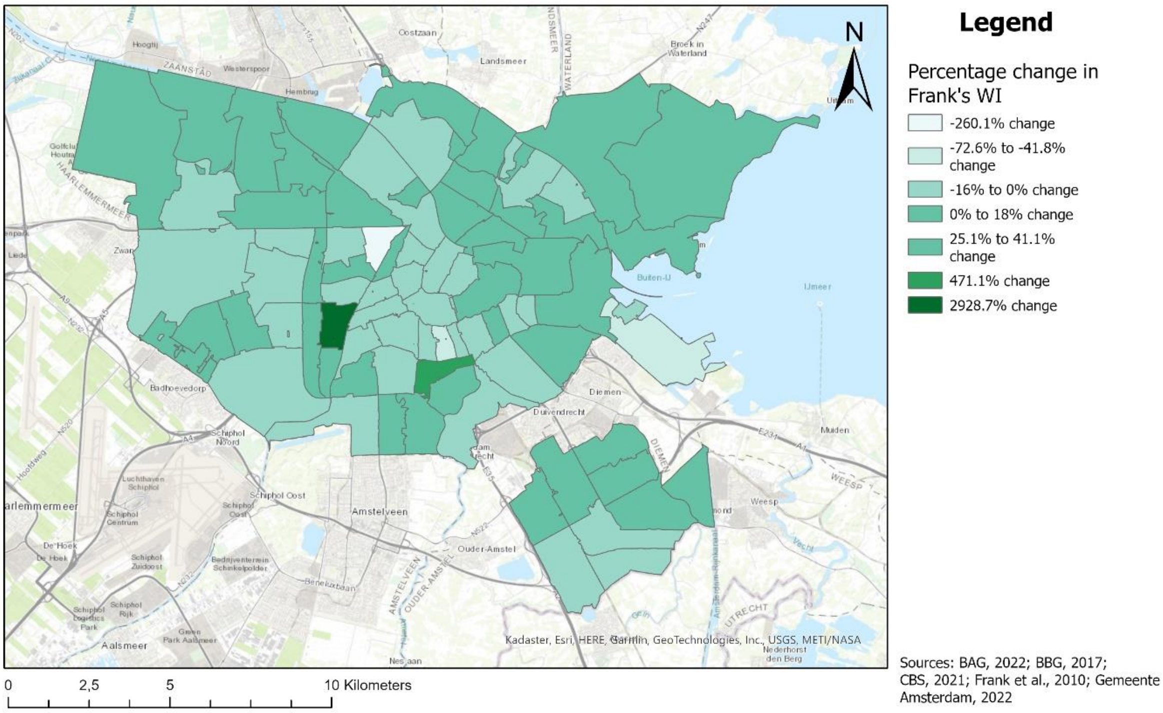


Figure 10: The impact of 10% weight increase of intersection density, measured in percentage change of the Graz WI Z-score.

Impact of 10% extra weight assigned to intersection density measured in percentage change of Graz Walkability Index

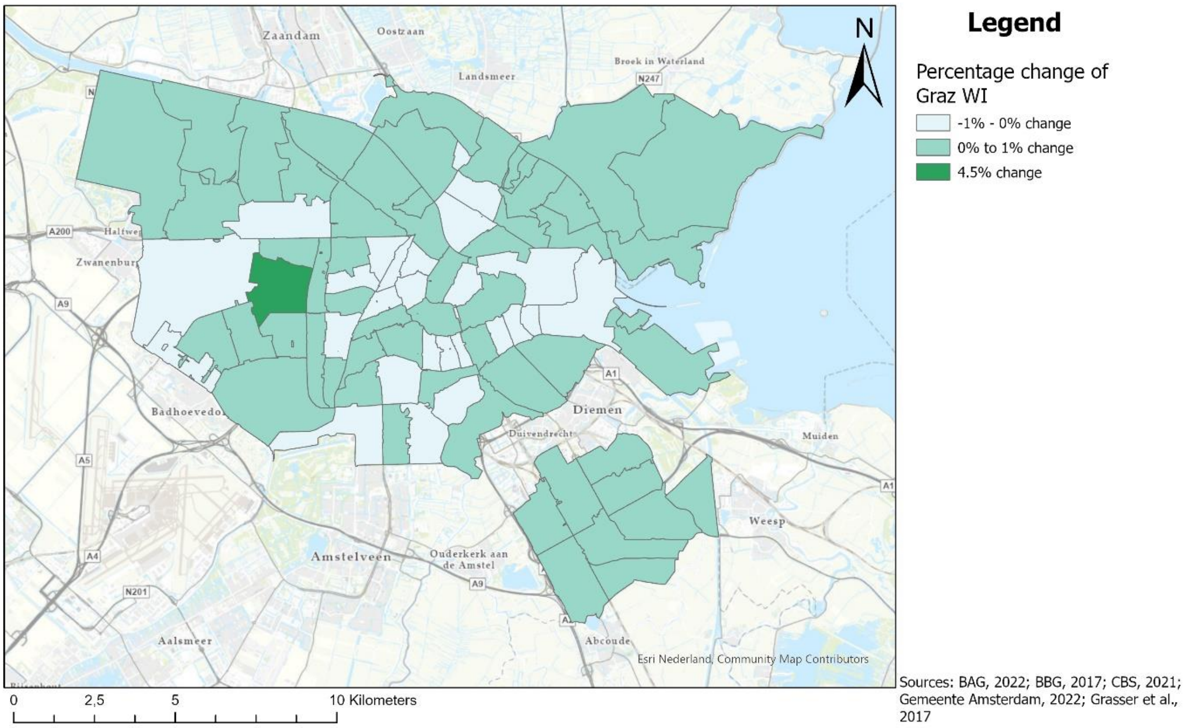


Figure 11: The impact of 10% weight increase of residential density, measured in percentage change of the Graz WI Z-score.

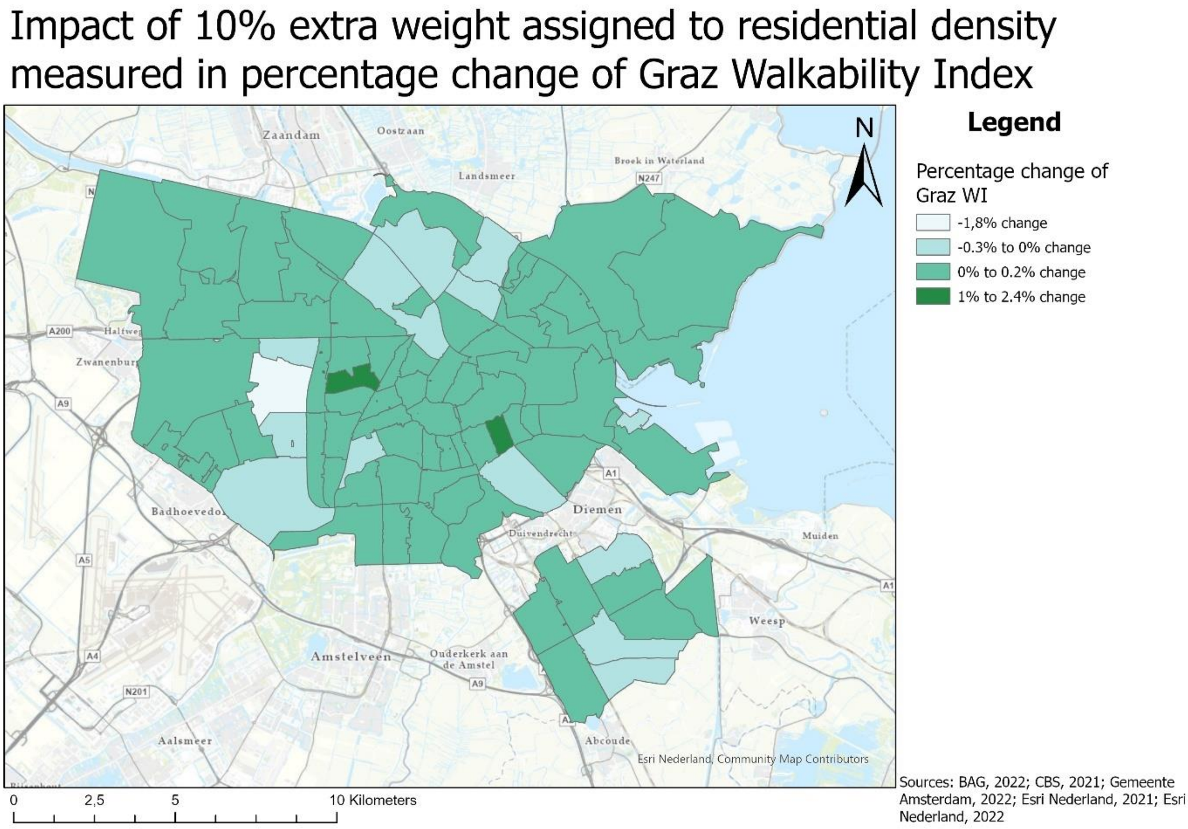


Figure 12: The impact of 10% weight increase of Shannon Entropy land use mix, measured in percentage change of the Graz WI Z-score.

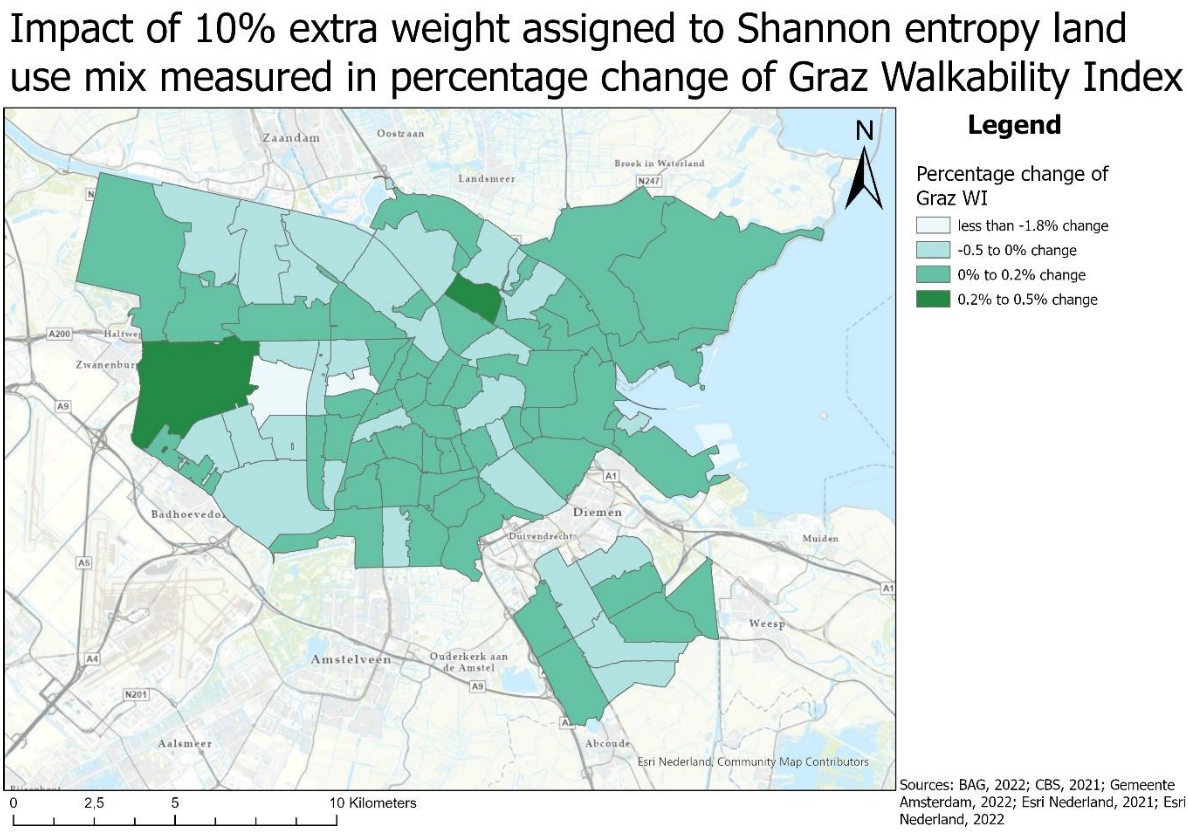


Figure 13: The impact of 10% weight increase of population density, measured in percentage change of the Urban WI Z-score.

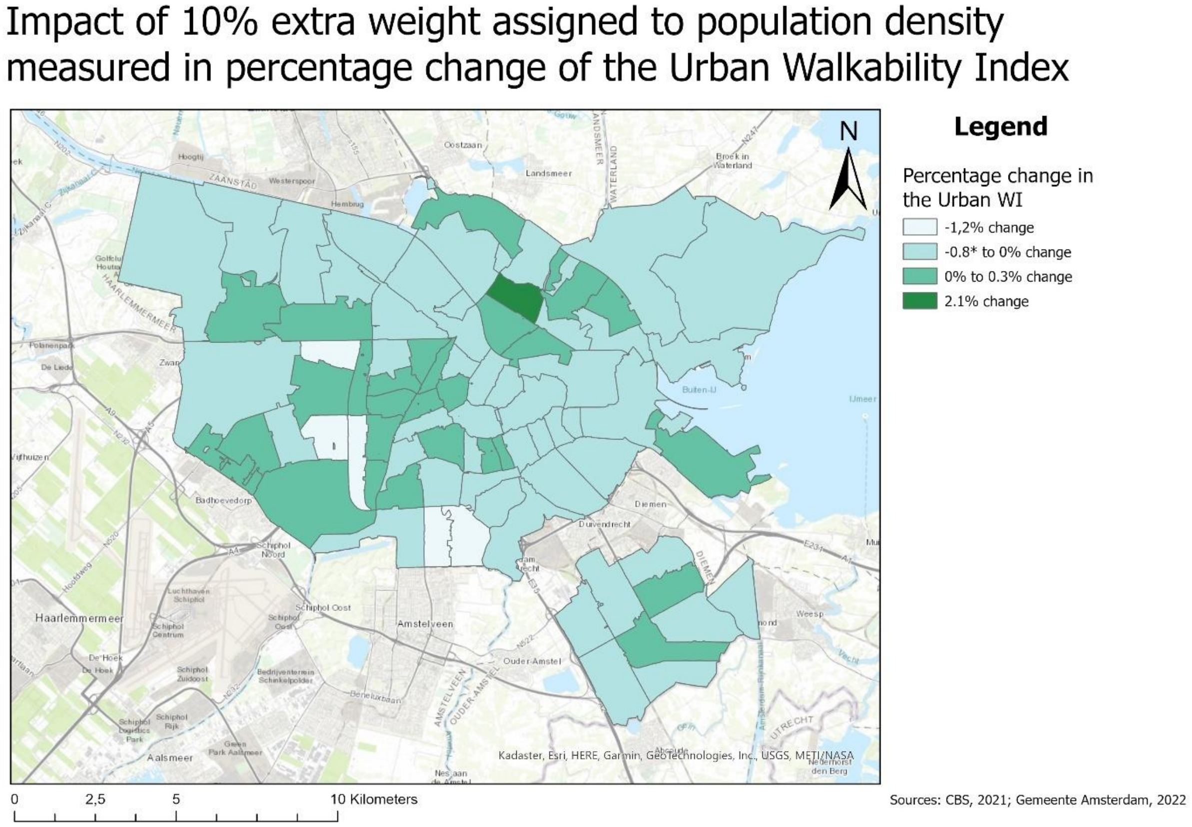


Figure 14: The impact of 10% weight increase of residential density, measured in percentage change of the Urban WI Z-score.

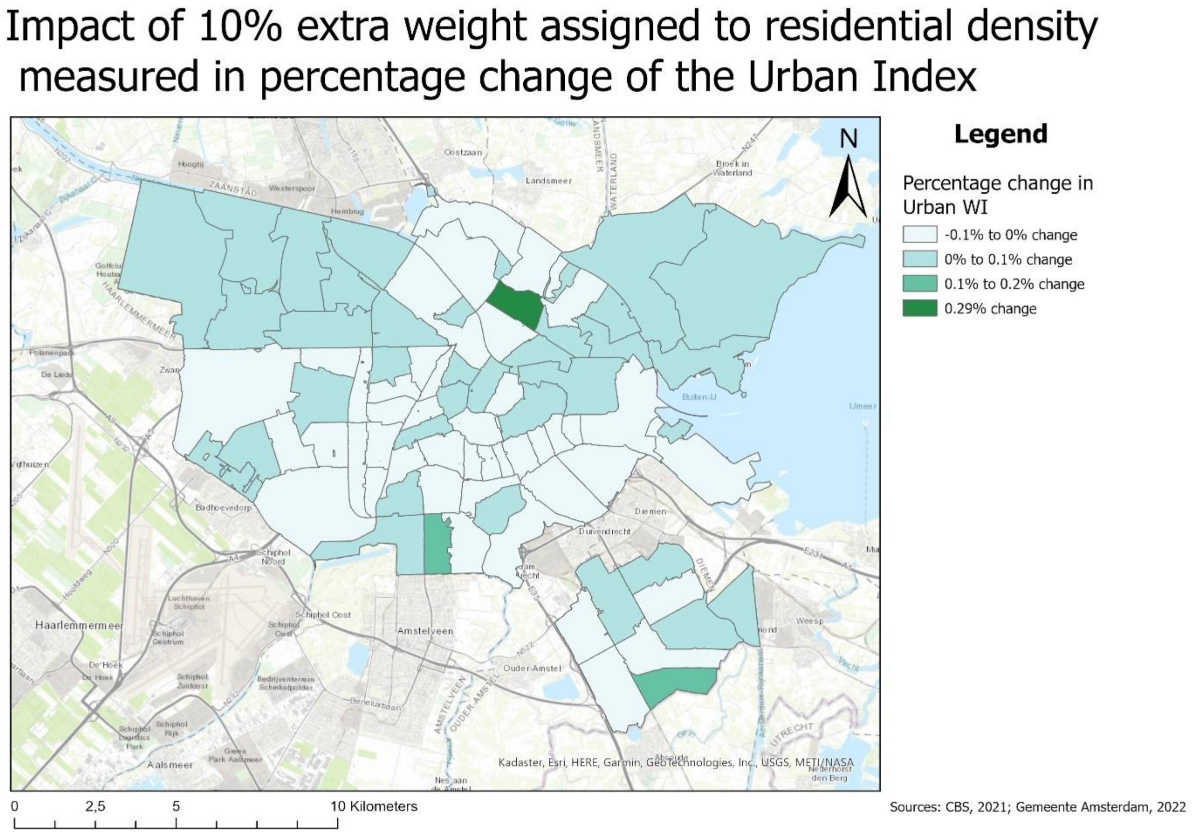


Figure 15: The impact of 10% weight increase of retail and services availability, measured in percentage change of the Urban WI Z-score.

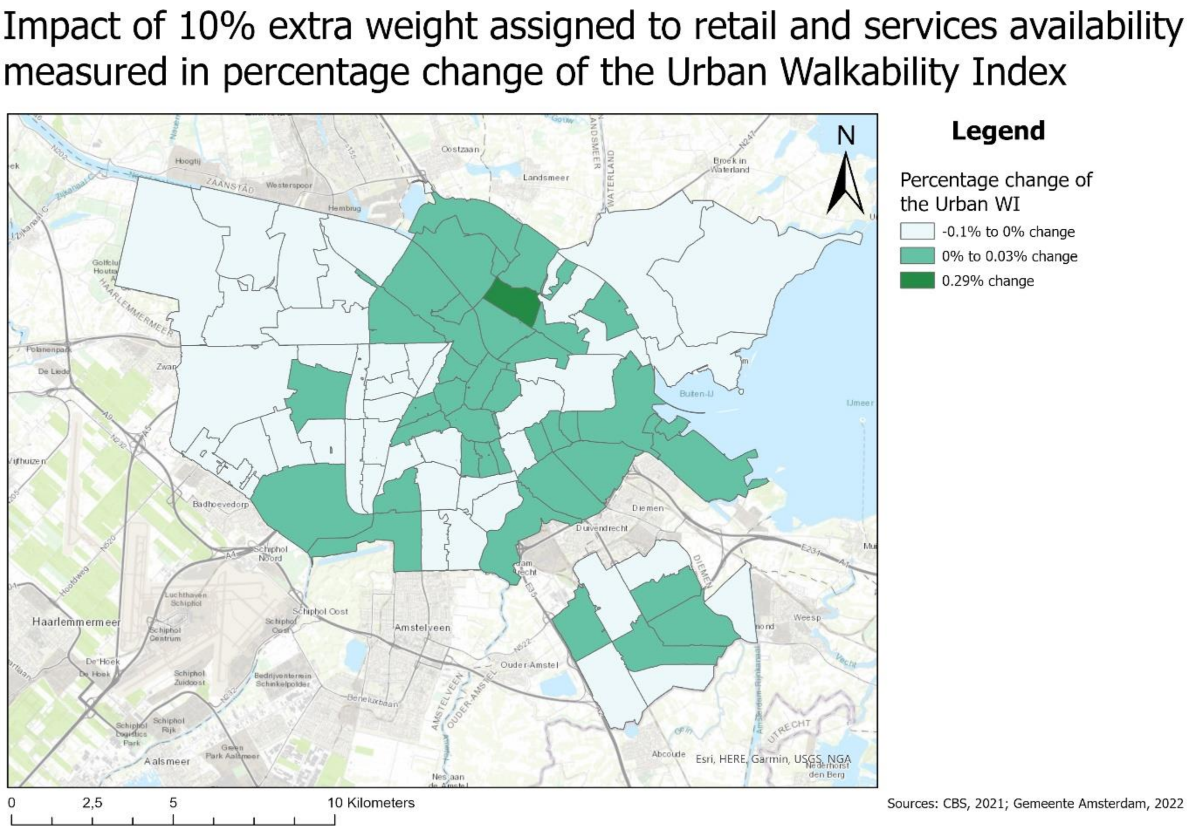
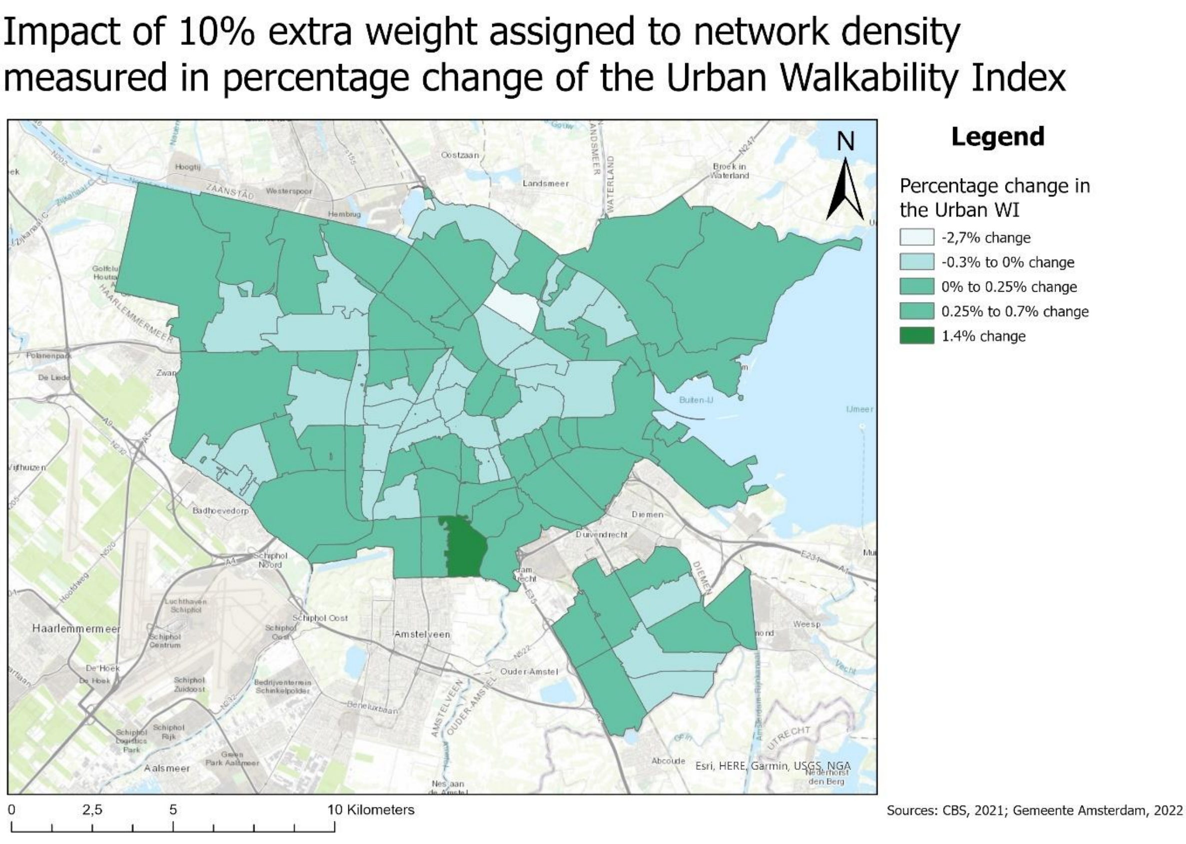


Figure 16: The impact of 10% weight increase of intersection density, measured in percentage change of the Urban WI Z-score.



Sensitivity analysis type 2

Figure 17: Map showing three (or more) way intersection density, which is one of the variables in Frank's WI.

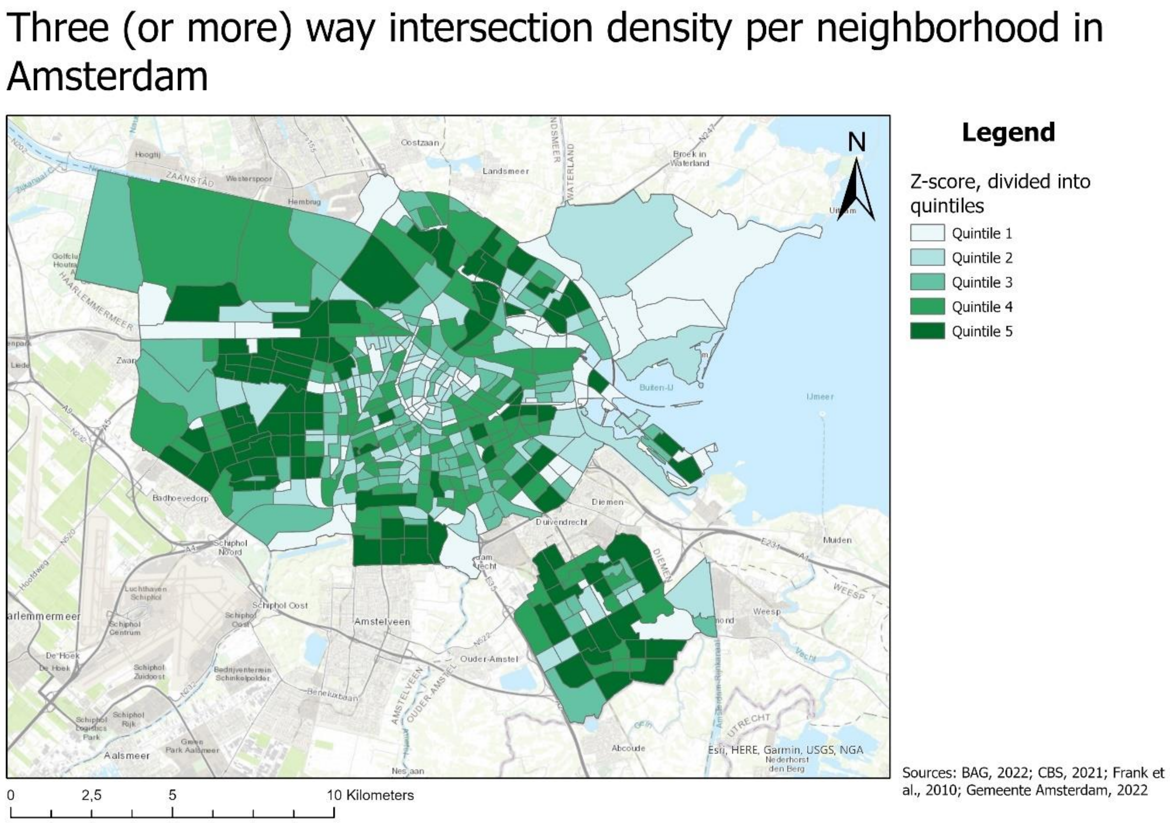


Figure 18: Map showing the SA performed on residential density, which is one of the variables in Frank's WI.

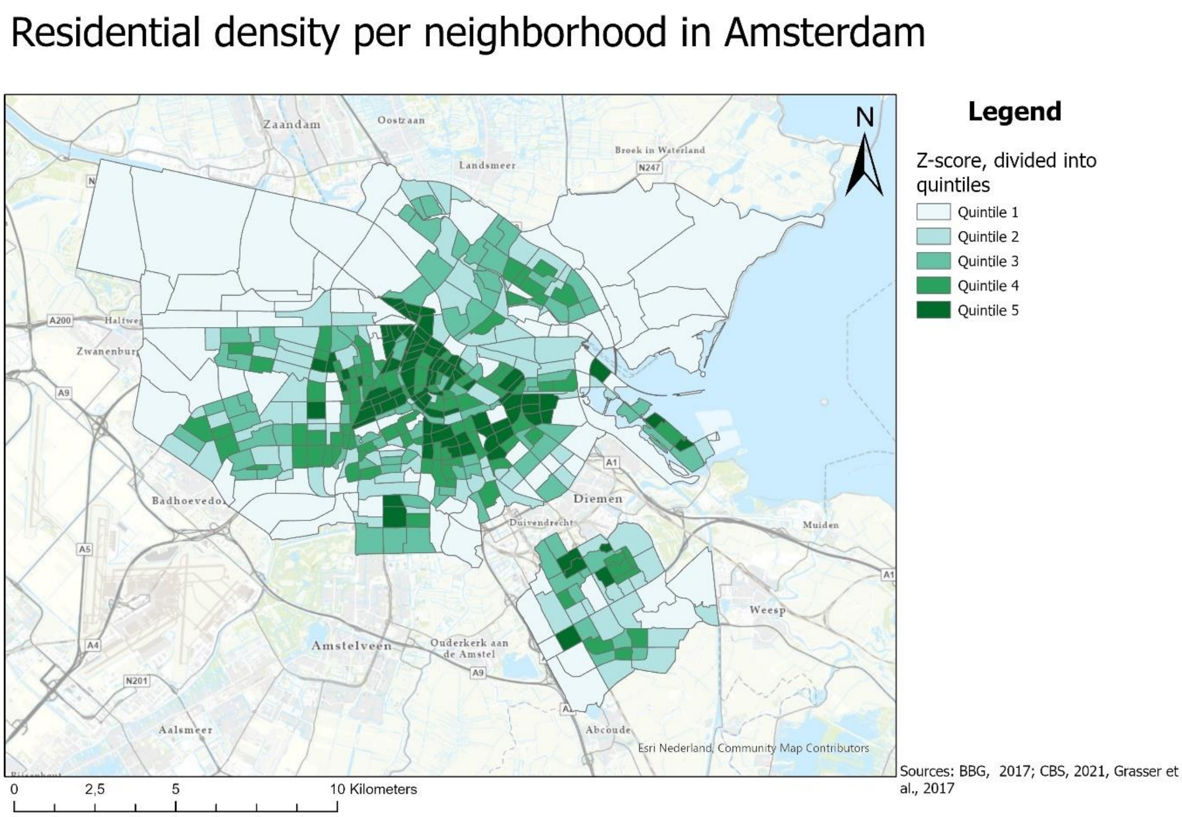


Figure 19: Map showing the SA performed on retail FAR, which is one of the four variables in Frank's WI.

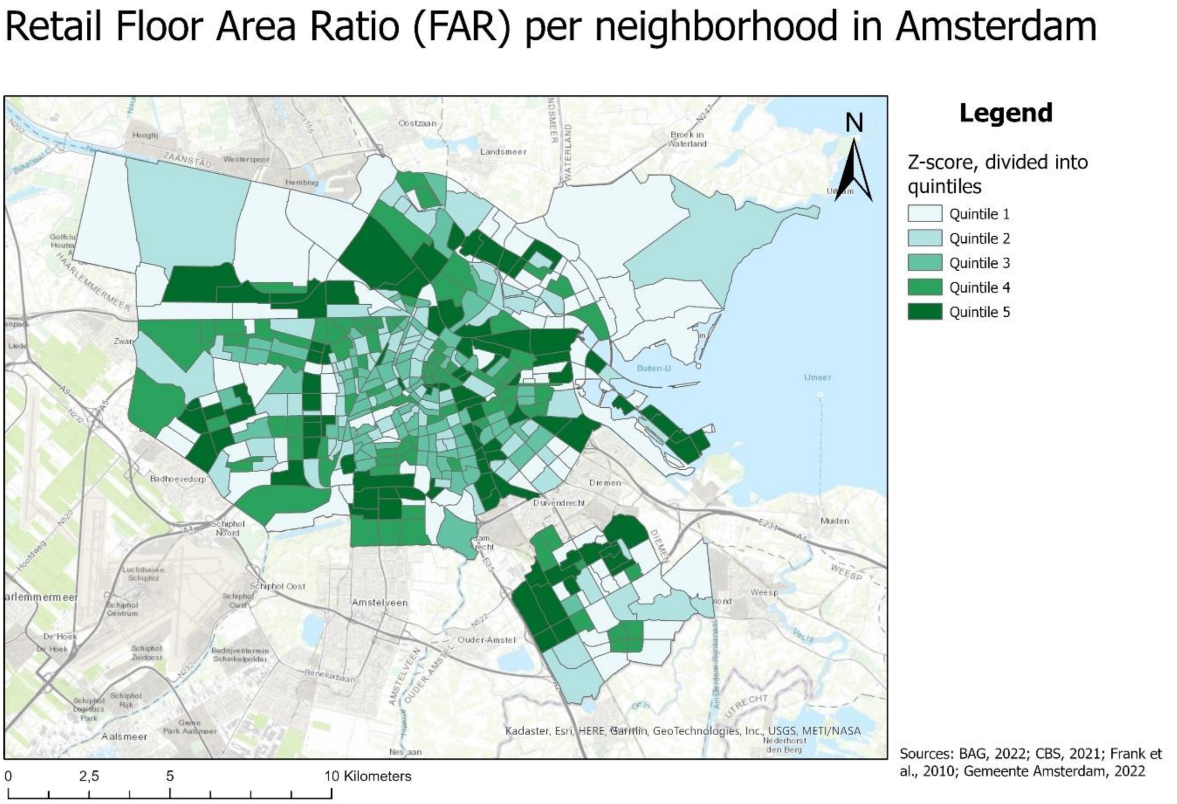


Figure 20: Map showing the SA performed on the Shannon Entropy land use mix, which is one of the four variables in the Frank WI.

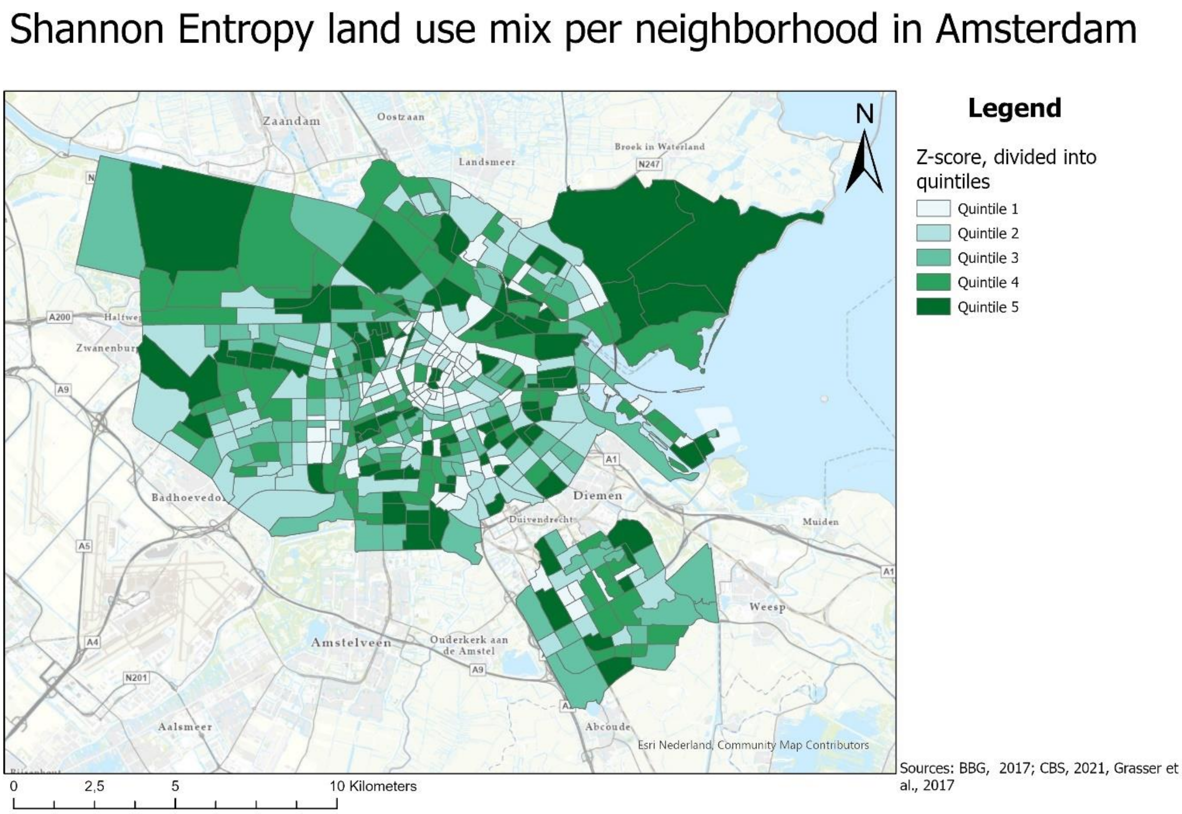


Figure 21: Frank's WI score per neighborhood, divided into quintiles.

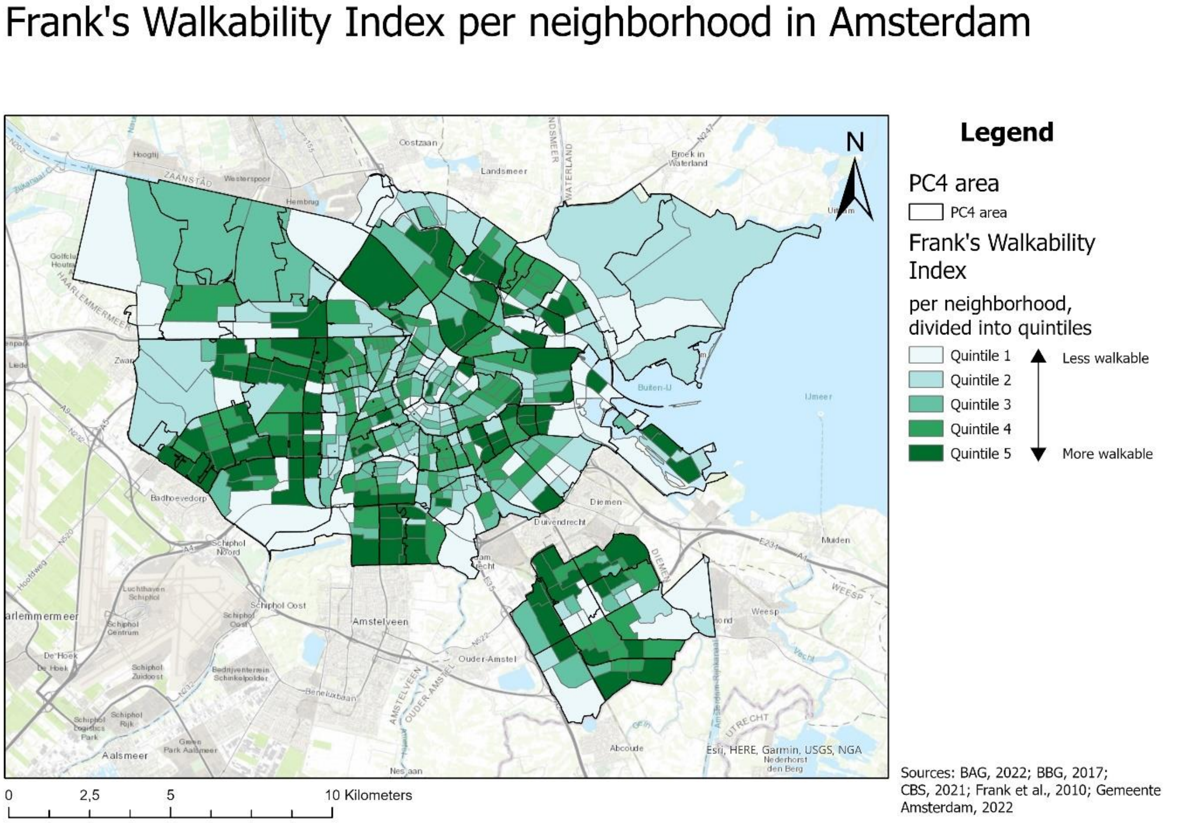


Figure 22: Map showing the SA performed on intersection density, which is one of the three variables in the Graz WI.

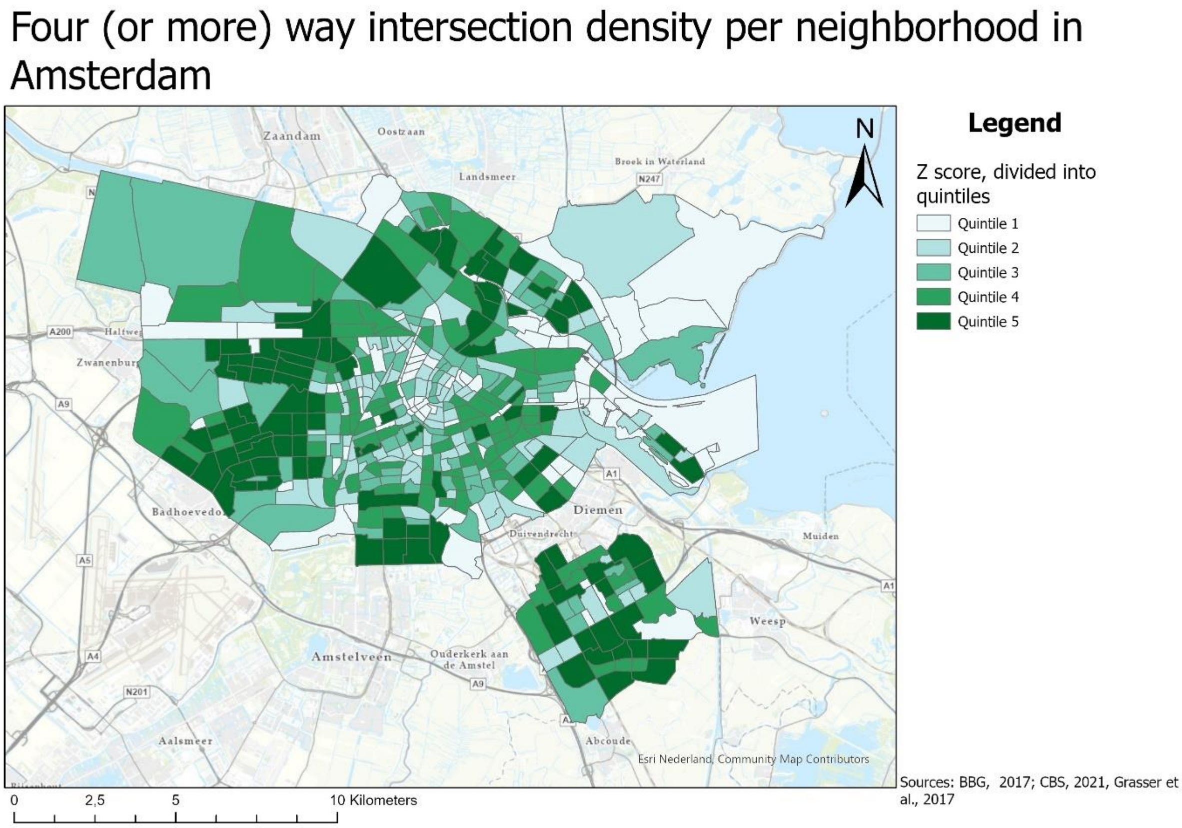


Figure 23: Map showing the SA performed on residential density, which is one of the three variables in the Graz WI.

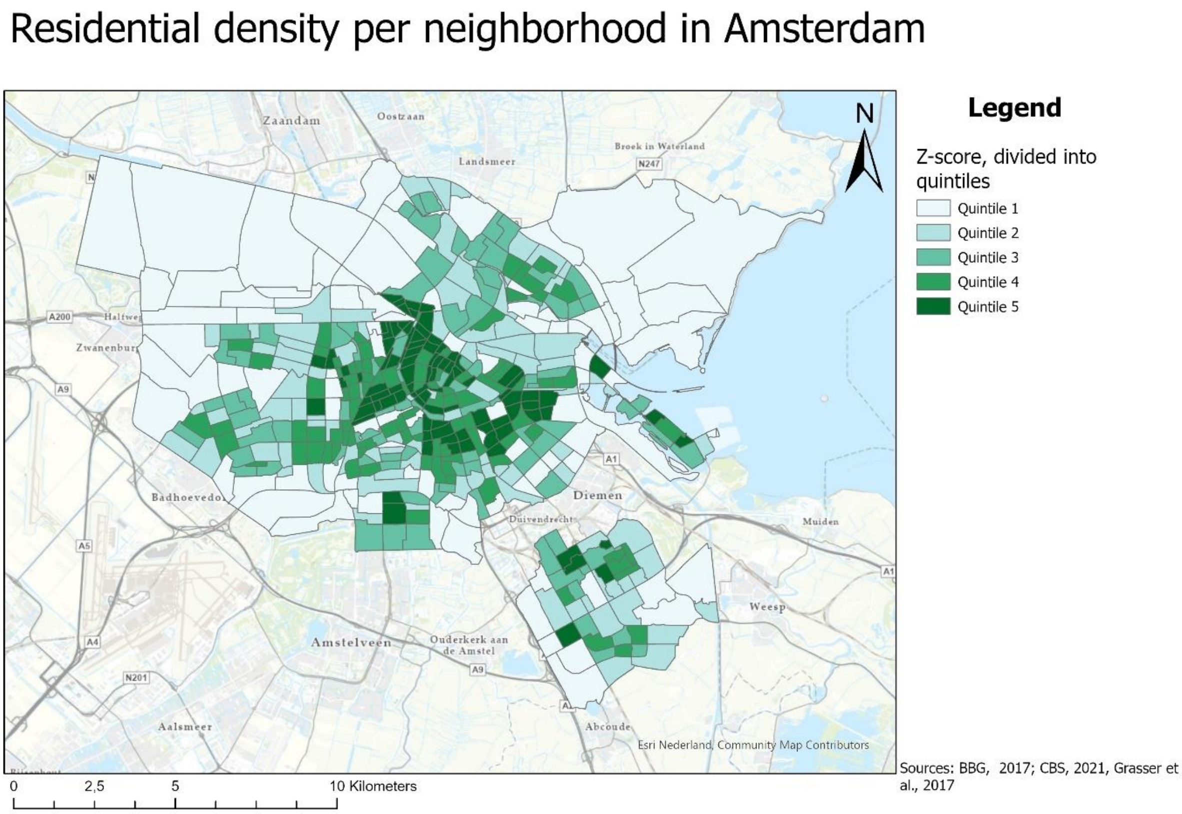


Figure 24: Map showing the SA performed on the Shannon Entropy land use mix, which is one of the three variables in the Graz WI.

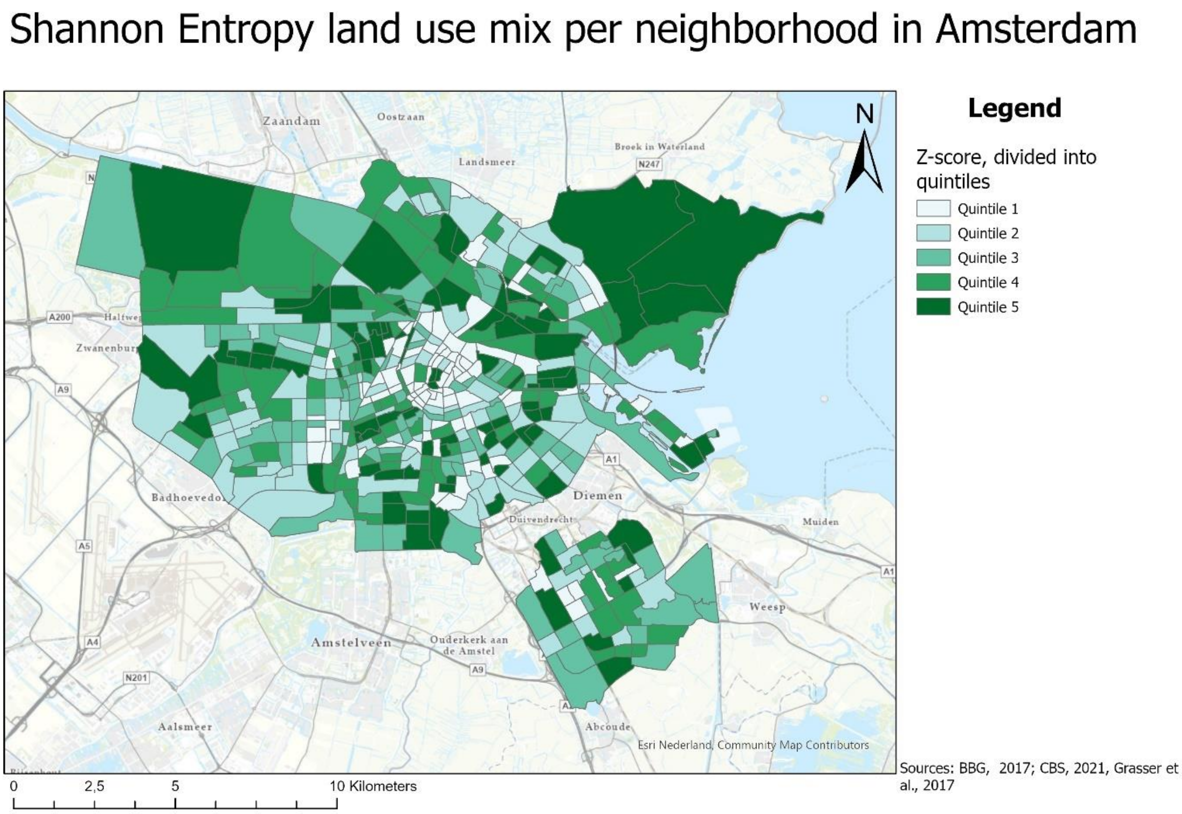


Figure 25: The Graz WI score per neighborhood, divided into quintiles.

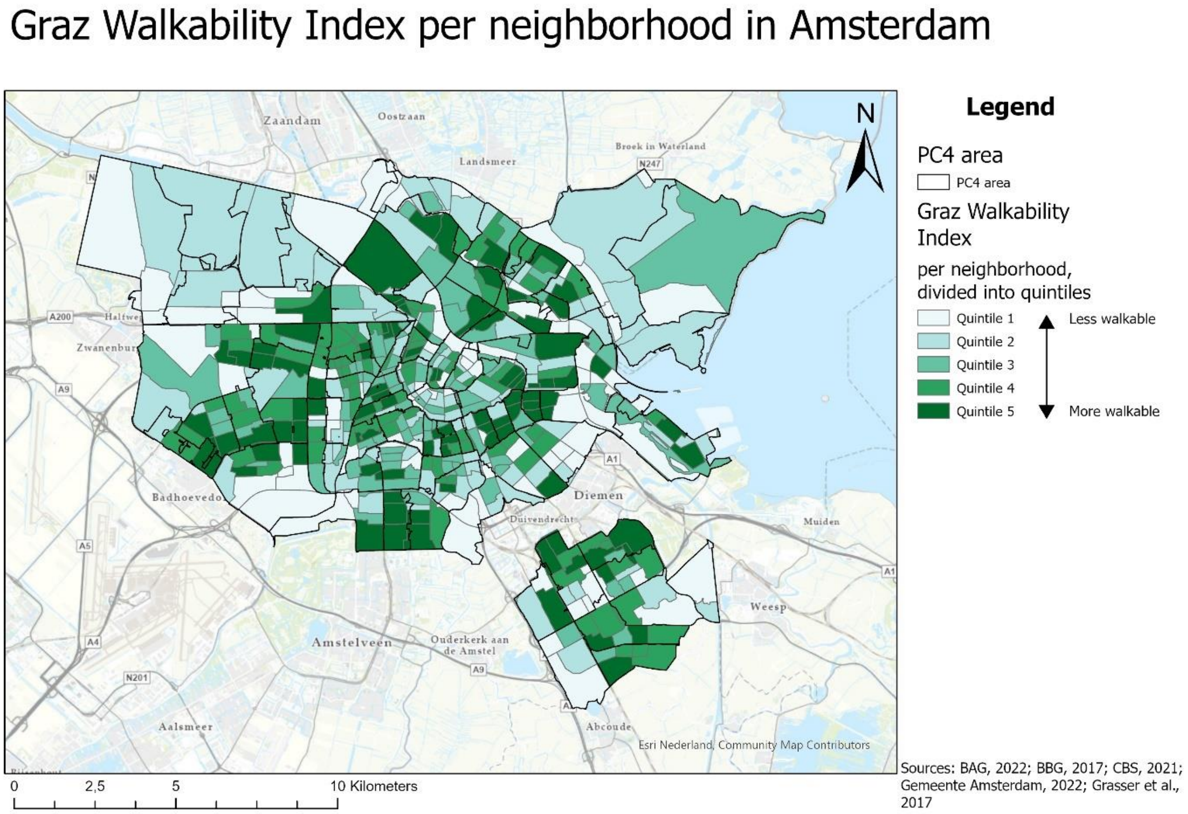


Figure 26: Map showing the SA performed on the residential density, which is one of the four variables in the Urban WI.

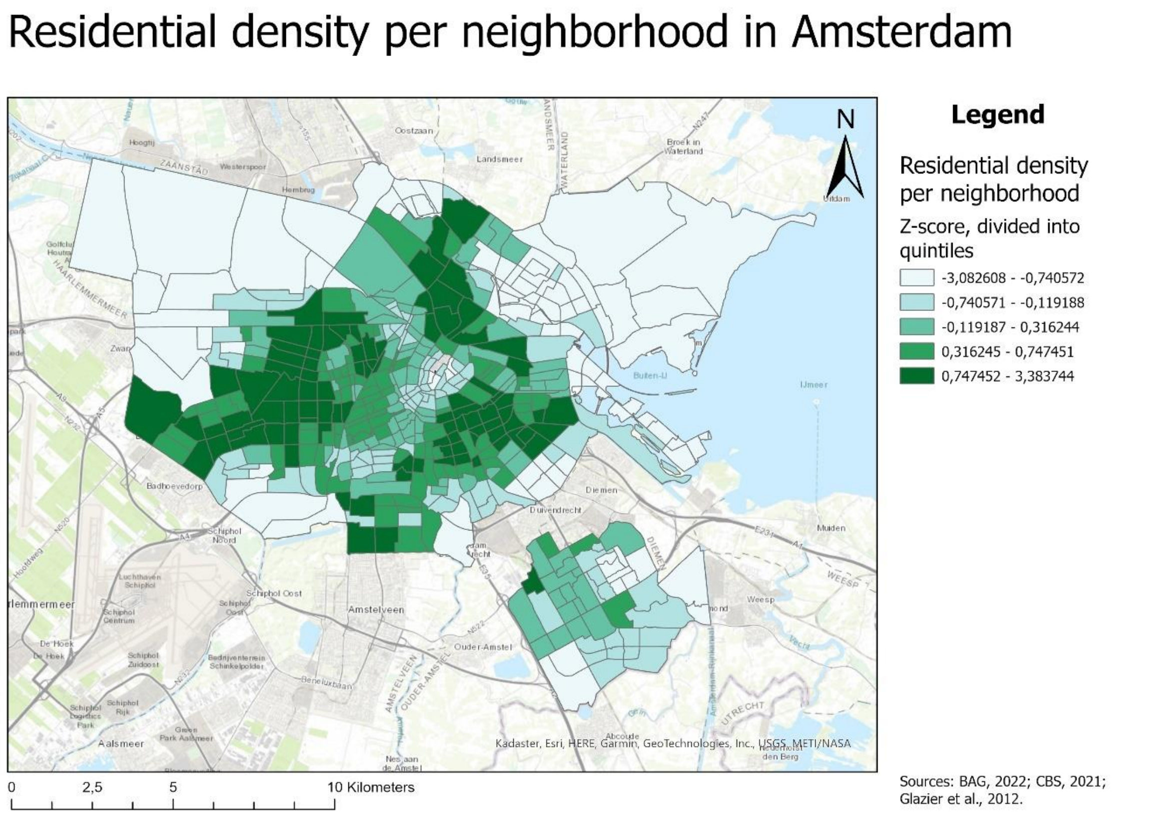


Figure 27: Map showing the SA performed on the residential density (in the neighborhood centroid buffer), which is one of the four variables in the Urban WI.

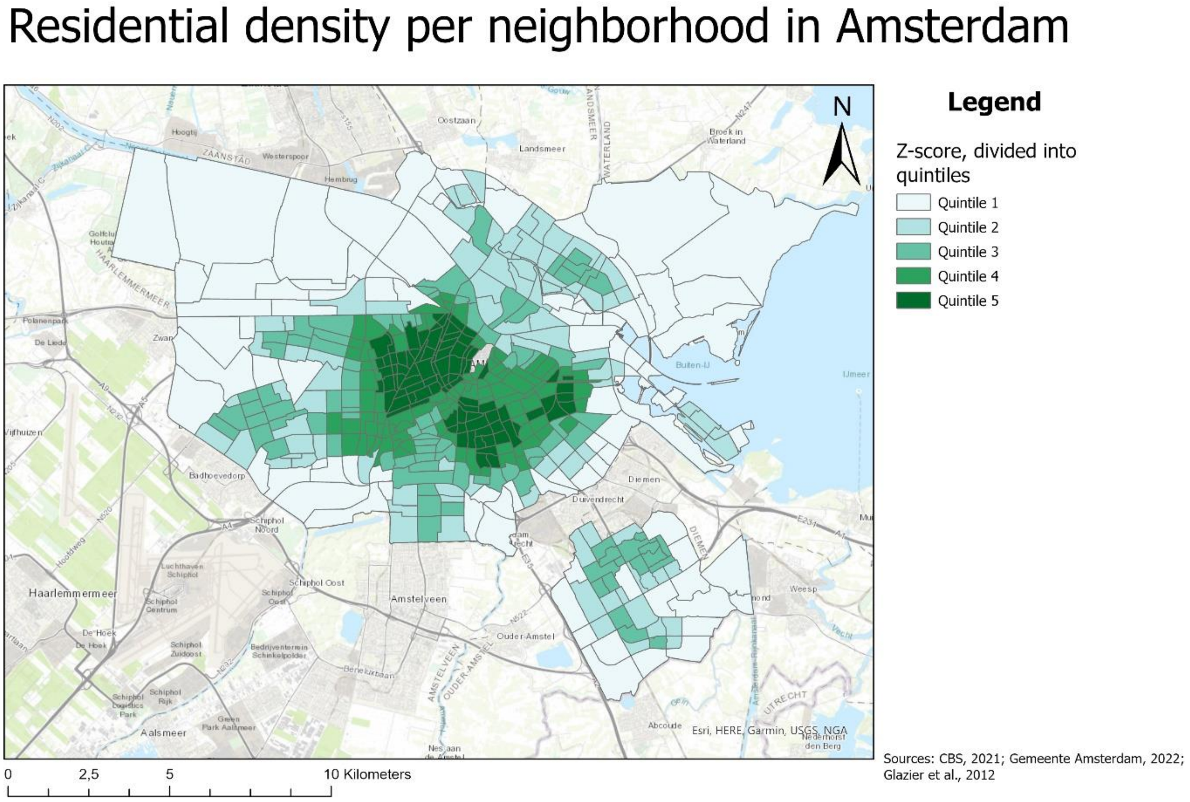


Figure 28: Map showing the SA performed on population density, which is one of the four variables in the Urban WI.

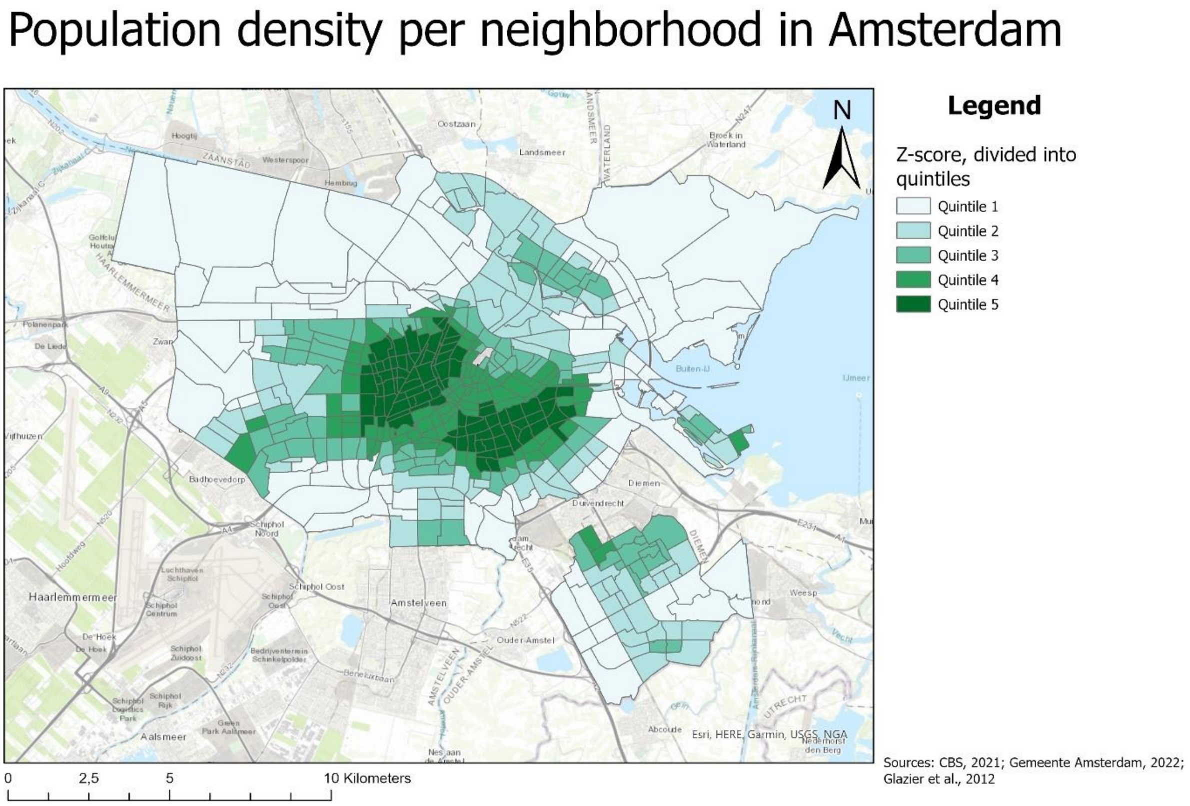


Figure 29: Map showing the SA performed on intersection density, which is one of the four variables in the Urban WI.

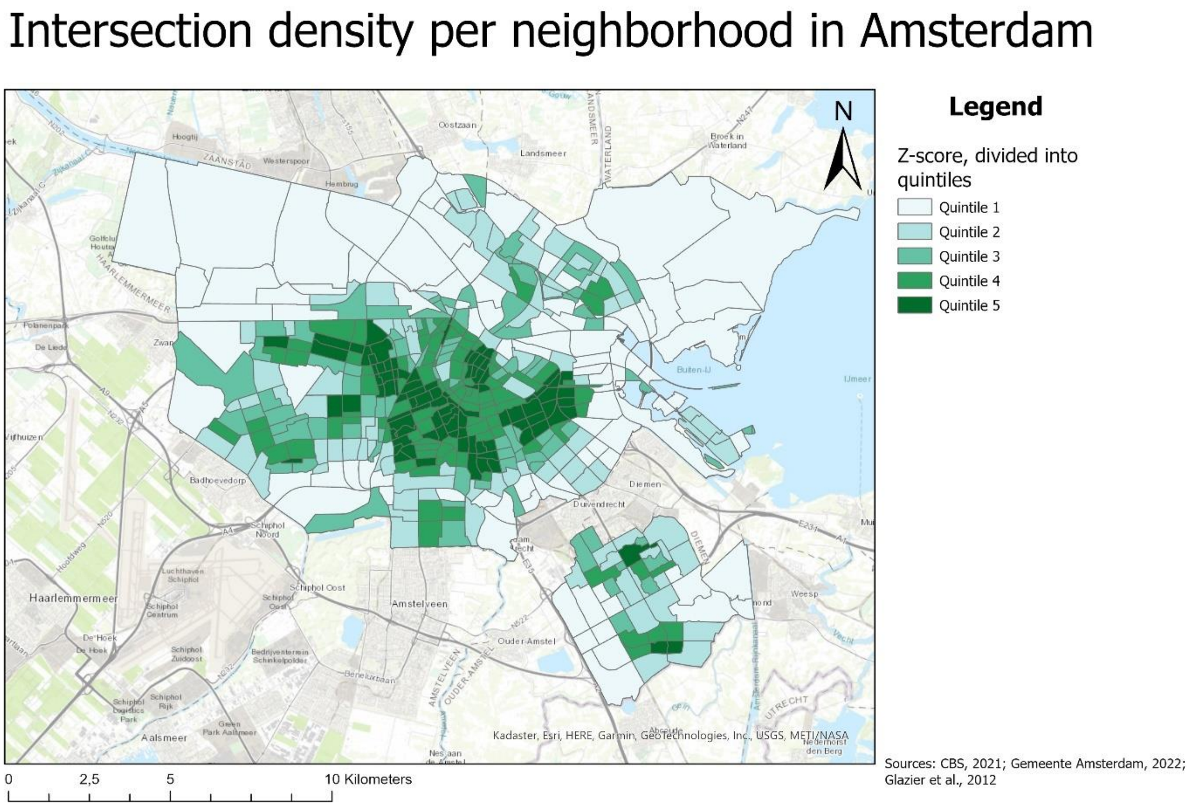


Figure 30: Map showing the SA performed on the availability of retail and services, which is one of the four variables in the Urban WI.

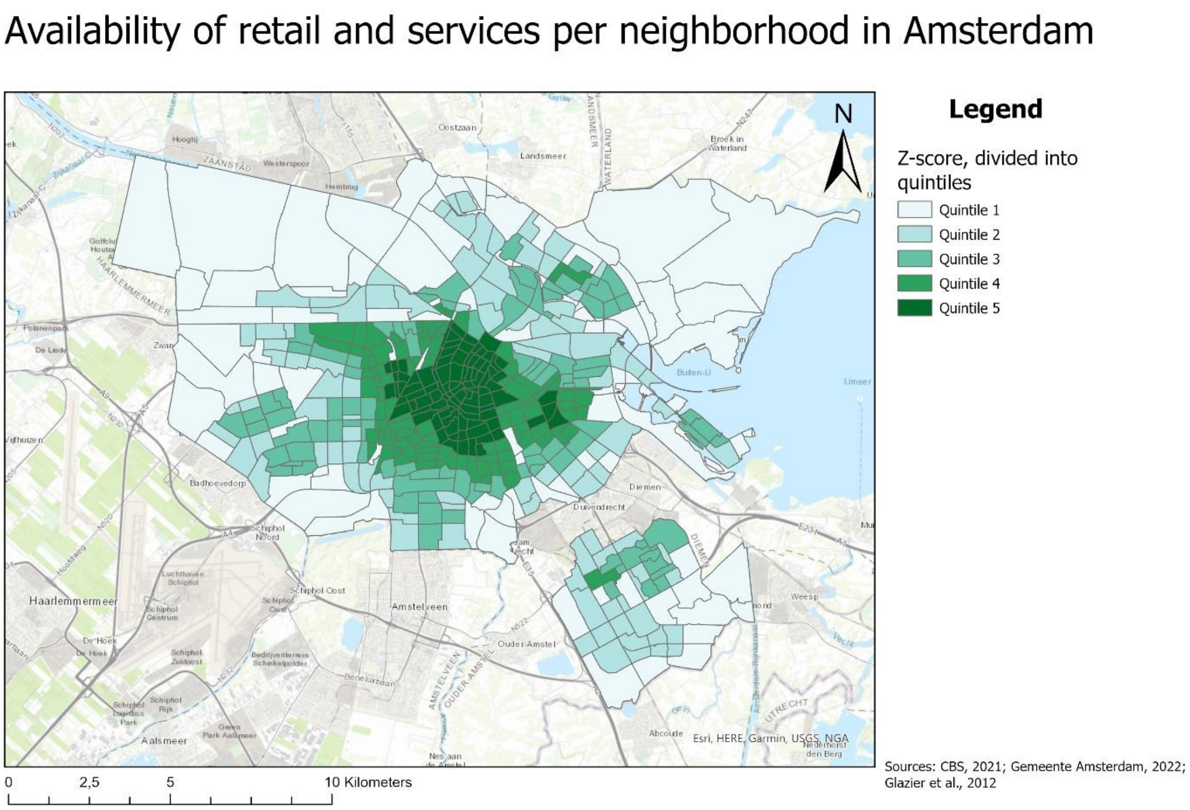
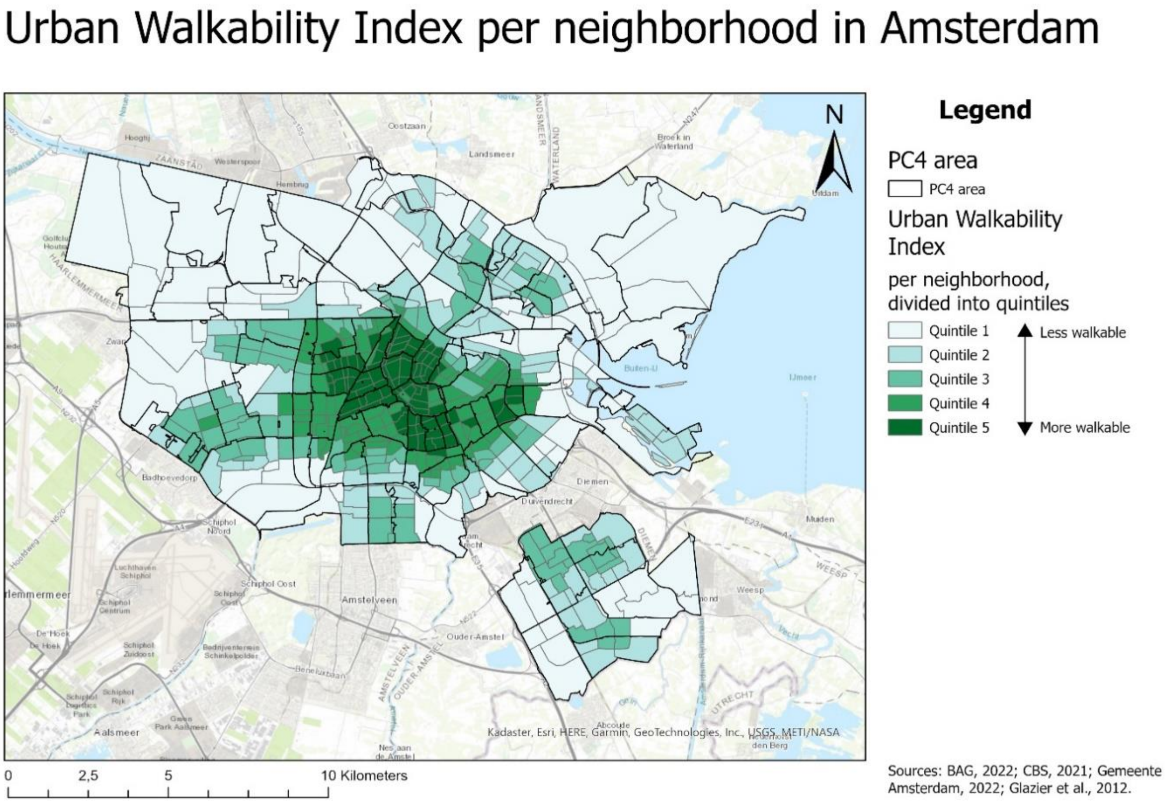


Figure 31: The Urban WI score per neighborhood, divided into quintiles.



Regression analysis

Figure 32: The newly constructed Amsterdam Walkability Index per PC4 area in Amsterdam.