ROYAL INSTITUTE
OF TECHNOLOGY

Using GIS to measure walkability: A Case study in New York City

Razmik Agampatian

Master's of Science Thesis in Geoinformatics
TRITA-GIT EX 14-002

School of Architecture and the Built Environment
Royal Institute of Technology (KTH)
Stockholm, Sweden

April 2014

Abstract

Obesity has become a global epidemic due to changes in society and in behavioral patterns of communities over the last decades. The decline in physical activity is one of the major contributors to the global obesity epidemic. Thus programs, plans and policies that promote walking could be a possible solution against obesity and its comorbidities. That is because walking is the simplest and most common form of physical activity among adults, regardless of age, sex, ethnic group, education or income level.

The characteristics of the built environment might be significant factors that affect people's decision to walk. Thus, measurable characteristics can assist in determining the extent to which the built environment affects the people. These characteristics can also provide indirect evidence of the state of population health for the area under study. Towards the analysis and assessment of potential associations between a number of measures of the built environment and walking, Geographic Information Systems have an increasing acceptance. Composite measures, also known as Walkability Indices, are a promising method to measure the degree to which an area provides opportunities to walk to various destinations.

The main objective of this research is to develop a method to model walkability drawing partially from previously developed Walkability Indices and walkability measures, and suggest eventually an improved Walkability Index composed of 6 parameters. These are: i) Residential Density, ii) Diversity – Entropy Index, iii) Connectivity, iv) Proximity, v) Environmental Friendliness, vi) Commercial Density – FAR. The chosen spatial unit of analysis is the Census Tract level. The method of buffering that defines spatial units around geocoded locations at a given distance is also employed in an attempt to suggest an improvement of previously used methods. The study area is New York City (NYC).

The results imply that Manhattan is the most walkable Borough, while Staten Island is the least walkable Borough. It is also suggested that NYC has a centripetal structure, meaning that the historical center and the entire island of Manhattan is more developed, and more walkable, followed by the adjacent areas of the neighboring Boroughs of Bronx, Brooklyn and Queens. The farthest areas of NYC's periphery are consistently found to have the lowest walkability. Additionally, neighborhoods that are extremely homogeneous in terms of land-use and do not include considerable number of commercial parcels score very low. Hence, Census Tracts that are mainly characterized by primarily industrial land-use or contain large transportation infrastructures (e.g. ports, airports, large train stations) or even large metropolitan parks display limited walkability.

The results and findings coincide to a satisfactory extent with the results of previous studies. However, the comparison is simple and barely based on easily observed patterns. As a result, the validity of the new Walkability Index might need further assessment due to limitations and lack of data.

All types of limitations have been identified including limitations in data and in methodology. Suggestions for further research include possible additional parameters that can be employed in our Walkability Indices (e.g. crime rate, and separate parameter for parks and green areas) and further research whether the components of a Walkability Index should be weighted or not. In general, Walkability Indices are promising GIS applications that still need further research and development.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Hans Hauska for his help and constructive comments.

I also gratefully acknowledge the guidance and comments given to me by Dr. Yifang Ban, Professor of Geoinformatics at Dept. of Urban Planning and Environment, KTH Royal Institute of Technology.

I would also like to thank NYC Department of City Planning; NYC Department of Information Technology and Telecommunications; NYC Department of Parks and Recreation; and Metropolitan Transportation Authority for the data provision.

My deepest gratitude goes to my family and my life companion Elena Kiourktsi for their unconditional love and support throughout my life, this thesis would have been simply impossible without them.

Table of Contents

List of Figures

List of Tables

1. Introduction

In 1997 the World Health Organization (WHO) formally recognized obesity as a global epidemic. As of 2008 the estimates regarding obesity were that at least 500 million adults are obese. Among the countries with the highest rates of obesity are the USA, Australia and Canada. The spread of obesity epidemic can be attributed to the profound changes in society and in behavioral patterns of communities over the last decades, leading especially in developed countries to less physically demanding activities (WHO, 2007). Numerous studies have described the interrelation between physical activity and improved health, indicating that moderate but regular physical activity is maybe the best method against obesity and its comorbidities (Bassuk and Manson, 2005).

Although the mortality rate of obese adults may not differ significantly from that of adults with a normal weight, the higher risk of a number of noncommunicable diseases due to obesity eventually contributes significantly to the total health burden and leads to reduced life expectancy (de Lusignan et al., 2006). Another important aspect of obesity is the economic consequences. For example, in the USA annual direct medical expenditures attributable to obesity are estimated to be as high as $147 billion per year (Finkelstein et al., 2009). As a result, the total annual economic costs associated with obesity might eventually exceed $215 billion.

Walking is the simplest and most common form of physical activity among adults, regardless of age, sex, ethnic group, education or income level (Saelens et al., 2003). Thus, a potential solution to deal with obesity epidemic could be programs, plans and policies that promote walking (Curran et. al. 2006). The characteristics of the built environment might be a significant promoting factor for people’s decision to walk, as those characteristics are related to travel patterns, primarily by impacting proximity between destinations, and directness of travel between these destinations. Understanding the potential impact of the built environment on walkability requires relevant, easy-to-comprehend, and reliable measurable characteristics (Brownson et. al. 2009). These characteristics can show the effect of the built environment to the people. In other words, they can provide indirectly evidence of the state of population health for the area under study.

The increasing use of Geographic Information Systems (GIS) has made them an essential part of health research. Nowadays, GIS techniques are being utilized more frequently by the public health sector. In particular GIS has been used to assess potential associations between a number of built environment characteristics and walking (Butler et al., 2011). GIS takes into account the physical location of areas, boundaries, people, and services, as well as types of land use and natural features. Thus, the spatial approach of GIS can facilitate researchers by providing them the ability to create maps, measure distances and travel times, as well as define the extent and nature of spatial relationships.

One of the latest advancements in GIS methods and techniques that are used to measure walkability is the employment of composite measures called Walkability Indices (WI). WIs are expected to measure the degree to which an area provides opportunities to walk to various destinations (Manaugh et al., 2011). This is achieved by measuring both form and content of neighborhoods, which eventually captures the interrelation of various built environment characteristics, minimizes the effect of spatial collinearity and facilitates the communications of results (Brownson et al., 2009). Composite measures of walkability are also expected to provide more consistent predictors of walking behavior than single component measures (Vargo et al., 2011).

1.1 Research Objectives

The main objective of this research is to develop a new method to model walkability drawing partially from previously developed WIs and walkability measures and to suggest eventually an improved Walkability Index. A common method to construct measures of the built environment is to use pre-defined spatial units (i.e. Census Tracts). Another common method is the use of buffers to define spatial units around geocoded locations at a given distance. In the present study the method of pre-defined spatial unites is basically used. Yet the method of buffering is also integrated in the aforementioned method of pre-defined spatial unites, in an attempt to suggest a hybrid method as an improvement of previously used methods.

The study area is the entire area of New York City (NYC). Previous attempts to calculate walkability in NYC have been made from various studies and projects, such as the projects WalkScore (WalkScore, 2011) and Walkshed (Azavea, 2010). However, WalkScore and Walkshed calculate walkability based mainly on proximity and do not take into account other environmental characteristics. While in our method, we suggest a different approach and we attempt to calculate walkability employing a new Walkability Index. The resulting walkability map of the present study is going to be compared and contrasted with the suggested walkability maps of the aforementioned projects in a simple attempt to validate the results of this study.

1.2 The Structure of the Thesis

A short introduction to the thesis is presented in Chapter 1, followed by a review of some of the commonly used measures of walkability in Chapter 2. The study area and the various data used in this thesis are described in Chapter 3. The walkability index of this study and all six components are identified, defined and discussed with an example to clarify the key concepts in chapter 4. In chapter 5, at first the six individual components are addressed independently of the composite measure to have a better insight of the contribution of each parameter to the final result, and then the general results of this study's Walkability Index are presented and discussed more thoroughly. A short conclusion is drawn in Chapter 7.

2. Background

In this chapter, the context of walkability is going to be discussed. A comprehensive literature review that addresses relevant issues such as obesity, physical activity and measures of walkability is going to be presented in order to acquire a better understanding of the interrelation of all the aforementioned terms.

This chapter is structured in a way that allows the reader to acquire a general understanding of obesity, its negative repercussions on public health, as well as its social and economic burden in general. The review begins by presenting an overall picture of obesity epidemic in the US and it gradually focuses on New York City (NYC), where obesity is also a problem. The study area of this project is New York City. The following section discusses the benefits of physical activity as a possibly effective solution, and the potential of promoting walking in reducing obesity rates. Then, the term walkability is introduced, followed by a thorough presentation of selected walkability measures. The final section is dedicated to the presentation of previous relevant studies and project for the case of NYC.

2.1 Obesity

Obesity is commonly measured by Body Mass Index (BMI) and people are considered obese when their BMI is equal or higher than 30. The Body Mass Index is a measure of an adult's weight in relation to his or her height, in particular the adult's weight in kilogram divided by the square of his or her height in meters (kg/m^2). For children, the BMI is calculated similarly with the BMI of adults, but no thresholds are set when considering overweight or underweight children. Instead, children's BMI is compared to typical values for other children of the same age (WHO, 2003).

Obesity was formally recognized as a global epidemic by the World Health Organization (WHO) in 1997 and as of 2008 the estimates regarding obesity were that at least 500 million adults are obese. The USA is among the countries with the highest rates of obesity. Obesity used to be considered a health issue of the developed countries. Now, the only remaining region of the world where obesity is not common is sub-Saharan Africa (Haslam et al, 2005).

Obesity rates in the United States have increased significantly over the last two decades. The spatiotemporal change of obesity trends are given according to the United States Centers for Disease Control and Prevention, as presented in Figure 2.1 (CDC, 2012). From Figure 2.1, it can easily be deduced that there is a rising epidemic of obesity from 1990 to 2010. While in 1990 obesity rates were lower than 15%, in 2000 they increased over 15%, and in 2010 obesity rates increased over 20% and in some states they even exceeded 30%.

According to a report for the US National Center of Health Statistics, more than one-third of adults (35,7%) and almost 17% of youth were obese in 2009–2010 (Ogden et al., 2012). Although there are signs of slowing or even leveling off of the increase of the rates (Flegal et al, 2010) it is important to continue tracking obesity and its social, economical and health impacts, because of the high prevalence of obesity.

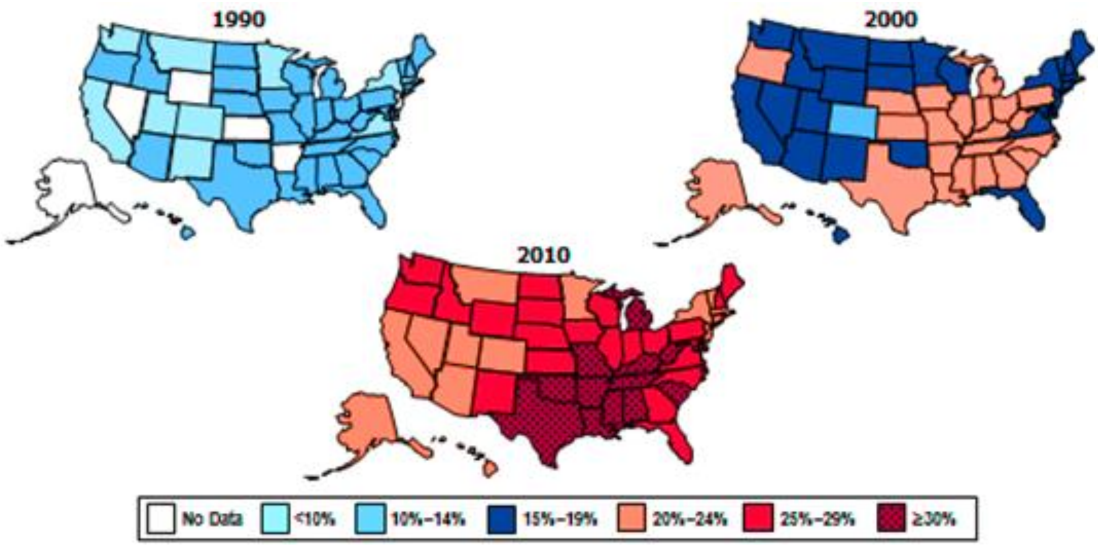


Figure 2.1 Obesity Trends (BMI>30) Among U.S. Adults (after CDC 2012)

2.1.1 Impacts of Obesity on Health

Many epidemiological studies have consistently shown that obesity is associated with increased risks of morbidity, disability and mortality (Visscher et al., 2001). A study revealed that the impact of obesity on mortality is nearly as important as that of cigarette smoking (Peeters et al., 2003). Generally, obesity increases the risk of a number of health conditions, the more important of which are several cardiovascular diseases, various types of cancer, adverse lipid concentrations, and type 2 diabetes. In fact, it is the severity and the duration of obesity that contribute to the risk of comorbidities. Although the mortality rate of obese adults may not differ significantly from that of adults with a normal weight, the higher risk of a number of noncommunicable diseases due to obesity eventually contributes significantly to the total health burden and leads to reduced life expectancy (de Lusignan et al., 2006).

Even obese children show raised levels of susceptibility for many of the aforementioned diseases (WHO, 2007). For example Cook et al (2003) showed that 4% of adolescents and approximately 30% of overweight adolescents in the United States met the criteria for the metabolic syndrome, which dramatically increases the possibility to develop type 2 diabetes and get inflicted by cardiovascular diseases in the future.

2.1.2 Economic Consequences of Obesity

According to Thompson and colleagues, obese people (BMI above 30 kg/m2), in comparison with people of normal weight (BMI of 20.0–24.9 kg/m2), had 36% higher annual health care costs, while overweight people (BMI of 25.0–29.9 kg/m2) had 10% higher annual health care costs (Thompson et al., 2001). Numerous studies have attempted to estimate the economic consequences of obesity. However, estimating the total cost of obesity is not a simple task and cost estimates differ among studies, depending mainly on the data and methods employed. Most of the studies describe the medical costs associated with obesity (direct costs), while some also take into account loss of productivity (indirect costs).

Productivity effects may be categorized into at least four different types: absenteeism, presenteeism, disability, and premature mortality. Absenteeism is defined as the productivity costs due to employees being absent from work for obesity-related health reasons and presenteeism as the decreased productivity of employees while at work. Although it is difficult to estimate the total loss of productivity costs, it is assumed to be substantial and probably higher than $66 billion annually for the US (Hammond et al., 2010). Additionally, annual direct medical expenditures attributable to obesity are estimated to be as high as $147 billion per year (Finkelstein et al., 2009). As a result, in the United States the total annual economic costs associated with obesity might eventually exceed $215 billion.

2.1.3 The Case of New York City

According to the New York City (NYC) Community Health Survey of 2011 conducted by the NYC Department Of Health and Mental Hygiene (DOHMH) almost one in four New Yorkers is obese, and more than half of the population is overweight or obese, as can be seen in Figure 2.2.

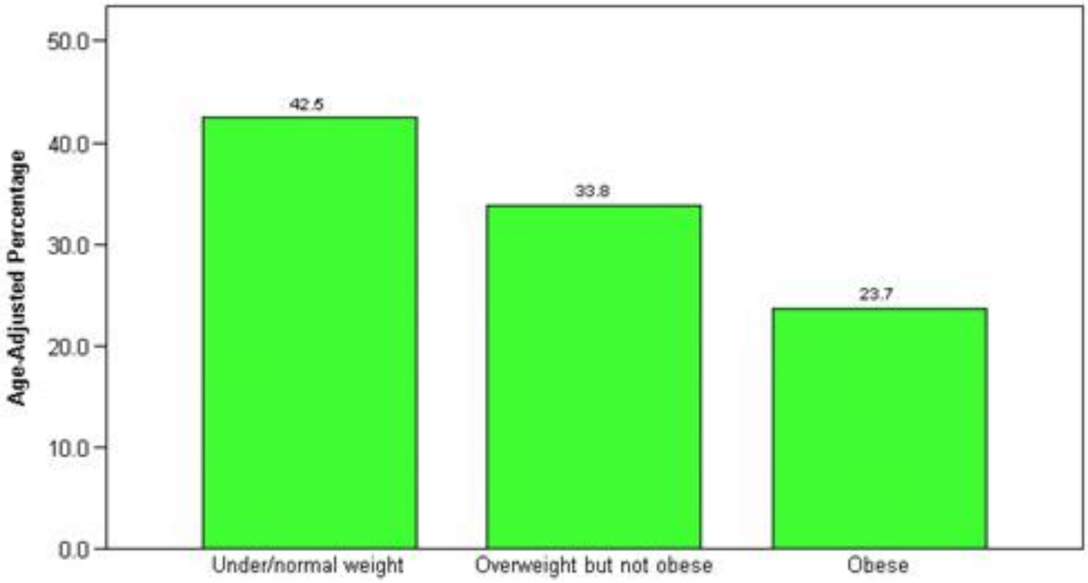


Figure 2.2 Overweight and Obese in NYC in 2011 (after NYC DOHMH, 2013)

Among children, obesity rates are even higher – almost 33% of the children in NYC are obese (NYSDOH, 2011). Since 2002, when the NYC DOHMH started releasing the results of NYC community health surveys, obesity rates in NYC have increased steadily, as revealed in Figure 2.3. In 2002, the obesity rate was 18,2% in 2002, but in 10 years it has increased to 23,7%.

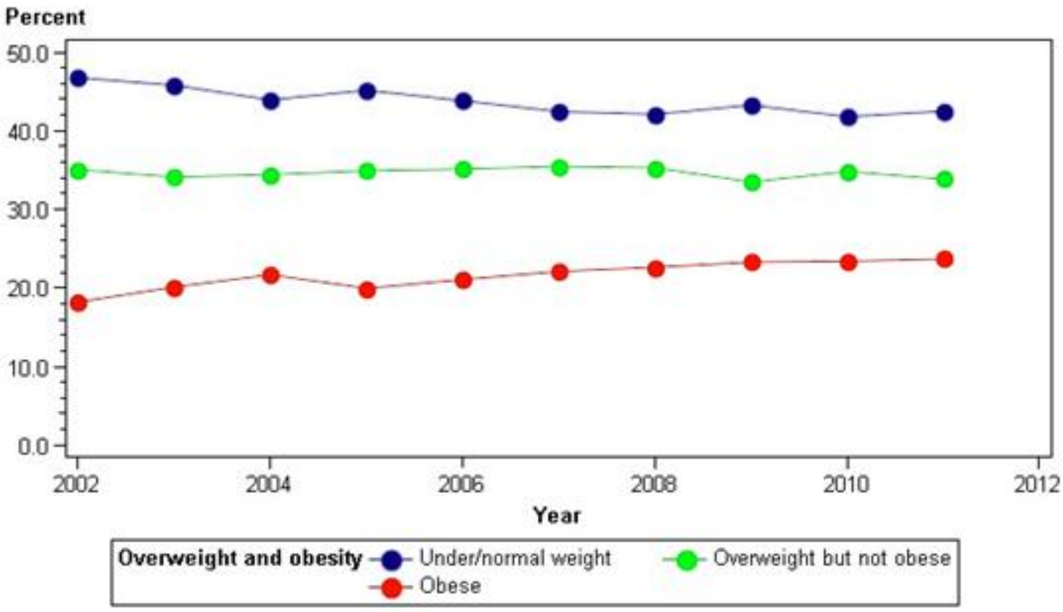


Figure 2.3 Overweight and obesity by Year - Trends in NYC (after NYC DOHMH, 2013)

On the other hand, the percentage of those who are overweight but not obese has decreased by 1,2. In 2002 it was 35% and in a decade dropped to 33,8%.

The spatial distribution of obesity in NYC as presented in the choropleth map of Figure 2.4 is interesting as it will be discussed later on. It can be easily seen that obesity rates are fairly low in Manhattan and quite high in Bronx.

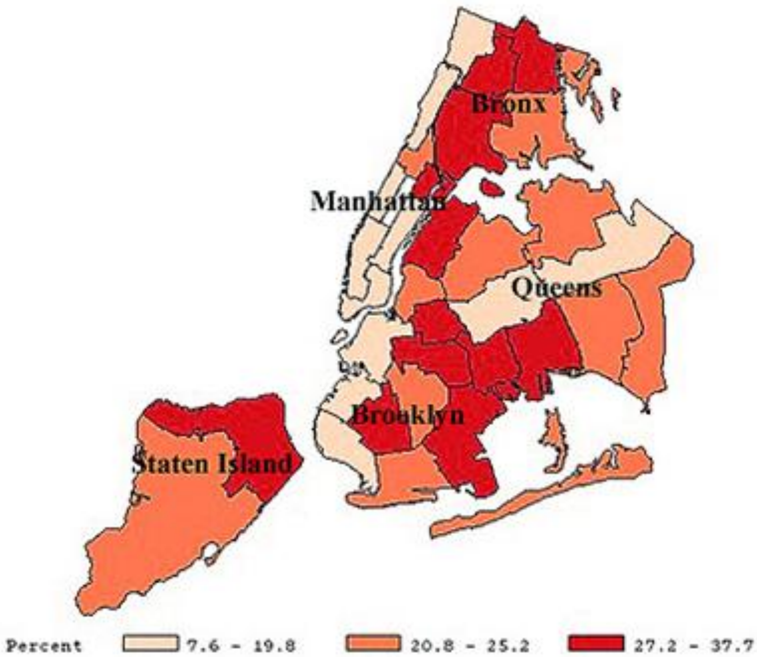


Figure 2.4 Percentage of obese adults by neighborhood of NYC in 2011 (after NYC DOHMH, 2013)

The New York State Department of Health estimated that New York State Health Care System needs more than $7.6 billion every year to treat obesity-related illnesses and conditions. The cost to care obesity-related health problems is expected to reach $136.3 billion over the 10-year period from 2011 to 2020 (The Lewin Group, 2010).

2.1.4 Benefits of Physical Activity on Health

The spread of the obesity epidemic can be attributed to profound changes in society and in behavioral patterns of communities over the last decades. While genes are important in determining a person's susceptibility to weight gain, economic growth, modernization, urbanization and globalization of food markets are some of the forces believed to underlie the epidemic. At the same time, considerable changes have been observed especially in developed countries towards less physically demanding activities. The increased use of motorized transport, technology in the home, and more passive leisure pursuits is another fact that also leads to less physical activity (WHO, 2007).

A large number of studies have described the interrelation between physical activity and improved health, providing adequate evidence that moderate but regular physical activity is maybe the best method against obesity and its comorbidities. For example, Bassuk and Manson (2005) have found that physically active persons have a significantly lower risk of developing heart disease and type 2 diabetes. According to the World Health Organization (WHO, 2010) physical inactivity is the fourth-leading risk factor for global mortality and is responsible for 6% of deaths globally – around 3.2 million deaths per year. In another report of WHO it is clearly stated that (WHO, 2006):

"...it is recommended that individuals engage in adequate levels [of physical activity] throughout their lives. Different types and amounts of physical activity are required for different health outcomes: at least 30 minutes of regular, moderate-intensity physical activity on most days reduces the risk of cardiovascular disease and diabetes, colon cancer and breast cancer. Muscle strengthening and balance training can reduce falls and increase functional status among older adults. More activity may be required for weight control."

2.2 Walking: Simple But Healthy

Walking is the simplest and most common form of physical activity among adults, regardless of age, sex, ethnic group, education or income level (Saelens et al., 2003). It is a rather inexpensive form of exercise, does not require learning new skills and can also be used for transportation purposes. Besides, walking is the most sustainable form of transportation per se, as it contributes to reductions in air pollution and has the potential to reduce the rates of respiratory diseases associated with air pollution, by reducing reliance on the automated transport at the same time (Frank et al., 2007).

Walking promotes also social life and public participation by providing opportunities for face-to-face contact and casual interaction, all of which subsequently are proven to improve mental health and well-being (Robertson et al., 2012). A vibrant, economically viable and safe community needs people on streets and in public places. A walkable environment can also provide significant health benefits and independence to specific groups such as children and third age people who rely more on their local neighborhoods (Berke et al., 2007).

2.2.1 Walkability and Neighborhood Characteristics

Neighborhood characteristics might significantly affect people's decision to walk because those characteristics are related to travel patterns basically by impacting directness of travel between destinations and proximity between these destinations. For example, when common destinations such as shops, grocery stores, post offices, schools and daycare stations are situated within the close vicinity of a neighborhood, people are more likely to prefer to reach their destinations on foot or by bicycle, instead of driving or be driven. Besides, a neighborhood characterized by higher population densities tends to support a richer variety of shops and services in the neighborhood, while there is a higher possibility for increased ridership and higher quality transit, encouraging people to walk to and from transit stops. In the same way, a residential area that has a street network of reduced traffic speeds tends to become more walkable as it becomes more pedestrian-friendly. Those are examples of some neighborhood features that alone or in combination can contribute to the walkability of a neighborhood (Tomalty et al., 2009).

Understanding the potential impact of the built environment on walkability requires relevant, easy-to-comprehend, and reliable measurable features (Brownson et. al. 2009). These measurable characteristics can assist in determining how much the built environment affects the people. These measures can also provide indirectly evidence of the state of population health for the area under study.

Over the last 2 decades, there has been considerable progress regarding measuring walkability, various different measurable features of the built environment have been incorporated to models, and different approaches have been developed.

The first method is based on interviews or self-administered questionnaires. Questionnaires can potentially reveal the extent to which individuals perceive various elements of the built environment and how a person experiences a neighborhood. This method is considered as "subjective" because two unique individuals may perceive the same environment differently. The most commonly assessed measurable environmental features of perception are land use, traffic, aesthetics, and neighborhood safety from crime (Tomalty et al., 2009; Brownson et al., 2009).

The second approach uses built environment characteristics obtained by systematic observations or audits that quantify the environmental attributes of an area, including the presence or absence of features hypothesized to affect physical activity.

Audit tools are used for measuring and assessing physical features through direct observation, and include one or more measurable characteristics, such as: land use (e.g. commercial space), streets and traffic (e.g. pedestrian crosswalk), sidewalks (e.g. presence, width, and continuity of sidewalks), public space/amenities (e.g. presence of benches), architecture or building characteristics (e.g. building height), parking/driveways (e.g. presence of parking lot(s)), maintenance (e.g. presence of litter), and indicators related to safety (e.g. presence of graffiti) (Brownson et al., 2009; Pelletier, 2009; Gauvin et al., 2005).

Finally, the third method uses geospatial databases and Geographic Information Systems (GIS) in order to assess or develop relevant indicators that measure walkability. A GIS "is a facility for preparing, presenting, and interpreting facts that pertain to the surface of the earth. This is a broad definition… a considerably narrower definition, however, is more often employed. In common parlance, a Geographic Information System or GIS is a configuration of computer hardware and software specifically designed for the acquisition, maintenance, and use of cartographic data." (Tomlin, 1990).

The increasing potential of GIS has made it an essential part of health research and GIS techniques are being utilized more frequently by the public health sector. In particular GIS has been used to assess possible associations between a number of built environment features and physical activity or walking (Butler et al., 2011). GIS takes into account the physical location of areas, boundaries, people, and services, as well as types of land use and natural features. Thus, the spatial approach of GIS can facilitate researchers by providing them the ability to create maps, measure distances and travel times, as well as define the extent and nature of spatial relationships. In the following sections a number of walkability measures and GIS techniques on walkability is going to be presented and discussed.

2.2.2 Measurable Neighborhood Characteristics

The measurable neighborhood characteristics can be categorized into:

- Density
- Diversity
- Proximity
- Connectivity
- Environmental Friendliness
- Walkability Indices (WI)

Walkability Indices (WI) are a distinct category of composite measures that measure walkability by incorporating a number of various parameters.

Density

Density or compactness is defined as the amount of activity found in an area and can be measured in terms of population, housing unit, or employment density. High density implies compact land development, reduced travel distances between departure sites and destination sites, and decreased dependence on motorized transportation (Feng et al., 2010). Hence, density is considered as an essential measure which is highly correlated with walking. Population density is one of the most commonly cited measures in the literature. Gross population density (population per total land area) and net residential density (e.g. residential units per residential acre) are also other commonly used measures of density (Brownson et al., 2006). Generally, it is recommended that, where possible, net density should be preferred, as it excludes other land uses. That being said, residential density is important because it serves as a proxy for other urban form factors, and is especially important when measuring in larger geographic scales or in cases of insufficient data (Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009). Finally, the retail floor area ratio (FAR), also known as commercial density, is an alternative measure of density that can also be used as an indicator of walkability and in conjunction with land use mix (LUM). Table 2.1 shows some selected measures of Density. For each measure there is a summarized description of the rationale behind the use and application of each measure, and a brief example of how these measures are calculated.

Table 2.1 Measures of Density

1. Population Density

Population Density can potentially affect walkability positively as it implies that the more densely populated an area is the higher levels of Land Use Mix (LUM), and urbanization and centralization are expected to occur (Tomalty et al., 2009; Saelens et al., 2003; Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009).

Example of measure:

- Crude = no. of people per unit area.
- Net = no. of people per unit residential area.

2. Household Density

Household Density can potentially affect walkability positively. High Household Density implies that more commercial and recreational areas are expected to be built in the vicinity of dense housing areas (Saelens et al., 2003; Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009).

Example of measure:

- Crude = no. of households per unit area.
- Net = no. of households per unit residential area.

3. Employment Density

Employment Density can potentially affect walkability positively. Higher levels of employment are observed in areas with higher development of urban industry. Thus, a nearby population is expected to work at these industries. These industries/jobs are easily accessible by walking, as they are expected to be close to residential areas (Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009).

Example of measure:

- No. of employees per unit area.

4. Retail Density

Retail Density can potentially affect walkability positively when high levels of retail store clusters are observed in an area as this infers that the area is also characterized by residential locations (Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009).

Example of measure:

- No. of Retail locations in a given area.

5. Establishment Density

Higher number of establishments is expected to affect walkability positively as higher numbers of establishments suggest an urbanized area where many locations are accessible on foot (Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009).

Example of measure:

- No. of establishments per unit area.

6. Retail Floor Area Ratio (FAR)

The larger the retail floor area of an area, the more stores a shopping centre will have. Thus, more people are expected to prefer walking, which in turn affects walkability positively (Brownson et al., 2009; Robitaille et al., 2009; Feng et al., 2010).

Example of measure:

- Ratio of building square footage (retail floor area) to land area square footage.

Diversity

The term diversity refers to the spatial arrangement of land use that affects the type and nature traveling. A well-mixed land use supports and offers a large number of various services in the near vicinity of an area, shortening trip distances and making walking more attractive (Fend et al., 2010). Another advantage of mixed land use is that it can provide higher visual variety and interest for pedestrians (Forsyth et al., 2006). Diversity is quite an abstract characteristic to be measured, but it is frequently taken into account when measuring walkability and assessing physical activity. In theoretical terms, a multifunctional environment is expected to reduce travel times between origin and destination, and to improve proximity, which in turn promotes physically active means of transportation (Robitaille et al., 2009). Finally, high diversity is associated with lower car ownership and use, as well as reduced emissions (Song et al., 2005). Table 2.2 shows some selected measures of Diversity. For each measure there is again a summarized description of the rationale behind the use and application of each measure, and a brief example of how these measures are calculated.

Table 2.2 Measures of Diversity
1. Land Use Mix (LUM)
When in a given area there is a rich mix of non-residential zones-usages to residential zones-usages, then a high proportion of people is expected to prefer walking to a diverse number of locations, which in turn affects walkability positively (Tomalty et al., 2009; Tucker et al., 2009; Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • Ratio of non-residential zones-usages to residential zones-usages.
2. Mean Entropy Index
The higher value the Mean Entropy Index of an area has the more diverse the area and the more walkable the area is expected to be, as diversity tends to attract people to walk through the area, thus affecting walkability positively (Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • $Mean\ Entropy\ Index = -\sum_k \frac{\sum_j [\frac{P_j \ln(p_{jk})}{\ln(J)}]}{K}$ (2) k: Number of actively developed hectares within each census tract. P_{jk}: Proportion of land-use type j within a half mile (1/2) radius of development area surrounding the k^{th} hectare.
3. Dissimilarity Index
When in a given area, the groups of residential to non-residential or urban to natural are more evenly distributed, then walkability in the area is expected to be affected positively (Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009; Feng et al., 2010). Example of measure: • $Dissimilarity\ Index = \sum_k \frac{1}{K} \sum_{i=1}^{8} \frac{X_{ik}}{8}$ (3) k: Number of actively developed hectares within each census tract. X_{ik}: 1 if the central active hectare's use differs from that of a neighboring hectare, 0 otherwise. 1/8 point is assigned to each adjacent hectare. The final score is between 0 and 1.

4. Entropy Index
The entropy index depicts the intensity of land-use diversity of a given area and it could affect walkability either positively or negatively. If there is a low level of redundancy (meaning the Entropy Index has a value close to zero) a positive walkability result is expected as the given area is characterized by diverse land-use. If there is a high level of redundancy (meaning the Entropy Index has a value close to one) walkability will be affected negatively (Brennan Ramirez et al., 2006; Tucker et al., 2009; Brownson et al., 2009; Robitaille et al., 2009; Feng et al., 2010). Example of measure: • $Entropy\ Index = -\sum_{i=1}^{n}\left[\left(\frac{P_{ij}}{P_j}\right)\ln\left(\frac{P_{ij}}{P_j}\right)\right] / \ln n$ (1) n: Number of land-use clusters. P_{ij}: Number of property assessment units i in zone j. P_j: Sum of property assessment units 1 to n in zone j. Entropy Index varies between 0 and 1 where 0 = Maximum specialization. 1 = Maximum diversification
5. Percentage of non-residential buildings
When in an area, the percentage of non-residential buildings is higher, it is more likely that people are not within walking distance of non-residential locations, because residential and non-residential areas tend to be separate. Thus, walkability is affected negatively. However, if there is a large residential area near a area of high percentage of non-residential building, walkability is expected to be affected positively (Robitaille et al., 2009). Example of measure: • Number of residential buildings in a given area divided by the total number of buildings in the area.

Connectivity

Connectivity is defined as the measure that quantifies the degree to which roads, sideways, pedestrian walkways and trails are connected (Marhall, 2005). A high connectivity is expected to ease the transportation and travel between places, as a well-connected network is expected to offer shorter and many alternate routes, which in turn affects walkability positively. The grid pattern is considered as the archetypal high connectivity network, where streets cross each other at right angles and the urban environment is characterized by small rectangular blocks and numerous intersections.

On the other hand, neighborhoods characterized by a considerable number of dead-ends and fewer intersections, blocks or sidewalks are regarded to be less supportive of walking (Feng et al., 2010). Measures of connectivity are basically related to the layout of transportation infrastructure and the physical design. A considerable number of studies have examined the association between various measures of street connectivity and have concluded that many measures of connectivity and walkability are positively associated, as it is going to be presented in the next section (Tomalty et al., 2009; Vargo et al., 2011; Robitaille et al., 2009; Pelletier, 2009; Berrigan et al., 2010). Table 2.3 shows some selected measures of Connectivity.

Table 2.3 Measures of Connectivity

1. Types of Streets (e.g.: 3-way, T, highway)

The type of street can affect walkability negatively or positively. When a road is a side street, then there are more opportunities for a pedestrian to cross the road and less traffic is expected to disrupt walking. On the contrary, when the road is a major highway, pedestrians are expected to have fewer opportunities to cross it and considerable traffic is expected to disrupt walking (Robitaille et al., 2009; Pelletier, 2009; Berrigan et al., 2010).

Example of measure:

- Categorization of streets by type in a given area

2. Intersection count or density

A larger number of intersections in a given area is expected to provide higher street connectivity and larger variety in walking itineraries, which in turn affects walkability positively (Tomalty et al., 2009; Vargo et al., 2011; Robitaille et al., 2009; Pelletier, 2009; Berrigan et al., 2010).

Example of measure:

- Number of intersections in a given area

3. Four-way intersections per unit land area (raw intersections, 10m, 15m buffers)

The more four-way intersections there are in a unit land area the more desirable the area tends to become for walking as the streets are connected better, and thus affecting walkability positively (Robitaille et al., 2009; Pelletier, 2009).

Example of measure:

- Number of intersections per unit land area

4. Alpha Index (Ratio of the number of actual circuits to the maximum number of circuits)

When the number of actual circuits (i.e. a series of point, stops or places in an itinerary) is higher with relation to the maximum number of circuits, then a street section has increased diversity and connectivity, which consequently affects walkability positively. A zero score means that the network has no itineraries, while a score of one defines a network with a maximum number of itineraries (Robitaille et al., 2009; Berrigan et al 2010).

Example of measure:

- $$Alpha\ Index = \frac{(L-V)+1}{(2V-5)} \qquad (4)$$
  L: Number of segments in a network
  V: Number of nodes (intersections)
  The Alpha index represents the level of possible itineraries included in a given network.

5. Connectivity Index

Derived by dividing the total number of street segments (street lengths between intersections) by the total number of street nodes (intersections or dead-ends). The higher value the index has, the more choices the travelers have, allowing for more direct connections between any two points. A higher level of connectivity between segments and street nodes means increased route choice, and infers higher walkability. A perfect grid network receives a score of “1.5” (Robitaille et al., 2009; Berrigan et al 2010).

Example of measure:

- Total number of street segments divided by total number of street nodes

6. Gamma Index (Ratio of the number of links in the network to the maximum possible number of links between nodes)
A higher number of actual links compared to the maximum number of links infers better connectivity and it is expected to affect walkability positively. A zero score means that none of the intersections are connected, while a score of one corresponds to a network where all possible segments are linked to all of the possible intersections (Robitaille et al., 2009; Berrigan et al 2010). Example of measure: • $Gamma\ Index = \frac{L}{3(V-2)}$ (5) L: Number of segments in a network V: Number of nodes (intersections) The Gamma index represents a measure of network connectivity

Proximity

Proximity is associated with the number and variety of destinations within a specified distance of a given area. It is a function of both density and diversity. When proximity is higher and higher directness occurs between points of destinations, it is more probable for the distance between destinations to decrease, which consequently reduces the use and need of cars. It is reasonable to assume that when the distances between destinations is equal or less than 1 km, driving is more likely to be substituted by walking (Vargo et al., 2011). In a considerable number of studies, close proximity to parks, pathways, trails, schools and recreational facilities have consistently been correlated with walking and physical activity in general (Berke et al., 2007; Tucker et al., 2009; Robitaille et al., 2009; Lovasi et al., 2008; Curran et al., 2006).

Proximity with regard to retail establishments is also considered to be important (Krizek and Johnson, 2006). According to Krizek and Johnson (2006), retail establishment includes the following general categories:

- Food and beverage stores
- Health and personal care stores
- Clothing and clothing accessory stores
- Sporting goods, hobby, book, and music stores, general merchandise stores
- Miscellaneous stores (e.g. used merchandise, pet, art, tobacco etc)
- Food services and drinking places.

Additionally, it is suggested that highly walkable areas tend to support higher levels of public transit service and ridership (Park, 2010; Frank et al., 2010). As a result these areas are less depended to cars. Proximity with regard to public transit stops is commonly used to asses and test any association with physically active transportation (Tomalty et al., 2009; Vargo et al., 2011; Forsyth et al., 2008; Frank et al., 2010). Table 2.4 shows some selected measures of Proximity. For each measure there is again a summarized description of the rationale behind the use and application of each measure, and a brief example of how these measures are calculated.

Table 2.4 Measures of Proximity
1. Distance between point of origin & closest destination
The closer the distance between a point of origin and a point of destination, the more likely it is that a person will walk that distance, affecting walkability positively (Feng et al., 2010; Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • $A_i^a = (\min\lvert d_{ij}\rvert)$ (6) A_i^a: Distance between spatial unit i and the closest service d_{ij}: Distance between spatial unit i, and spatial unit and service j
2. Total distance between point of origin and all destinations
The closer the average distance is from all destinations with relation to the point of origin, the more expected it is to walk to these destinations, implying that walkability is affected positively (Feng et al., 2010; Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • $A_i^d = \sum_j d_{ij}$ (7) A_i^d: Total distance between spatial unit i and all services j d_{ij}: Distance between spatial unit i, and spatial unit and service j
3. Average distance between point of origin and a number of destinations
The closer the average distance from a number of destinations with relation to the point of origin, the more expected it is to walk to these destinations, implying that walkability is affected positively (Feng et al., 2010; Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • $A_i^c = \sum_j d_{ij}/n$ (8) A_i^c: Average distance between spatial unit i and a "n" number of services d_{ij}: Distance between spatial unit i, and spatial unit and service j n: Number of services included in the analysis
4. Proportion of residents within walking distance of defined diverse uses
When the percent of the population that is within walking distance of defined diverse uses is high, the expectance that people will walk to these destinations increases, which consequently affects walkability positively (Curran et al., 2006). Example of measure: • Residents within walking distance of an area of diverse uses divided by all residents in given area
5. Hectares of parks and playgrounds per/capita
The higher levels of hectares of park and playground exist per capita, the less crowded the park will be. Additionally, the larger the park is, the more accessible it is from multiple locations. Under these condition, walkability is expected to be affected positively (Tomalty et al., 2009; Curran et al., 2006; Brennan Ramirez et al., 2006; Oliver et al., 2007; Brownson et al., 2009). Example of measure: • Number of hectares of park and playground divided by the population of a given area. • Park area (km^2)/1000

6. Proximity to schools
A school in the close vicinity of a residential area infers that walking to it is easier and more likely to occur, affecting walkability positively (Oliver et al., 2007; Brownson et al., 2009). Example of measure: • Distance from a given point of origin to the closest school.
7. Density of food outlets in a given area
The more densely situated food outlets are, the larger variety of outlets an area will have. This is expected to attract more people to walk in that area, affecting walkability positively (Brownson et al., 2009). Example of measure: • Number of food outlets in a given area • Number of outlet per km^2
8. Proximity to food outlets
Same logic with "proximity to schools". The shorter the distance to the closest food outlet is, the higher likelihood there is that residents in the area will be within walking distance of the outlet, which in turn implies that walkability is affected positively (Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • Distance from a given point of origin (e.g. average) to closest food outlets
9. Food stores per 10,000 people
The more food stores per 10,000 people there are, the more likely people are to be within walking distance to one, affecting walkability positively (Brownson et al., 2009). Example of measure: • Number of food stores in an area divided by people in the area: X Where "X" is multiplied by 10.000: X* 10.000
10. Number of supermarkets within 1000 meters
When the number of supermarkets within a 1000 meters buffer around a location is high, people have a better accessibility to the supermarkets, and they can walk easily to those places, affecting walkability positively (Larsen et al., 2008). Example of measure: • Number of supermarkets within 1000 m from the point of origin. (Network analysis can be employed to create a "service area" buffer of 1000 meters around each supermarket)
11. Distance to nearest transit stop
The closer the distance to the nearest transit stop, the more likely people are to use public transit and walk to and from the transit stop, affecting walkability positively (Forsyth et al., 2006; Shay et al., 2009; Brennan Ramirez et al., 2006; Brownson et al., 2009; Larsen et al., 2008). Example of measure: • Find the closest transit stop from a given point of origin

12. Number of transit stops
The higher the number of transit stops in a given area, the more places in an area travelers can access on foot. Additionally a high number of transit stops suggests high urbanization of the area. Walkability is expected to be affected positively (Robitaille et al., 2009). Example of measure: • Total number of stops in the given area
13. Retail points, service points, schools and jobs within walking distance to transit stops
The higher the number of retail points, service points, schools and jobs within walking distance of transit stops, the more likely it is that a larger majority of travelers will use public transportation means and walk to go to their destinations points from the transit stops, affecting walkability positively (Brownson et al., 2009). Example of measure: • All retail points, service points, schools, and jobs within a chosen area and within walking distance from transit stops
14. Distance to closest recreational facility
When recreational facilities are within walking distance and the closer these facilities are, the more likely an individual is to prefer walking to those facilities, affecting walkability positively (Brennan Ramirez et al., 2006; Brownson et al., 2009). Example of measure: • Distance from point of origin to closest recreational facility

Environmental Friendliness

Another important aspect that determines whether people will walk into an area or not is, how "friendly" and attractive the area is. Several studies have documented the relation of walkability with regard to safety, aesthetics of the surroundings, existence of sidewalks and accessible recreational facilities such as parks and walking trails (Brownson et al., 2009; Pelletier, 2009; Robitaille et al., 2009). For example, effective street design, which refers to the scale and design of sidewalks and roads, and their management (e.g. traffic signaling, calming design for speed and volume regulation) can affect walkability positively. The presence of grassed open spaces with trees and flowers or public art and other attractive natural, architectural or historical features can also increase peoples' interest to walk through neighborhoods with these characteristics (Brennan Ramirez et al., 2006). Safety is also an element that can affect walkability. In particular, high crime rates are expected to reduce walkability of an area (Brownson et al., 2009; Tomalty et al., 2010; Alfonzoa et al., 2008). Another important factor is whether a neighborhood supports and offers a balanced variety of transport modes (e.g. public transit, cycling, walking etc). In some cases, the use of street lights that reduce night-time glare might also be interesting to assess (MMAHO and OPPI, 2009). In general, measures of environmental friendliness can be grouped into three categories: i) Comfort; ii) Cleanliness; and iii) Safety. Measures of these categories include the following (please note that some measures could appear in multiple categories):

• Comfort: presence of cross-walks, sidewalk buffers, number of traffic lanes, street width, block length, sidewalk width, traffic circles, curb bulb-outs, speed bumps/humps, pavement treatments, posted speed limits.

• Cleanliness: percentage of street segments without visible litter, graffiti or dumpsters.

• Safety: crime rates, presence of graffiti, windows facing street, street lighting, abandoned or vacant buildings, and rundown buildings, indicators of loitering, alcohol or drugs, and gang activity.

Table 2.5 shows some selected measures of Environmental Friendliness. For each measure there is again a summarized description of the rationale behind the use and application of each measure, and a brief example of how these measures are calculated.

Table 2.5 Measures of Environmental Friendliness
1. Sidewalk Length
The longer a sidewalk is without interruption, the friendlier for walking the area becomes, suggesting more positive walkability (Forsyth et al., 2006; Robitaille et al., 2009; Park 2008). Example of measure: • Measure sidewalk length (by unit area if appropriate)
2. Sidewalk Width
Sidewalk width is expected to affect walkability positively as wide sidewalks can allow more pedestrians to walk, while keeping them at safe distance from cars (Park 2008). Example of measure: • Measure width of sidewalk
3. Average or median census block area
Large average or median block areas imply less connected road network, affecting walkability negatively (Forsyth et al., 2006; Robitaille et al., 2009; Berrigan et al., 2010). Example of measure: • The sum of lengths of all blocks within a chosen area divided by the number of blocks
4. Percentage of street segments with visible litter, graffiti or dumpsters
A littered area with graffiti is expected to be less desirable as a walking area, affecting walkability negatively (Brownson et al., 2009; Pelletier, 2009). Example of measure: • Street segments with litter, graffiti or dumpsters in a given area
5. Number of Traffic lanes
A high number of traffic lanes suggests a large traffic flow making the surrounding area less desirable for walking and, consequently, affecting walkability negatively (Park, 2010; Pelletier, 2009). Example of measure: • Number of traffic lanes in a given area
6. Sidewalk to road ratios
A good ratio of sidewalk to road coverage connotes a more desirable the area for walking, affecting walkability positively (Tomalty et al., 2009). Example of measure: • Number of sidewalks divided by number of roads

7. Median housing age
Housing age can affect walkability indirectly. Old neighborhoods could become rundown and debased, becoming gradually less safe areas for walking, which in turn affects walkability negatively. In some cases though, even new housing groups have the potential of also being unsafe for walking (Robitaille et al., 2009). Example of measure: • Identify age of houses and calculate median age
8. Traffic speed limits
When the traffic speed limits are high, the cars move faster, suggesting higher difficulty and less safety for pedestrians to cross roads. This is expected to affect walkability negatively (Park, 2010; Pelletier, 2009). Example of measure: • Note posted speed limits
9. Bus Stop / Subway Stations Density
The more densely situated bus stops and subway stations are, the more positively walkability will be affected, because higher public transit suggests well developed urban infrastructure. A well developed urban infrastructure is also expected to be easier accessible to larger group of the population (Brownson et al., 2009; Robitaille et al., 2009). Example of measure: • Calculate density by counting the number of bus stops / subway stations in a given area
10. Proportion of commercial parcels with paid parking, side, front, street parking
The more parking places are available near points of interest the more likely people are expected to leave their car at a location and walk to their desired destination, affecting walkability positively (Pelletier, 2009). Example of measure: Number paid parking, side, front and street parking in a given area
11. Crime rates
Higher crime rates suggest that an area is less attractive for walking, affecting walkability negatively (Brownson et al., 2009). Example of measure: • Measure level of crime in an area

Walkability Indices

The next evolutionary step of walkability measures were composite indices. Researchers suggested the use of composite measures motivated by the fact that neighborhood characteristics are often correlated with one another (Feng et al., 2010). For example, grid patterns, sidewalks, and public transit stations usually coexist in old parts of cities or traditional pre-World War II neighborhoods. These areas are usually characterized by high density and diversity. This, along with the fact that transportation mode choice is highly affected by numerous environmental factors, motivated the use of composite measures in measuring walkability (Feng et al., 2010).

Thus by taking into account both form and content of neighborhoods, walkability indices (WI) are expected to indicate the degree to which an area is pedestrian oriented and attractive for walking. WI are expected to capture the interrelation of various environmental characteristics which in turn are expected to minimize the effect of spatial collinearity, and facilitate the communication of results (Brownson et al., 2009). Composite indices vary by the components they include; they might use different scale and/or different computation methods but they are still valid and suitable measures of walkability (Frank et al., 2010). Finally, composite measures of walkability are considered to be more consistent predictors of walking behavior than single component measures (e.g. Density, Connectivity etc) (Vargo et al., 2011). Table 2.6 shows some selected Walkability Indices. For each Walkability Index there is a short description of the components and how the Index is calculated.

Table 2.6 Walkability Indices
1. Walkability Index I (Lachapelle et al., 2011)
Walkability index is calculated at the block group level across each region using the sum of the z-scores of: i. Net residential density ii. Intersection density iii. Retail floor area ratio iv. Entropy based measure of land use mix.
2. Walkability Index II (Frank et al., 2010)
Walkability index is calculated at the block group level across each region using the sum of the z-scores of: i. Net residential density: No. of residential units per acre designated for residential use within a neighborhood buffer. ii. Commercial density (or Retail Floor Area Ratio): Amount of area designated for commercial use within a neighborhood buffer, using a ratio of commercial floor area to commercial land area. iii. Land use mix (mixed use index): The evenness of square footage distribution across residential, commercial (including retail and services), entertainment, and office development within a neighborhood buffer. iv. Street connectivity: Number of street intersections in a neighborhood buffer.
3. Walkability Index III (Frank et al., 2009)
This walkability index is based on: i. Net residential density: Ratio of residential units to the land area devoted to residential use per block group. ii. Retail floor area ratio: Retail building floor area footprint divided by retail land floor area footprint. iii. Land use mix: The mix measure considered 5 land use types: residential, retail, entertainment, office and institutional. Values were normalized between 0 and 1, with 0 being fully homogenous use and 1 indicating a completely even distribution of floor area across the five uses. iv. Intersection density: Ratio between the number of true intersections (three or more legs) to the land area of the block group in acre. The four calculated values were normalized using a z score = [(2x"z-intersection density") + ("z-net residential density")+("z-retail floor area ratio")+("z-land use mix")

4. Walkability Index IV (Doyle et al., 2006)
Walkability index is calculated at the block group level across each region using the sum of the z-scores of: i. The negative of average block size, which should be positively related to connectivity ii. The percent of all blocks having areas of less than 0.01 square miles iii. The number of 3-, 4-, and 5-way intersections divided by the total number of road miles In order to make the measures comparable, they were converted to z-scores then added these values to derive the final walkability measure.
5. Activity-Friendly Index (AFI) (Glazier et al., 2007)
The Activity Friendly Index (AFI) measures how conducive neighborhoods are to walking, bicycling and other types of physical activity. The AFI consists of the following five variables: i. Car ownership per household (values reversed) ii. Population density per km^2 of residential area iii. Density of all retail services per 10,000 people iv. Average distance from residential points to the nearest five retail locations (values reversed) v. Rates of drug-related and violent crime rate The values of each of the five variables of AFI were standardized to the range of zero to ten. Then the standardized values of the five variables were added together (equally weighted) and divided by five. As a result, the AFI scale ranged from zero to ten, where zero represents the least, and ten represents the most activity-friendly conditions within a neighborhood.

Walkability Index I is a simpler version of Walkability Index III. The difference is that in Walkability Index I all four components have equal weights. While in Walkability Index III Intersection Density is given a weight factor of two based on prior evidence that street connectivity has a strong influence on non-motorized travel choice. Both Walkability Index I and III calculate residential density as a ration of residential units to the land area devoted to residential use. This measure describes quite effectively the residential density but there might be accuracy issues because there is no clear evidence whether for instance the residential units are multi-storey block of flats or houses. The intersection density measure that is used in these two Indices (Walkability Index I and III) is based only on true intersections, namely only on intersection with more than three legs, but it includes intersections that are not walkable such as intersections on highway interchanges. As a result there might be an overestimation regarding the density of the intersections. The Land Use Mix component of Walkability Index I and III considers only five land use types which are the most important ones, but they do not include all different possible land use types such as parking lots, vacant land etc. Thus, the measure of diversity might be a bit underestimated.

Walkability Index II is similar to Walkability Index III but it employs a different method in terms of spatial unit. It does not uses predefined spatial units as in Walkability Index I and III. Instead the spatial units are neighborhoods that are defined by drawing a 1-kilometer street network buffer (representing a 10- to 15-minute walking distance) from each postal code centroid of the study area. Thus, a walkability map of higher resolution is achieved. All these three Walkability Indices (I, II and III) incorporate measures of density, diversity and connectivity but there are not measures of environmental friendliness and proximity. While density combined with diversity can give an indirect estimation of proximity, it might be useful to consider a direct measure of proximity too when measuring walkability.

Walkability Index IV focuses only on indirect or direct measures of connectivity. The number of true intersection divided by the total number of road miles is clearly an alternative measure of intersection density. On the other hand a low average block size and a higher number of blocks having areas of less than 0.01 square miles clearly infers a denser road network that gives more alternatives to the walkers to choose their route. Walkability Index IV could be seen as an attempt to search different ways to measure and model connectivity rather than a composite measure of Walkability.

The Activity Friendly Index does not contain any measure of diversity, but it employs a measure of proximity (average distance from residential points to the nearest five retail locations) and a measure of environmental friendliness (rates of drug-related and violent crime rate). The residential density is calculated based on population density which might raise some accuracy issues because population might occur even in non residential areas. To measure residential density more accurately, household density is probably more suitable as it will be discussed in the chapter of Methodology. Car ownership per household is more relevant to measure how conducive is an area to various types of physical activity and it does not focus only on walkability. Hence, it is rarely, if not never, used in composite measures that measure strictly walkability. This Index is referred in this report for the lax relevance to walkability indices and because most of its components can be used in Walkability Indices.

2.3 New York City, Obesity and Walkability: Previous Studies and Modeling Attempts

During the last 5 years several studies related to obesity and walkability for NYC have been published. Black et al. attempted to measure the relations between neighborhood food availability, opportunities and barriers for physical activity, income and racial composition with obesity in NYC, controlling for individual level factor. Their study revealed that obesity rates ranged from 6,8% to 31,7%, widely varying between neighborhoods. They also concluded that several neighborhood-level factors significantly affected obesity. In particular, the researchers found that availability of supermarkets and foodstores, fitness facilities, percent of commercial land use area and income are significantly associated with obesity in NYC (Black et al., 2010).

Rundle et al (2009) examined the relation of neighborhood food environments with body mass index (BMI) and obesity. Controlling neighborhood walkability in NYC revealed that differences in the neighborhood food environment contribute significantly to disparities in obesity. In particular, the density of BMI-healthy food outlets (supermarkets, fruit and vegetable markets, and natural food stores) was inversely associated with BMI, meaning that access to the aforementioned stores is linked with lower prevalence of obesity and lower BMI.

Lovasi et al (2013) examined possible relations between modifiable neighborhood characteristics, active transportation and health problems related to insufficient physical activity for the NYC area. They attempted to evaluate whether aesthetically “friendly” neighborhoods with sidewalk cafés, street trees and clean sidewalks, and fewer safety hazards (pedestrian-auto fatalities and homicides) are associated with active transportation. Their findings revealed that those living near sidewalk cafés were 10 % more likely to report active transportation. Besides, higher homicide rates affected physically active transportation negatively. Their conclusion was that measures to prevent homicides and investments in aesthetic amenities are expected to promote active transportation.

A very promising project of the last years is the "Walk Score". The "Walk Score" project calculates walkability for NYC and all major cities of USA, Canada, Australia, and New Zealand. Its main objective is to aid the home-buying process. The "Walk Score" walkability map is based on a set of two different types of data:

- Walking routes and distances to amenities
- Road connectivity metrics

The final product of "Walk Score" is determined by a set of customizable amenity categories, weights, and distance decay functions.

In a nutshell, the algorithm which is used in "Walk Score" creates a score for a group of selected amenities and facilities based on the street network analysis and the given weight of each amenity. In the calculation, a decay function is applied, that dictates the score between the closest zone around an amenity (a radius of 400 m) and the theoretically accepted farthest distance around the amenity (2,4 km). Then the score is penalized by 0-10% according to an estimated pedestrian friendliness that is based on intersection density and average block length (WalkScore, 2010). Figure 2.5 presents a screen dump of WalkScore's walkability map.

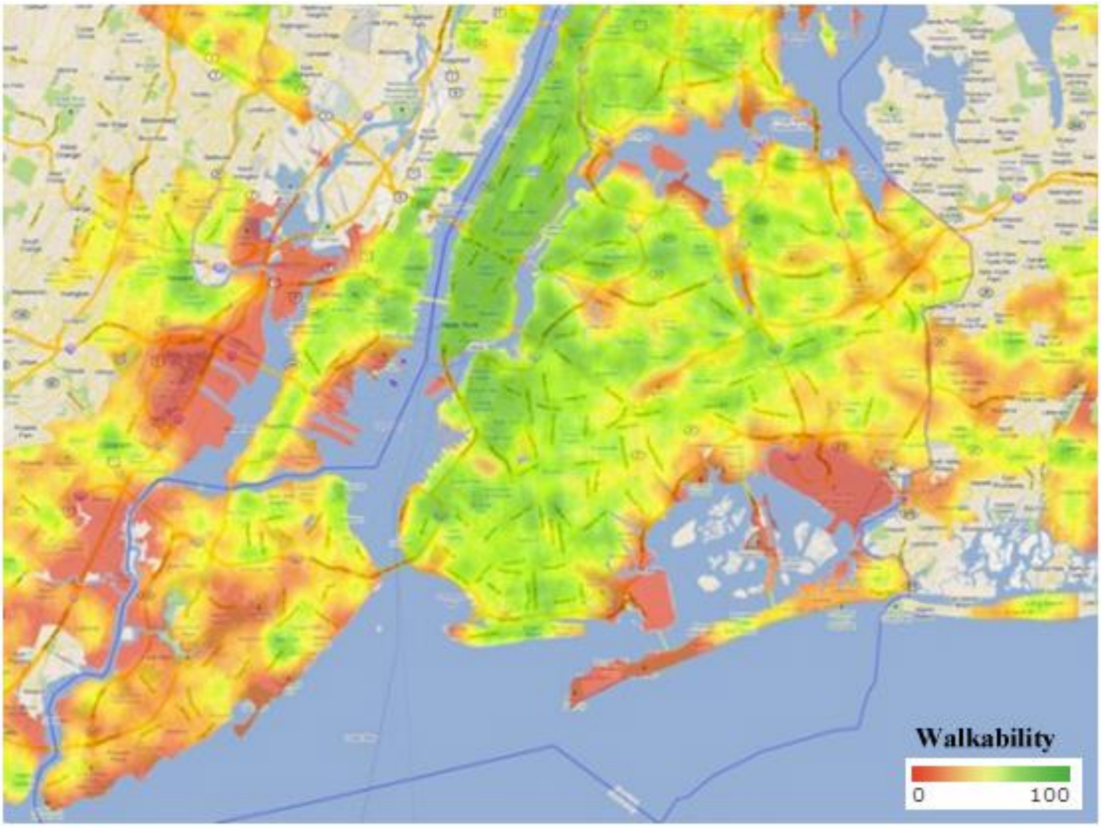


Figure 2.5 WalkScore NYC walkability map (after WalkScore, 2010)

Ogle (Azavea, 2010) developed a web tool named "Walkshed", suggesting a slightly different calculation of walkability in NYC. Walkshed is a method that scores a location according to the diversity and quantity of amenities within a buffer of one-mile radius. Ogle suggested some improvements to the “Walk Score” and thus developed “Walkshed NYC” to demonstrate that it is possible to improve the estimation of walkability.

The main idea is to calculate walkability by using “friction-based” distance calculations instead of simple network analysis, as in the case of project “Walk Score”. This is achieved by determining the theoretical “friction” a pedestrian would encounter. For example, in a raster map of an entire city, interstates, rivers and other various barriers that would hinder a pedestrian correspond to cells with high friction, while parks or areas with high street connectivity correspond to cells with very low friction.

Another aspect of “Walkshed” is that it includes a “Decision Tree” prioritization tool, which allows users to prioritize locations by setting custom weighted geographic preferences. In other words, a user can select his/her preferred facilities among a given set of facilities and prioritize his/her selections by giving custom weights for each facility. The end product is a map that demonstrates the walkability of an urban area. Figure 2.6 shows Walkshed’s NYC walkability heatmap.

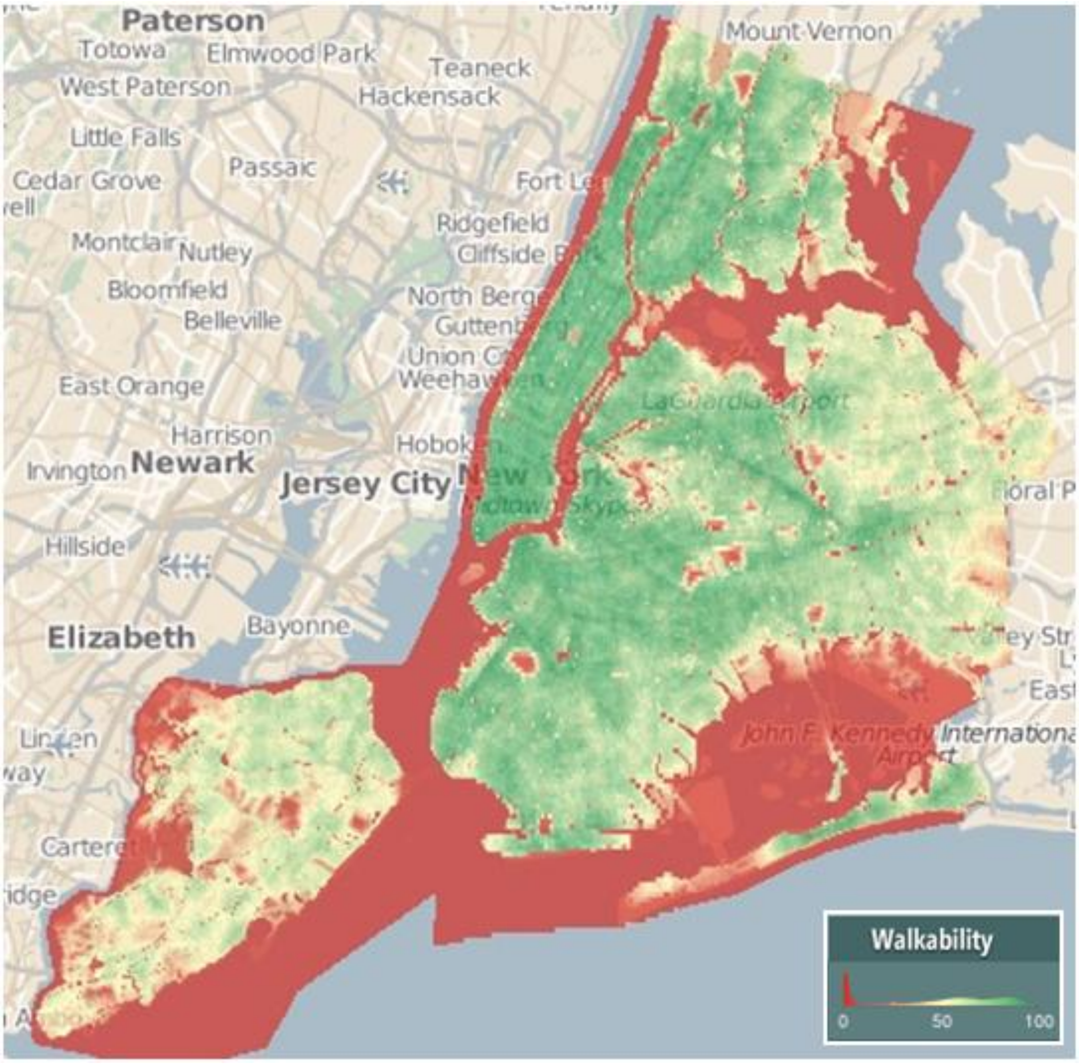


Figure 2.6 Walkshed's NYC walkability map (after Azavea, 2010)

The resulting Walkability maps of “Walk Score” and “Walkshed NYC” are quite similar which is expected because the approach of both studies is very similar to the core. In both methods the main idea is to build an algorithm that will estimate walkability based only on criteria of proximity through a simple network analysis. The differences can be identified only in the way these algorithms are optimized when it comes to consider different types of barriers or environmental features that disturb walking.

Due to these differences the resulting Walkability map of “Walkshed NYC” has a higher “contrast” compared to “Walk Score’s” map, meaning that the changes between the different zones of walkability in “Walk Score’s” map are smoother and more gradual, while in “Walkshed NYC” are more abrupt. Regarding the walkability scores of each Borough, Manhattan is found to be highly walkable in both methods, but Brooklyn and Bronx are considered to be slightly walkable according to the calculations of “Walkshed NYC”. On the other hand, Queens and Staten Island are found to be less walkable in “Walk Score” compared to the map of “Walkshed NYC”. So there is no general pattern that can lead us to a conclusion whether any of these methods overestimates or not walkability. In general, most of the areas that scored very low seem to coincide in both methods.

The major drawback of these methods is that they focus mainly on proximity and do not consider other measures of walkability such as Diversity and Density or consider indirect measures of connectivity (network analysis) or environmental friendliness (pedestrian friendliness on “Walk Score”; friction on “Walkshed NYC”). But they avoid the limitations when using predefined spatial units by using buffers around geocoded points of interest and facilities. Most of the Walkability Indices do not incorporate direct proximity measures and use predefined spatial units in their analysis. In the present study a more holistic approach is attempted that includes at least one parameter for each major category of Walkability measures (Density, Diversity, Proximity, Connectivity and Environmental Friendliness) and combines the two different methods of analysis: buffering and predefined spatial units. It is an attempt to suggest an improved “hybrid” method to measure Walkability.

3. Study Area and Data Description

This chapter is dedicated to the general description of the study area and the data that was used in the analysis.

3.1 Study Area

New York City is selected as the area to study walkability. The area under consideration consists of five boroughs – The Bronx, Brooklyn, Manhattan, Queens, and Staten Island - and 2165 Census Tracts (CT). According to the U.S. Census Bureau, Census Tracts are small, relatively permanent statistical subdivisions of a county or equivalent with respect to population characteristics, economic status, and living conditions. Census Tracts have an average population of about 4,000 people (U.C.B., 2010). The latest census estimated the population of NYC to 8.336.697 (U.C.B., 2012), which is distributed over a land area of 783,8 km^2, making New York the most densely populated city of the USA.

Located on one of the world's largest natural harbors, NYC is probably the world's pivotal financial center. It exerts significant influence upon commerce, media, art, entertainment, culture, fashion, research, technology, and education. NYC has a considerable number of colleges and universities including Columbia University, New York University, and Rockefeller University, while an area of over 110 km^2 form the city's municipal parkland that includes Central Park, Prospect Park, Flushing Meadows–Corona Park, Forest Park, and Washington Square Park.

New York City's public bus fleet and commuter rail network are the largest in North America (M.T.A., 2012), while the Subway provides continuous 24/7 service. That said, NYC is the most public transport oriented city of US (Weinberger et al., 2010). As a matter of fact more than half of the households in NYC (52%) do not have a car. The percentage is higher for Manhattan (78%) (Weinberger et al., 2010). 21% of the everyday trips in the city are done by bicycle or on foot (Erlbaum, 2005). The combination of low motorized transport use, the high population density, and the high public transit utility renders NYC the most energy efficient city in the US (Jervey, 2006). "Walk Score" ranked NYC in the first place in the list with the most walkable cities in USA (WalkScore, 2011). However, despite the efficient public transport system, roads are a defining feature of NYC, such as Manhattan's street grid plan. Figure 3.1 shows NYC with the 5 boroughs.

Figure 3.1 Study area - NYC's 5 Boroughs

3.2 Data Description

Various types of geographic data and demographic data were used in this study. The data were used in the calculation of parameters of our Walkability Index which is going to be discussed in chapter 4. The demographic data was obtained from the NYC Department of City Planning (DCP) and contained information about the total population and the number of households of each Census Tracts.

Initially, all the geographic data was vector data. Table 3.1 shows the geographic coordinate system and the projected coordinate system of the data. The data is from 2009 or newer, they were available from the NYC GIS consortium.

The geographic data is grouped into four groups. The first group includes various points of interest, facilities and institutions (e.g. sport facilities, subway entrances, theaters, green markets etc). These are mainly in the form of geocoded points, and whenever the initial data were polygons, they were converted into points. The second group consists of various zonal data, such as primary commercial zones of NYC or parks and outdoor recreational areas. The third group contains any other type of data, such as the footprint of the buildings, polygons that delineate the roadbed or the sidewalks etc. The fourth group contains data that had to be generated. In the following sections, all data will be discussed in detail.

Table 3.1. Data Coordinate Systems and Projection	
Projected Coordinate System:	NAD 1983 StatePlane New York Long Island FIPS 3104 Feet
Projection:	Lambert Conformal Conic
Geographic Coordinate System:	GCS North American 1983
Datum:	D North American 1983

3.2.1 Points of Interest and Facilities

Various facilities and points of interest were taken into account in the analysis in several steps, as will be discussed more thoroughly in the next chapters. The data was retrieved from the database of:

- NYC Department of City Planning (DCP)
- NYC Department of Information Technology and Telecommunications (DOITT)
- NYC Department of Parks and Recreation (DPR)
- Metropolitan Transportation Authority (MTA)

A list and a short description of this type of data can be found in Table 3.2.

Table 3.2 Facilities and Points of Interest

Facilities, institutions and points of interest	Source	Type
Museums	NYC DOITT	Points
Theaters	NYC DOITT	Points
Senior Centres	NYC DOITT & NYC DCP	Points
Education Facilities	NYC DOITT & NYC DCP	Points
Universities and Colleges	NYC DOITT	Points
Libraries	NYC DCP	Points
Public Safety or Criminal Justice Facilities	NYC DCP	Points
Hospital/Nursing Home or Ambulatory Program Facilties	NYC DCP	Points
Chemical Dependency Facilties	NYC DCP	Points
Mental Health Services	NYC DCP	Points
Mental Retardation or Developmental Disabilities Services	NYC DCP	Points
Daycares or Residential Facilities for Children	NYC DOITT & NYC DCP	Points
Facilities for Seniors	NYC DOITT & NYC DCP	Points
Food Programs or Residential Facilities for Adults/Families	NYC DCP	Points
Transportation Facilities	NYC DCP	Points
Waste Management Facilities	NYC DCP	Points
Green Markets	NYC DOITT	Points
Post Officies	NYC DOITT	Points
Subway Entrances	NYC DOITT	Points
Bus Stops	MTA	Points
Sport Facilties	NYC DPR	Polygons

In Table 3.2. some of the categorized facilities are groups of facilities and institutions of similar function. For instance, “Education Facilities” contain public schools, private or parochial schools, and post-secondary institutions. “Public Safety or Criminal Justice Facilities consists of police stations and other NYPD facilities, fire departments, courts, and criminal justice facilities (such as correctional facilities or probation facilities). Regarding the data retrieved from the NYC DCP, the grouping of facilities that is suggested by the NYC DCP was also used in the analysis of this study without any changes. The data retrieved from the NYC DPR was used to form the "Sport Facilities" group that contains almost any type of large sport facility, court or field (pools, baseball fields, soccer-football fields, tennis, handball and basketball courts, playgrounds etc).

3.2.2 Zonal Data

The zonal land-use information that was available as free accessible GIS data for NYC were:

i) Vector data (shapefile) that contained all the areas characterized as parks or outdoor recreational areas.
ii) Vector data (shapefile) that contained all the areas characterized as primary commercial zones.
iii) Vector data (shapefile) that contained all the parking lots.
Table 3.3 shows the source and the type of these data.

Table 3.3. Zonal Data

Zoning & Land-Use	Source	Type
Primary commercial zones	NYC DCP	Polygons
Parks and outdoor recrational areas	NYC DPR	Polygons
Parking lots	NYC DOITT	Polygons

3.2.3 Auxiliary Data

This group of data consists of vector data that were used as they are, without any modification, and cannot be included in the previous groups; or auxiliary data that were used to construct and prepare other forms of data that were needed. Table 3.4. shows all these data, their source and their type.

Table 3.4. Auxiliary Data

Vector Data	Source	Type
Roadbed	NYC DOITT	Polygons
Sidewalks	NYC DOITT	Polygons
Street centerlines	NYC DOITT	Polygons
Railroads	NYC DOITT	Polygons
Buildings	NYC DOITT	Polygons
NYC census blocks	NYC DCP	Polygons
NYC census tracts	NYC DCP	Polygons

The shapefiles "Railroads", "Buildings" and "NYC census blocks" were used as auxiliary data combined with data from the other two aforementioned groups to create the land-use map of NYC. The "Roadbed", "Sidewalks" and "Street centerlines" were used basically in various steps of the analysis that will be discussed in the fourth chapter "Methodology". Finally, the shapefile "NYC CT" was used to define the spatial units of analysis.

3.2.4 Data Created for the Study

Except for the data that were available from the NYC GIS consortium two additional layers were needed that could not be obtained due to cost limitations. As a result they had to be generated based on the available data and sources. The first layer was a vector dataset with points for each street junction of the NYC road network. The second layer was a land-use map. The land-use map incorporated nine classes of land-use type based on the official land-use type classification of the NYC planning authority. These maps were available only as non-downloadable web maps. The land-use classes that were used in the analysis are:

- Residential
- Open Space and Outdoor Recreation
- Parking Lots
- Vacant Lots
- Institutions and Facilities
- Commercial
- Transportation
- Industrial
- Mixed Use

3.3 Spatial Units of Analysis

Attempts to calculate walkability should utilize the most detailed spatial unit available, because the physical environment of urban areas is almost never homogenous and is commonly characterized by various forms. Larger and less detailed spatial units usually account for lower accuracy with higher variability among urban landscapes, while smaller and more detailed spatial units provide potentially higher accuracy in the analysis of built environment characteristics. In the study, the chosen spatial unit of analysis is the Census Tract. It would have been more preferable to utilize the census blocks spatial unit, but blocks level was used to create the NYC land-use map. The benefit of using predefined spatial units – such as Census Tracts – is that it allows for incorporation of socioeconomic data (income, demographic, household counts, etc.). While such data is not used in this study, future research could incorporate the results of this study for a more thorough insight into the obesity epidemic.

4. Methodology

In chapter 2, it was discussed that many neighborhood and built environment characteristics can potentially affect walkability and walking behavior. Walkability indices are an effective way to measure the degree to which an area provides opportunities to walk to various destinations (Manaugh et al., 2011). The walkability index of this study draws partially from previous attempts to measure walkability (basically from Walkability Index III) and incorporates six parameters. Two measures of density (Residential Density, and Retail Floor Area Ratio (FAR)) and one parameter for each general category of walkability measures: diversity, proximity, connectivity, and environmental friendliness. Each parameter is identified and defined, with an example to clarify the key concepts.

4.1 Preparation of Land-use Map and Street Junction Map

For the method of this study, two additional layers were needed, that weren't available for free or weren't in a useful format. The first layer was a map with points that indicated the street junctions of the NYC road network. This layer was created based on the "Street centerlines". The second layer was a land-use map. This map was created with the help of all available zonal data of the second group of data, combined with the first group of data. The vector data "Buildings" and "NYC census blocks" were also used to define the borders of different land-use areas.

The land-use map was built in several steps, based on the web maps "ZoLa" (NYC DCP, 2013) and "OASIS NYC map" (OASIS CUR, 2010) which provide very detailed online land-use maps for the entire area of NYC. Figure 4.1 shows a flow chart with the different steps and procedures of the land use map generation. First, the entire area of NYC was defined as "Residential" because, as a matter of fact, the sovereign land-use type of NYC is residential. Then some parts of NYC were redefined as "Commercial", "Parks and outdoor recreation", and "Parking lots", according to the available zonal data. Then, similarly to a building process, the crude land-use map started to be "honed", meaning to become more accurate, by adding features and redefining smaller areas as "Institutions-Facilities" with the assistance of the first group of data "Facilities and Points of Interest" combined with the vector data "Buildings". The basic concept was to locate the buildings via the data of the first group and then use their footprint to redefine the areas within the borders of the footprints as "Institutions-Facilities". Finally, by manually setting the type of the land-use to a large number of polygons taken from the vector data "NYC census blocks", a land-use map was generated that might not be 100% accurate compared to the web maps "ZoLa" and "OASIS NYC map", but it has still acceptable accuracy for the needs of the study. Both "ZoLa" and "OASIS" maps are based on the same data. Hence, both maps have the same accuracy. The final outcome of this process can be seen in chapter 5. Figure 4.2 shows a part of the created land-use map and the same mapped area of the more accurate map of "OASIS" project.

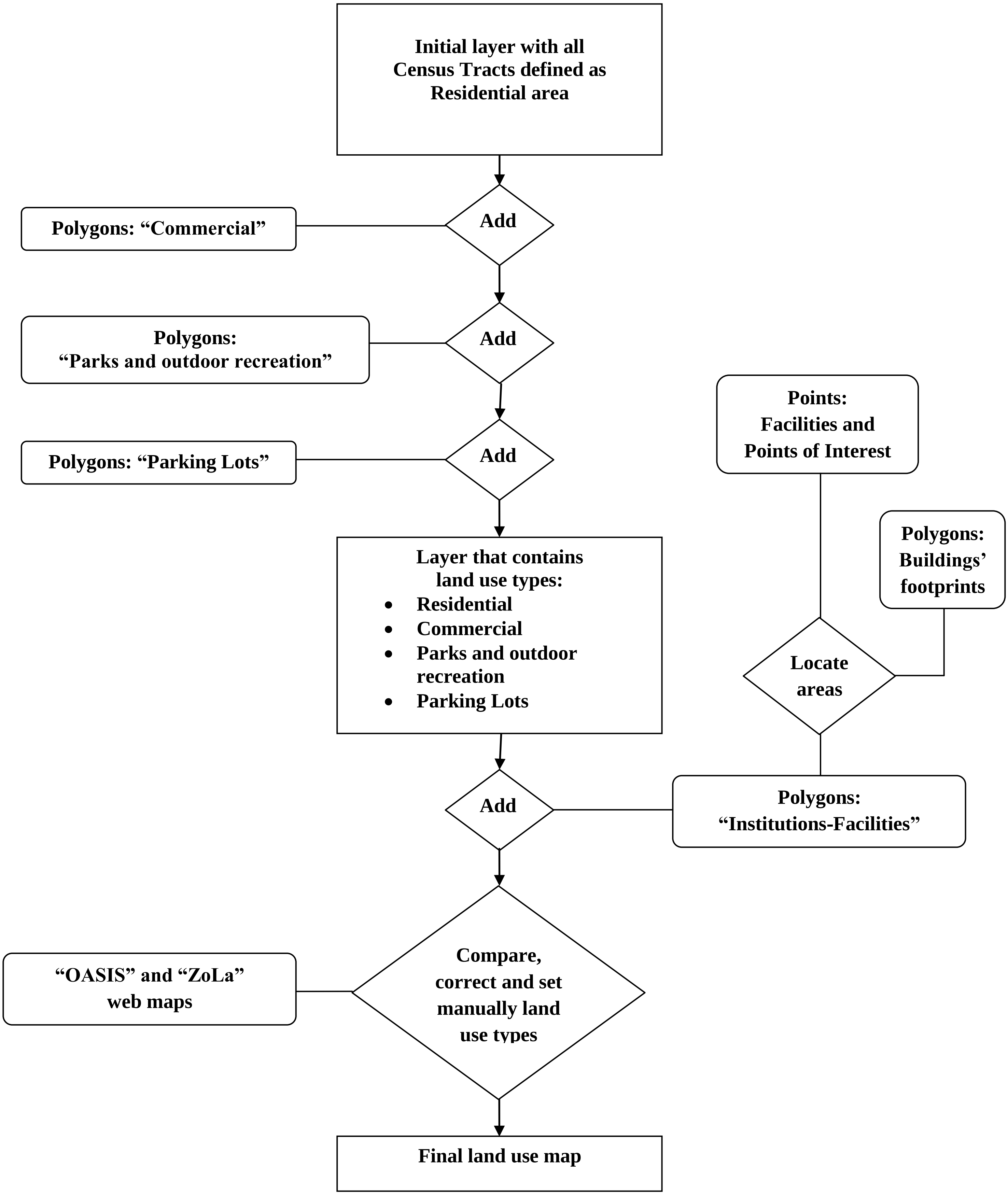
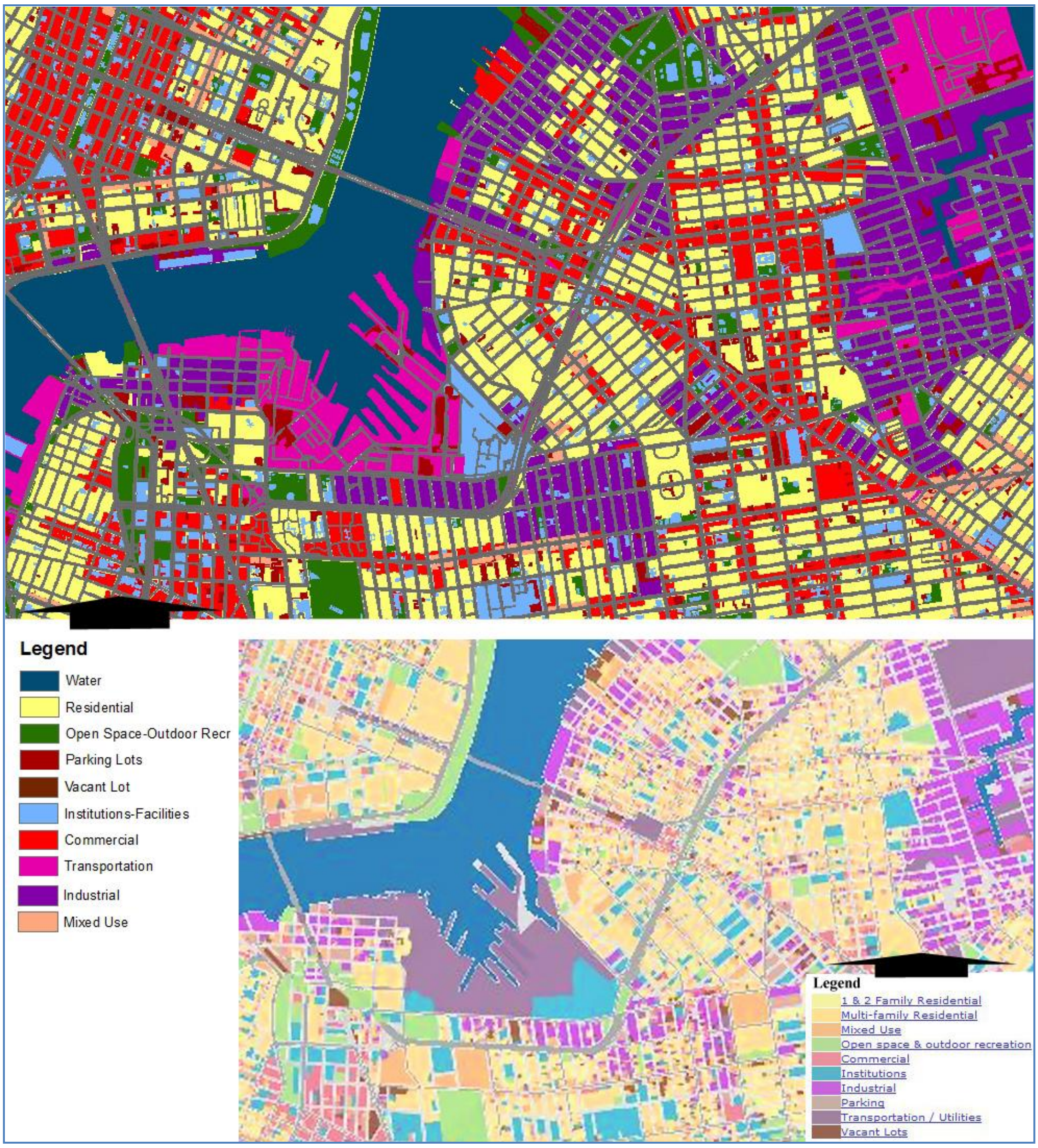


Figure 4.1Flow chart with the different steps and procedures of the land use map generation



The upper part of the figure is a part of the generated land-use map for the needs of the study, while the lower part shows the same area according to the "OASIS" land-use map (OASIS, 2010).

Figure 4.2 Comparison of the created land-use map and the OASIS land-use map

4.2 Density – Household Density

According to a number of studies Household Density is an appropriate indicator of walkability as it indicates general density patterns (Saelens et al., 2003; Brennan Ramirez et al., 2006; Brownson et al., 2009; Robitaille et al., 2009). This means that a neighborhood with high Household Density implies that more commercial and recreational sites will be built near "dense" housing areas, as there are more people to support those services, and urbanization and centralization will occur. Additionally, denser development patterns in general infer that there are fewer parking lots available, which in turn renders the environments to be more conducive to active transportation. Household Density was chosen instead of Population

Density. This choice was made because households are a residential unit of one or more people who live together and may consist of a single family or some other grouping of people, and as a consequence are more strongly related to measures of residential density such as Household Density. In other words, population might occur in non-residential areas but households rarely occur outside residential areas. Choosing Household Density is an effective way to avoid problems in the results were population occur in regions where no residential areas exist, making the analysis more difficult.

In this study, Household density is calculated for each Census Tract as bellow:

- *Household Density = Number of households / km^2 of residential land*

4.3 Diversity

It can be strongly argued that although Diversity is a rather abstract measure, it is an important component when assessing the walkability of an area (Brennan Ramirez et al., 2006; Tucker et al., 2009; Brownson et al., 2009; Robitaille et al., 2009; Feng et al., 2010). This study incorporates the factor of Diversity by employing the Entropy Index that illustrates the diversity of the built environment. As described before, the Entropy Index can affect walkability either positively or negatively.

If there is a low level of redundancy (meaning the Entropy Index has a value close to 0) a positive walkability result is expected, because the given area is characterized by diverse land-use. If there is a high level of redundancy (meaning the Entropy Index has a value close to 1) walkability will be affected negatively. Figure 4.3 shows an example of an area on the island of Manhattan with high land-use diversification. On the other hand, Figure 4.4 shows the area of Kennedy airport as an example of an urban area characterized by a rather high specialization.

- $Entropy\ Index = -\sum_{i=1}^{n}\left[\left(\frac{P_{ij}}{P_j}\right)\ln\left(\frac{P_{ij}}{P_j}\right)\right] / \ln n$ (1)

n: Number of land-use clusters.

P_{ij}: Number of property assessment units i in zone j.

P_j: Sum of property assessment units 1 to n in zone j.

Entropy Index varies between 0 and 1 where
0 = Maximum specialization.
1 = Maximum diversification.

In this study 9 different land-use classes are taken into account. These are: Residential, Commercial, Mix Use, Facilities-Institutions, Industrial, Transportation, Parking Lots, Vacant Land, and Open Space-Outdoor Recreational. As there are some considerably large areas identified as “Vacant Land”, it was chosen to be kept as a distinct class in our analysis and not to be merged with the type “Parking Lots” into one class. Besides, “Parking Lots” are relatively smaller and have different characteristics compared to “Vacant Land”.

An issue that needed special attention was that some Census Tracts did not contain all the various types of land-use. This means practically that various complications would occur due to a number of zeros, especially if these zeros were to be contained in a natural logarithm (see equation (1)). In order to avoid these complications in the calculations, whenever in a census tract the percentage of a specific land-use type was zero that number was set to be nearly zero (i.e. $1 x 10^{-20}$).

Figure 4.3 Example of an urban area with high land-use diversification

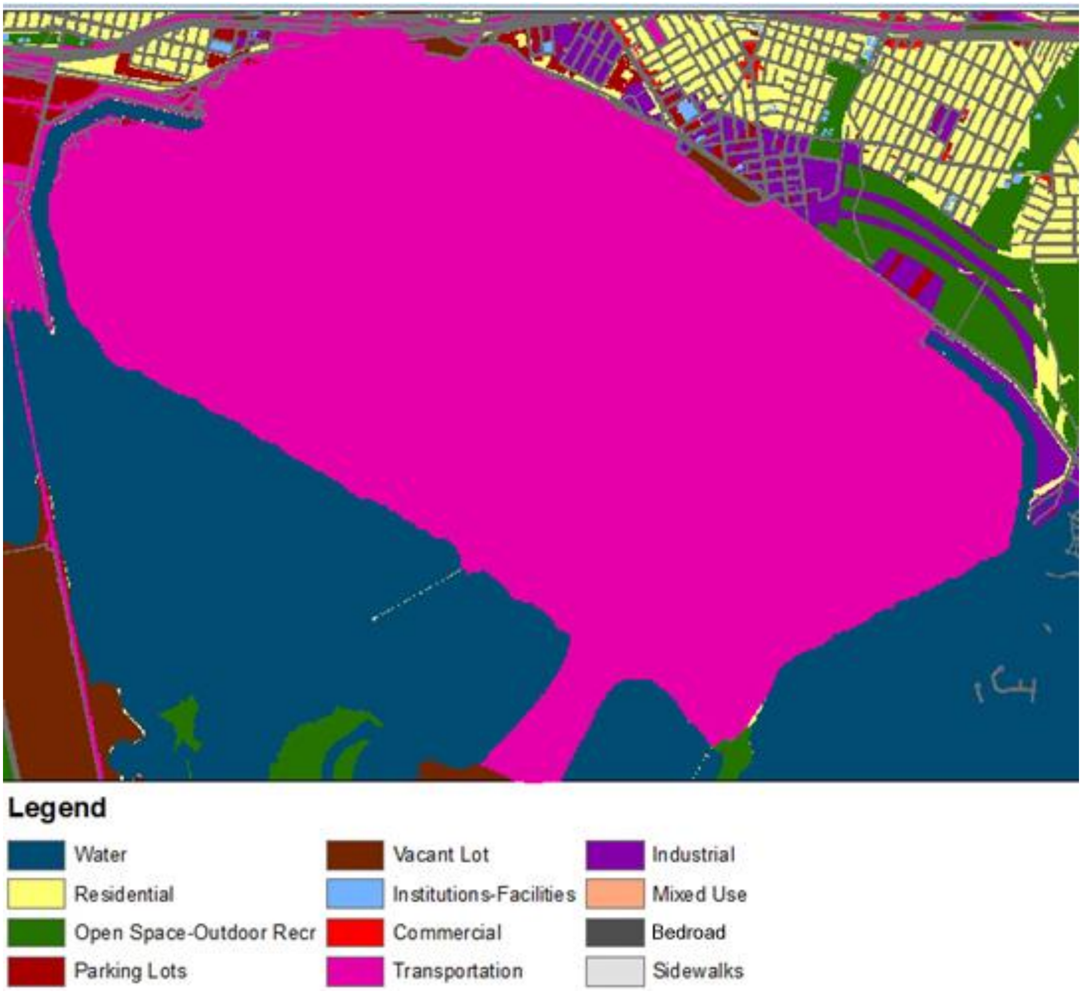


Figure 4.4 Example of an area with low land-use diversification

This area is basically the J.F. Kennedy international airport of NYC

4.4 Connectivity

The third measurable neighborhood characteristic included in the walkability index is street connectivity. In this study Connectivity was determined using the Street centerline vector data, taking into account only the number of true intersections (three or more legs) and excluding street intersections on highways or interchanges because they are not suitable for being walked along. The selection of true intersections was done manually. Any crossroads that are too close, that means closer than 15 m, are merged and treated as one intersection. The distance of 15 m is recommended in other studies too (Frank et al., 2009; Dobesova and Krivka, 2012). An example of true intersections can be seen in Figure 4.5. Similarly to walkability indices I and III (refer to chapter 2), connectivity was calculated for each Census Tract as bellow:

- *Connectivity = number of true intersections / km*^{2} *of land area*

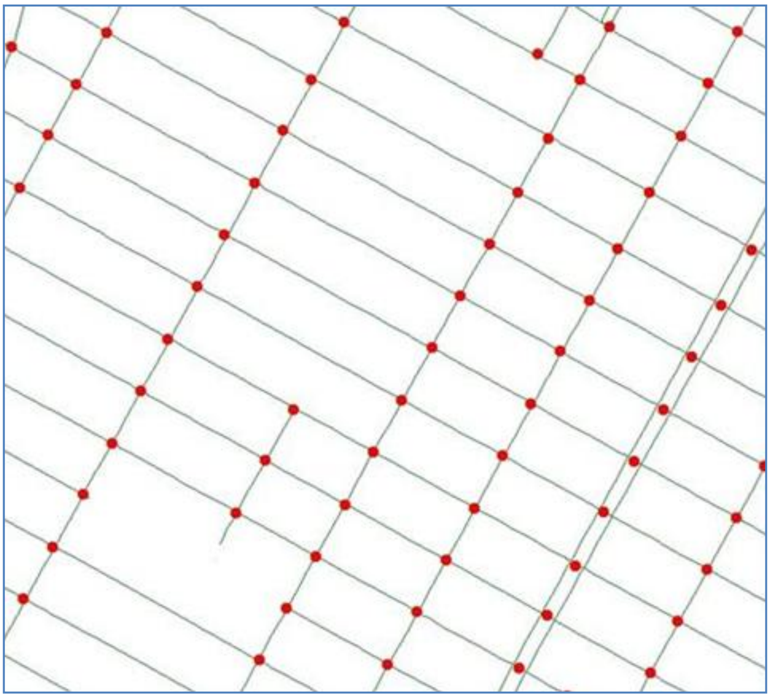


Figure 4.5 Example of true intersection

4.5 Proximity

Proximity describes the number and variety of destinations within a specified distance (buffer) of any location. This measure was calculated by mainly using the first group of data ("Facilities and points of interests", see chapter 3.2.1). Thus, in the analysis a set of points of interest and facilities was taken into account. The employed method consists of four steps:

1. A set of destinations is chosen to be included in the analysis. The selection depends on the availability. Vector polygons are converted to points. If there are polygons that define large areas, such as parks, these polygons are converted to a multiple number of points. Eventually, each point corresponds to a destination which is given a specific weight and a specific buffer distance according to the importance of the destination. The weights and distances can be seen in Table 4.1. The selection of the weights and distances is discussed in the next paragraphs.
2. Network analysis is employed and a service area (buffer) is generated around all destination types. The network analysis is based on the street centerline data (see chapter 3). As a result the NYC street network is depicted with simple lines for each road and it does not include any additional information regarding the type of the road, the number of lanes or any type of street barriers.
3. The buffer polygons are converted to raster format. The cells inside the buffers are given a value of 1, while the cells outside the buffers have a value of 0. At this step, each destination type corresponds to a layer that contains the newly generated buffers.
4. The layers are multiplied by the assigned weight of the destination types and they are added, resulting to a raster surface where the cells can have a value between 0 and 34 - the maximum value a cell can take if all the buffers coincide (which is equal with the sum of all weights). Then zonal statistics is applied according to the zones of the

Census Tracts. Thus the Census Tracts take the mean value of the cells that are within them.

Table 4.1. Weights and Buffer Distance per Facility-Destination Type		
Facility-Destination Type	Weight	Buffer Distance (m)
Bus Stops	2	400
Subway Entrances	4	600
Theaters	1	600
Green Markets	3	800
Parking Lots	2	400
Parks	4	700
Sport Facilities	3	400
Post Office	2	700
Schools	3	400
Universities and Colleges	2	400
Libraries	1	600
Day Care	2	400
Senior Centers	2	400
Outpatient Centers	3	400

The destination types were chosen according to their availability as geocoded data and their likelihood to be drivers of walking. For example, Frank et al. (2010) found that residing in a neighborhood with parks and open spaces in the vicinity increases the chances of an adult walking a home-based discretionary trip (e.g. shopping, recreation, etc.) by almost two times. The Cerrin et al. (2007) survey responses also revealed that people chose frequently to walk to parks. Another study showed that nearby grocery stores and schools increase walking too (Lee and Moudon, 2006). Particularly, schools have been consistently correlated with physical activity in children (Lovasi et al. 2008; Curran et al., 2006; Robitaille et al., 2009; Grigsby-Toussaint et al., 2011).

The weights and distances were assigned according to suggestions of previous studies when possible. The weights indicate the relative importance of destination types to one another. Thus having a grocery store nearby is 3 times as important as having a theater, because they have been found in surveys to be drivers of walking (Lee and Moudon 2006), as well as the most common walking destination (Moudon et al., 2006; Cerrin et al., 2007). Additionally, as third age people rely more on their local neighborhoods (Berke et al., 2007) and walkability is rather important for the quality of life of older adults (Friedman et al., 2011), the service points-buffers of senior centers were weighted with double importance than other destination points. Destination types that were not found in other studies were given a subjective weight.

Distances commonly used to analyze proximity include 400 meters, 600 meters and sometimes more (Krizek and Johnson, 2006). These distances are commonly accepted distances that are believed to be usually walked according to the type of destination and the cause of trip. The distances also infer the relative importance of destination types to one another, as they indicate an increased probability to attract walkers even if they reside at a farther distance. In Figure 4.6 an example of buffers-service areas based on network analysis can be seen.

Figure 4.6 Example of buffers-service areas generated with network analysis

4.6 Environmental Friendliness – Sidewalk Roadbed Ratio (SRR)

In addition to the characteristics of household density, diversity, connectivity and proximity, several studies have documented the importance of pedestrian oriented design towards a friendly environment for walking. Safety, pleasant surroundings, accessible recreational facilities such as walking trails, and sidewalks are strongly related with a higher likelihood for walking. In this study, the coverage of sidewalks is accounted as a measure of environmental friendliness. Thus a good ratio of sidewalk coverage to road coverage infers a more desirable area for walking, which in turn affects walkability positively.

The ratio of sidewalk coverage in square meters (m^2) to the street-roadbed coverage in square meters (m^2) was calculated for each census tract. The higher the ratio is, the more walkable the area is expected to be. Figure 4.7 shows an example of the sidewalk and roadbed coverage.

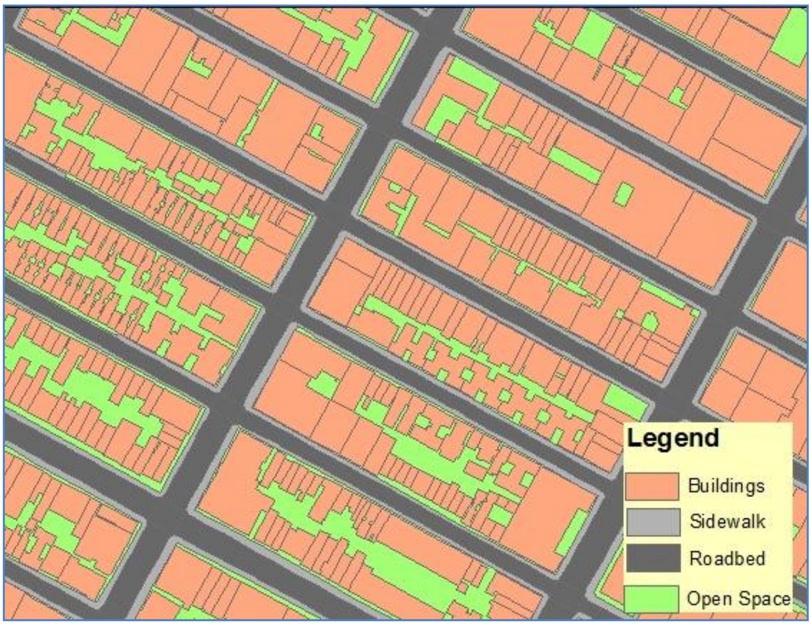


Figure 4.7 Example of roadbed coverage and sidewalk coverage

4.7 Commercial Density - Retail Floor Area Ratio (FAR)

The sixth component to be included in the walkability index is the retail Floor Area Ration (FAR). This measure is the ratio of the coverage of the retail building floor area footprint to the coverage of the total commercial land area. Although this simple measure has potential flaws due to limited availability of data with acceptable accuracy, it can reveal the general characteristics of a commercial area (Leslie et al., 2007).

Figure 4.8 Retail building floor area footprint to total commercial land area

Thus, a ratio equal to or close to one implies that less space is used for parking lots and the distance between other retail outlets is shortened. A visual representation of the FAR can be seen in Figure 4.8.

4.8 Calculating Walkability Index

The Walkability index was obtained by simply adding the partial walkability scores of each parameter after converting them to z-scores. Thus, each Census Tract got a score and a rank according to their respective characteristics that promote or hinder walkability. The normalization to z-scores was needed as none of the aforementioned measures-parameters is calculated in the same unit. The z-score of a raw score X is calculated according to the equation 8:

- $Z = \frac{X-\mu}{\sigma}$ (8)

  Z: z-score

  X: raw score

  μ: the mean of the raw scores of all Census Tracts

  σ: the standard deviation of the raw scores of all Census Tracts

The absolute value of Z indicates how far the raw score deviates from the mean in units of the standard deviation. Besides, Z-score is negative when the raw score is below the mean and positive when above.

Then the z-scores of each component were simply added according to the Equation 9 to provide the final walkability score of the block groups of each Census Tract.

- $WI = RD + EI + Con + Prox + SRR + FAR$ (9)

  WI: Walkability Index

  RD: Residential Density

  EI: Entropy Index

  Con: Connectivity

  Prox: Proximity

  SRR: Sidewalk Roadbed Ratio

  FAR: Retail Floor Area Ratio

5. Results and Discussion

In this chapter the results of this study are presented and discussed. At first the generated NYC land-use map is presented. The chapter continues with the six individual components of walkability that are addressed independently of the composite measure to have a better insight of the contribution of each parameter to the final result. Then the general results of this study's Walkability Index are presented and discussed more thoroughly. There is also a section where possible links between Walkability and Obesity are addressed. In the end of this chapter, some thoughts concerning GIS methods in calculating walkability are also addressed.

5.1 New York City Land-Use Map

Figure 5.1 shows the resulting NYC land-use map.

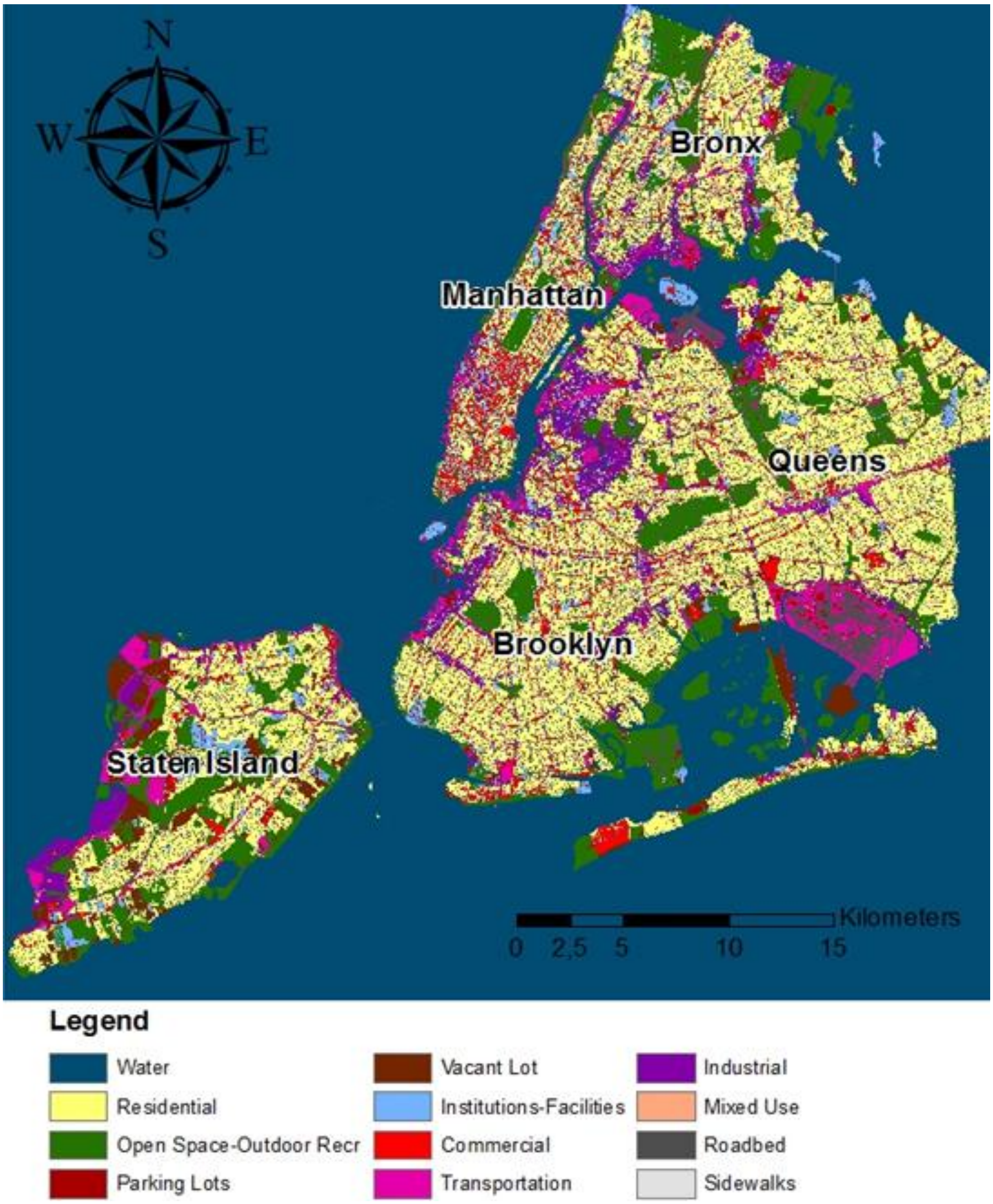


Figure 5.1 Land Use Map of NYC

Some general conclusions that can be drawn from Figure 5.1 are:

- Manhattan and South Manhattan are characterized by large areas with primarily commercial activities.
- A large number of facilities and institutions can be found in Manhattan.
- Large parts of Bronx, Brooklyn and Queens are characterized by primarily residential areas.
- The Southwest coastline of Bronx is characterized as primarily industrial.
- The West coastline of Queens is also primarily transportation or industrial area. A large harbor and a major train station can be found there.
- South Queens is mainly characterized by the infrastructure of the J.F. Kennedy international airport of NYC
- Brooklyn and Queens are mainly residential regions where commercial activity can be found mainly along some major avenues.
- Industrial activity is also evident along the West coastline of Brooklyn.
- A large part of Staten Island – mainly the West and central Staten Island – is characterized by primarily non-residential area.

5.2 Household Density

The Residential Density was calculated as the number of the total households divided by the amount of residentially designated land area (in m^2) within the Census Tracts blocks-group. Figure 5.2 shows which areas of NYC are considered as more residentially dense compared to other areas. From Figure 5.2 it is evident that the island of Manhattan is by far the most residentially dense NYC Borough. This was expected and can be explained by three possible reasons:

i. Manhattan is found to be the most densely populated area of NYC because it has a distinctive landscape with a large number of skyscrapers and huge buildings.
ii. Manhattan has large areas that are primarily used for commercial purpose, which in turn leaves less land to be used for residential purpose. As a consequence, the remaining residential areas tend to become more densely populated.
iii. Manhattan is the historic center of NYC which traditionally attracts a plethora of various activities, institutions and facilities that leave less area to be used for residential purpose. Less available residential land implies higher residential density.

At this point, it should be noted that the residents of these densely populated areas live in a region where a considerable number of various activities, institutions and facilities are expected to be in walking distance.

Staten Island scores naturally very low, because it is characterized by a wide industrial landscape. The same applies for the eastern part of Queens, while the South part of Brooklyn demonstrates a sparse residential development because these areas are mainly suburban areas with one or two family residences. Additionally, areas with large metropolitan parks – such as the Central Park of Manhattan - score very low for obvious reasons.

Another interesting element is the fact that NYC seems to have a typical centripetal structure. The residentially densest center is Manhattan, then all the adjacent areas of the neighboring Boroughs such as Southwest Bronx, North Brooklyn and East Queens follow in terms of residential density. The entire periphery scores remarkably low in terms of walkability.

To sum up, higher instances of residential density imply that there exists a population that can support other uses like commercial, retail or recreational. The densest area is the centre of NYC, followed by the closest adjacent areas of Bronx, Brooklyn and Queens. The most distant areas from the center score systematically low, and Staten Island demonstrates a remarkably low residential density.

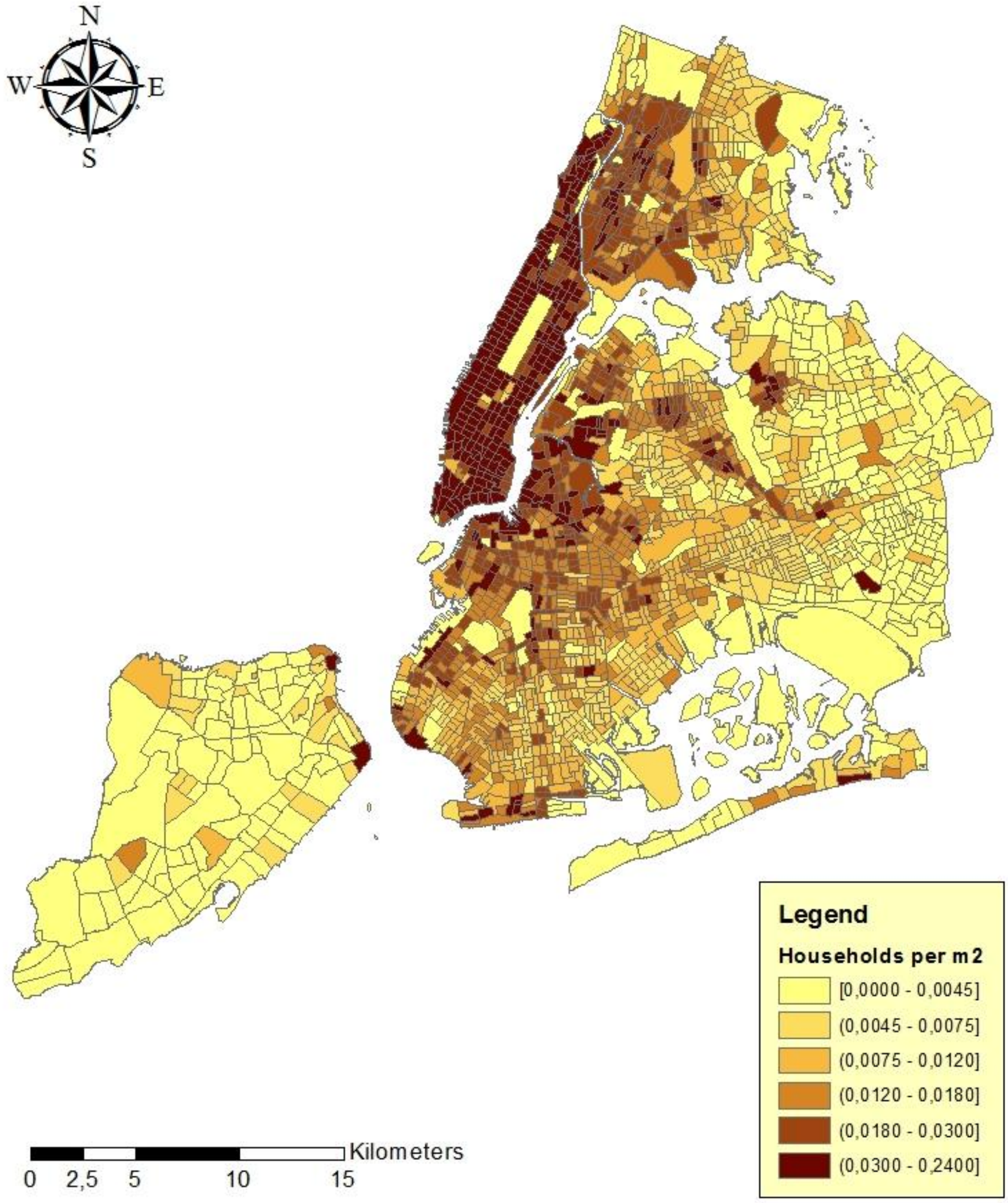


Figure 5.2 Residential Density Map of NYC - Households per amount of residential area

5.3 Diversity

As discussed in previous chapters, mixed land use can potentially support various uses into relative proximity, thereby shortening trip distances and promoting walking. To capture the homogeneity or heterogeneity of land use within the study area, the Entropy Index was employed and calculated for each Census Tract. A value close to one indicates a very heterogeneous area and a value close to zero implies decreased diversity. Figure 5.3 shows the ranking of all Census Tracts of NYC in terms of diversity.

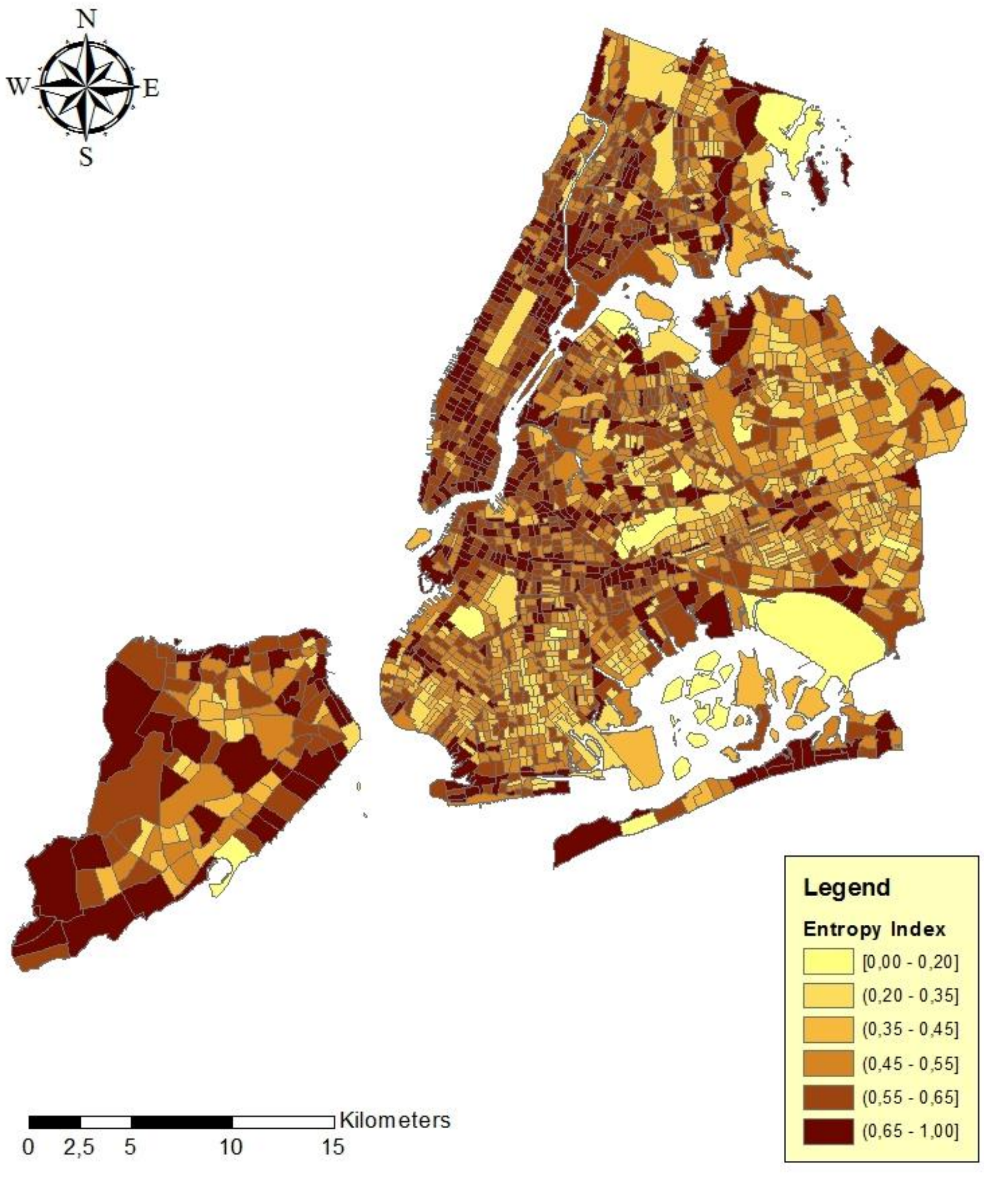


Figure 5.3 Diversity Map of NYC - Entropy Index

A major conclusion which can be drawn from Figure 5.3 is that Census Tracts that contain considerable percentage of commercial land tend to attract other types of land-uses as well, rendering the broader area of the Census Tract more heterogeneous. This can be seen by comparing Figure 5.3 with Figure 5.1. Census Tracts that contain large metropolitan parks score consistently low, because they are very homogeneous. Another obvious conclusion is that large Census Tracts tend to demonstrate increased diversity; a certain drawback of this measure which might have been avoided if a better spatial analysis unit was available for use. This is quite evident in the case of Staten Island (see Figure 5.3), where Census Tracts contain bigger amounts of land. Finally, the Census Tract that contains the J.F. Kennedy International airport of NYC scored close to zero, as expected (see example of homogeneity in Figure 4.2).

5.4 Connectivity

Connectivity affects the ease of travel between origin and destination and shows the degree to which roads, pedestrian walkways and trails are connected so that moving from place to place by walking is relatively easy. Connectivity was calculated as intersection density per m^2. This measure was preferred, instead of counting only the number of intersections within a Census Tract, because the geographic size of one Census Tract might be twice or even triple the size of another. Given this, normalizing the number of intersections of each Census Tract by the total area of Census Tract, was considered more appropriate for the analysis.

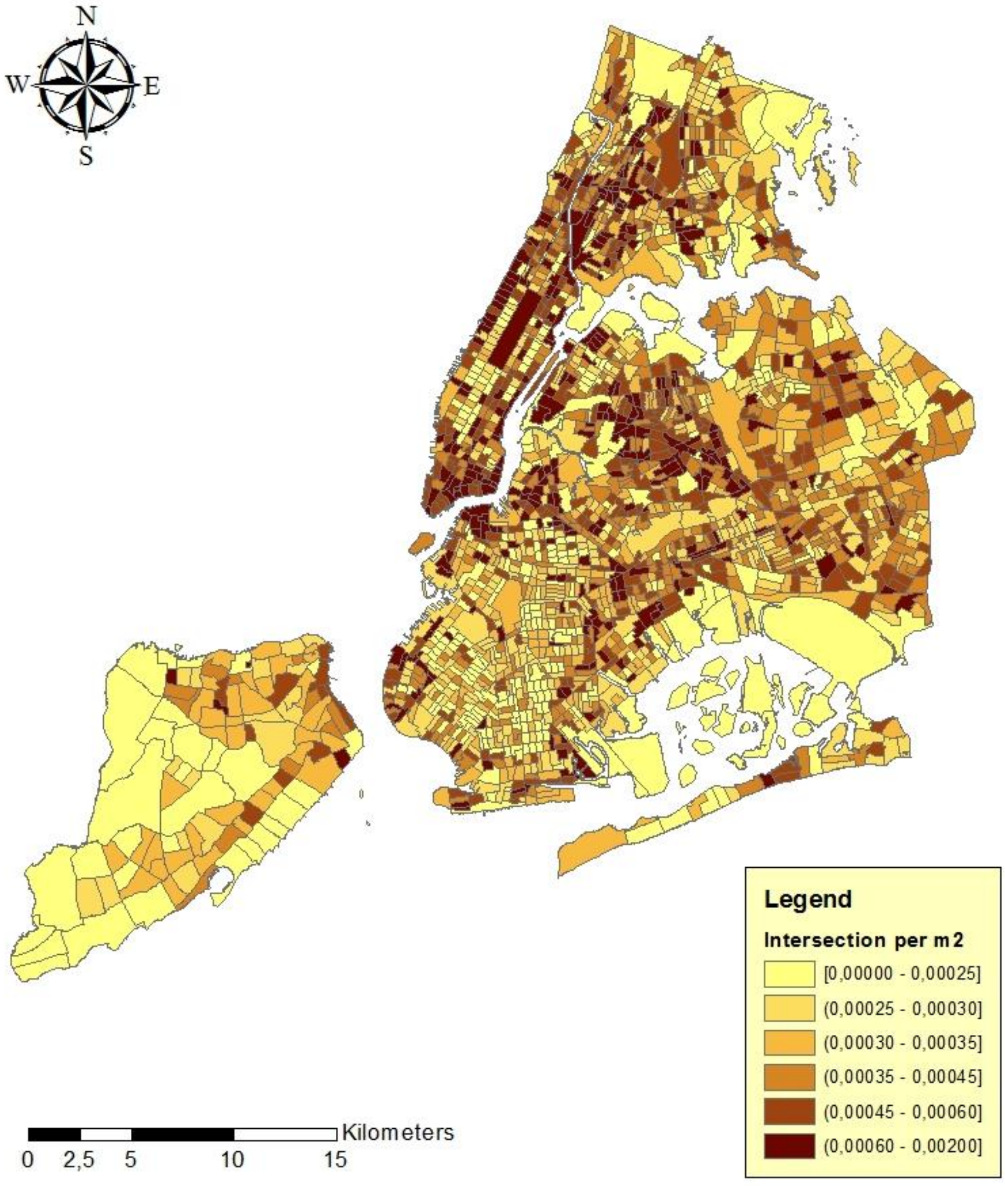


Figure 5.4 Connectivity Map of NYC - Intersection per m2

Figure 5.4 shows the intersection density of NYC. It can be easily seen from Figure 5.4 that Census Tracts that are sparsely populated also demonstrate decreased intersection density. A remarkable exception is the Census Tract of Central Park that scores unexpectedly high. This can be attributed to the fact that this metropolitan park has a rather developed network of roads.

5.5 Proximity

The map in Figure 5.5 shows the final result of the proximity calculations. This map has some apparent similarities with the map in Figure 5.2. The centripetal structure of NYC is once again obvious.

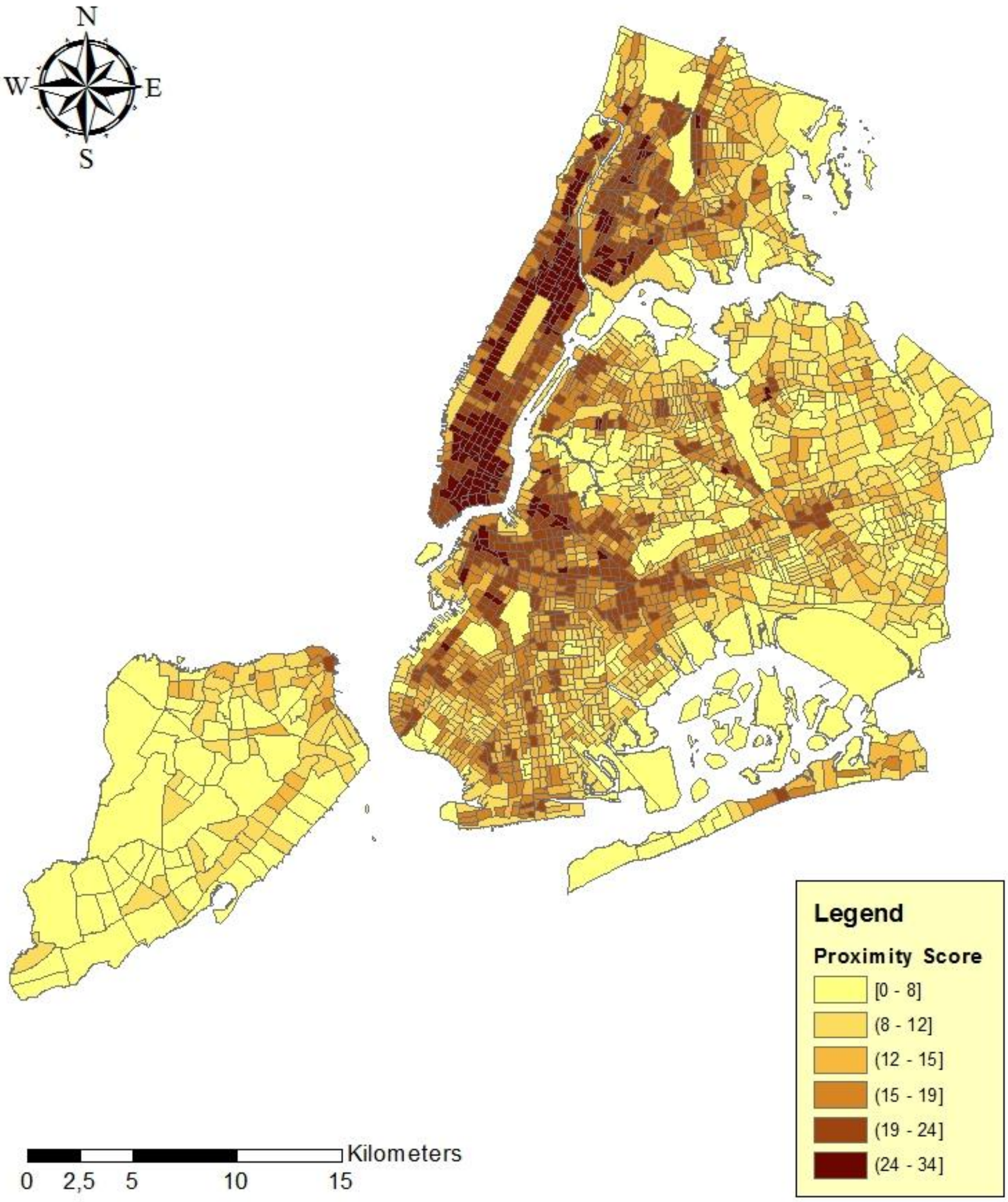


Figure 5.5 Proximity Map of NYC - Proximity Score

Proximity scoring is considerably higher in Manhattan, followed by the neighboring adjacent areas of West Bronx, North Brooklyn and West Queens. Additionally, the periphery of NYC including almost the entire area of Staten Island scores very low.

The explanation of this outcome has a logic similar to that of residential density. As proximity and directness between destinations increase, distance between destinations is expected to decrease. Subsequently, as the distance from place to place decreases, walking trips will more likely substitute driving trips. The island of Manhattan has by far the most facilities and supports more activities such as recreation and retail than any other part of NYC. On the other hand, homogeneous areas such as metropolitan parks, ports, airports or heavy industrial areas are expected to score very low as these places do not support various uses and interests, rendering them less attractive for walking.

An unexpected result is that places that are adjacent to metropolitan parks are considered privileged and score comparatively high. Census Tracts that coincide with the area of the metropolitan parks are expected to attract comparatively less trip makers than other facilities or recreational places. In other words, although parks and outdoor recreational sites admittedly support walking very much, proximity measures eventually the likelihood of walking to a place and not the accessibility or supportiveness of walking. This limitation of the measure is going to be discussed further in the next sections.

5.6 Environmental Friendliness for Walking

In addition to the characteristics of residential density, diversity and street pattern connectivity, the existence of adequate sidewalk and walking trails is considered to be rather important and is expected to promote physical activity and walking as discussed above. Figure 5.6 shows the result of analyzing an aspect of environmental friendliness in terms of sidewalk coverage to roadbed coverage ratio.

Some apparent conclusions that can be drawn from the map in Figure 5.6 are that Census Tracts whose use is primarily residential, commercial or both commercial and residential (mixed) score comparatively high. On the other hand, parks and highly homogenous areas which are characterized basically as industrial or are dedicated mainly to transportation infrastructure score comparatively very low. The reason why Census Tracts that contain large metropolitan parks score very low can be easily attributed to the fact that parks usually do not have sidewalks.

Finally, the centripetal structure of NYC, although less obvious, is still existent once again. Manhattan, North Brooklyn, West Bronx and Queens include the most Census Tracts that score higher than 0,65. On the other hand, the entire Staten Island does not include Census Tracts that score higher than 0,55, meaning that the industrial character of this Borough has left an “indelible mark” in the structure and development of the broader area of the island, affecting eventually a large number of other neighborhood characteristics too.

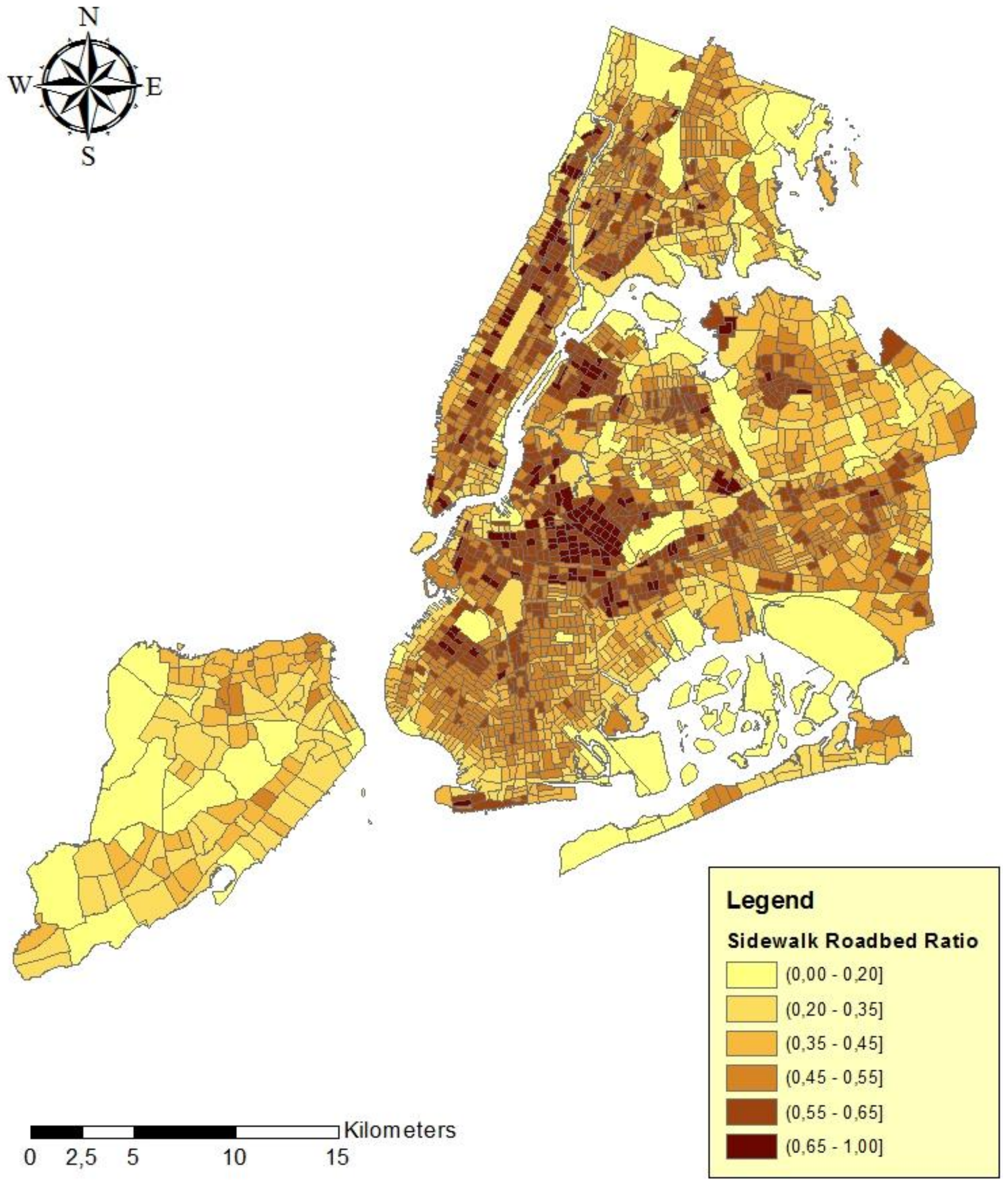


Figure 5.6 Environmental Friendliness for Walking Map of NYC -Sidewalk Roadbed Ratio

5.7 Commercial Density – Retail Floor Area

The retail floor area ratio (FAR), also known as commercial density, is an alternative measure of density indicator. FAR is also an indicator of pedestrian-oriented design and is frequently used in conjunction with land use mix (LUM). As explained in previous chapters, high values of FAR imply that a higher percentage of a commercial parcel is occupied by retail buildings, while lower numbers indicate that a considerable amount of land is used for parking.

In this study, FAR was calculated as the ratio of retail buildings' footprint in m^2 to the total area of commercial land in m^2. Figure 5.7 shows the commercial density map of NYC. As expected, higher values exist in Census Tracts whose land is characterized mainly as primarily commercial and contain malls or shopping centers. However, according to Handy et al. (2002) walking trips to local independent shops are more common. The Census Tracts of Manhattan and Census Tracts that contain sections of major shopping streets fall in the aforementioned category. On the other hand Census Tracts characterized by a predominance of industrial development or Census Tracts that coincide with the major metropolitan parks of NYC score rather low, as expected.

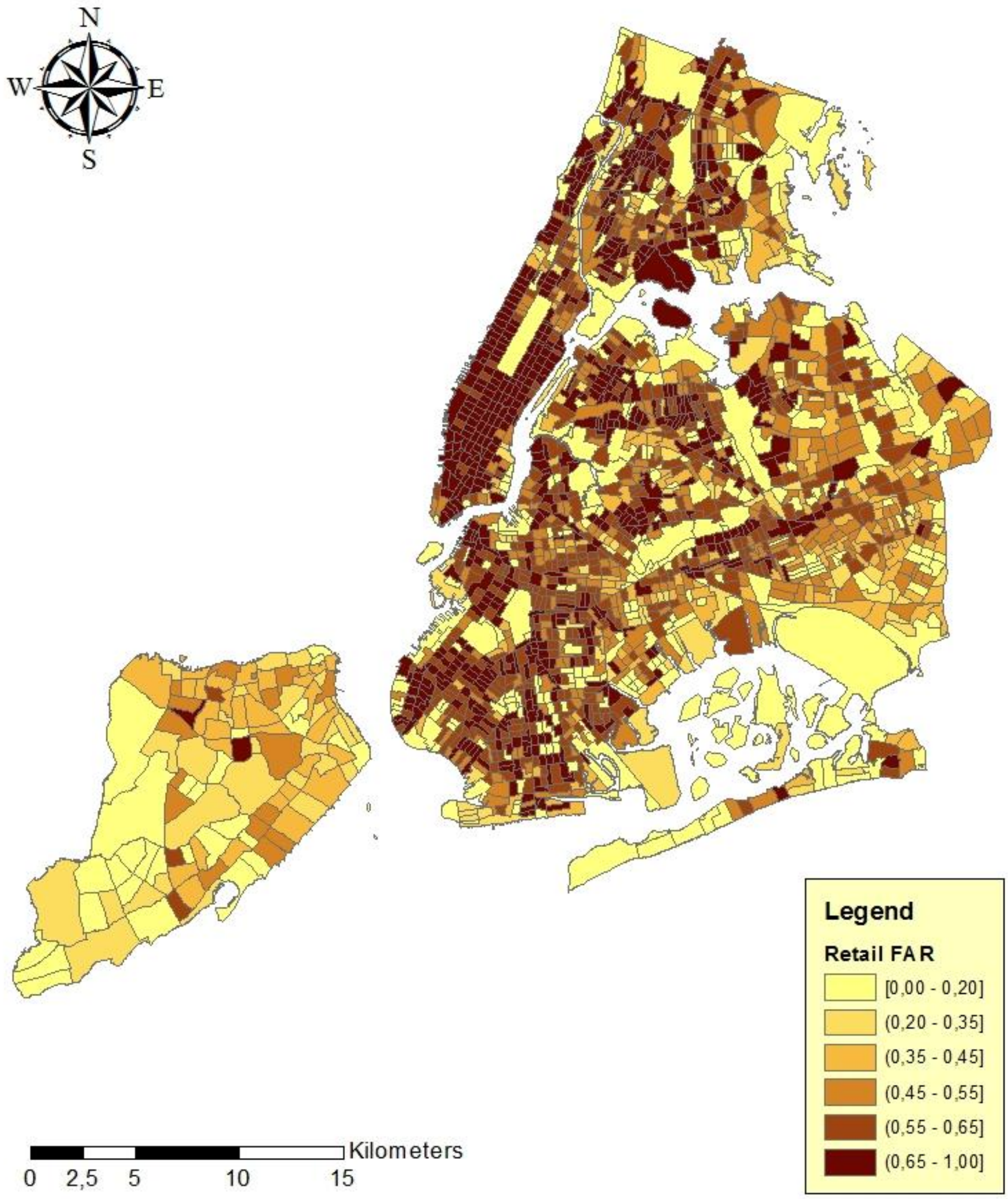


Figure 5.7 Commercial Density Map of NYC - Retail FAR

5.8 Results – Walkability of NYC

The Walkability Index was calculated based on the six walkability measures. The final result in Figure 5.8 shows the calculated walkability map of NYC.

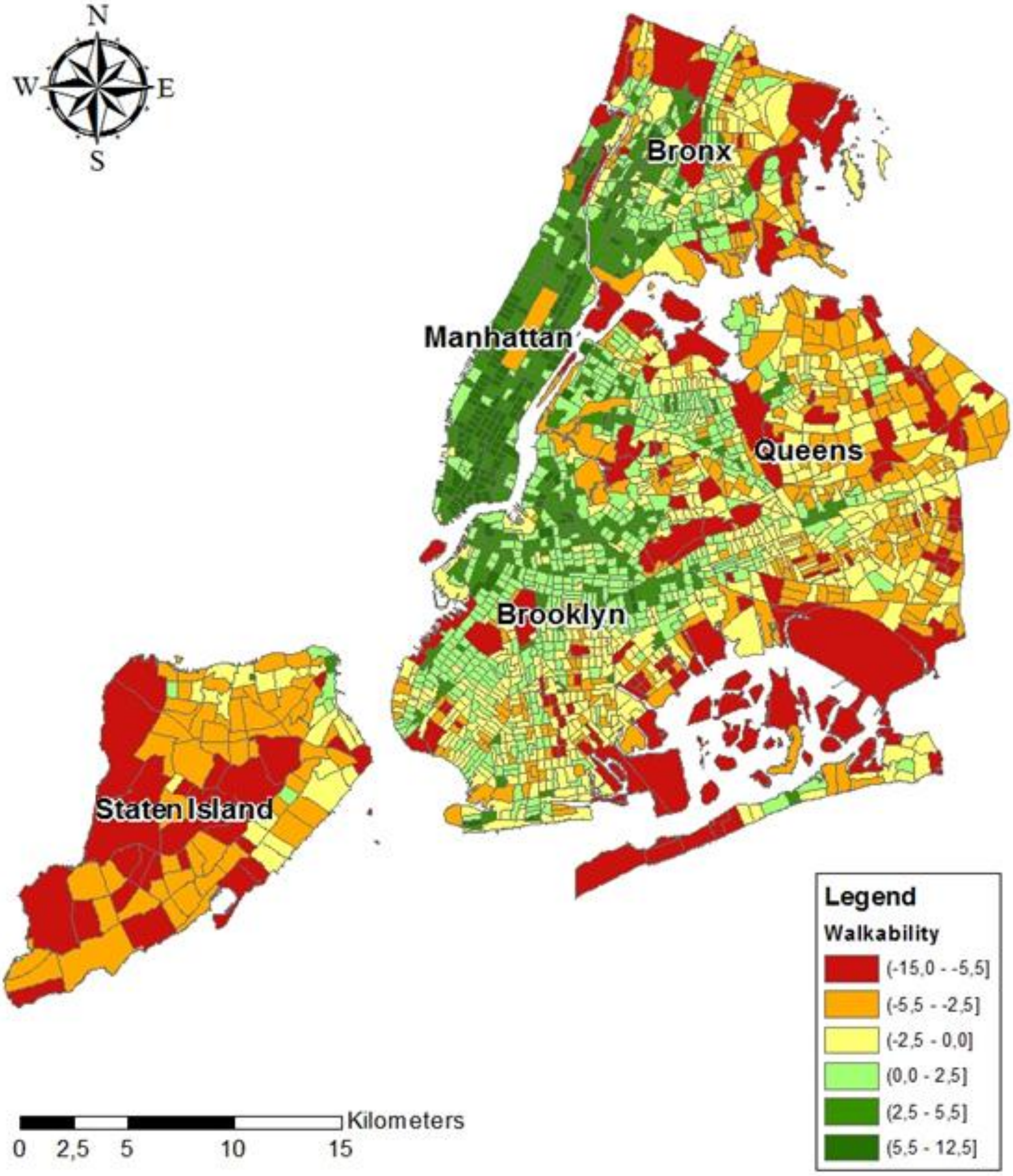


Figure 5.8 Walkability Map of NYC

The most walkable area of NYC is by far the island of Manhattan, followed by the adjacent areas of North Brooklyn, South and Southwest Bronx, and West Queens. Staten Island is the least walkable Borough. East, South and Southwest Queens, North Bronx and Southeast Brooklyn scored also very low. Census Tracts that contain large sections of metropolitan parks have also negative walkability, while Census Tracts with commercial land are fairly walkable . Areas that are primarily industrial or have large parts occupied by transportation infrastructure are characterized by rather negative walkability. The results indicate clearly a centripetal development of NYC, suggesting at the same time that the older, more central areas are more walkable, while the periphery, the suburbs and more recently developed areas are less supportive of walking.

It is also interesting to compare the resulting walkability heatmap of this study with the maps of WalkScore (Figure 2.5) and Walkshed (Figure 2.6). The reader will easily notice obvious similarities between these maps – the greenest parts and the red parts of all maps coincide, suggesting that the results of this study are fairly accurate. In particular, Staten Island was found in all cases the least walkable Borough, while Manhattan was found to be the Borough that can sustain and promote walking the most. It is also quite interesting to compare the findings with the Figure 5.1. It can be seen that parks and large recreational open areas are suggested by all studies – WalkScore, Walkshed and present one – to be the least walkable sites of NYC. This might be attributed to the fact that all walkability methods are mainly calculating the likelihood of an area to be walked rather than the accessibility of an area by pedestrians.

Another finding is that highly industrialized areas and areas with large transportation infrastructure (ports, airports, major train stations etc) are regarded to be unsupportive for walking by all methods, which is quite a logical and expected outcome. Neighborhoods that contain considerable amount of commercial land are suggested to be preferred for walking. Finally, one more interesting similarity between the maps in Figure 2.5, Figure 2.6 and Figure 5.8 is that East and South Queens and South Brooklyn, which constitute the suburban regions of NYC are less supportive to walking.

Some differences can be found in the region of North Bronx that in the present study is suggested to be less walkable compared to which WalkScore and Walkshed. Another obvious difference that highly affects the appearance of the overall results is the use of different spatial units and different approaches in the calculations, as both WalkScore and Walkshed do not use any predefined spatial unit. Instead they employ mainly the method of buffers around geocoded points.

Apparently, more thorough and deeper comparative analysis would have been useful to search possible numerical agreement between the results of each method. However, the numerical data of the previous projects are not accessible and thus any comparison is limited basically to noticeable and easily observed similarities and differences between Figures 2.5, 2.6 and 5.8 without any further statistical analysis.

5.9 Linking Obesity with Walkability

In order to compare the estimated walkability of NYC with the obesity rates, the average obesity rate of each neighborhood was estimated based on the NYC DOHMH surveys of every year from 2002 to 2011 (NYC DOHMH, 2013). This is done because the obesity percentages vary a lot from year to year. The resulting choropleth map of average obesity percentage can be seen in Figure 5.9 in the next page beside a smaller map of walkability in order to compare and contrast the two maps easier. In this new obesity map six classes are used (the same number of classes is also used in the NYC walkability map) to make comparisons easier.

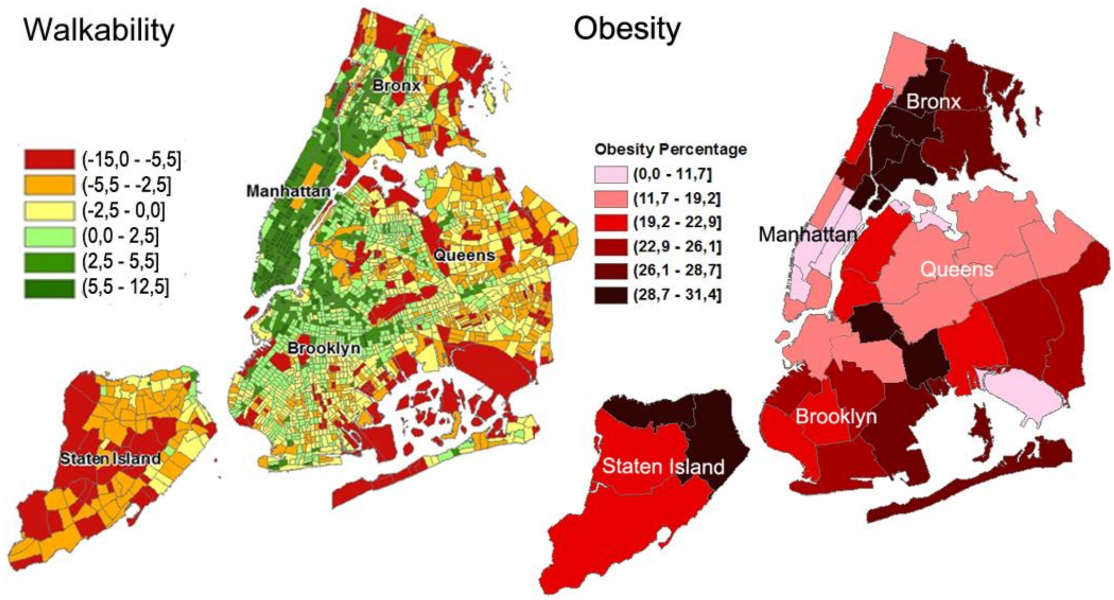


Figure 5.9 Comparison of Walkability map and Average Obesity Rate map

From Figure 5.9 it can be seen that it is difficult to find any profound relation between walkability and obesity percentages. The only clear consistency can be found in Manhattan, where the higher walkability levels seem to coincide with low obesity percentages. In Northwest Brooklyn where a region of high walkability is found, the obesity rates are comparatively low. In South Brooklyn where walkability decreases, obesity rates seem to increase. In the case of Bronx and Queens no clear correlation can be found. For Staten Island, it seems that walkability and obesity are inversely related. North Staten Island is more walkable compared to South Staten Island but the higher rates of obesity occur in the North of the island.

So, it is quite risky to try to control the results of this study for possible correlations with the obesity spatial epidemic of NYC of Figure 5.9 or even of Figure 2.4. Firstly, the spatial units of this study are much smaller than those used by the New York City Department of Health and Mental Hygiene (NYC DOHMH). Secondly, the data collected from NYC DOHMH on obesity vary considerably from year to year and can be eventually misleading.

The findings and results of NYC Walkability map can be used though for further analysis in future research taking advantage of the fact that walkability was computed for each Census Tract. For example, the results can be analyzed and searched for correlations with the socioeconomical data of the Census, to reveal possible social patterns and provide an insight to other possible aspects of walkability.

5.10 Limitations

The method of this study draws heavily from previous studies on walkability, suggesting modification and improvements. However, limitations still exist and deeper understanding of the parameters might be required. In this section some major limitation and issues are addressed for future research.

Pre-defined spatial units (i.e. Census Tracts) were used in this study to construct measures of the built environment. In some cases these pre-defined units could possibly be misleading. For example, some of the Census Tracts on Staten Island are very big, which in turn leads to decreased accuracy. Another issue is that the boundaries of Census Tracts might often be critically restrictive in the analysis. For example, a small Census Tract characterized by a homogeneous land-use might be surrounded by Census Tracts with heterogeneous land-uses, rendering the analysis potentially false or misguiding. People who live in the small homogeneous Census Tract might not have access to many facilities in their Census Tract, but as their Census Tract is small, they can easily walk to the neighboring heterogeneous Census Tracts to find their desired facilities. This means, that although they live in a homogenous Census Tract, they still live in a heterogeneous region. A possible solution to enhance accuracy has been given by some studies which have employed location information (e.g. addresses, postal codes) so as to define unique areas for each trip maker (Butler et al., 2011; Feng et al., 2010; Brownson et al., 2009).

In the present study, buffers were also used in an attempt to suggest a "hybrid" method that combines the buffering technique with the use of predefined spatial unites to suggest an improved method. The use of buffers combined with network analysis is a common method to define spatial units around geocoded locations at a given distance. Although this method can potentially provide a more representative assessment of the neighborhood built environment, there still exist important limitations. For example, different types of barriers such as rivers, fences, walls, multilane roads with heavy traffic are frequently present in the landscape preventing or making walk along the shortest part difficult (Biba et al., 2010). Such kind of information is not usually included in network databases because the cost and the time to collect such information are rather prohibitive (Biba et al., 2010). However, ways to integrate barriers to a model designed to measure accessibility is essential, because they are recognized to increase considerably walking distance. Promising examples are the methods applied by the projects WalkScore (WalkScore, 2010) and Walkshed (Azavea, 2010) that use a penalty system or friction-distance system respectively to adjust estimated walkability as discussed in chapter 2.

Many researchers might frequently face problems regarding GIS data layers. For example, GIS data layers might not be available for certain geographic areas or they can be rather expensive to acquire (Brownson et al., 2009). Due to economic limitations the official land-use maps of NYC's authorities could not be obtained. So the land-use map had to be generated as explained in chapter 3. The constructed land-use map has lower accuracy and precision than the official land-use maps of NYC's authorities. Another limitation of the present study is the fact that the calculations of the Proximity component do not include other destinations and points of interest suggested by other similar studies (e.g. restaurants, bars, super markets), as it was difficult to acquire this type of data due to time and cost constraints.

Another possible limitation of Walkability Indices that needs special attendance is the fact that all the parameters are equally weighted. This is something that future researchers should address when trying to find methods to optimize walkability modeling. In this study another source of uncertainty is the weights and distances that were used for the calculation of Proximity parameter, because some of these factors were set subjectively (e.g. post offices, theaters, libraries), as discussed in detail in chapter 4.5.

The benefits of Walkability Indices have been addressed several times in this report, but the use of the indices raise methodological concerns regarding validity, reliability, and generalizability, because many of these composite measures are developed for a specific setting (Feng et al., 2010; Frank et al., 2010). Composite measures might be proven less useful when assessing the impact of an intervention, because it is difficult to identify the specific component that should be given the highest priority for change (Fent et al., 2010).

Another issue that needs further research is the computation of Diversity variables. Christian et al. (2011) examined different entropy based computations of Diversity that were used in the development of Walkability Indices (WI) and their association with walking behavior. Different combinations of land uses in the Diversity calculations of WI were found to be relevant for different types and amounts of walking, suggesting the need for further research to define and identify alternative or complimentary measures of the environment.

5.11 Validating the Walkability Index

Systematic filed observation (audits) and self-reported information will help improving and assessing the validity of the proposed WI but, as Duncan et al. (2011) show, the systematic field observations can be very laborious, time demanding, often require specialized training and are remarkably costly. Regarding self-reported measures of neighborhood walkability, several problems might arise in terms of reliability, validity, low or biased response and an incompetent sample of respondents.

Another issue that needs attendance is that observed relationships cannot be interpreted as definitely causal. This means that available evidence establishes correlations between walking and neighborhood attributes or built environment. Yet correlation is not the same as causality (Kenny, 1979). There is still a debate among researchers whether self-selection explains and validates the observed correlation, and if for examples people who prefer to walk, eventually decide to live in more walkable neighborhoods (Handy et al., 2006). Beside this a considerable number of research reviews address the importance of using specific measures of the built environment for specific types of physical activity, and expand their criticism stressing that studies on built environment and walking do not frequently take into account the possibility that walking substitutes for other forms of physical activity (Saelens et al., 2008).

5.12 Further Development of Walkability Methods

The built environment can be measured based on a variety of data sources, with different computational methods and under various characterizations of the environment. In particular in the proposed Walkability Index of the present study further improvement could have been achieved if there was a separate component for parks or if an additional parameter of Environmental Friendliness was used that takes into account crime rates. Another issue that has to be address from future researchers is the lack of high quality data on public health. In our study we tried to search possible links between walkability and obesity rates. Yet this attempted remained incomplete due to insufficient information on the spatial epidemic of obesity in NYC.

Although GIS based measures have evolved and improved over the years, there is still wide variability in how GIS variables are constructed to represent built environment concepts (Feng et al., 2010). In particular, the lack of standardization among built environment definitions might be challenging for future researches and might prevent synthetic analysis. Thus, it might eventually hinder the identification of environmental characteristics and the search for evidence that might provide a better understanding of physical activity behavior (Butler et al., 2011).

Spatial extent, source of data, and the number and range of places compared across studies is frequently so variable, that apparently there are no two studies that evaluated built environment metrics handling the aforementioned features in the same way (Feng et al., 2010). Last but not least is the fact that variations in terminology also cause further complications when it comes to choose the appropriate methodology.

In general, it should be acknowledged though that limitations have been identified for all types of built environment measures. Additionally, thanks to the existing measures, rapid advancements have been achieved in understanding the relation between the built environment and physical activity in a variety of settings and populations. Nevertheless, most measures of the built environment are considered to be first-generation measures and further development is needed. Cross- and inter-disciplinary collaboration is probably the key to deal with the challenge of standardizing measures and terminology, improving the technical quality of measure and ensuring relevance for diverse populations.

It is important to measure objectively built environment indicators that can provide clear information about the scale of walkability in a community. GIS can be very assistive in the analysis of objective measures of the built environment. A better understanding on how land use may influence walking behaviors can assist public health practitioners identify their needs, facilitate research, monitor health outcomes, suggest possible interventions and help planning authorities in decision and policy making. Furthermore, GIS can also help to convey complex built environment and land use information in a user-friendly approach, providing to the wider public a tool to assess their quality of life.

6. Conclusion

In this study a new Walkability Index was developed. Walkability Indices measure the degree to which an area provides opportunities to walk to various destinations. This Walkability Index (WI) draws partially from the WI that was developed by Frank et al. (2010). Thus, it contains all 4 parameter suggested by Frank et al. plus 2 new parameters. The 6 parameters are:

- Residential Density
- Diversity – Entropy Index
- Connectivity
- Proximity
- Environmental Friendliness
- Commercial Density – FAR.

In order to construct measures of the built environment, the pre-defined spatial unit of Census Tract is employed. Buffers that define spatial units around geocoded locations at a given distance are also used and integrated in the calculations in an attempt to suggest a hybrid method that combines buffering with pre-defined spatial unites.

The study area is New York City (NYC). It was found that Manhattan is the most walkable Borough and Staten Island the least walkable Borough. NYC was found to have a centripetal structure, meaning that the historical center and the entire island of Manhattan is more developed and more walkable. The regions of Bronx, Brooklyn and Queens that are adjacent to Manhattan were found to have a bit lower walkability compared to Manhattan. The farthest areas of NYC’s periphery have consistently scored the lowest walkability. Neighborhoods that are extremely homogeneous in terms of land-use and do not include considerable number of commercial parcels were also found to score very low. Hence, Census Tracts that are mainly characterized by primarily industrial land-use or contain large transportation infrastructure (e.g. ports, airports, large train stations) or large metropolitan parks display limited walkability.

The results and findings seem to coincide with the results of previous studies. However, the comparison is simple and due to lack of data based on easily observed patterns. A number of limitations is acknowledged, followed by suggestions for future research. The particular limitations of this study are:

- Inaccuracies and precision issues in the generated land-use map of NYC
- Lack of data, subjective assignment of a number of factors (weights and distances) and simple network analysis in the calculation of Proximity parameter
- Limitations in the final result due to the use of the pre-defined spatial units of Census Tracts.
- Equal weights for all 6 parameters of the WI.

The validity of the new WI remains to be assessed. Future researches can incorporate the results of this study and search for correlations with the socioeconomical data of the Census. Addition of more parameters such as a separate parameter for parks or a parameter related to crime rates might improve our WI and therefore could be a subject for further research.

References

Alfonzo, M., Boarnet, M.G., Day, K., Mcmillan, T., Anderson, C.L., (2008). *The Relationship of Neighborhood Built Environment Features and Adult Parents' Walking*. Journal of Urban Design 13, 29–51.

Azavea . (2010). *About Walkshed*. Available: http://walkshed.org/nyc. Last accessed 1st June 2013.

Bassuk, S.S., Manson, J.E., (2005). *Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease*. J. Appl. Physiol. 99, 1193–1204.

Berke, E.M., Koepsell, T.D., Moudon, A.V., Hoskins, R.E., Larson, E.B., (2007). *Association of the built environment with physical activity and obesity in older persons*. Am J Public Health 97, 486–492.

Berrigan, D., Pickle, L.W., Dill, J., (2010). *Associations between street connectivity and active transportation*. International Journal of Health Geographics 9, 20.

Biba, S., Curtin, K.M., Manca, G., (2010). *A new method for determining the population with walking access to transit*. International Journal of Geographical Information Science 24, 347–364.

Black, J.L., Macinko, J., Dixon, L.B., Fryer, G.E., Jr, (2010). *Neighborhoods and obesity in New York City*. Health Place 16, 489–499.

Brennan Ramirez, L.K., Hoehner, C.M., Brownson, R.C., Cook, R., Orleans, C.T., Hollander, M., Barker, D.C., Bors, P., Ewing, R., Killingsworth, R., Petersmarck, K., Schmid, T., Wilkinson, W., (2006). *Indicators of activity-friendly communities: an evidence-based consensus process*. Am J Prev Med 31, 515–524.

Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., Sallis, J.F., (2009). *Measuring the Built Environment for Physical Activity*. American Journal of Preventive Medicine 36, S99–S123.e12.

Butler, E.N., Ambs, A.M.H., Reedy, J., Bowles, H.R., (2011). *Identifying GIS measures of the physical activity built environment through a review of the literature*. J Phys Act Health 8 Suppl 1, S91–97.

C.D.C. - Centers for Disease Control and Prevention. (2012). *Obesity and Overweight*. Available: http://www.cdc.gov/obesity/downloads/obesity_trends_2010.ppt. Last accessed 12th Mar 2013.

Cerin, E., Saelens, B.E., Sallis, J.F., Frank, L.D., (2006). *Neighborhood Environment Walkability Scale: Validity and Development of a Short Form.* Medicine & Science in Sports & Exercise 38, 1682–1691.

Christian, H.E., Bull, F.C., Middleton, N.J., Knuiman, M.W., Divitini, M.L., Hooper, P., Amarasinghe, A., Giles-Corti, B., (2011). *How important is the land use mix measure in* *understanding walking behaviour? Results from the RESIDE study*. International Journal of Behavioral Nutrition and Physical Activity 8, 55.

Cook, S., Weitzman, M., Auinger, P., Nguyen, M., Dietz, W.H., (2003). *Prevalence of a metabolic syndrome phenotype in adolescents: findings from the third National Health and Nutrition Examination Survey, 1988-1994*. Arch Pediatr Adolesc Med 157, 821–827.

Curran, A., Grant, J., Wood, M., (2006). *Indicators for community action: built environment and community health*. Journal of Rural and Community Development. 2, 59-74.

De Lusignan, S., Hague, N., Van Vlymen, J., Dhoul, N., Chan, T., Thana, L., Kumarapeli, P., (2006). *A study of cardiovascular risk in overweight and obese people in England*. European Journal of General Practice 12, 19–29.

Dobesova, Z., Krivka, T., (2012). *Walkability Index in the Urban Planning: A Case Study in Olomouc City.*

Doyle, S., Kelly-Schwartz, A., Schlossberg, M., Stockard, J., (2006). *Active Community Environments and Health: The Relationship of Walkable and Safe Communities to Individual Health*. Journal of the American Planning Association 72, 19–31.

Duncan, D.T., Aldstadt, J., Whalen, J., Melly, S.J., Gortmaker, S.L., (2011). *Validation of Walk Score® for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas*. International Journal of Environmental Research and Public Health 8, 4160–4179.

Erlbaum, N., (2005). *Assessment of 2001 New York State NHTS add-on data using empirical and auditable data sources*. Journal of Transportation and Statistics 8, 1.

Feng, J., Glass, T.A., Curriero, F.C., Stewart, W.F., Schwartz, B.S., (2010). *The built environment and obesity: a systematic review of the epidemiologic evidence*. Health Place 16, 175–190.

Finkelstein, E.A., Trogdon, J.G., Cohen, J.W., Dietz, W., (2009). *Annual medical spending attributable to obesity: payer-and service-specific estimates*. Health Aff (Millwood) 28, w822–831.

Flegal, K.M., (2010). *Prevalence and Trends in Obesity Among US Adults, 1999-2008.* JAMA: The Journal of the American Medical Association 303, 235.

Forsyth, A., Schmitz, K.H., Oakes, M., Zimmerman, J., Koepp, J., (2006). *Standards for environmental measurement using GIS: toward a protocol for protocols*. Journal of Physical Activity & Health 3, S241.

Frank, L.D., Devlin, A., Johnstone, S., van Loon, J., (2010). *Neighbourhood design, travel, and health in Metro Vancouver: using a walkability index - executive summary*. Vancouver: University of British Columbia.

Frank, L.D., Saelens, B.E., Powell, K.E., Chapman, J.E., (2007). *Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity?* Soc Sci Med 65, 1898–1914.

Frank, L.D., Sallis, J.F., Saelens, B.E., Leary, L., Cain, K., Conway, T.L., Hess, P.M., (2009). *The development of a walkability index: application to the Neighborhood Quality of Life Study*. British Journal of Sports Medicine 44, 924–933.

Friedman, D., Parikh, N.S., Giunta, N., Fahs, M.C., Gallo, W.T., (2011). *The influence of neighborhood factors on the quality of life of older adults attending New York City senior centers: results from the Health Indicators Project*. Quality of Life Research 21, 123–131.

Gauvin, L., Richard, L., Craig, C.L., Spivock, M., Riva, M., Forster, M., Laforest, S., Laberge, S., Fournel, M.-C., Gagnon, H., Gagné, S., Potvin, L., (2005). *From walkability to active living potential: an "ecometric" validation study*. Am J Prev Med 28, 126–133.

Glazier, R.H., Booth, G.L., Institute for Clinical Evaluative Sciences in Ontario, (2007). *Neighbourhood environments and resources for healthy living a focus on diabetes in Toronto*. Institute for Clinical Evaluative Sciences, Toronto, ON.

Grigsby-Toussaint, D.S., Chi, S.-H., Fiese, B.H., (2011). *Where they live, how they play: Neighborhood greenness and outdoor physical activity among preschoolers*. International Journal of Health Geographics 10, 66.

Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., (2002). *How the built environment affects physical activity: views from urban planning*. American journal of preventive medicine 23, 64–73.

Handy, S., Cao, X., Mokhtarian, P.L., (2006). *Self-Selection in the Relationship between the Built Environment and Walking: Empirical Evidence from Northern California*. Journal of the American Planning Association 72, 55–74.

Hammond, R., Levine, (2010). *The economic impact of obesity in the United States. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 285.

Haslam, D.W., James, W.P.T., (2005). *Obesity*. Lancet 366, 1197–1209.

Jervey, B., (2006). The *big green apple: your guide to eco-friendly living in New York City*. Insiders' Guide, Guilford, Conn.

Kenny, D.A., (1979). *Correlation and causality*. Wiley, New York.

Krizek, K.J., Johnson, P.J., (2006). *Proximity to trails and retail: effects on urban cycling and walking*. Journal of the American Planning Association 72, 33–42.

Krizek, K.J., Johnson, P.J., (2006). *Proximity to trails and retail: effects on urban cycling and walking*. Journal of the American Planning Association 72, 33–42.

Lachapelle, U., Frank, L., Saelens, B.E., Sallis, J.F., Conway, T.L., (2009). *Commuting by public transit and physical activity: where you live, where you work, and how you get there*. Journal of physical activity & health, 8, 1, 72-82.

Larsen, K., Gilliland, J., (2008). *Mapping the evolution of "food deserts" in a Canadian city: Supermarket accessibility in London*, Ontario, 1961–2005. International Journal of Health Geographics 7, 16.

Lee, C., Moudon, A.V., (2006). *The 3Ds + R: Quantifying land use and urban form correlates of walking*. Transportation Research Part D: Transport and Environment 11, 204–215.

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

Lovasi, G.S., Moudon, A.V., Pearson, A.L., Hurvitz, P.M., Larson, E.B., Siscovick, D.S., Berke, E.M., Lumley, T., Psaty, B.M., (2008). *Using built environment characteristics to predict walking for exercise*. International Journal of Health Geographics 7, 10.

Lovasi, G.S., Schwartz-Soicher, O., Neckerman, K.M., Konty, K., Kerker, B., Quinn, J., Rundle, A., (2012). *Aesthetic Amenities and Safety Hazards Associated with Walking and Bicycling for Transportation in New York City*. Annals of Behavioral Medicine 45, 76–85.

Manaugh, K., El-Geneidy, A., (2011). *Validating walkability indices: How do different households respond to the walkability of their neighborhood?* Transportation Research Part D: Transport and Environment 16, 309–315.

Marshall, S., (2005). *Streets & patterns*, 1st ed. ed. Spon, London ; New York. 120.

MMAHO and OPPI - Ministry of Municipal Affairs and Housing, Ontario Professional Planners Institute., (2009). *Planning by design: a healthy communities handbook*. Toronto: Ministry of Municipal Affairs and Housing.

Moudon, A.V., Lee, C., Cheadle, A.D., Garvin, C., Johnson, D., Schmid, T.L., Weathers, R.D., Lin, L., (2006). *Operational definitions of walkable neighborhood: theoretical and empirical insights*. Journal of Physical Activity & Health 3, S99.

MTA - Metropolitan Transportation Authority. (2012). The MTA Network - Public Transportation for the New York Region. Available: http://www.mta.info/mta/network.htm. Last accessed 10th Jul 2013.

NYC DCP – New York City Department of City Planning. (2013). *ZoLa Zoning and Land Use NYC.* Available: http://gis.nyc.gov/doitt/nycitymap/template?applicationName=ZOLA. Last accessed 20th Jul 2013.

NYC DHOMH - New York City Department of Health and Mental Hygiene., (2013). *Epiquery: NYC Interactive Health Data System - [Community Health Survey 2011]*. Available: http://nyc.gov/health/epiquery. Last accessed 26th Mar 2013.

NYSDOH - New York State Department Of Health., (2011). *Obesity Prevention*. Available: http://www.health.ny.gov/prevention/obesity/. Last accessed 26th Mar 2013.

OASIS CUR- Open Accessible Space Information System Center of Urban Research. (2010). *NYC Land-Use map.* Available: http://www.oasisnyc.net/map.aspx. Last accessed 20th Jul 2013.

Ogden, C.L., Carroll, M.D., Kit, B.K., Flegal, K.M., (2012). *Prevalence of obesity in the United States, 2009–2010*. NCHS data brief, no 82. Hyattsville, MD: National Center for

Health Statistics. Available: http://www.cdc.gov/nchs/data/databriefs/db82.pdf. Last accessed 10th Apr 2013.

Oliver, L.N., Schuurman, N., Hall, A.W., (2007). *Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands*. Int J Health Geogr 6, 41.

Park, S., (2008). *Defining, measuring, and evaluating path walkability, and testing its impacts on transit users' mode choice and walking distance to the station*. ProQuest.

Peeters, A., Barendregt, J.J., Willekens, F., Mackenbach, J.P., Al Mamun, A., Bonneux, L., NEDCOM, the Netherlands Epidemiology and Demography Compression of Morbidity Research Group, (2003). *Obesity in adulthood and its consequences for life expectancy: a life-table analysis*. Ann. Intern. Med. 138, 24–32.

Pelletier, A., Paquin, S., Chartrand, A., Montréal, Q., n.d. (2009). *Are all walking audits equivalent? A comparison of three walking audits in a Montreal neighbourhood*.

Robertson, R., Robertson, A., Jepson, R., Maxwell, M., (2012). *Walking for depression or depressive symptoms: A systematic review and meta-analysis*. Mental Health and Physical Activity 5, 66–75.

Robitaille, É., Laguë, J., Institut national de santé publique du Queebec, Direction du développement des individus et des communautees, (2009). *Indicateurs géographiques de l'environnement bâti et de l'environnement des services influant sur l'activité physique, l'alimentation et le poids corporel*. Direction du développement des individus et des communautés, Institut national de santé publique du Québec, [Montréal, Qué.].

Rundle, A., Neckerman, K.M., Freeman, L., Lovasi, G.S., Purciel, M., Quinn, J., Richards, C., Sircar, N., Weiss, C., (2009). *Neighborhood food environment and walkability predict obesity in New York City*. Environ. Health Perspect. 117, 442–447.

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*. Ann Behav Med 25, 80–91.

Saelens, B.E., Handy, S.L., (2008). *Built Environment Correlates of Walking: A Review*. Medicine & Science in Sports & Exercise 40, S550–S566.

Shay, E., Rodriguez, D.A., Cho, G., Clifton, K.J., Evenson, K.R., (2009). *Comparing objective measures of environmental supports for pedestrian travel in adults*. International Journal of Health Geographics 8, 62.

Song, Y., Rodriguez, D., (2005). *The measurement of the level of mixed land uses: a synthetic approach*. Chapel Hill: Carolina Transportation Program.

The Lewin Group ., (2010). *Bending the Health Care Cost Curve in New York State: Options for Saving Money and Improving Care*. Available: http://nyshealthfoundation.org/uploads/resources/health-care-cost-curve-technical-appendix-july-2010.pdf. Last accessed 26th Mar 2013.

The Mori Memorial Foundation. (2009). Global Power City Index 2009. Available: http://www.mori-m-foundation.or.jp/english/research/project/6/pdf/GPCI2009_English.pdf. Last accessed 10th Jul 2013.

Thompson, D., Brown, J.B., Nichols, G.A., Elmer, P.J., Oster, G., (2001). *Body Mass Index and Future Healthcare Costs: A Retrospective Cohort Study*. Obesity 9, 210–218.

Tomalty, R., Haider, M., Smart Growth BC, (2009). *BC sprawl report: walkability and health 2009*. Smart Growth BC, Vancouver, B.C.

Tomlin, C.D., (1990). *Geographic Information Systems and Cartographic Modeling*. Englewood Cliffs: NJ: Prentice-Hall. xi.

Tucker, P., Irwin, J.D., Gilliland, J., He, M., Larsen, K., Hess, P., (2009). *Environmental influences on physical activity levels in youth*. Health Place 15, 357–363.

UCB - United States' Census Bureau. (2010). Geographic Terms and Concepts - Census Tract . Available: http://www.census.gov/geo/reference/gtc/gtc_ct.html. Last accessed 10th Jul 2013.

UCB - United States' Census Bureau. (2012). Geographic Terms and Concepts - Census Tract . Available: http://quickfacts.census.gov/qfd/states/36/3651000.html. Last accessed 10th Jul 2013.

Vargo, J., Stone, B., Glanz, K., (2012). *Google walkability: a new tool for local planning and public health research?* Journal of Physical Activity and Health 9, 689-697.

Visscher, T.L., Seidell, J.C., (2001). *The public health impact of obesity*. Annu Rev Public Health 22, 355–375.

WalkScore (2010). *Walk Score Methodology*. Available: http://blog.walkscore.com/wp-content/uploads/2010/12/WalkScoreMethodology.pdf. Last accessed 1st June 2013.

WalkScore. (2011). *Most walkable large cities - New York*. Available: http://www.walkscore.com/NY/New_York. Last accessed 10th Jul 2013.

Weinberger, R., Kaehny, J., Rufo, M., (2010). *US Parking Policies: an overview of management strategies.*

WHO - World Health Organization. (2003). *Obesity and Overweight.* Available: http://www.who.int/dietphysicalactivity/media/en/gsfs_obesity.pdf. Last accessed 12th Mar 2013.

WHO - World Health Organization. (2010). *Global Recommendations on Physical Activity for Health*. Available: http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf. Last accessed 20th Mar 2013.

WHO - World Health Organization., (2006). *Diet and physical activity: a public health priority*. Available: http://www. who.int/dietphysicalactivity/en. Last accessed 20th Mar 2013.

WHO - World Health Organization., (2007). Current prevalence and trends of overweight and obesity. In: Branca, F., Nikogosian, H., Lobstei, T., *The challenge of obesity in the WHO European Region and the strategies for response*. Denmark: World Health Organization. 1-11.

Reports in Geodesy and Geographic Information Technology

The TRITA-GIT Series - ISSN 1653-5227

13-001 **Mohammad Amin Alizadeh Khameneh.** Tree Detection and Species Identification using LiDAR Data. Master of Science thesis in Geodesy No. 3127. Supervisor: Milan Horemuž. January 2013.

13-002 **Duc Khanh Ngo**. Relief Planning Management Systems - Investigation of the Geospatial Components. Master of Science thesis in Geoinformatics. Supervisor: Tuong Thuy Vu and Yifang Ban. February 2013.

13-003 **Hulumtaye Kefyalew**. Investigation of the use of Laser Scanning for Deformation Monitoring. Master of Science thesis in Geodesy No. 3128. Supervisor: Milan Horemuž. June 2013.

13-004 **Annmari Skrifvare**. Change Detection in Stockholm between 1986 and 2006 using SPOT Multispectral and Panchromatic Data. Supervisor: Yifang Ban. June 2013.

13-005 **Axel Bronder and Erik Persson**. Design, Implementation and Evaluation of a Mobile GIS Solution for a Land Registration Project in Lesotho. Master of Science thesis in Geoinformatics. Supervisors: Gyözö Gidófalvi and Jakob Riise. June 2013.

13-006 **Camilla Isaksson.** HTML5-based Travel Habit Application: Investigation of Vector Web Mapping Possibilities. Master of Science thesis in Geoinformatics. Supervisor: Gyözö Gidófalvi. June 2013.

13-007 **Nuosuola Li.** Large-Scale Realistic Macro-Simulation of Vehicle Movement on Road Networks. Master of Science thesis in Geoinformatics. Supervisor: Gyözö Gidófalvi. June 2013.

13-008 **Sara Jonasson.** Jämförelse av objekt-orienterad ECHO med Pixelbaserad MLC för klassificering av fjärranalysdata. Supervisor: Hans Hauska. September 2013.

13-009 **Maria Valasia Peppa.** Precision analysis of 3D camera. Master of Science thesis in Geodesy No. 3129. Supervisor: Milan Horemuž. September 2013.

13-010 **Sofie Ekström.** Geografisk Tillgänglighet i Hälsovalet Gävleborg. Master of Science thesis in Geoinformatics. Supervisor: Takeshi Shirabe and Johanna Runarson. September 2013.

13-011 **Rui Zhu.** Moving Object Trajectory Based Intelligent Traffic Information Hub. Master of Science thesis in Geoinformatics. Supervisor: Gyözö Gidófalvi. November 2013.

13-012 **Jing Sun**. Semi-automated rapid damage assessment using high-resolution satellite imagery: a case study of the 2008 Wenchuan earthquake, China. Master of Science thesis in Geoinformatics. Supervisors: Thuy Vu and Yifang Ban. December 2013.

14-001 **Muhammad Zaheer.** Kinematic orbit determination of low Earth orbiting satellites, using satellite-to-satellite tracking data. Master of Science thesis in Geodesy No. 3130. Supervisor: Milan Horemuž. March 2014.

14-002 **Razmik Agampatian.** Using GIS to measure walkability: A Case study in New York City. Master of Science thesis in Geoinformatics. Supervisor: Hans Hauska. April 2014.

TRITA-GIT EX 14-002

ISSN 1653-5227

ISRN KTH/GIT/EX--14/002-SE