Centre for Geo-Information

Thesis Report GIRS-2016-01

The impact of local food environments on diet

Do neighbouring food retailers influence what you eat?

Barbara Sienkiewicz

Date 29.01.2015

WAGENINGEN UNIVERSITY
WAGENINGEN UR

The impact of local food environments on diet

Do neighbouring food retailers influence what you eat?

Barbara Sienkiewicz

Registration number 88 01 25 760 020

Supervisors:

dr. ir. Ron van Lammeren

dr. ir. Anouk Geelen

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

29.01.2015

Wageningen, the Netherlands

Thesis code number: GRS-80436

Thesis Report: GIRS-2016-01

Wageningen University and Research Centre

Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements and Foreword

The process of writing this master thesis would be very difficult without the help of my supervisors: dr. ir. Ron van Lammeren and dr. ir. Anouk Geelen. Thank you for your support and for making my work possible (and easier). Our meetings gave me a lot of ideas and helped me to develop the appropriate strategy. I am very grateful for that. I also want to thank you for reading my drafts and giving me valuable comments. Finally, I want to thank you for all of your suggestions and moral support.

This project would not be feasible without diet data. NQplus data I obtained from Division of Human Nutrition (with the help of dr. ir. Anouk Geelen) was really crucial. I could not investigate a Dutch case without it. Therefore, I want to thank you for providing it.

Working on this project was a challenge, but with the support of my friends I succeeded in finishing. Thank you for having faith in me.

I am aware that this study is just the beginning of investigating local food environments as a possible factor in influencing diet in the Netherlands. Nevertheless, I am sure it has potential and can be used to create a healthier environment for Dutch citizens. Diet is an important aspect of our lives because it influences our health. Society gets sicker and sicker every day (also because of diet). I wish that this process would reverse one day.

Abstract

Geographic Information Systems (GIS) have improved our understanding of the variations of food location availability and accessibility and its relation with diet and weight. However, studies on this topic are relatively new to Europe. Therefore, investigating how local food environments influence the diets of Europeans is of interest. To my knowledge, no studies investigating this problem in the Netherlands have been conducted. Consequently, this study is the first one investigating how the spatial distribution of food locations affects Dutch citizens. This investigation involved GIS methods to study the local food environment by calculating the density and proximity of food retailers. The methods used include: Euclidean distance, network distance, clustering to CBS neighbourhoods and kernel density. Three main variables of diet: DHD (Dutch Healthy Diet) index, calorie intake and BMI (Body Mass Index) - were investigated in relation to a food environment (density and proximity of a certain retailer).

Main findings indicate that there may be a relationship between BMI and the following retailers: restaurants, cafes, grocers and supermarkets, and takeaways. It was found that people with high and low BMI are clustered. High BMI clusters (obese people) lived closer to grocery stores and supermarkets than people from low BMI clusters (normal weight). Besides that, people from normal weight clusters lived in places with higher densities of restaurants and cafés than in places where obese people lived. It can be concluded, then, that the more restaurants and cafés there are in your neighbourhood, the less likely you are to be obese. It was also found that the further away your closest meal delivery, convenience store, takeaway, grocery store or supermarket is, the less likely you are to gain weight and become obese/overweight. The accuracy of these assumptions, however, can be discussed. Certainly, it has to be investigated further because there is a chance that restaurants and cafes are not responsible for healthier diets.

These results indicate that the spatial configuration of food retailers is influencing diet in the Netherlands. However, the strength of this relation is unknown. Therefore, it is suggested to investigate this problem further, using larger groups of people and new techniques like GPS tracking. This may help us to understand this relation better.

Keywords: *GIS, local food environment, BMI, nutrition, diet quality, food availability, Dutch Healthy Diet Index, spatial analysis, eating behaviour, residence characteristics, spatial configuration of food environment*

Table of Contents

List of figures

List of tables

List of abbreviations

BMI – Body Mass Index

DHD-index – Dutch Healthy Diet index

DQI – Diet Quality Index

DQI-P – Diet Quality Index for Pregnancy

kcal – kilocalorie

MAUP - Modifiable Areal Unit Problem

MSOA – geographical areas across England known as Middle Super Output Areas

OSM – Open Street Map

POI – Points of Interest

SpatCon D – Spatial Configuration of Diet Specification

SpatCon F – Spatial Configuration of Food Retailers

1 Introduction

1.1 Context and background

The rates of obesity have been rapidly growing in the past years (Kosti and Panagiotakos, 2006), and it is unlikely that it has stopped. This "epidemic" has started to become a major issue, especially in developed countries. Increased availability of food in combination with sedentary lifestyles are suspected to be main risk factors of this problem (Banwell et al., 2005). Theoretically, weight gain is caused by greater energy consumption than expenditure (Burgoine et al., 2011). The important question is: what causes the increase in calorie intake? The answer to this question is not easy because the obesogenic environment is multi-dimensional. According to Burgoine et al. (2011), to understand obesity we have to fully understand "the way individuals choose to behave within, interact with and react to, the diverse range of environments". Therefore research on the diverse range of dimensions is crucial.

Table 1 Examples of GIS measures used in studies on diet and obesity

What was measured:	Measured by:
Distance from each child's home to the closest fast food restaurant	(Larsen et al., 2015)
Presence of small food store (within 100 m) or supermarket (within 1000m) of residence	(Bodor et al., 2008)
Presence of BMI-healthy and BMI-unhealthy stores within 800 m of participants' residences	(Jennings et al., 2011)
Distance to nearest food store	(Jago et al., 2007)
Travel time to the nearest supermarket and convenience store along the road network	(Pearce et al., 2008)
Distance to closest food store	(Timperio et al., 2008)
Density of food retailers per square mile per neighbourhood	(Gibson, 2011)
The number and type of food stores within the census tract	(Gustafson et al., 2011)

The local food environment has been proven to be an independent predictor of individual food choice and diet quality in developed countries (McKinnon et al., 2009; Moore et al., 2008). As mentioned by Caspi et al. (2012), most of the research in the category of local food environment and diet studies used

Geographic Information Systems (GIS)-based measures. The fact that these measures are widely used shows that GIS concepts and techniques can contribute to research on local food environments. Some examples of measures mentioned are shown in Table 1.

A food environment (or foodscape – term used by Pearce et al. (2007)) includes features like the proximity to or density of food retailers in people's neighbourhoods (Canada, 2013). It is being investigated if and how these features can influence diet. We know that our diet is directly connected to our health. Researchers linked unhealthy diets to diseases like cancer (Willett and Trichopoulos, 1996), diabetes (Swinburn et al., 2001), hypertension (Taubes, 1997), birth defects (Botto et al., 2000), heart disease (Yen, 1998) and many more. An unhealthy diet is probably caused by many factors. One of them may be the local food environment. This assumption has already been investigated by researchers. Bodor et al. (2008) investigated the relationship between the local food environment and consumption of fruits and vegetables. They found that greater fresh vegetable availability within 100m of a residence was a positive predictor of vegetable intake. The relationship between fast-food outlet access and consumption of food or nutrients was examined by Fraser et al. (2012). They found that the relationship between the accessibility of outlets and consumption do vary over space. The influence of the environment (defined as 'walkability', food availability and deprivation) on Body Mass Index (BMI) and fruit and vegetable consumption was investigated by Burgoine et al. (2011). Their findings suggest that a few elements of both walkability and food availability are significantly associated with BMI and fruit and vegetable intake.

There were also studies focused on a specific age group, e.g. children. Fraser and Edwards (2010) researched the association between childhood overweight/obesity and the density/proximity of fast food outlets in relation to the child's residential postcode. This study found that a higher density of fast food outlets was significantly associated with the child being obese. The relationship between neighbourhood food outlets and weight status/dietary intake of children was investigated by Jennings et al. (2011). They found that the availability of BMI-healthy outlets in neighbourhoods was associated with lower body weight.

Most of the studies were performed outside Europe. For that reason, we still do not know how the food environment affects the dietary patterns of Europeans. The European food environment may differ from the ones outside Europe because of different street patterns, different distance perception, or different walking behaviour. Therefore it could be worth investigating.

Few studies were already performed in Europe. Most of them were conducted in England. All of them were investigating the existence of food deserts (neighbourhoods with poor access to healthy foods (Boone-Heinonen et al., 2011)). Clarke et al. (2002) studied Cardiff and Leeds/Bradford (UK) area. They identified 6 food deserts there. Donkin et al. (1999) investigated the same problem in London where a few deserts were also found. Cummins et al. (1999) used the area of Glasgow and did not find any food deserts there.

No studies on the local food environment of the Netherlands were found. Therefore this will be a first study investigating if the Dutch local food environment influences Dutch Healthy Diet (DHD) index, kcal intake and BMI. The obtained results will be compared with those from other studies. Additionally, the methodology developed in this study can be used again in similar studies in the future.

1.2 Problem definition

The influence of the spatial distribution of food retailers on the dietary intake was not yet studied in the Netherlands. It may differ from the associations found in USA because, as it was mentioned in paragraph 1.1, the local food environment may be different there. Besides that, diet behaviour in USA may differ from the European one. Therefore, this study is meant to investigate the relation between these two components (diet and food environment) in the Netherlands.

In order to investigate it, two components have to be defined:

- spatial configuration of individual's diet specification,
- spatial configuration of food retailers on a local food environment level.

It is crucial to find appropriate variables describing those two components.

Defining a local food environment in the Netherlands may be challenging because the topic of food environment is relatively new in Europe. However, looking at studies from other countries (especially outside Europe) and analysing their methods can be helpful.

1.2.1 Spatial configuration of diet specification

Dietary intake can be described by energy intake or intake of specific nutrients/ foods, but also by diet quality. When dietary intakes get attached to the homes of (study) respondents, they become spatial. In previous studies, different diet variables have been investigated. Consumption of fruits and vegetables was used by Shearer et al. (2014), whereas Fraser (2012) has used fast food consumption, and Boone-Heinonen (2011) has investigated food environment impact on diet quality. In general, dietary patterns have been studied in many countries. In the Netherlands, the Dutch Healthy Diet Index (DHD-index) can be used to study diet quality (Lee, 2014). The index combines components on:

- vegetables,
- fruit,
- fish,
- dietary fibre,
- saturated fatty acids,
- trans fatty acids,
- sodium,
- alcohol.

Scores for each component range between 0 (no adherence) and 10 (complete adherence) points. Consequently the range of the DHD index is 0-80 (the higher the index, the better the quality of the diet).

1.2.2 Spatial configuration of food retailers

The second required component of the local food environment can be defined in many ways. Literature describes two main approaches to measure a local foodscape: a density approach and a proximity approach. A **density approach** quantifies the spatial density of available food outlets using the buffer method, kernel density estimation or spatial clustering (Charreire et al., 2010). A **proximity approach** estimates the distance from the study respondent to the closest food outlet by measuring distances or travel time (Charreire et al., 2010). To obtain the most accurate possible measurements (for both proximity and density), street network dataset is often used. The locations of streets are used to calculate distances between participants and the closest retailer. Besides that, road network is used to create network buffers (see 2.3.2 Spatial representation of the food retailers). Street network is publicly available data that can be easily obtained from open sources. Researchers obtained it mostly from municipalities or by using OpenStreetMap (OSM).

Booth et al. (2001) have suggested that food availability and accessibility may also be important determinants of dietary intake and related health outcomes. Therefore those measures were also used in previous studies. **Availability** was applied as: store presence (Gustafson et al., 2011), store density (Murakami et al.,

2009) or variety; **accessibility** was calculated as the distance between store and participant's home (Michimi and Wimberly, 2010), or as travel time on that distance (Pearce et al., 2008). Mentioned variables were used to find the associations between food environment and individual dietary behaviours.

Relationships between the local food environment and weight status (Jago et al., 2007), and perceived availability of healthy food and weight status (Moore et al., 2008) were also investigated. In all of those studies, the addresses of respondents and retailers were geocoded and used as references for GIS analyses (Charreire et al., 2010).

Glanz et al. (2005), with their food outlet oriented approach, divided the food environment into three groups: the community nutrition environment, the organizational nutrition environment and the consumer nutrition environment. **The community nutrition environment** is composed of the number, type, location, and accessibility of food outlets such as grocery stores, convenience stores, fast-food restaurants, and full-service restaurants. **The organisational nutrition environment** includes all food outlets within settings, such as schools and workplaces. **The consumer nutrition environment** is what consumers encounter in and around places where they buy food. Glanz et al. (2005) suggested that the associations of these 3 environments with diet patterns should be analysed separately.

1.2.3 Limitations

Previously completed studies encountered some issues: availability of data on food retailers was often limited or some methods did not work as intended. A few examples of the problems that researchers encountered are listed in Table 2.

A few things can be done to avoid possible issues. The accuracy of the data can be checked by ground-truthing. This study uses a few different data sources so if one of the datasets has low accuracy, one that is more accurate can easily replace it. Neighbourhood can be defined not only as an administrative unit, but also as an area covering places that can be easily encountered by study participants (by using the location of their homes and creating service areas). In order to achieve this, a street network is necessary. It makes the defined neighbourhood more accurate than zip-code areas.

Similarly to previous studies, some of the issues cannot be avoided, but I will try to minimalize their influence in order to achieve the best possible results.

Table 2 Limitations encountered by scientists

Limitation	Study
Inaccurate geolocations (the coordinates of the zip code centroid, or zip plus four centroid)	(Laraia et al., 2004)
Sampling errors and non-coverage of some areas (especially less-urbanized areas), different areas were covered in different time periods	(Sturm and Datar, 2005)
Incorrect buffer (the distance should be increased or decreased based on a person's physical ability to walk to the store and carry grocery bags)	(Algert et al., 2006)
Variations in urban form interfered with the statistical analysis	(Frank, 2006)
Neighbourhood predefined as the administrative unit (census areas)	(Pearce et al., 2006)
Taking into account the geographic position of supermarkets but not their characteristics	(Apparicio et al., 2007)
Small number of outlets (limited the ability to compare two areas using statistical test)	(Latham and Moffat, 2007)
Too homogenous research area (similar exposure to food stores, particularly supermarkets)	(Bodor et al., 2008)
Not considering obstacles other than geographic access - for example, financial barriers	(Pearce et al., 2008)

1.3 Research questions

The goal of this study is to investigate the impact of the local food environment on diet and weight in the Netherlands by using the Dutch case study. For that reason the following research questions have been formulated:

1. How to spatially express human diets?
2. How to spatially express the food environment?
3. What approach can be used to study the impact of the food environment on diet patterns?
4. Does a Dutch food environment influence dietary patterns?

1.4 Structure of the report

The report is organized in following way. Chapter 2 is a review of selected papers investigating the relation between a local food environment and human diets/weights. It includes the overview of the diet variables used in previous studies together with types of food retailers that were investigated. It also describes the methods of analysing relationships between diet and food environment and the results of it. Chapter 3 describes the methodology used in this study. Chapter 4 presents the results obtained by using methods described in chapter 3. Finally, Chapter 5 includes conclusions, discussion and recommendations for future studies.

2 Review

2.1 Introduction

A review of articles on local food environment is an important part of this study. Articles have been reviewed in order to identify and describe different methodological procedures, which can be used in order to assess the spatial accessibility of food outlets and their influence on people's diets. Some of the articles used in the following review were mentioned in Problem definition (subchapter 1.2).

Analysed literature included 35 articles from years 2002 – 2015. All of them used GIS methods to identify local food environment. Articles were collected via Scopus database and via Global Search on Wageningen University Library Website. First search queries used keywords: food environment, GIS, geographic information systems, diet, food, dietary patterns, and foodscape. Later "snowball method" was used in order to find more relevant articles. The overview of the results from reviewed studies is placed in Appendix II.

Found studies were conducted in a few countries. Table 3 shows the number of articles per country. 57,1 % of studies were conducted in USA. From European studies, only English and German ones were found (together they are 22,9 % of all reviewed studies). No Dutch studies have been found.

Table 3 Number of studies per country

Country	Number of studies	Percent of all studies
USA	20	57.1 %
England	7	20.0 %
Canada	4	11.4 %
New Zealand	2	5.8 %
Germany	1	2.9 %
Japan	1	2.9 %

In this review, the definitions of local food environment will be explored and analysed. Identified GIS methodologies will be used to support the methodology of this study. Additionally, dietary patterns used by researchers will be investigated. Finally, the results of all studies will be compared with a focus on the relationship between diet and food environment to identify the most suitable methods to explain this relationship.

2.2 Human diets

2.2.1 Variables

The aim of this study is an investigation of the food environment's impact on dietary intake and patterns in the Netherlands. In order to study this, variables of diet have to be defined. Reviewed studies used 27 diet descriptive variables, which I categorized into 5 groups: food intake, nutrient intake, indices, intake frequency (nutrients or foods) and other variables. Each group is explained in separate paragraphs below.

Food intake

In this category, the intake of specific items was measured. The intake of junk food, soda, sweets and salty snacks was measured by LeDoux & Vojnovic (2014). Shearer et al. (2014) measured healthy and unhealthy food intake. The intake of total consumed food (g/1000 kcal) was investigated by Murakami et al. (2009). Other investigated food items were: fruits and vegetables (Gustafson et al., 2011), grains (Laraia et al., 2004), and fast food (Boone-Heinonen et al., 2011). Fruit and vegetable intake was investigated most often (9 studies). Fast food intake was measured in 3 studies.

Sometimes, instead of measuring actual intake, the frequency of eating was investigated. Pearson et al. (2005) have measured frequency of fruit and vegetable intake. The frequency of consuming food was measured by Murakami et al. (2009).

Nutrient intake

In this group, the intake of specific nutrients was investigated. The intake of folate, iron, and calcium was used by Laraia et al. (2004). Buck et al. (2013) measured the intake of carbohydrates, and simple sugars. Fibre intake was investigated by Gustafson et al. (2011).

Indices

Diet can be also described by indices. Laraia et al. (2004) calculated diet quality index for pregnancy (DQI-P), diet quality index (Healthy Eating Index - 2005) was measured by Casagrande et al. (2011). Healthy and unhealthy diet scores were used by Smith et al. (2013). From these indices, DQI was measured most often (4 times).

Other variables

I created the category "other variables" for variables that do not fit into any category. This group includes variables like energy intake, body fat or portion size. The first variable was investigated by Shearer et al. (2014). Murakami et al. (2009) was interested in portion size, cooking methods, and general diet behaviour. Interestingly, parent encouraging healthful eating was also used as a variable (Berge et al., 2014).

Summary

Diet can be described by many variables. Reviewed studies used a wide range of diet variables. In order to have a clear insight into them, they were categorized according to the characteristics or type of food that they were describing. Each of the variables was used in one or more of the studies to investigate their relation with local food environment.

2.2.2 Spatial representation of diets

Different datasets describing diet were used. Most studies obtained data from other studies. Only researchers from 8 studies (22,86 % of all investigated studies) were collecting data by themselves. In that case, special questionnaires were created and all participants had to fill them.

Sample sizes varied from 102 random households (Bodor et al., 2008) to 1,477,828 adults with obesity (Michimi and Wimberly, 2010). Different groups of participants were measured. Some studies were focused on children (Sturm and Datar, 2005) while others studied adults (Gibson, 2011). Sometimes very specific groups of adults were studied, e.g. Laraia et al. (2004) investigated dietary behaviour of pregnant women. Most of the studies used locations of participant's homes (Morland and Evenson, 2009; Murakami et al., 2009; Spence et al., 2009) but a few studies used locations of schools (Smith et al., 2013).

2.3 Food retailers

2.3.1 Variables (food retailer types)

Food can be purchased in many places. Main categories are retailers where you can eat (e.g. restaurants, snack bars, cafés) and those where you can only purchase (e.g. supermarkets, grocery shops, convenience stores). Only eight of the studies were focused on a particular retailer's type. Laraia et al. (2004) were focused on supermarkets. Inagami et al. (2006) were investigating nearby groceries. A few years later, Inagami et al. (2009) investigated restaurant impact. Finally, Fraser et al. (2012) and Buck et al. (2013) looked at fast food restaurants. Most of the studies used more than one type of retailer in their research. Table 4 shows the overview of the store types used. Supermarkets and grocery stores were used most often (57,1 % of reviewed studies).

In twelve studies, retailers were categorized. Morland et al. (2002) categorized stores into chain stores and locally owned grocery stores. Healthy food outlets and unhealthy food outlets categories were created by Janevic et al. (2010). Other categories were:

- small, medium and large supermarkets (Michimi and Wimberly, 2010),
- BMI-healthy, BMI-intermediate and BMI-unhealthy (Jennings et al., 2011),
- fast-food outlets, other unhealthy outlets and mixed food outlets (Cetateanu and Jones, 2014).

In the first categorization, retailer's sizes were taken into account. In the second, Jennings et al. (2011) categorized supermarkets and fruit and vegetable stores as BMI-healthy. Takeout/fast-food outlets, and convenience stores were classified as BMI-unhealthy. In the third group (BMI-intermediate) were: non-fast-food restaurants and any other food shops. The third categorization used unhealthy and mixed food retailers.

Table 4 Retailer's types used in reviewed studies

Retailers type	Number of studies	Percentage of studies (n=35) which used this type
supermarkets	20	57,1 %
grocery stores	20	57,1 %
convenience stores	19	54,3 %
full service restaurants	18	51,4 %
fast food restaurants	17	48,6 %
bakeries	6	17,1 %
takeaway	4	11,4 %
meat stores	4	11,4 %
cafeterias	4	11,4 %
fish stores	3	8,6 %
chain stores	3	8,6 %
supercentres	2	5,7 %
speciality food stores	2	5,7 %
pizza	2	5,7 %
non cooperate owned stores	2	5,7 %
bars/taverns	2	5,7 %
warehouse club	1	2,6 %
small stores	1	2,6 %
sandwich stores	1	2,6 %

This sub-chapter was meant to create foundations for the methodology of this study. The overview of methods used in previous research will be helpful in developing the methodology for a Dutch study case. At that point, approaches and recommendations used before will be essential. Food environment may differ between countries. Therefore, it is recommended to compare food environments from countries where studies were conducted and find differences and similarities with the Dutch food environment. This knowledge will help to choose methods that can be used in this study.

2.3.2 Spatial representation of the food retailers

The local food environment and its association with health or diet has been widely investigated the last few years in observational research. Most of these studies used Geographic Information Systems technology as a measurement and analysis tool (Caspi et al., 2012). Measurements commonly used were **store density** or **proximity to food retailers** (Charreire et al., 2010). Both of these measurements have been calculated in a few ways. Table 5 presents the overview of the used methods.

Table 5 Overview of used GIS methods

GIS measure of food accessibility		Number of studies
Density		
Buffer	*Circular buffer*	8
	Network buffer	7
Kernel density		1
Spatial clustering		13
Proximity		
Euclidean distance		8
Distance by road or street		11
Population weighted distance		1
Modelling travel time		1

Density

In reviewed studies, the density of food retailers was measured using **circular buffer** (Bodor et al., 2008; Jago et al., 2007; Seliske et al., 2009), **network buffer** (Jennings et al., 2011; Larsen et al., 2015; Shearer et al., 2014), **kernel density** (Buck et al., 2013) or **spatial clustering** (Mehta and Chang, 2008; Morland et al., 2002; Sturm and Datar, 2005). Circular buffer method uses a circle area with a centre in investigated location (which often is the participant's home) and counts how many points of interest (e.g. restaurants) it overlaps. The network buffer method does the same count but for an area defined by the network (it uses actual walking access, rather than using a straight line distance, which is the case in circular buffer). The difference between circular and network buffer is presented in Figure 1.

Most often, density of retailers was measured by clustering food retailers to census tract or zip-code district. It means that predefined neighbourhoods, like zip-code districts, were used to count the number of retailers within each of them. When the density for each neighbourhood was calculated, it could be joined with the participant record based on her or his neighbourhood.

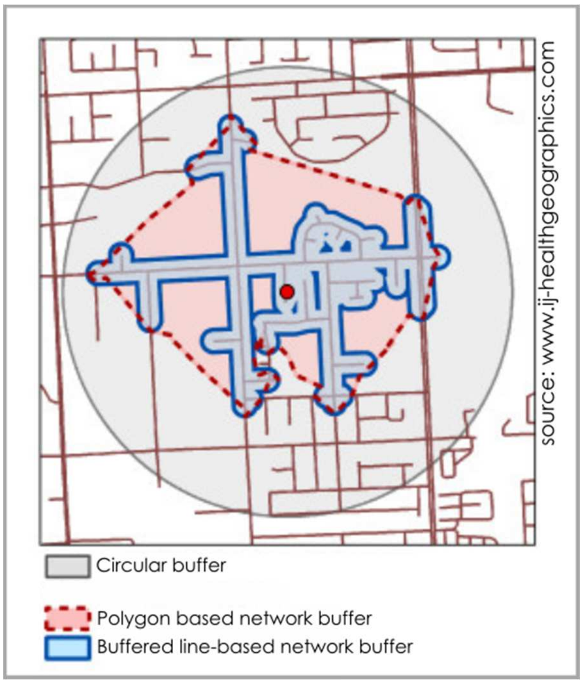

Figure 1 Network buffer and circular buffer comparison

From an accuracy perspective, the network buffer approach seems the most reliable of the mentioned methods because it defines the "walkable" neighbourhood area within the distance specified as radius. Therefore the use of the network buffer should more accurately capture access to food retail (Lasen et al., 2015). Kernel density also has a high potential because it adds weight to the area depending on intensity of investigated points, but also the distance between points, creating an "intensity map". Buck et al. (2013) have used kernel density of junk food supplies per service area. Their study was the only one using kernel density (in reviewed studies group).

Proximity

Proximity was measured as **Euclidean distance** (Cerin et al., 2011; Clark et al., 2014; Fraser and Edwards, 2010), **network distance** (Laraia et al., 2004; Morland and Evenson, 2009; Pearson et al., 2005), **population weighted distance** (Michimi and Wimberly, 2010) or by **modelling travel time** (Pearce et al., 2008). Two first measures (Euclidean distance and network distance) are explained in Figure 2. The most commonly used was network distance. Beside that, the tendency of using network analysis (for calculating network proximity) was observed in most recent studies (from 2011 and later). This can be explained by highest accuracy (distance by street/road is more accurate than straight line distance) but also by technology development. Modelling travel time could also be a good method to use, but it requires more detailed data and longer computation time.

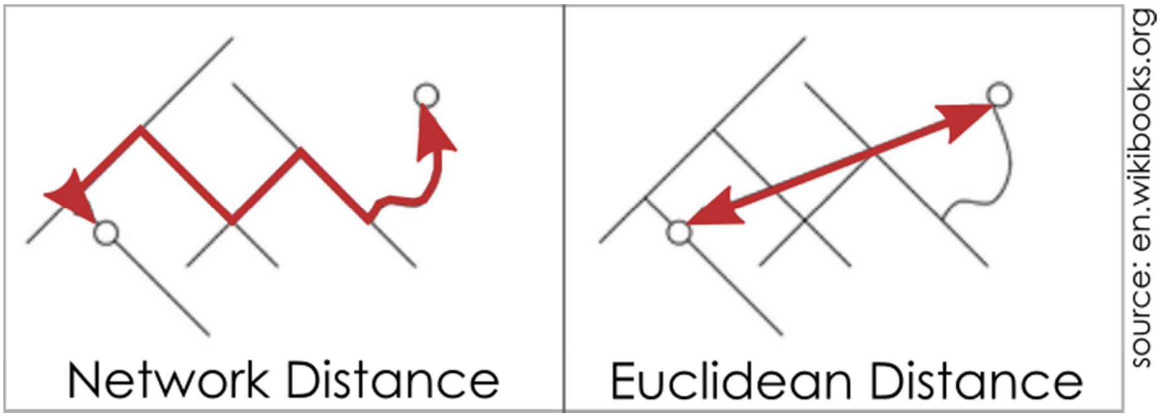

Figure 2 Network Distance and Euclidean Distance compared

2.4 Spatial relationships between diet and food retailers

2.4.1 Defining neighbourhood

As mentioned by Charreire (2010), defining criteria for appropriate geographic boundaries of a neighbourhood has proved to be challenging. Some of the studies used **predefined neighbourhood** like:

- postcode district (area defined on the basis of postcodes)
- census tract (region defined for the purpose of taking a census)
- county
- mesh blocks (a small geographic unit used in the census in Australia and New Zealand)
- Middle Super Output Area (MSOA – geographic units used for small area statistics in England)

All of them clustered food retailers to boundaries of those neighbourhoods. Another option to define a neighbourhood was **buffer**. In this case, participant or school locations were buffered. As mentioned in subchapter 2.3.2, two types of buffer were used: a circular buffer and a network buffer. Unfortunately there was no clear strategy of what radius sizes should be used. It is still unknown what buffer sizes should be used in this type of research. The range of radius distances used was inconsistent (balances between 100m and 8km). Some of the researchers explained choosing a specific buffer size by selecting distances equal to minutes which were needed to e.g. 15 minutes' walk. Laska et al. (2010) have chosen 800m buffer size as equal to 10 minute walking, Smith has chosen 400m as representation of 5 minute walking. In this study 800 and 1600 m buffers will be used. They are equal to 10 and 20 minutes walking distance, respectively.

Food environment is multidimensional. Penchasky & Thomas (1981) outlined 5 dimensions of it: availability, accessibility, affordability, acceptability, and accommodation. Availability is a presence of a certain retailer in the neighbourhood. Accessibility refers to the location of the food retailers and ease of getting to that location. Affordability refers to prices and people's perception. Acceptability refers to people's attitudes about their local food environment, and accommodation refers to how well a local food environment adapts to local resident's needs. In reviewed articles, the first four dimensions were used with an emphasis on availability and accessibility.

2.4.2 Statistical analysis

It is probable that there are other variables which could influence the diet. Therefore, most of the studies included additional factors in their statistical analysis.

Table 6 Additional characteristics of participants, used in investigated studies

Variable	Number of studies
age	24
gender	22
education level	18
race/ethnicity	17
income	14
area deprivation index	7
marital status	6
car ownership	5
employment status	5
physical activity	5
smoking	3

Table 6 shows an overview of these variables with the count of the studies that used them. As it is shown, the ones used most often were: age, gender and educational level (LeDoux and Vojnovic, 2014). Race and income (Gustafson et al., 2011) were also used in more than a half of the studies.

Reviewed studies have used different statistical methods to analyse possible associations between local food environment and diet/obesity. The most often-used methods were: **General Linear Regression** models (Murakami et al., 2009; Pearson et al., 2005) and **Multivariate Linear Regression** models (Bodor et al., 2008; Buck et al., 2013; Gustafson et al., 2011). **Multilevel Linear models** were also commonly used (Inagami et al., 2006; Janevic et al., 2010; Pearce et al., 2008). Additionally, **Simple Pearson's correlation** (Fraser and Edwards, 2010), **Least-**

Squares Regression (Sturm and Datar, 2005) and **Binominal Regression** (Morland et al., 2006) models were applied. The statistical methods used in reviewed studies had been counted (Table 7). The most commonly used were Multilevel Linear Regression and Binomial Linear Regression models.

Table 7 Overview of the methods used in reviewed studies

Model	How many studies had used it
Multilevel Linear Regression	13
Binomial Logistic Regression	9
General Linear Regression	6
Multinomial Logistic Regression	4
Least-squares Regression	2
Poisson Regression	2
Geographically Weighted Regression	1

2.4.3 Results

Most of the reviewed studies investigated the relationship between local food environment and diet or weight status. 82,11 % of the reviewed studies found at least one statistically significant association between diet/obesity and the local food environment. I categorized the results according to food retailer type, which was found as influential. Discovered associations will be explained in separate paragraphs below.

Supermarkets

In a few studies, associations between diet and locations of supermarkets were found. Laraia et al. (2004) found that distance to nearest supermarket can influence DQI-P (Diet Quality Index for Pregnancy): women living at a distance greater than 4 miles from a supermarket were more likely to have lower DQI-P than women living within 2 miles from supermarket.

Morland et al. (2006), by using multilevel regression models, and Morland & Evenson (2009), by using binomial logistic regression, found that the presence of supermarkets can be associated with a prevalence of obesity. The prevalence of obesity was lower in areas that had supermarkets and higher in areas with small grocery stores or fast food restaurants. Finally, distance to supermarkets was positively associated with obesity prevalence in metropolitan areas (Larsen et al., 2015; Michimi and Wimberly, 2010).

Mentioned studies have found that supermarkets may influence the diet. However, not all studies have found supermarkets influential. Boone-Heinonen et al. (2011) have found that greater supermarket availability was generally unrelated to diet quality. Besides that, Michimi et al. (2010) have found that distance to supermarket had no associations with obesity. Finally, an access to supermarkets was not related to vegetable intake (Pearce et al., 2008; Michimi et al., 2010, Boone-Heinonen et al., 2011).

Restaurants

Only two studies had found an association between diet/obesity and restaurant locations. Mehta & Chang (2008) found that a higher density of full-service restaurants was associated with lower weight status. It was later confirmed by Inagami et al. (2009), who found that a high concentration of restaurants **influences BMI values**. However, their results were opposite to those found by Metha & Chang: higher restaurant density is associated with higher BMI value.

Fast food restaurants

Associations with fast food restaurants were also found. Distance to the nearest fast food restaurant was negatively associated with fat vegetables consumption (e.g. avocado, olives or soybeans, Jago et al., (2007)) and also negatively associated with BMI (Block et al., 2011).

Density of fast food retailers was also influential. Mehta & Chang (2008) found a positive association between fast food density and a higher ratio of fast-food to full service restaurants and BMI. Fraser & Edwards (2010) have found a positive relationship between the density of fast food outlets per area and obesity. Finally Boone-Heinonen et al. (2011) found that fast food consumption was related to fast food availability.

Presence of fast food retailers within the neighbourhood was also important. It was positively associated with a prevalence of obesity (Morland and Evenson, 2009).

Six studies found that fast food outlets' spatial distribution can influence diet behaviour. However, some studies found no associations between fast food restaurants and BMI. The example of this is a study by Block et al. (2011) where no consistent relation between access to fast-food restaurants and individual BMI was found. This was also confirmed by Fraser et al. (2010).

Grocery stores

Another retailer type is grocery store. This type of retailer also was influential. Multilevel Linear Regression conducted by Inagami (2006) proved that the densities of the grocers were positively associated with BMI. Gibson (2011) had found the same associations by using only small grocery stores. The presence of grocery stores was positively associated with the prevalence of overweight, obesity, diabetes and hypertension (Morland et al., 2006). Besides that, distance to the closest grocery retailer was positively associated with healthy diet scores (Smith et al., 2013).

Contradictory results were found in a study by Boone-Heinonen et al. (2011). They have investigated the relationships between grocery store availability and diet outcomes. The results they obtained were mixed.

Convenience stores

The next group included convenience stores. Here the influence of the proximity is enhanced. Laraia et al. (2004) found that distance to nearest convenience store was negatively associated with mean DQI-P. The distance to the convenience stores was also negatively associated with BMI (Berge et al., 2014) and positively associated with fruit and vegetable consumption (Shearer et al., 2014). Additionally, the presence of convenience shops was found influential, causing a higher prevalence of obesity and overweight (Morland et al., 2006). However, Pearce et al. (2008) have found that the consumption of the recommended daily intake of fruit was not associated with living in a neighbourhood with better access to convenience stores.

Supercentres

Supercentres (shopping malls) are big complexes of shops. They are very popular in USA. This type of retailer was investigated in only one of the reviewed studies. This study found that individuals with a supercentre in their census tract weighed more than individuals without one. Those who lived in a census tract with a supercentre and a convenience store consumed fewer servings of fruits and vegetables (Gustafson et al., 2011).

Unhealthy food outlets

In a few studies, retailers were categorized into two groups: healthy and unhealthy. These studies also found some associations with diet. In the case of unhealthy retailers, only density was found influential. Positive association was found between overweight and obesity and a number of unhealthy outlets (Cetateanu and Jones, 2014; Jennings et al., 2011). The same influence was found in a study focused on children (Cetateanu and Jones, 2014).

Healthy food outlets

Retailers categorized as healthy were also found important. Janevic et al. (2010)

have found that healthy food outlets were associated with obesity and that the lack of healthy food outlets was associated with pre-pregnancy weight more than 95 kg. A positive association between the availability of healthy food and higher BMI was found by Casagrande (2011) and Jennings (2011).

Small food stores

A group of small food stores was influencing the weight of the respondents as well as the diet. Jago et al. (2007) found that the distance to the nearest small store was positively associated with high fat vegetable consumption. Prevalence of obesity in areas with small grocery stores was higher than what was found by Morland and Evenson (2009).

Food outlets in general

Food retailers were also investigated as one group. Clark et al. (2014) have found that both distance to and density of food outlets were associated with dietary quality of adolescents. They also found that every 100m increase in distance to the nearest food outlet of any type was associated with a decrease in DQI score for girls. Spence et al. (2009) were also investigating a few retailer types at once. They found that the lower the ratio of fast food restaurants and convenience stores to grocery stores and vendors near home, the lower the odds of being obese.

In the previous paragraphs, a number of detailed food retail definitions have been implicitly used. These definitions could play an interesting role in the collection of food retailer data of your case study and the discussion of your results.

2.5 Relevancy of methods for this study

The main concern of this review was identifying the methodologies used in studies on a local food environment's influence on diet. Methodologies were compared and analysed in terms of (diet) variables, spatio-analytical (GIS) methods and study results regarding the spatial dimension of the food environment.

All of the reviewed studies used spatial data. Two main datasets were identified: dietary patterns and local food environment. The first one describes diet together with the location of the participant. The second one describes food retailers that are located in a close neighbourhood of the participant's home. The first one is often predefined, but the second is constructed in the study.

The key question here is: how to correctly define the neighbourhood? Studies have shown that it can be defined in many ways. The method depends also on the country where the study was conducted. US-American studies, for example, widely used spatial clustering for defining store density. In these studies, neighbourhoods were predefined (zip code areas or census tracts). This tactic seems not to be the perfect option because zip code areas/census tracts often differ in size between each other. Pearce et al. (2006) mentioned that predefined neighbourhood often limited the studies. They also mentioned that currently available GIS methods offer better solutions so neighbourhood can be defined differently (e.g. by network buffer).

The overview of GIS methods indicates that making use of a street network might have a high potential. It is because street network seems to be the most appropriate available method that it will be used in this study as well.

The results suggest that retailers should be measured as **one group** (all retailer types as one group) and as **separate groups** (each retailer type separately). Both methods were sufficient in finding associations with a diet or obesity. This study will use both mentioned approaches. Different food retailer types will be investigated as separate groups, but there will also be an additional group where all retailers will be included.

There were several limitations mentioned in the reviewed studies. Most of them were related to the self-reporting character of the variables. In the case of BMI, it has been proven that the Body Mass Index values are often underestimated. Other were low responding rate or inaccurate data.

2.6 Lessons learned

Based on the results, it can be concluded that the research on the relation between diet and local food environment has already been successful but needs a continuation. It has been found that spatial distribution of food retailers may influence diet of both children and adults, but we still need stronger proof. This proof may help policy makers and spatial planners in situating retailers. The restriction in fast food density can be particularly beneficial. Also, the number of healthy food retailers should increase in order to prevent the growth of the number of obese people.

Diet and retailer variables

Variables which were used in investigated studies were concerning diet and food retailers in their neighbourhoods. The first group included the characteristics of people's diets (e.g. intake of fruits and vegetables, intake of kcal). They were geocoded by using the location of a study participant's (or diet holder's) home, or by using the location of the school the child attends. The second group (food retailers) included locations of places where food can be purchased (e.g. restaurants, supermarkets, groceries, etc.).

Spatial configuration of diets and retailers

Locations of diet holders and food retailers were used to investigate spatial configurations of variables they included. Spatial Configuration of Diet (SpatCon D) was constructed by using locations of diet holders. It allowed an investigation of how different diet variables were distributed over space. Consequently, it helped to find out if the variables are clustered, or if there are any spatial patterns behind them. Spatial Configuration of Food Retailers (SpatCon F) helped to understand the spatial patterns of food retailers. In the case of this group, retailers were investigated separately and together. It helped locate the places of high retailer density.

Methods recommended for the Netherlands

Some of the methods used in investigated studies can be applicable in studying this problem in the Netherlands. Therefore, this review contributed to methodological choices. These choices are explained below.

In order to calculate **proximity**, network distance and Euclidean distance will be used. These two methods were most commonly used in reviewed studies because of the possibilities they provide. They will be used because distances between diet holder's location and retailers are crucial in order to find the relationship between diet and food environment. Using Euclidean distance, Jago et al. (2007) found a relation between distance to small food store and consumption of fruit, juice and vegetables. Network distance was used by Timperio et al. (2008). They discovered that the further a child lived from the supermarket and fast food, the greater the likelihood of consuming more vegetables.

Density will be calculated by using network buffer and spatial clustering. Network buffer was previously used by Larsen et al. (2015). In their study, they found that people living in the area of healthy food outlets were less likely to become overweight or obese. In a study by Morland et al. (2006) where retailers were clustered into census tracts, the results showed that the presence of supermarkets was associated with a lower prevalence of obesity. In the same study, presence of convenience stores proved to be associated with higher prevalence of overweight and obesity.

3 Methodology

3.1 Introduction

Methodology of this study has been divided into 3 main phases: constructing spatial configuration of diet, constructing spatial configuration of food retailers, and investigating spatial relations between them. These phases and data used in them are described in 6 following subchapters:

- 3.2 Case study area,
- 3.3 The construction of SpatCon D,
- 3.4 The construction of SpatCon F,
- 3.5 Spatial relation between SpatCon D and SpatCon F.

Mentioned chapters explain the process of obtaining, preparing and investigating datasets, and also analysing spatial relations between them.

Diet and food environment datasets were crucial elements of this study. Before constructing SpatCon D and SpatCon F, datasets had to be downloaded and prepared. Actions applied in this phase are presented in a flow chart in Figure 3. They are more precisely explained in chapters 3.3 The construction of SpatCon D and 3.4 The construction of SpatCon F.

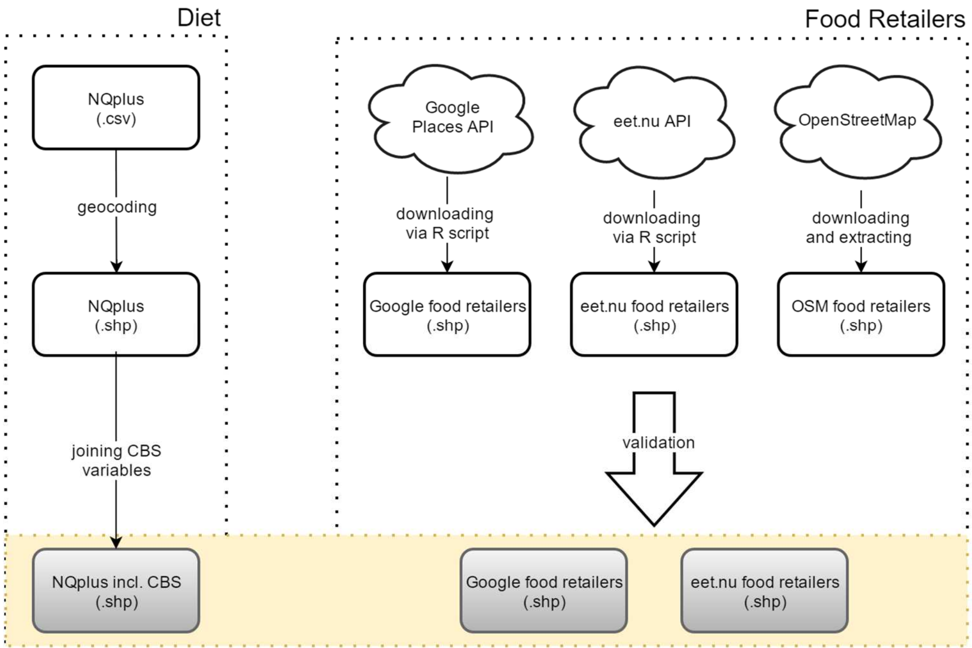

Figure 3 Flow diagram of methods used to obtain and prepare data

When datasets were prepared, SpatCon D and SpatCon F could be constructed and analysed. The last step was to investigate spatial relationships between them. Investigated relations have been presented in Figure 4. It included 3 datatypes: points (vector), polygons (vector) and raster.

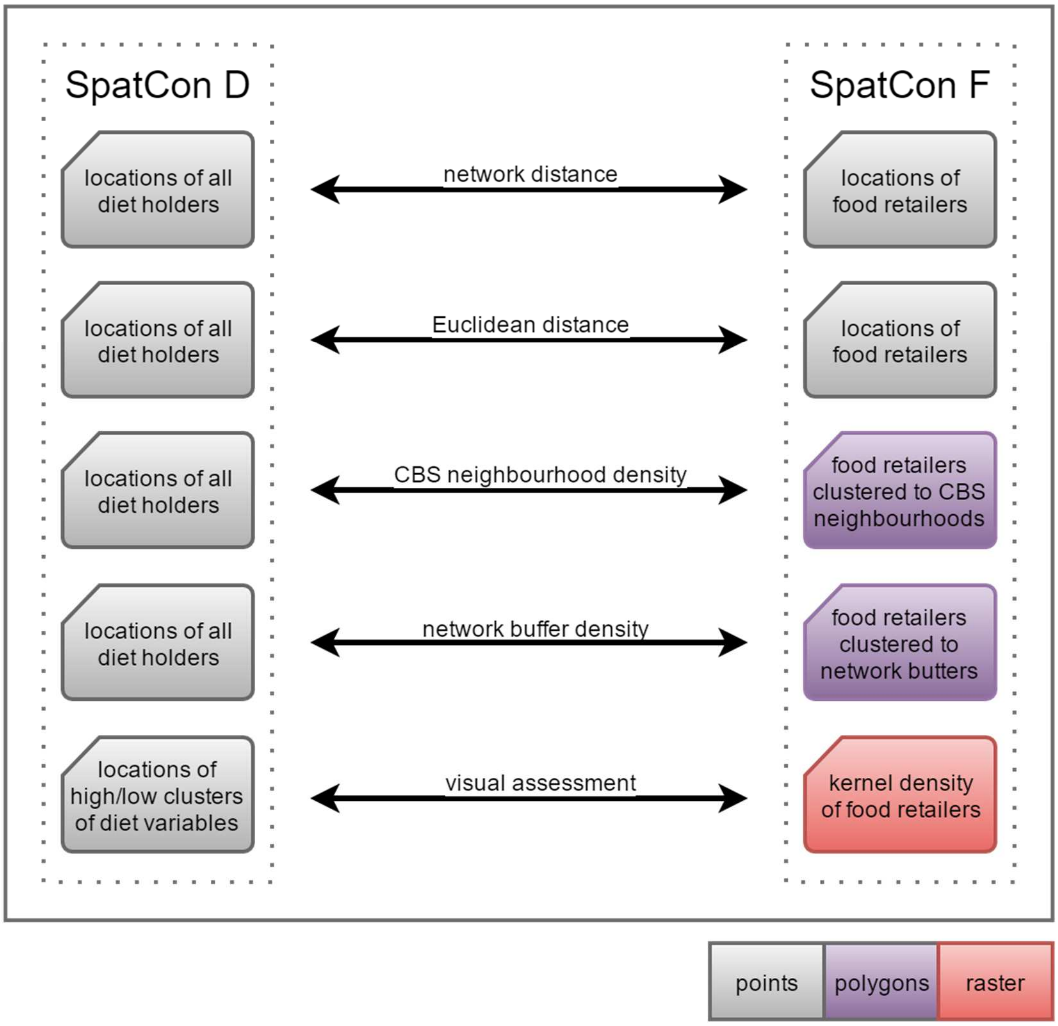

Figure 4 Diagram of investigated spatial relationships between SpatCon D and SpatCon F

3.2 Case study area

The study area is a Dutch case study. It covers the area of NQplus study, i.e. Wageningen, Arnhem, Veenendaal, Ede and Renkum (see Figure 5). This area is located in a province Gelderland.

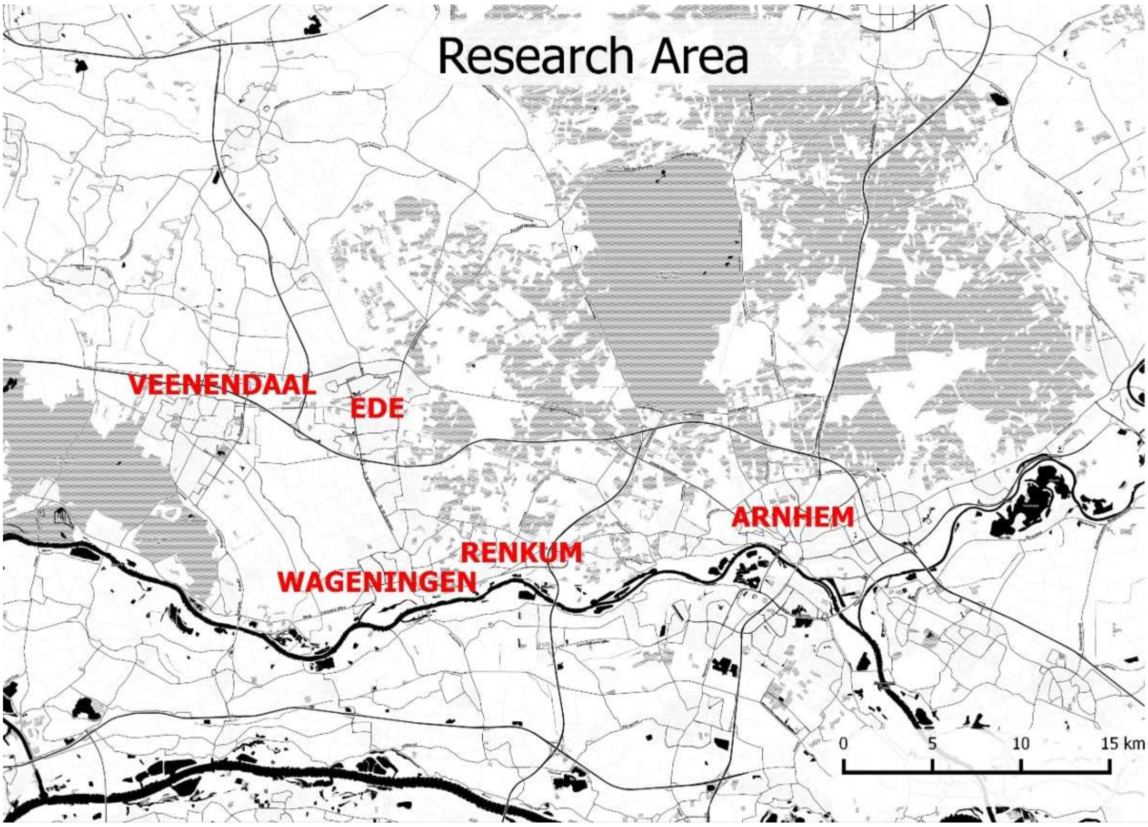

Figure 5 Research area

3.3 The construction of SpatCon D

3.3.1 Preliminary diet data (NQplus)

Dietary pattern data is a necessary element of this study. This data was provided by the Division of Human Nutrition of Wageningen University. The dataset was created during the NQplus project, which is an epidemiological study focused on investigating associations between diet and health outcomes. 2049 people living in the area of Wageningen, Arnhem, Veenendaal, Ede and Renkum (WAVER) were recruited as study participants. Inclusion criteria were the ability to speak and write Dutch and age between 20 and 70 years (van Lee, 2014). Participants were recruited by sending invitations to randomly selected inhabitants from cities mentioned above. Baseline measurements consisted of physical examination (e.g. blood pressure and body weight measurement), a fasting venipuncture, 24-hour urine collection and questionnaires on lifestyle (e.g. physical activity, and smoking), history of disease and demographics. Dietary intake was assessed using multiple 24-hour recalls (24hR) and food frequency questionnaires. The data used in this study was based on one of these questionnaires (FFQ).

The result of NQplus study was data that contains characteristics of participants' diets (DHD-index, calorie intake, nutrient content etc.), baseline BMI and BMI after one year. As a determinant of the location, postal codes of participants' residence places were used. Since locations are a crucial element of this study, all rows which were missing postcodes or which had incorrect postcodes were excluded. The final dataset contains 1956 participants. Their locations are displayed in Figure 1. The process of exclusion (missing and incorrect data) is explained in section 3.1.2.

Table 8 Characteristics of the NQplus participants

Characteristic	n	%
Gender	**1956**	**100**
Female	932	47,65
Male	1024	52,35
Age	**1953**	**100**
20-30	145	7,42
31-40	250	12,80
41-50	405	20,74
51-60	581	29,75
61-70	559	28,62
71-77	13	0,67
Education	**1946**	**100**
low	140	7,19
medium	593	30,47
high	1213	62,33
Weight status (baseline)	**1955**	**100**
underweight	13	0,66
normal	836	42,76
overweight	819	41,89
obese	287	14,68
Weight status (after 1 year)	**1525**	**100**
underweight	9	0,59
normal	693	45,44
overweight	614	40,26
obese	209	13,70

Characteristics of participants are presented in Table 8. Not all variables were filled for all participants, which is why the table includes separate n values for each category.

3.3.2 Geocoding

In order to calculate the distances from people's homes to the food facilities, locations of both places were required. In case of NQplus participants, the determinant of locations was postcodes and house numbers. In order to convert them into in longitude and latitude, the open source web Postcode API (http://www.postcodeapi.nu) was used. This API offers the information from BAG (Basisregistraties Adressen en Gebouwen) database. This information includes WGS 84 and RD coordinates. In this study, WGS 84 coordinates were used. Obtained locations of participants are shown in Figure 6.

The Postcode API was used to convert postcodes of participants into coordinates. For that purpose, a Python script written by Stefan Jansen was used (it is freely available at https://github.com/steffex/pyPostcode). An additional script was written in order to obtain locations of the NQplus participants. It takes a .csv file containing IDs of the participants as attributes; their postcodes and house numbers as input. The output table contains the same information but with two additional columns (longitude and latitude).

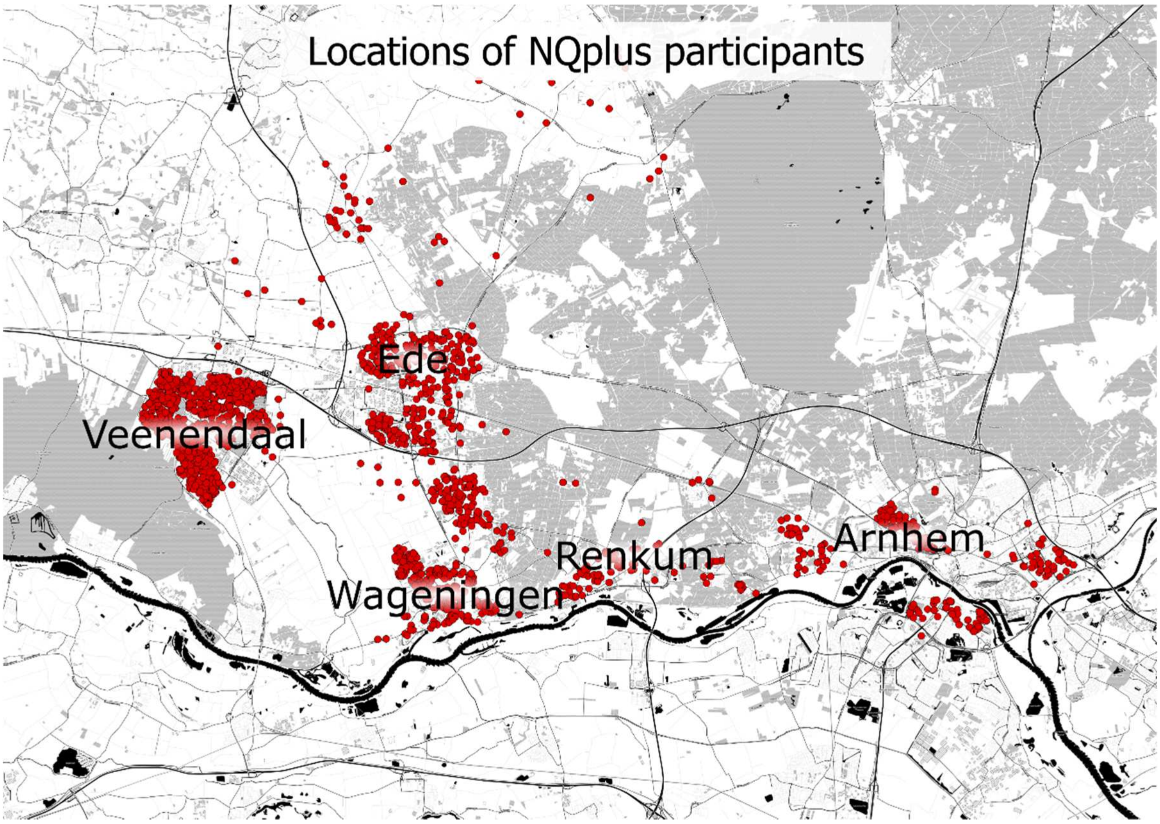

Figure 6 Locations of NQplus participants (red points)

First, all postcodes and house numbers are changed into the desirable format (empty spaces and unnecessary elements are removed). When the table is ready, search is conducted. However, some of the postcodes were incorrect.

These locations were automatically removed by the script. Therefore, the output contains all participants who filled in their locations (longitude and latitude) correctly.

The created .csv file was imported into QGIS and saved as point shapefile. After visual inspection of data, it appeared that some of the points are located far from the research area (the case when the postcode exists in postcode.nl database but it is not correct). These points were considered as incorrect and were manually removed from the file.

3.3.3 CBS data join

Statistics Netherlands (Centraal Bureau voor de Statistiek) is responsible for collecting and processing data in order to publish statistics to be used in practice, by policymakers and for scientific research. In addition to its responsibility for (official) national statistics, Statistics Netherlands also has the task of producing European (community) statistics. The information Statistics Netherlands publications incorporate a multitude of societal aspects, from macro-economic indicators such as economic growth and consumer prices, to the incomes of individual people and households.

Table 9 Overview of the CBS datasets

Field name (NL)	Field name (ENG)	Year
Aantal inwoners	Population	2013
Gemiddeld inkomen per inkomensontvanger	Average income per income recipient	2010
Gemiddeld inkomen per inwoner	Average income per citizen	2010
Personenauto's per huishouden	Cars per household	2013
Grote supermarkt, gemiddelde afstand in km	Large supermarket, average distance in kilometres	2012
Grote supermarkt, aantal binnen 3 km	Large supermarket, some within 3 km	2012
Restaurant, gemiddelde afstand in km	Restaurants, average distance in kilometres	2012

In this study, various CBS information (see Table 9) will be used. All this information was obtained from CBS datasets from years: 2010, 2012 and 2013. Different years were used because not all variables were calculated by CBS for 2013 file (e.g. average income per income recipient was not calculated for 2010). Mentioned datasets contained data per neighbourhood.

Data from Centraal Bureau voor de Statistiek (Statistics Netherlands) was downloaded from website http://www.cbsinuwbuurt.nl/ ("Meer informatie" button/download kaartlagen/download data). Three datasets were obtained: Gemeente, wijk- en buurtkaart for each of the mentioned years. The reason to use datasets of three years instead of one was caused by a lack of information in the latest dataset. All information obtained was the latest possible. The dataset from 2013 contained population density and number of cars per household but lacked other data. Therefore, older datasets were used to harvest the rest of the data (see Table 12), which were considered important for this study.

Spatial join tool was used in order to join fields from CBS Neighbourhood datasets to NQplus participants attribute table. In other words, each of the participant rows got CBS attributes. Joined CBS data was used as variables in statistical analysis.

3.3.4 Creating service areas (network buffer)

This study uses network buffers to define neighbourhoods of NQplus participants. To create such a buffer radius, size needs to be specified. In this study, radii of 800m and 1600m were used to define food environments. According to previous studies (Algert et al., 2006; Spence et al., 2009), they are respectively equivalent to 10 and 20 minutes walking distance.

Neighbourhoods for all NQplus participants were created in ArcMap, using "Create Service Areas" from the network analysis toolset. Each of the created polygons contained participants' attributes like ID, BMI or DHD index. A separate file was created for both 800m and 1600m radius.

3.3.5 Moran's I

Moran's I is one of the spatial analysis tools in Arc Map. It calculates Moran's I index value and a Z score, which indicate statistical significance. In general, a Moran's Index value near +1.0 indicates clustering while an index value near -1.0 indicates dispersion. A statistically significant positive Z score means that similar values cluster spatially (high values are found closer together, and low values are found closer together). A statistically significant negative Z score means that similar values are spatially dispersed (high values are found far away from other high values, and low values are found far away from other low values).

In the case of this study, values used in Moran's I analysis were: BMI, DHD and kcal intake. First, Global Moran's I was used in order to investigate if clustering occurs for the specific variable. In the case that clustering was present, local Moran's I was used in order to locate high and low clusters (high cluster contain group(s) of points with high values, low cluster contain points with low values) and outliers. Outliers were filtered out and separate files were created for high and low clusters. Clusters attribute tables were joined with local food environment variables. Then, mean values of these variables were calculated per cluster type. These numbers characterized the average person from each cluster so it represents the group in general. The last step was comparing these variables between clusters. If the difference was big, it means that this variable may be influencing the high or low values.

3.4 The construction of SpatCon F

3.4.1 Obtaining retailers data

Google Places API

Google APIs, developed by Google, allow interaction with Google Services. The Google Places API is the API for Google Maps. It allows for requesting data about places of selected type by using URL address (examples will be listed further in this chapter). The output is an XML or JSON file. Depending on the search type used, this file contains basic or detailed data about places. From available data, this study uses the place name, its address, type and location (longitude and latitude). To obtain this data, a script in R language has been created. The aim was to have a reproducible and easy way of obtaining the data. There were 2 versions of script developed. The first version was created (together with Roeland de Koning) as a final project of a Geo-Scripting course. This script took city name, radius size and retailer type as an input. The script automatically locates the centre of the city, which is the centre of the search buffer, and afterwards searches for places of specified type within given radius.

Google Places API supports the user with a few types of search. In the mentioned script, the "search nearby" method was used. In this method a Nearby Search request is an HTTP URL of the following form:

https://maps.googleapis.com/maps/api/place/nearbysearch/**output?parameters**

where "output?parameters" is a location of the buffer centre, search radius, place type and API key which has to be generated via Google Developers website. The example search will look like:

https://maps.googleapis.com/maps/api/place/nearbysearch/**json**?**location=-33.8670522,151.1957362**&**radius=500**&**types=food**&**key=API_KEY**

First, the result format has to be specified (JSON or XML). After that, location of the search centre separated by coma, radius size (in meters) and retailer types have to be defined. In the end of the query, the API key (explained later) has to be added (key=*your_API_key*).

After filling all these variables, a search can be conducted. Using script helps to automate work and get data in a more desirable format. In this case, the output is .csv table with latitude, longitude, name, type and address of the place.

However, this method is not perfect. The "search nearby" method is limited to 60 places per search. This means that only the first 60 places per search are listed in results. It limited the data because each of the search buffers covered the whole city (often there was more than 60 places within a city). Therefore, the script had to be adapted.

The new approach was different. This time, a radar search method was used. This method's output includes less information than search nearby. Results include two variables:

- **the geometry field** containing geographic coordinates,
- **the place_id**, which can be used in a place details request to get more information about the place.

Thus, to obtain the name of the place, the address and additional information, a second search has to be conducted. This additional information is obtained by using a unique place id obtained via radar search. URL used for obtaining details has the following structure:

https://maps.googleapis.com/maps/api/place/details/**output?parameters**

The example search looks like:

https://maps.googleapis.com/maps/api/place/details/**?json&placeid=ChIJwSU9sbWyx0cRWHXKh0vLHQw&key=API_KEY**

Where **placeid** is a unique place identifier obtained via radar search.

In the script, user specifies place type, API key and location of .csv file containing buffers' details (id, longitude, latitude, radius in meters). The API key can be created at the Google Developers Console (https://console.developers.google.com) website.

Initial search is conducted within buffers. In order to conduct it, the locations of buffer centres and radius size are required. The centres of buffers were created in ArcMap 10.2.2 using Create Fishnet (Data Management) tool. This tool creates a fishnet of rectangular cells. The output can be polyline or polygon features. Additionally, by checking "Create Label points", an additional file is created with points in the centres of all grid cells. Figure 2 presents how search buffers cover grid cells. Search buffers should be small, so I decided to use 700m radius (different sizes were also investigated but this one was optimal). In the case of the larger buffer, some of the points could be not downloaded (because of Google limits).

Grid cells should be ~989,95x~989,95 m. I rounded it up to 990m. Afterwards I enlarged the buffer radius size to 701m. The explanation of this can be seen in Figure 7 and in the description below.

Red rectangles (Figure 5) symbolize fishnet created in ArcMap. Because, as it was mentioned before, if the search radius is supposed to be 700m, the diagonal of the rectangle should be radius x 2 (1400m). In the rectangle with diagonal 1400m, each of the sides is ~989,95m. This number was rounded up to 990m because of practical reasons. After the fishnet with these parameters was created, the centroids of rectangles were used as centres of buffers. Using radius size 700m, the full area was covered. Some parts were overlapping, but it was the only way to cover the full area with circles.

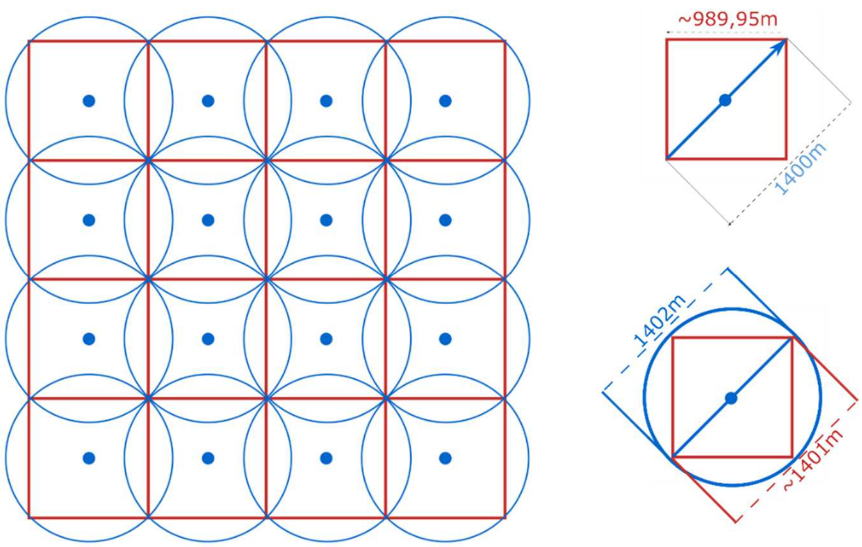

Figure 7 Search buffers covering grid cells

All centres of the buffers are saved into .csv file (with columns: id, lon, lat, rad) which is used to search for places and locations expressed in longitude and latitude. The function searches in each buffer separately and in the end merges all results into single .csv file containing coordinates and a place ID. The IDs of the places are used in the next search, where the details of a place are obtained. Each place's details are downloaded in a separate JSON file, which is used to search for places details. When all the details of a place are downloaded, they are merged with the initial file (containing only geometry and place ID). The final file (the output) contains the following columns: latitude, longitude, ID, address, place name and type. The file gets the name of the place type that was used in a search.

The result of the search is locations of all places from the Google database (one file per category). The number of retailers per retailer type are presented in Table 10. In total, 1317 food retailer locations were collected. This number is lower than a sum of all categories because some locations were categorized in more than one category, which semantically were the same location. To calculate how many locations ("all retailers" in Table 10) were found, "Delete Identical" tool in ArcMap was used. 371 points were deleted in this step.

This method still has some downsides. Firstly, only 1000 places per 24h can be requested. Secondly, the obtainable number of places per radar search request is limited to 200 places. This means that if there are more than 200 places within a search buffer, only the first 200 will be listed. I optimized the buffer size in order to not miss any data.

Store type	Number of points
Groceries and supermarkets	270
Bars	47
Meal takeaway	63
Meal deliveries	38
Restaurants	843
Convenience stores	15
Cafes	351
Bakeries	154
ALL RETAILERS	**1317**

Table 10 Number of points found within the research area

Google Places API has a better database than other sources used in this study. It contained more places comparing to OSM or eet.nu API. In conclusion it can be considered as the best available.

Eet.nu API

Eet.nu API allows the access to all restaurants in their database. Similarly to Google API, the request is done via URL. The output is a JSON file. The Venues API enables the user to find and retrieve information about restaurants. The API can be used to:

- retrieve information of a specific venue
- find restaurants with a list of IDs
- search the database for venues
- narrow results with filters
- find restaurants in a location
- find restaurants near a geolocation
- access the reviews of a specific venue

By using a restaurant's ID, detailed information about the restaurant can be obtained. Venue details contain information such as: a list of tags, images, menu's, opening hours and more. It is also possible to search for details of many restaurants within one request. Additionally, search by query is possible (e.g. adding 'query=Wageningen&tags=snackbar' to basic url will locate all snack bars in Wageningen).

R script has been created to obtain a restaurant's location within research area. The script downloads all restaurants in the Netherlands. First, the list of all cities available in eet.nu website is obtained. Afterwards, the list of tags is downloaded and kitchen related tags are selected. Finally, each of the restaurant types is searched within each city from the list.

An alternative solution would have been downloading venues from Arnhem, Wageningen, Veenendaal, Ede and Renkum, but this method would not cover the whole research area. Therefore, all restaurants from the Netherlands were downloaded and clipped with the boundary of the research area.

The result of each search is a json file. In a case when there are more than 100 places found, there will be more json files (each available created via a separate URL, containing maximally 100 places). Basic search URL allows the user to download the first page of results. The end of the file contains information about the number of pages ("total_pages") and the URL where the next page can be found (available under "next_page"). Script merges all these files into a single file and saves it as csv. When all restaurants from the whole country are downloaded, script merges them into one file with a name of the tag that was used in search (e.g. snackbar). Final folder contains separate files for each tag.

They were merged and saved into a point dataset in QGIS where they were clipped into a research area (before importing to QGIS, the duplicates of places were removed). The attribute table contains the name of the place, longitude, latitude, category, street, zip code, city and region.

Table 11 The overview of the eet.nu API data

Store type	Number of points	Store type	Number of points
African	1	Kebab	14
American	2	Lunchroom	45
Asian	6	Mediterranean	6
BBQ	2	Mexican	2
Bistro	4	Middle-east	1
Chinese	3	Pancake	14
Chinese-Indian	43	Pizza	33
Dutch	106	Regional	1
Eatery	37	Sandwiches	29
Egyptian	1	Snackbar	167
English	1	Spanish	1
Fish	2	Steak house	7
Food-vendor	23	Surinamese	2
French	43	Sushi	12
Fusion	3	Takeaway and delivery	10
Greek	10	Tapas	10
Grill	14	Thai	5
Ice-cream	18	Turkish	19
Indian	9	Vegetarian	3
Indonesian	8	Vietnamese	2
International	56	Western European	2
Italian	37	Wok	2
Japanese	6	**ALL RETAILERS**	**819**

Table 11 presents the number of points gathered via this source. Comparing to restaurants' locations obtained via Google Places, it has less points. A limitation of this source is the fact that only restaurants can be found here. The advantage is a more extensive tag list, containing 56 tags (e.g. bistro, Chinese, fish, fusion, ice-cream-parlor, kebab, kosher-depricated, pizza, regional, sandwiches, snackbar, steak-house, sushi, takeaway-and-delivery, tapas, vegetarian-depricated) and no limits concerning the number of places per search (in the case of Google Places API, there were some limits concerning the number of places which could be downloaded per day). Because of the variety of restaurant types, unhealthy food retailers can be filtered more easily.

In total, 819 restaurant locations have been obtained. The most common groups were: snack bars (167 places found) and Dutch (106 places found).

OSM POI

OpenStreetMap is a project where users create the maps. All the data is created voluntary and is open source, free to download and up-to-date (the file is updated few times a day). The downloaded shapefile for the Netherlands (netherlands-latest.shp.zip) downloaded from:

http://download.geofabrik.de/europe/netherlands.html
includes a Points Of Interest (POI) dataset.

Table 12 The overview of the OSM food retailers data

Store type	Number of points
Bakery	26
Bar	14
BBQ	4
Butcher	10
Cafe	31
Chinese	3
Chemist	18
Coffee	1
Confectionery	4
Convenience store	19
Fast food	58
Greengrocer	8
Ice cream	5
Kiosk	6
Restaurant	150
Seafood	1
Supermarket	88

Food retailers were obtained by extracting from the mentioned dataset the following categories: bakery, bar, BBQ, butcher, café, Chinese, chemist, coffee, confectionery, convenience store, fast food, greengrocer, ice cream, kiosk, restaurant, seafood. In order to create a file with retailers within the research area, the retailers file was clipped by using delineation of research area. Downloaded data was cropped into research areas. The number of points obtained is displayed in Table 12.

3.4.2 Validation of the points

In order to validate the locational accuracy of the data, validation of the points was conducted. For that purpose, area of Wageningen (covering ~ 5 % of the obtained points) was selected. The decision was made to exclude data from OSM because, compared to two other sources, this data contained significantly less points (only 15 retailers). The sample contained points from the centre of Wageningen: 73 in case of Google data and 43 in case of eet.nu data. Mentioned points are displayed in Figure 8. Streets containing those points (Figure 9) have been visited and investigated. The Garmin eTrex 30 GNSS receiver was used in order to create points describing locations of the encountered food retailers. Created points were exported to ArcMap and compared with Google Places and Eet.nu points.

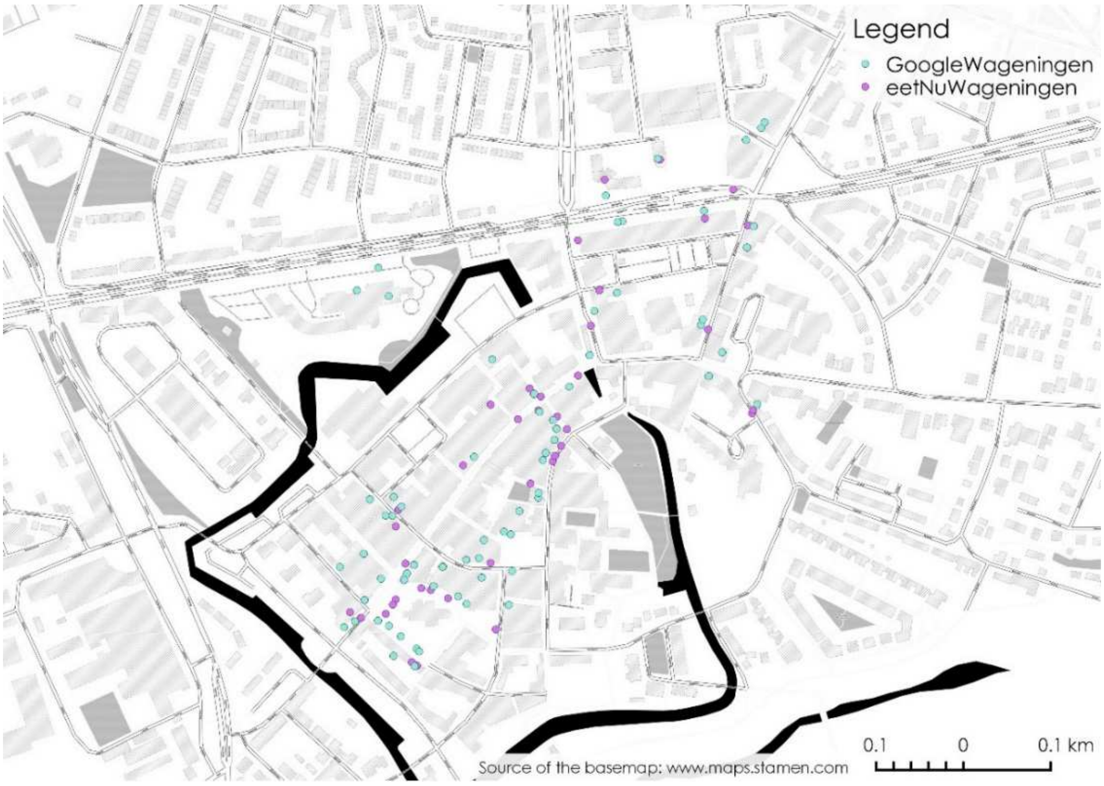

Figure 8 Points selected for validation

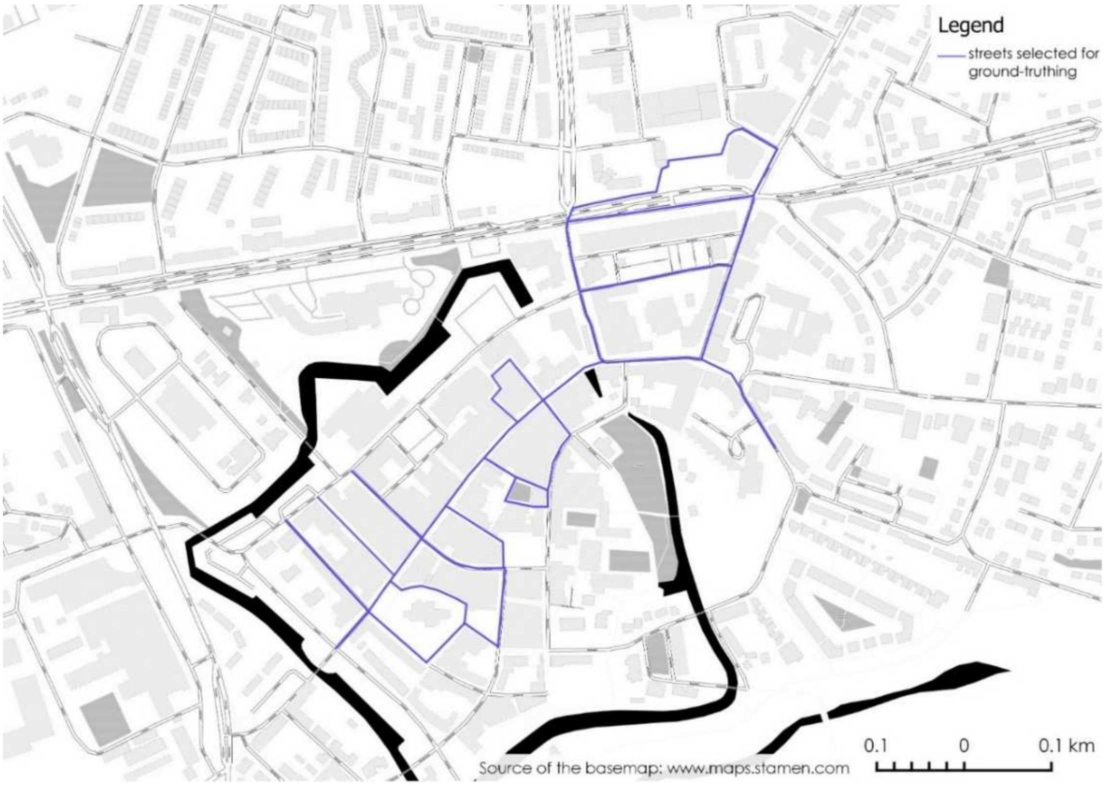

Figure 9 Streets selected for validation

Google Places

Sample data obtained via Google Places API included 73 points. Comparing these points with field work file led to following results:

- ~10% of the obtained places (7 points) do not exist anymore,
- 66 correct points cover 73,3% of places existing within area chosen for validation,
- 18 existing places were not included in Google data.

Mentioned numbers prove that Google Places API is not a perfect data source. Achieved accuracy could have been better, but unfortunately it cannot be improved. Therefore, Google Places dataset will be still used in the analysis.

Eet.nu

Sample cropped from eet.nu dataset contained locations of 43 restaurants from which:

- 3 places (7% from all obtained points) from this data do not exist anymore,
- 40 correct points covered 86,3% of the existing places within the area,
- 3 existing places were not included in eet.nu data.

This result is promising and proves that this data is most accurate from available datasets.

3.4.3 Kernel density of retailers

After identification of low and high clusters, its relation with local food environment was investigated. A kernel density tool was used at this point. This function estimates the intensity of a point across a surface by calculating the overall number of cases situated within a given search radius from a target point. The output is a raster layer (heatmap) with density values. Kernel density was used to create heatmap for every retailer type, but also for retailers in general.

3.5 Relationship between Spatcon D and Spatcon F

3.5.1 Density of retailers in the neighbourhood

Density was calculated using two approaches: clustering to neighbourhoods and counting retailers in network buffer. A separate file was created for each method.

Clustering to neighbourhoods

In this method, food retailers were clustered to neighbourhoods obtained from CBS data. New fields have been added to its attribute table where the number of each store type was added. Afterwards, all this information was joined to the participants attribute table using joining by spatial location. A file containing participants' information and density of each retailer type per participant neighbourhood was created.

Network buffer as neighbourhood

Network buffers created in ArcMap were used to calculate density of each food retailer type. For each retailer type (from both Google and eet.nu datasets) density was calculated. All values were added to the participants attribute table and saved as a separate file. Network dataset used in this step was OSM road dataset.

3.5.2 Proximity to closest retailers

Proximity was calculated using Euclidean distance and network distance. Separate files were created for each method.

Euclidean distance

Euclidean distance is a straight line distance between two points. It was calculated in QGIS by using Distance Matrix function which calculates distance to closest point(s). Distances to closest retailer were calculated separately for each store type from Google data, and for all restaurants and snack bars from eet.nu data. All this information was merged into one file containing all

Euclidean distances calculated for all participants.

Network distance

Network distance is a distance by road/street. In this study, it was calculated in ArcMap by using Network analysis closest facility tool. Similarly to Euclidean distance calculations, network distance was also calculated for all retailer types and saved in one shapefile in the end. Similarly to network buffer step, OSM network dataset was used.

3.5.3 Data analysis

Proximities and densities calculated in previous step were joined with information on participants and CBS data, resulting in 4 shapefiles: CBS neighbourhood density, network buffer density, Euclidean distance proximity, network distance proximity. Each file contains locations of participants with detailed attributes describing their local food environment by the method described in a file title.

Methods used to describe the relation between personal diet and local foodscape were: multiple regression models and comparison of local food environments of low (normal weight) and high (obese) cluster located by Moran's I.

Selection was motivated by literature review findings, but also by knowledge of GIS methods used in analysis of spatial patterns (e.g. kernel density or Moran's I). Mentioned GIS methods were not so common in reviewed studies, but useful for analysing spatial data. Therefore, they were used in this study.

Analysis was conducted in Minitab 17 and ArcMap 10.2 software.

Visual analysis

In order to investigate the relation between diet and food environment, kernel density of retailers was displayed together with clusters found with Moran's I and inspected visually.

Descriptive statistics

Descriptive statistics describe the main features of a collection (in this case characteristics of diet and local food environment). It provides simple summaries about the sample and about the observations that have been made. In this study, descriptive statistics of data were conducted in order to describe it before the regression analysis. It includes mean, standard deviation, minimum, median and maximum of all variables describing diet holders and their neighbourhoods.

Multiple regression analysis

The general purpose of multiple regression analysis is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. In case of this study dependent variables were: BMI, DHD and kcal intake. Predictors were characteristics of local food environment like shop density or proximity to the closest retailer. Regression was conducted in order to find if characteristics of local food environment can influence the diet. Statistically significant factors were used in further analysis.

For each of the four types of measurements (Euclidean distance, network distance, neighbourhood density, network buffer density) separate multiple regressions were performed. In each regression model, different additional variables were included. Additional variables included in analysis were:

- sex, age, education level and income (for models using BMI as dependent variable),
- age (for models using kcal intake as dependent variable),
- sex and age (for models using DHD index as dependent variable).

4 Results

4.1 Introduction

The results chapter includes the description of results obtained in steps presented in methodology chapter. It has been divided into 3 subchapters:

- 4.2 Spatial distribution of diets,
- 4.3 Spatial distribution of food retailers,
- 4.4 Spatial relationships between diet and food retailers.

They describe the output of the steps presented in chapters 3.3, 3.4, 3.5 (accordingly).

4.2 Spatial distribution of diets

4.2.1 Locations of study participants and characteristics of their diet

After geocoding, locations of diet holders could be presented on the map and investigated spatially. Figures 10 and 11 present locations of study participants with distinction of their DHD and kcal intake. It is a raw representation and it is hard to distinguish spatial patterns there. At this point, clustering to neighbourhoods became helpful. Figures 12 – 14 presents CBS neighbourhoods categorized according to mean kcal intake, DHD, and BMI (per neighbourhood). Highest kcal intake occurred in Veenendaal and Ede. BMI clustering showed highest values in Wageningen, Ede, Barneveld and Arnhem. Mean DHD index value did not show any spatial patterns. The disadvantage of this method is the influence of neighbourhoods containing only one participant. This neighbourhood automatically gets the value from this one participant (there is no average value). It disturbs the results, especially if this person has a high value of investigated variable.

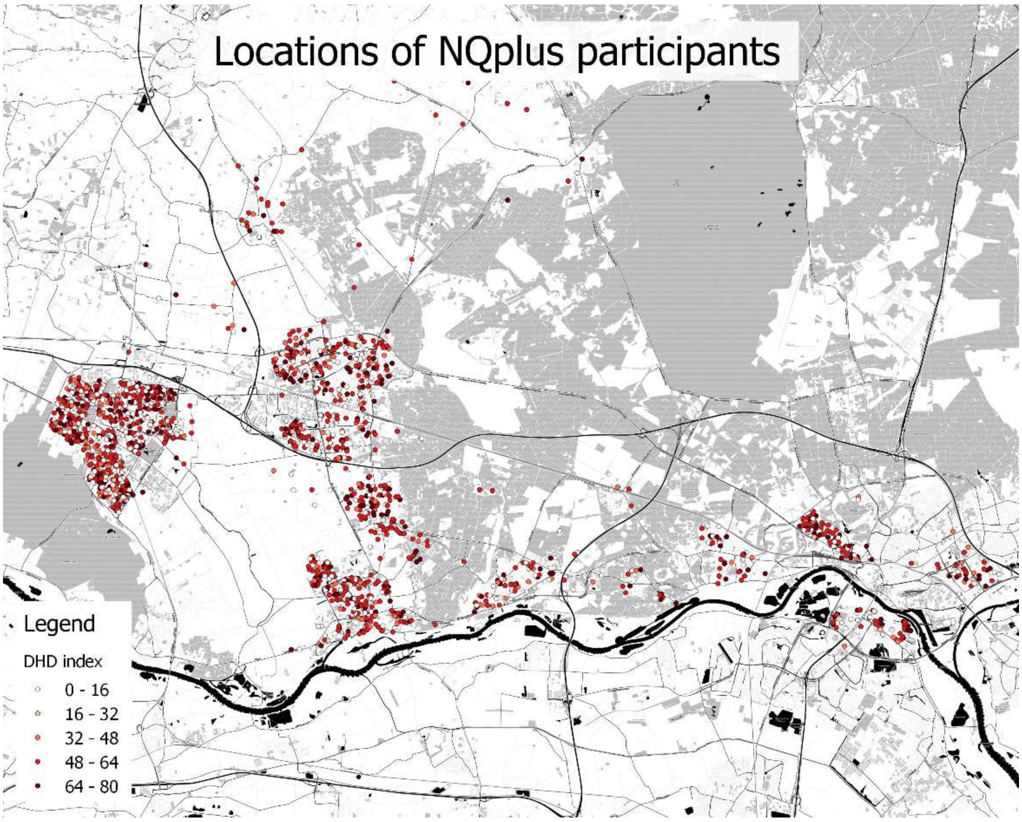

Figure 10 Spatial distribution of DHD index

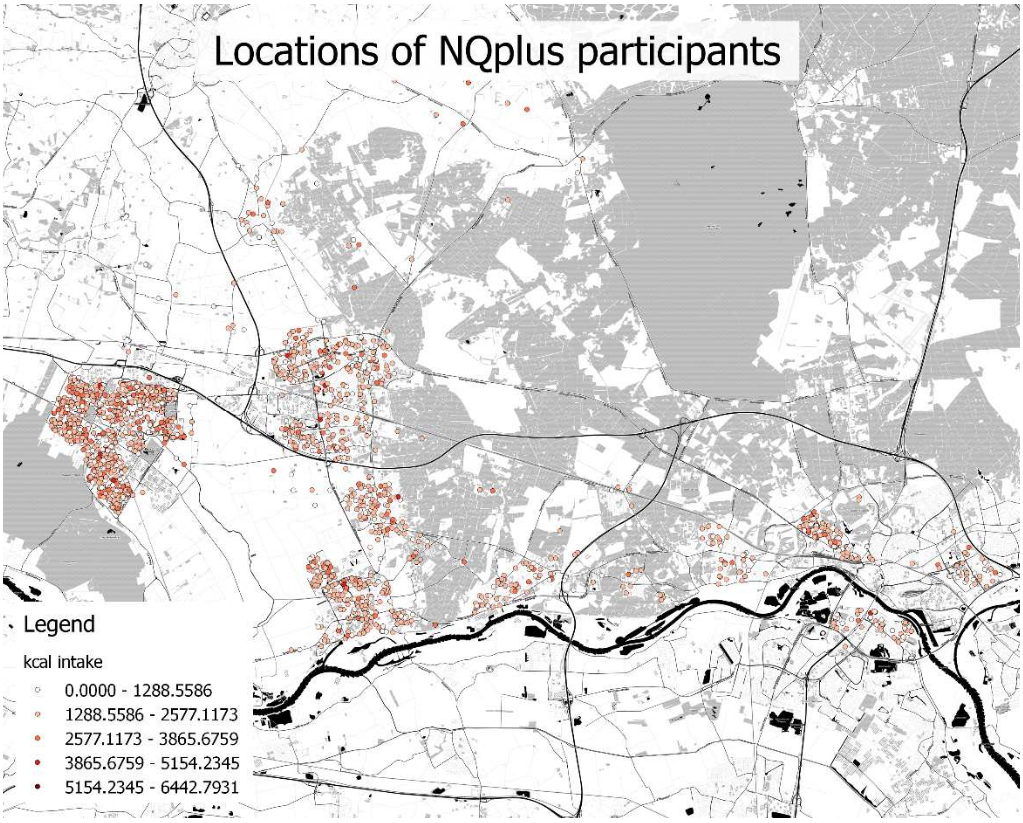

Figure 11 Spatial distribution of kcal intake

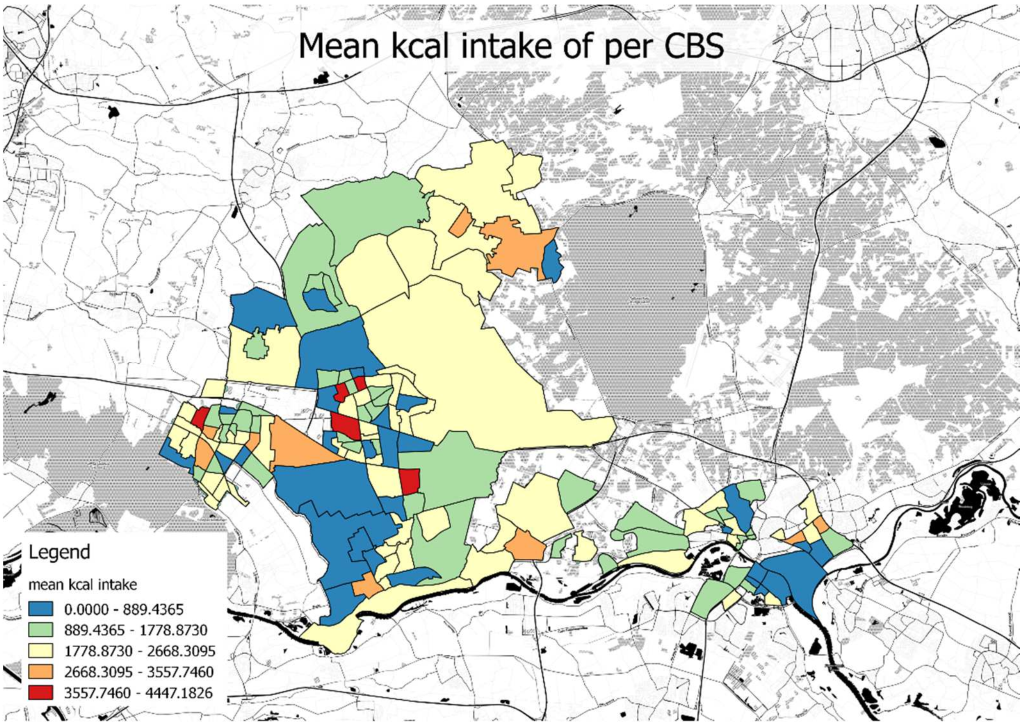

Figure 12 Mean kcal intake per CBS neighbourhood

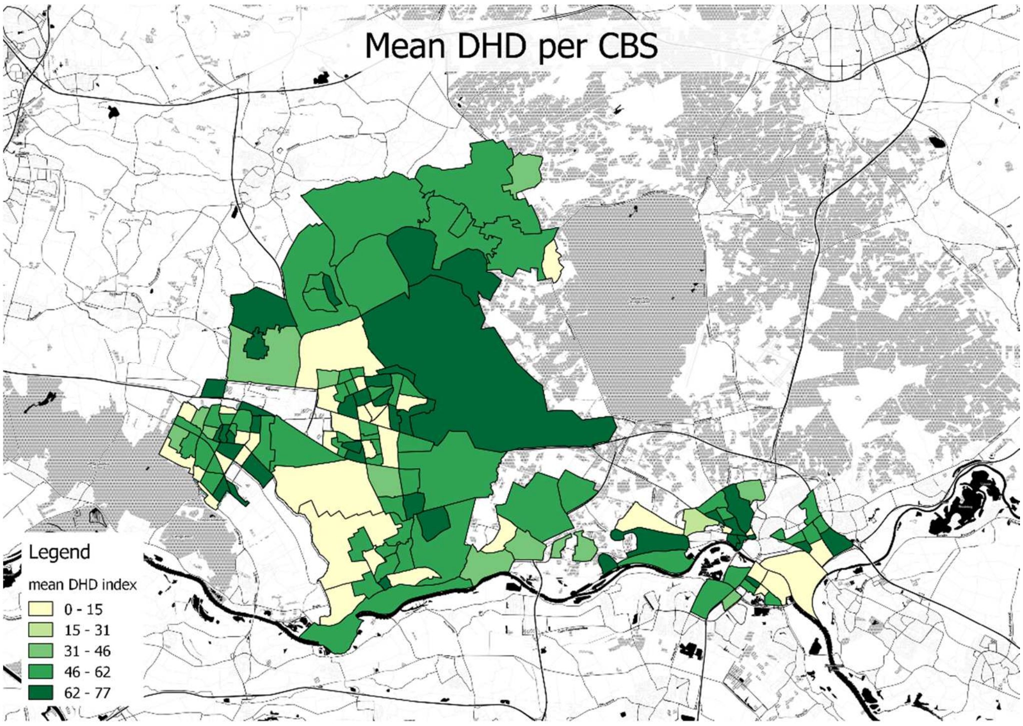

Figure 13 Mean DHD index per CBS neighbourhood

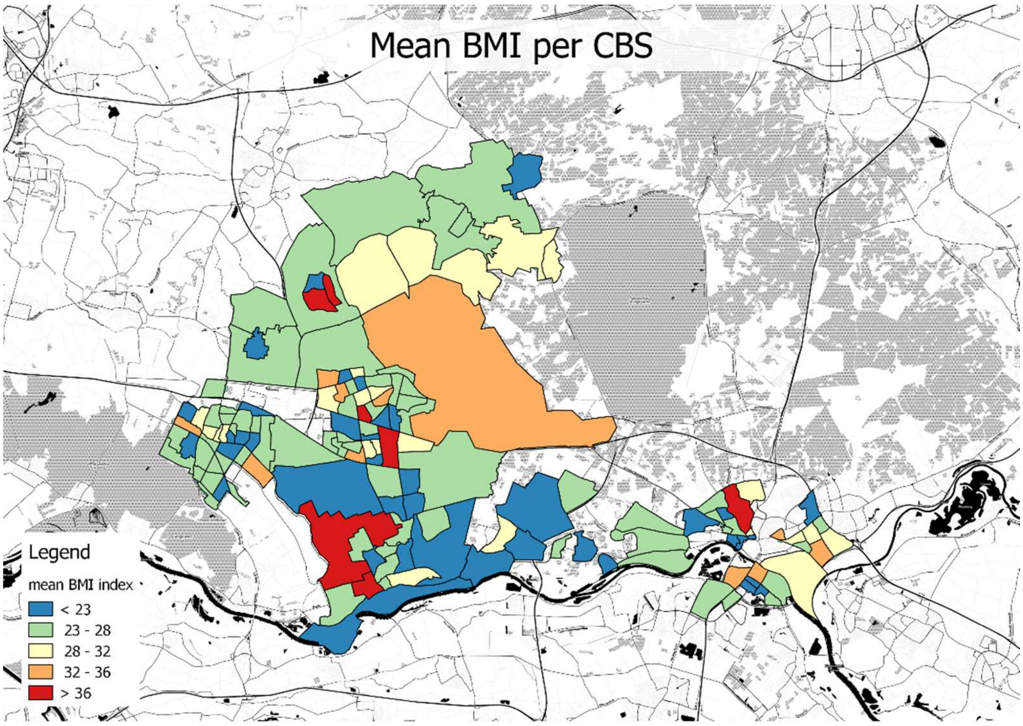

Figure 14 Mean BMI ped CBS neighbourhood

4.2.2 High/low clusters within SpatCon D

Global Moran's I

Spatial analysis of diet variables was conducted in ArcMap 10.2.2. It was done to investigate spatial clustering of BMI, DHD and kcal intake by using Global Moran's I. Analysis was conducted using inverse distance conceptualization (nearby neighbouring features had a larger influence on the computations for a target feature than features that are far away) and Euclidean Distance method. Table 13 displays results of this analysis. DHD and kcal intake did not show clustering pattern but BMI showed a significant clustering with z-score 2,12. This means that high and low values of BMI were grouped spatially, creating clusters. These clusters were mapped, using Cluster and Outlier Analysis (Anselin Locan Moran's I) tool in ArcMap. Clusters located in this step are presented in Figure 15.

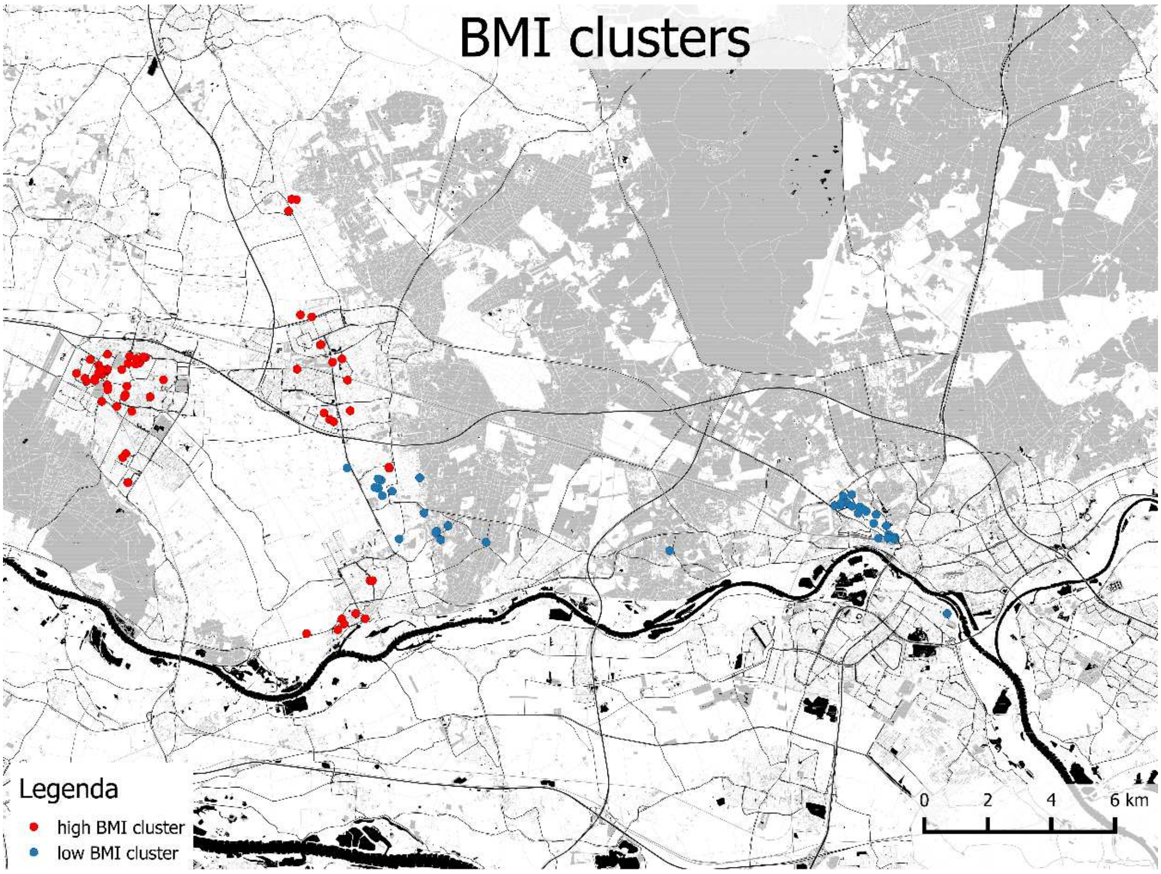

Figure 15 Low and high BMI clusters

All clusters occur within the cities that were expected because most of the participants live there. High BMI clusters occurred in Veenendaal, Ede and Wageningen, while low clusters occurred in Bennekom and Arnhem.

Table 13 Global Moran's I results

	BMI	DHD	Kcal intake
n	1953	1406	1272
Moran's Index	0,05	0,01	0,04
z-score	2,12	0,37	1,43
p-value	**0,03***	0,71	0,15

*significant

Characteristics of both low and high BMI cluster are displayed in Table 14. It includes mean values of BMI, DHD and kcal intake per cluster type and mean values of CBS descriptors (socio-economic variables) of neighbourhoods where clusters were located.

Table 14 Characteristings of BMI clusters

	BMI low (mean value)	**BMI high (mean value)**	**difference (absolute)**
BMI	20,95	34,70	**13,76**
DHD	59,56	39,39	**20,18**
kcal intake	1829,18	1926,88	97,70
Population Density	1875,26	2968,88	**1093,62**
Auto per Household	1,03	0,97	0,06
Income per Incomer	40.240	28.550	**11.690**
Income per Citizen	29.610	19.910	9.700

The low BMI cluster had mean BMI 20,95 (which is normal and healthy weight); while the high cluster's mean BMI was 34,70 (which represents obesity). The high cluster's DHD was ~20 points lower than in case of low cluster what means that overweight/obese people from the clusters do not follow a good, balanced diet (or have a worse diet than people from low cluster). Surprisingly, the reported kcal intake was similar in both clusters (the difference was only 97 kcal). It might be caused by under-reporting, which is more probable in the case of people with high BMI. This could cause the difference between kcal intake of normal weight people and obese people's kcal intake to be smaller.

High BMI clusters were located in more populated places (mean population density was ~ 1000 people higher than in case of low BMI cluster). People from the low cluster earn on average 11.690 more than people from the high cluster.

4.3 Spatial distribution of food retailers

4.3.1 Locations of food retailers

Based on data collection in April - July 2015, datasets containing locations of food retailers were created. Figures 16 and 17 present locations of food retailers obtained from Google and locations of restaurants obtained from eet.nu. It contains all food retailers (available in mentioned data sources) within 3 km from locations of diet holders. It is noticeable that most of the retailers are located in the city centres. It is also visible that most of the retailers are categorized as restaurants.

Figure 16 Locations of food retailers obtained from Google Places API

Figure 17 Locations of restaurants obtained from eet.nu API

4.3.2 Validation

As expected, data obtained via open sources is not ideal. Some of the existing places were not included in the dataset, and some of the non-existing places were still in the data. The question was: how many points were in these two groups? Points symbolizing places which do not exist were also labelled as incorrect. The validation tables can be seen in Appendix IV.

Validation revealed that Open Source data accuracy vary between sources. The accuracy of 73,3% (Google data) and 86,3% (eet.nu data) is still considered as a good one so both datasets were used in the further analysis. This decision was also dictated by no data alternatives.

4.3.3 Kernel density of food retailers

Kernel densities of food retailers were investigated in order to see how hotspots locations differ between retailer types. The results of this step are presented in Figure 18. As it is seen there, the patterns are similar for all retailer types. The retailers' hotspots are always located in the city centres.

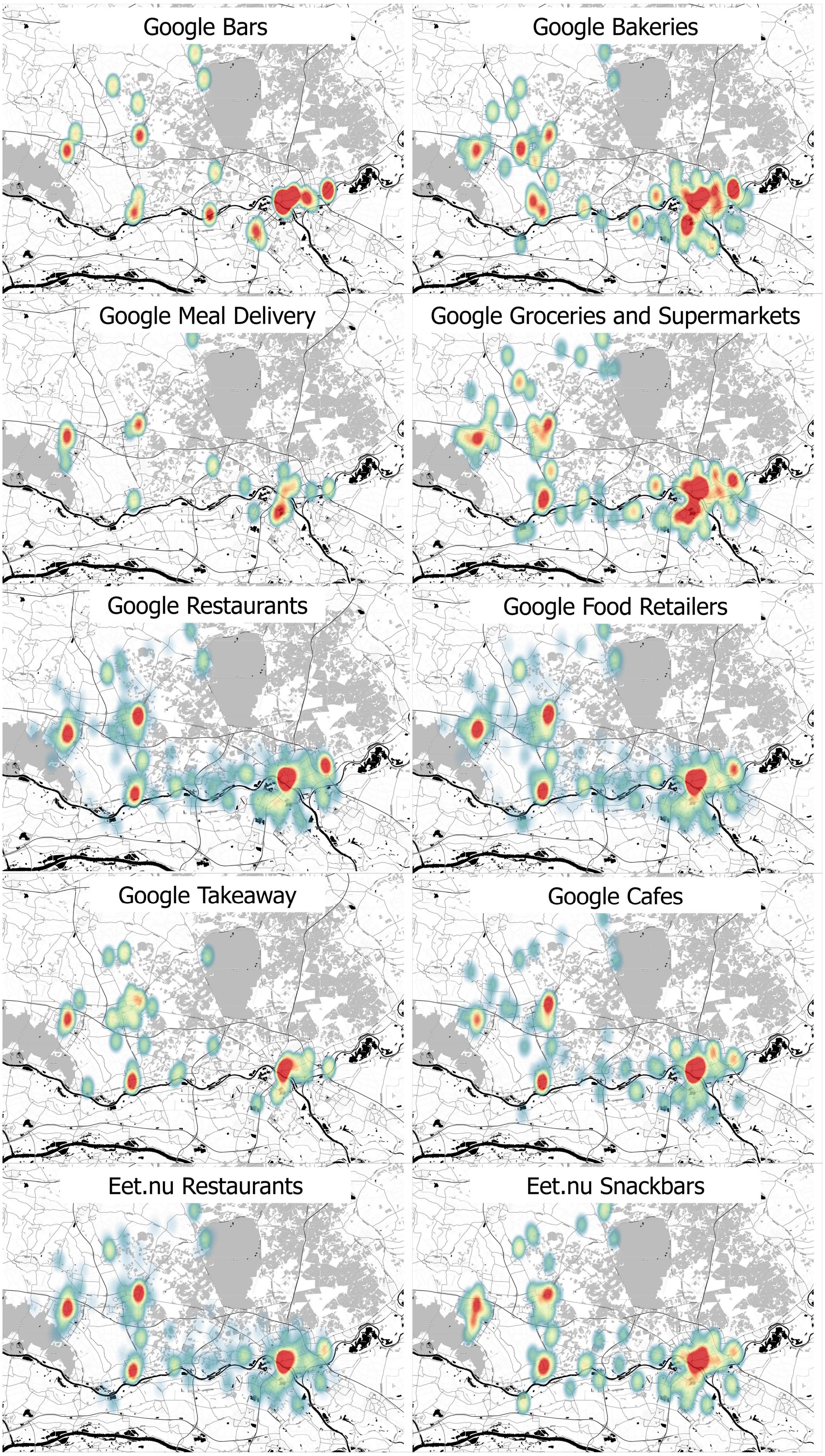

Figure 18 Kernel density of food retailers

4.3.4 Neighbourhood density of food retailers

Figures 19 and 20 present density of food retailers per CBS neighbourhood. Both restaurants collected from eet.nu and food retailers downloaded from Google show similar spatial patterns. High density neighbourhoods are most often located in the city centres. It is especially seen in the city centres of Veenendaal, Wageningen, Ede and Renkum. It would also probably be seen in case of Arnhem, but neighbourhoods which did not include the NQplus study participants were excluded in this analysis.

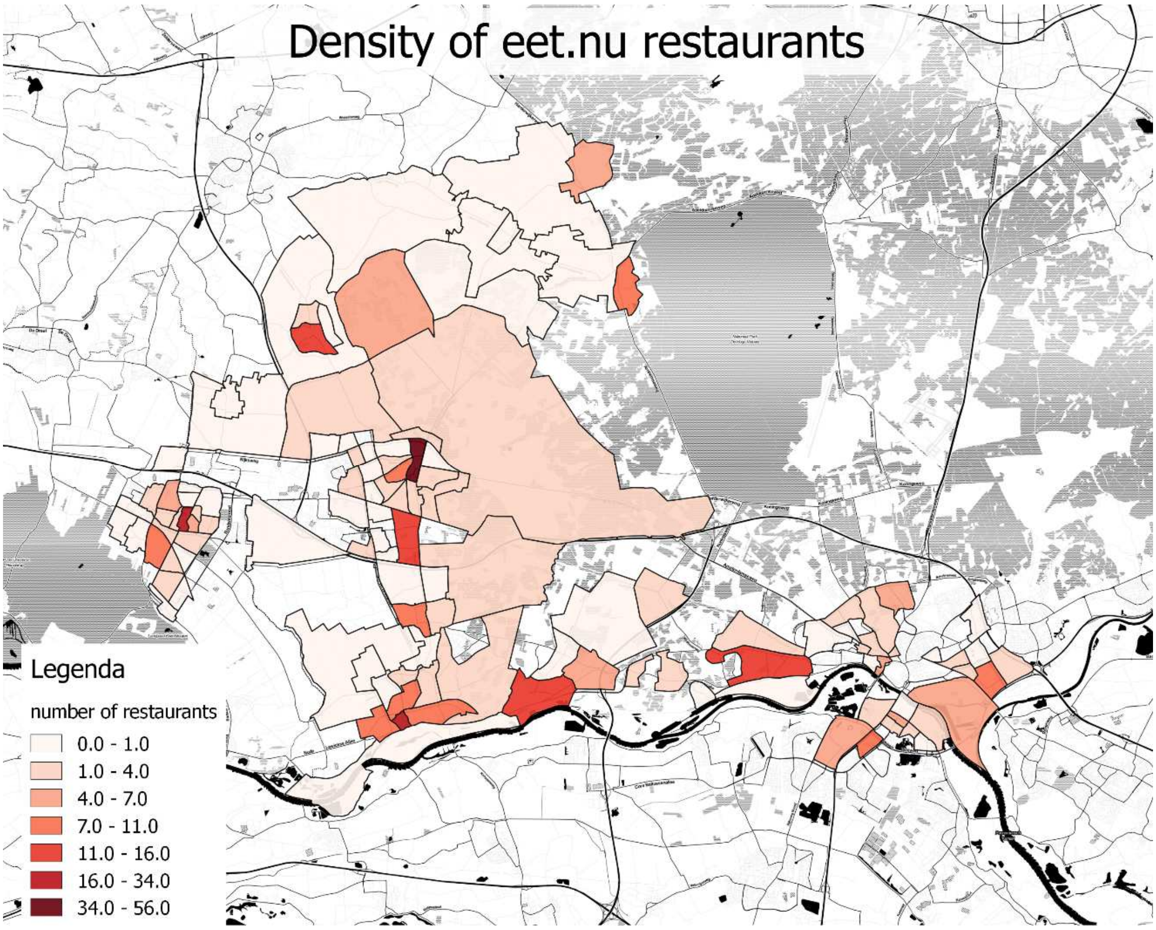

Figure 19 Density of eet.nu restaurants (per CBS neighbourhood)

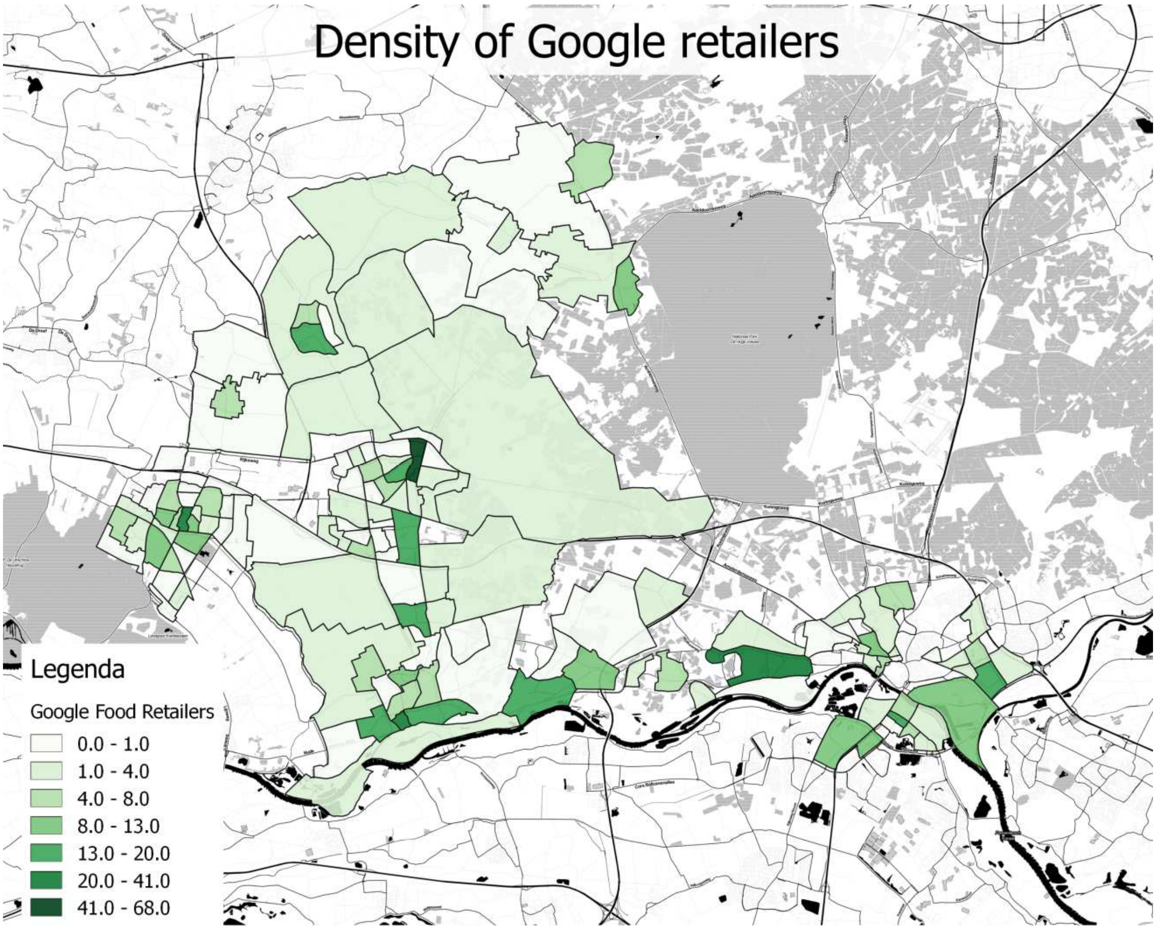

Figure 20 Density of Google food retailers (per CBS neighbourhood)

4.4 Spatial relationships between diet and food retailers

4.4.1 Characteristics of the BMI cluster's food environment

High and low BMI clusters have been saved as a separate file. Characteristics of local food environment (proximities and densities) calculated per participant were joined to these clusters' attribute tables. Mean distances and densities (per cluster) have been calculated. Table 15 contains the results of these calculations.

Table 15 Characteristics of the BMI clusters: mean variables of the local food environment

	BMI low (mean value)	BMI high (mean value)	difference (absolute)
Euclidean distance to:			
Restaurant CBS	0,74	0,77	0,03
Supermarket CBS	0,89	0,72	0,17
Bakery**	714,83	614,47	100,36
Bar**	1429,00	1382,29	46,71
Café**	484,62	584,95	100,33
Convenience Store**	2776,45	2042,73	**733,72**
Grocery store or supermarket**	797,08	473,00	324,07
Meal Delivery**	2034,66	1125,52	**909,14**
Restaurant**	467,67	366,41	101,25
Takeaway**	1339,83	790,79	**549,04**
Restaurant*	417,20	315,37	101,83
Snackbar*	708,85	453,06	255,79
Network Proximity to:			
Bakery**	895,20	783,44	111,76
Bar**	1721,39	1791,39	69,99
Café**	649,33	785,73	136,40
Grocery store or supermarket**	991,82	597,68	**394,14**
Restaurant**	412,27	421,23	8,97
Takeaway**	1584,26	1029,41	**554,85**
Restaurant*	524,76	453,13	71,63
Snackbar*	896,01	643,85	252,16
Neighbourhood Density of:			
Restaurants CBS (within 3km)	64,41	24,77	**39,65**
Supermarkets CBS (within 3km)	9,36	9,52	0,16
Bakeries**	1,36	0,25	1,11
Bars**	0,03	0,01	0,01
Cafes**	1,03	0,85	0,17
Convenience Stores**	0,05	0,04	0,01
Grocery store or Supermarket**	1,28	1,06	0,22
Meal Delivery**	0,00	0,18	0,18
Restaurants**	2,69	2,69	0,01
Takeaway**	0,03	0,30	0,27
Restaurants*	3,00	3,13	0,13
Snackbars*	0,49	0,78	0,29
Network distance density of:			
Bakeries within 800m**	2,26	1,06	1,20
Bakeries within 1600m**	7,49	4,18	3,31
Bars 800m**	2,13	0,22	1,91
Bars within 1600m**	4,90	0,69	4,21
Cafes within 800m**	8,97	2,96	6,01
Cafes within 1600m**	34,67	10,36	**24,31**
Grocery store or Supermarket 800m**	3,36	2,72	0,64

Grocery store or Supermarket 1600m**	14,54	8,87	**5,67**
Meal Delivery within 800m**	0,41	0,67	0,26
Meal Delivery within 1600m**	1,05	2,87	1,82
Restaurants within 800m**	16,49	8,3	**8,19**
Restaurants within 1600m**	55,92	32,01	**23,91**
Takeaway within 800m**	1,21	0,84	0,37
Takeaway within 1600m**	3,77	2,97	0,80
Restaurants within 800m*	19,08	9,19	**9,89**
Restaurants within 1600m*	59,15	35,54	**23,61**
Snackbars within 800m*	2,26	1,75	0,51
Snackbars within 1600m*	8,31	6,97	1,34

Data source: * eet.nu, ** Google

Investigation of differences between local food environment of high and low BMI cluster showed that:

- mean Euclidean proximity to the closest meal delivery was ~910 meters higher in case of low BMI cluster,
- mean Euclidean proximity to the closest convenience store was ~733 meters higher in case of low BMI cluster,
- mean Euclidean proximity to the closest takeaway was ~550 meters higher in case of low BMI cluster,
- mean network proximity to the closest takeaway was ~555 meters higher in case of low cluster,
- mean network proximity to the closest grocery store or supermarket was ~394 meters higher in case of low cluster,
- average person from low BMI cluster had ~40 restaurants more within 3km from home than person from high BMI cluster,
- average person from low BMI cluster had ~24 cafés, ~24 restaurants and ~6 grocery stores or supermarkets more within 1600 meters from home,
- average person from low BMI cluster had ~10 restaurants more within 800 meters from home.

These results show that people from the low BMI cluster (normal weight) live on average further away to food retailers like meal delivery, convenience store, takeaway, grocery store or supermarket. The density of food retailers like restaurants cafés, grocery store or supermarket in their neighbourhood (1600m from home) was higher than in case of the obese people cluster. It can be concluded that people from the normal weight cluster live in areas of higher density of restaurants and cafés, but their average distance to the closest meal delivery, convenience store, takeaway, grocery store or supermarket is higher than in case of obese people. It is unknown if these people are visiting cafés and restaurants more often than obese people (who have on average less of these places in their neighbourhood). Maybe they are going there instead of buying something in convenience stores or takeaway? Or the other way around—maybe obese people are buying more in convenience stores or takeaway because they live closer to them. It is hard to tell if and how local food environment affects peoples' BMI, but the apparent difference between food environment of normal weight people and obese people found in this study suggests that it may be influencing people's weight.

4.4.2 Relationship between BMI clusters and retailers kernel density

Figures 21-27 show kernel density of selected retailers together with location of BMI clusters. Selection of retailers was done according to Table 15. Selection included all retailers whose characteristics differ between low and high BMI clusters. The kernel density of retailers maps were already presented in Figure 18 (subchapter 4.3.3 Kernel density of food retailers). However, maps presented in this section included also locations of BMI clusters.

The reason for this step was to check if high and low BMI clusters are located differently with respect to chosen food retailers. However, the analysis did not result in a clearly visible relationship. Both clusters were situated similarly close to retailers' hotspots. The only retailers that showed some spatial relation with BMI clusters were **meal delivery** (Figure 26) and **takeaways** (Figure 27). In both cases, low BMI clusters are situated outside the hotspots of these retailers, while high clusters lay inside them.

In conclusion, most of the retailers do not seem to be spatially related between high/low clusters. The reason for this might be the (in)accuracy of retailers data (obtained from open sources) or the methods used. It might be also the case that these variables are not spatially related. Therefore, it should be further investigated.

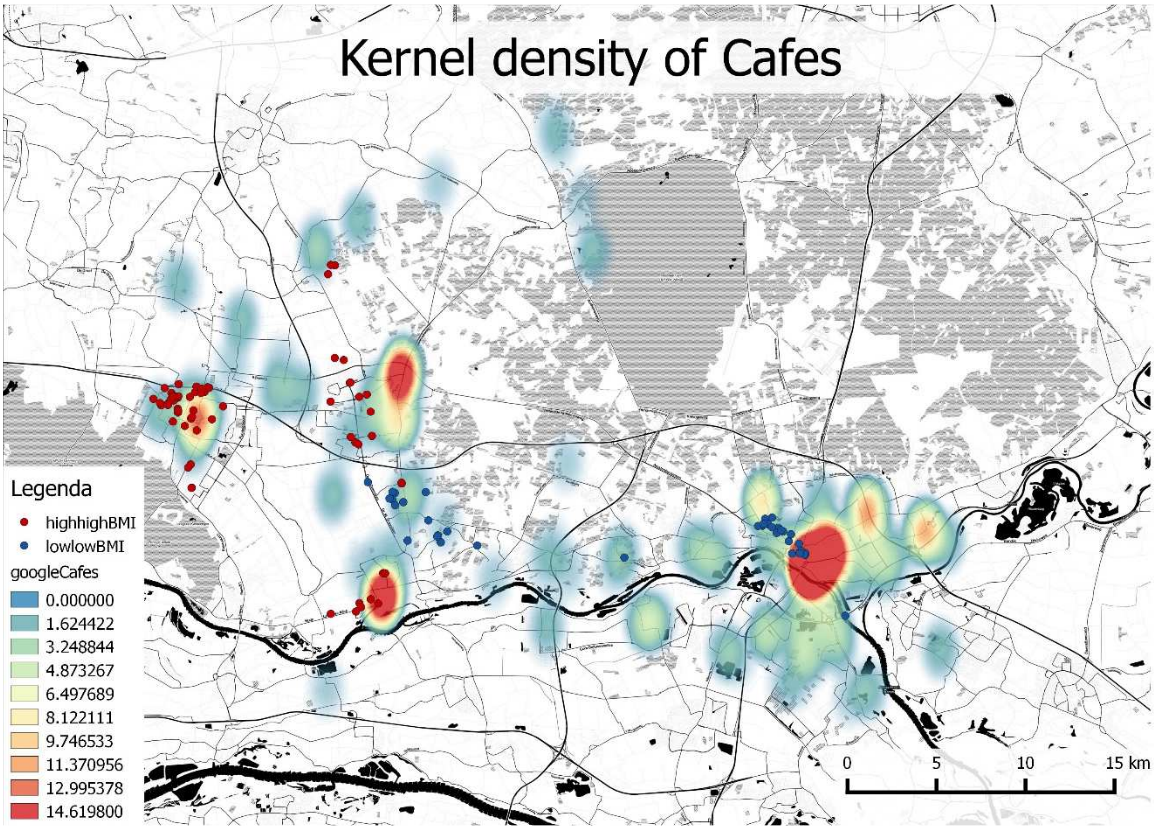

Figure 21 Kernel density of cafes

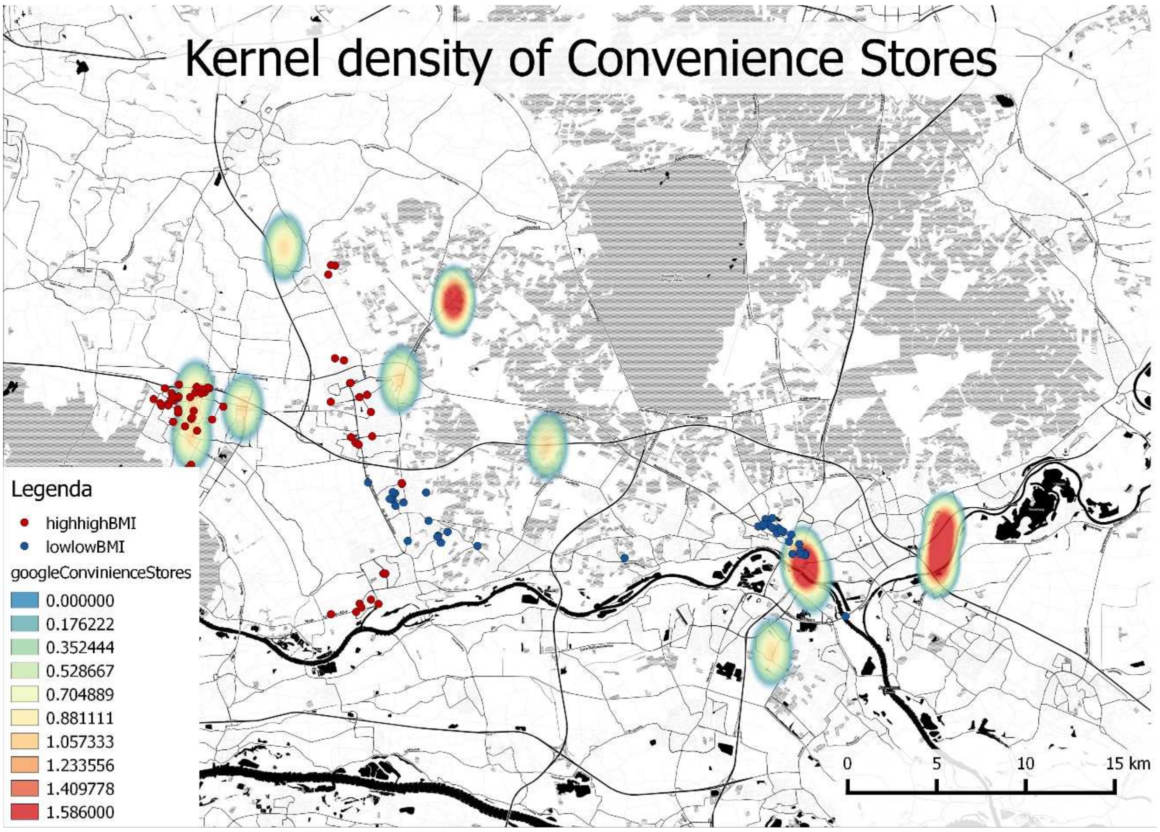

Figure 22 Kernel density of convenience stores

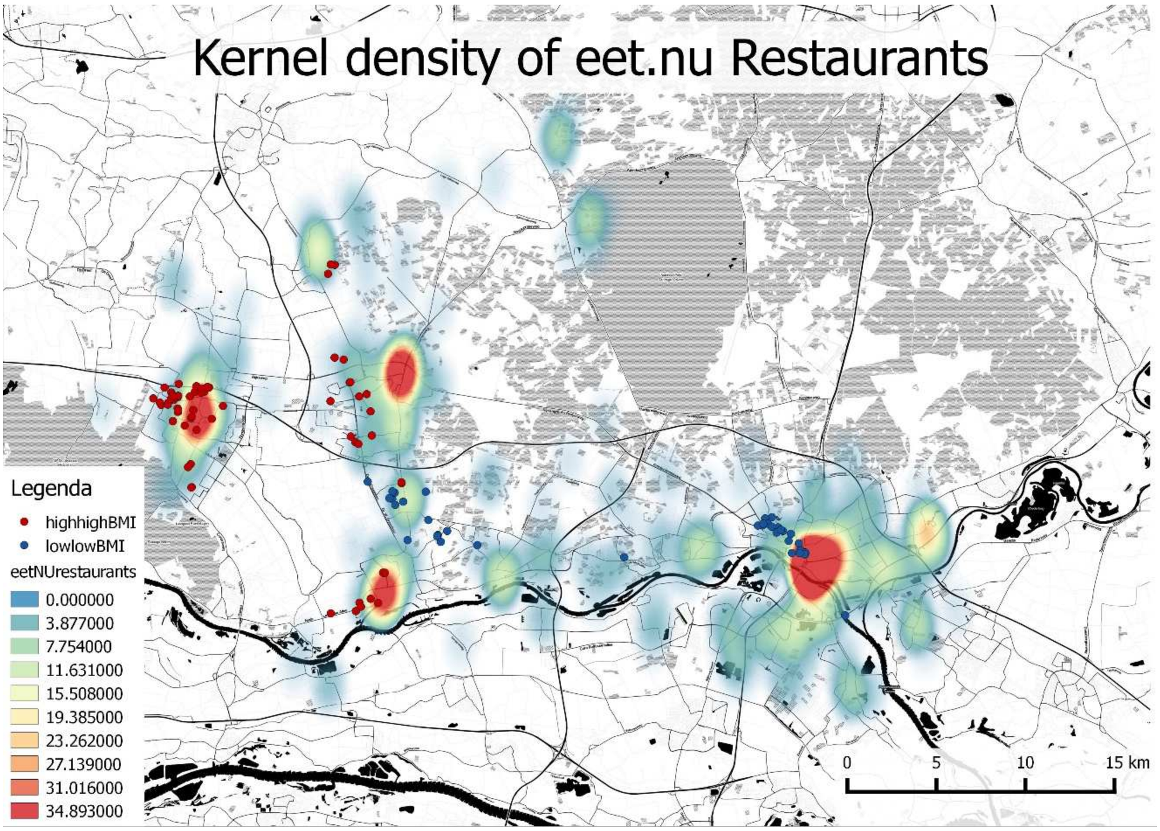

Figure 23 Kernel density of eet.nu restaurants

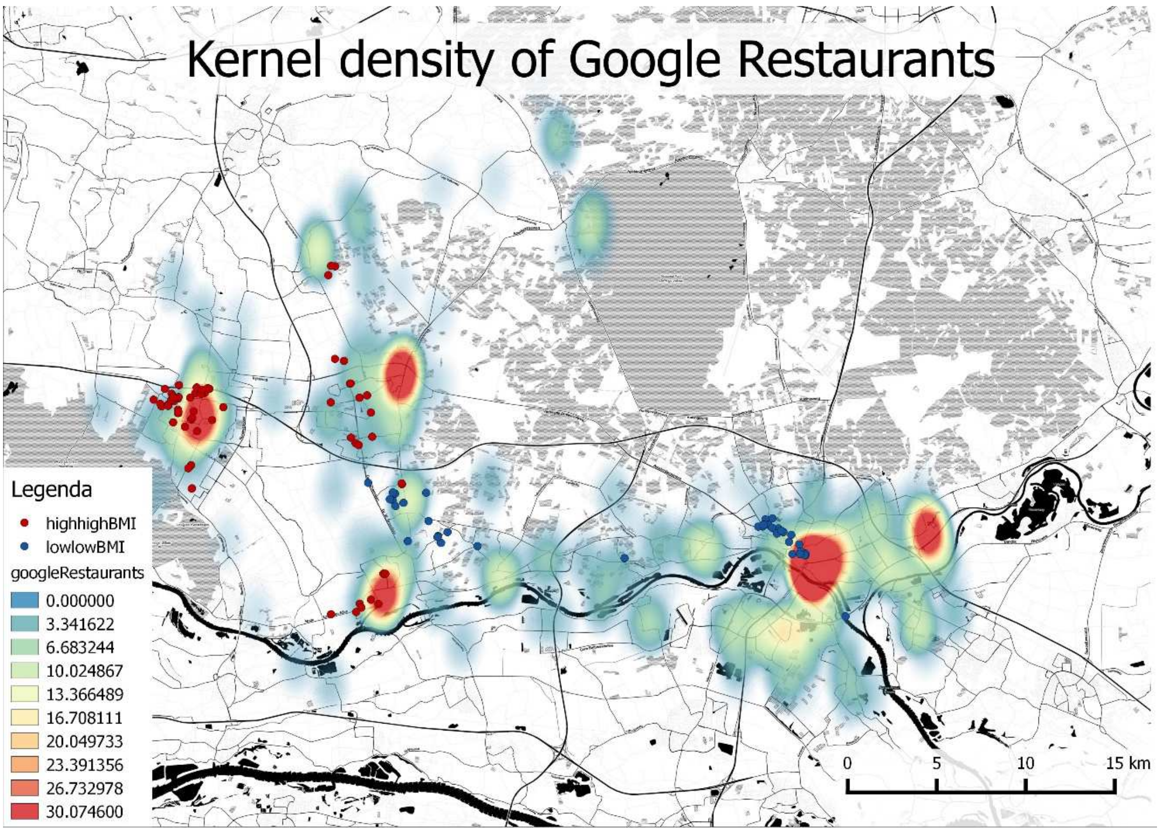

Figure 24 Kernel density of google restaurants

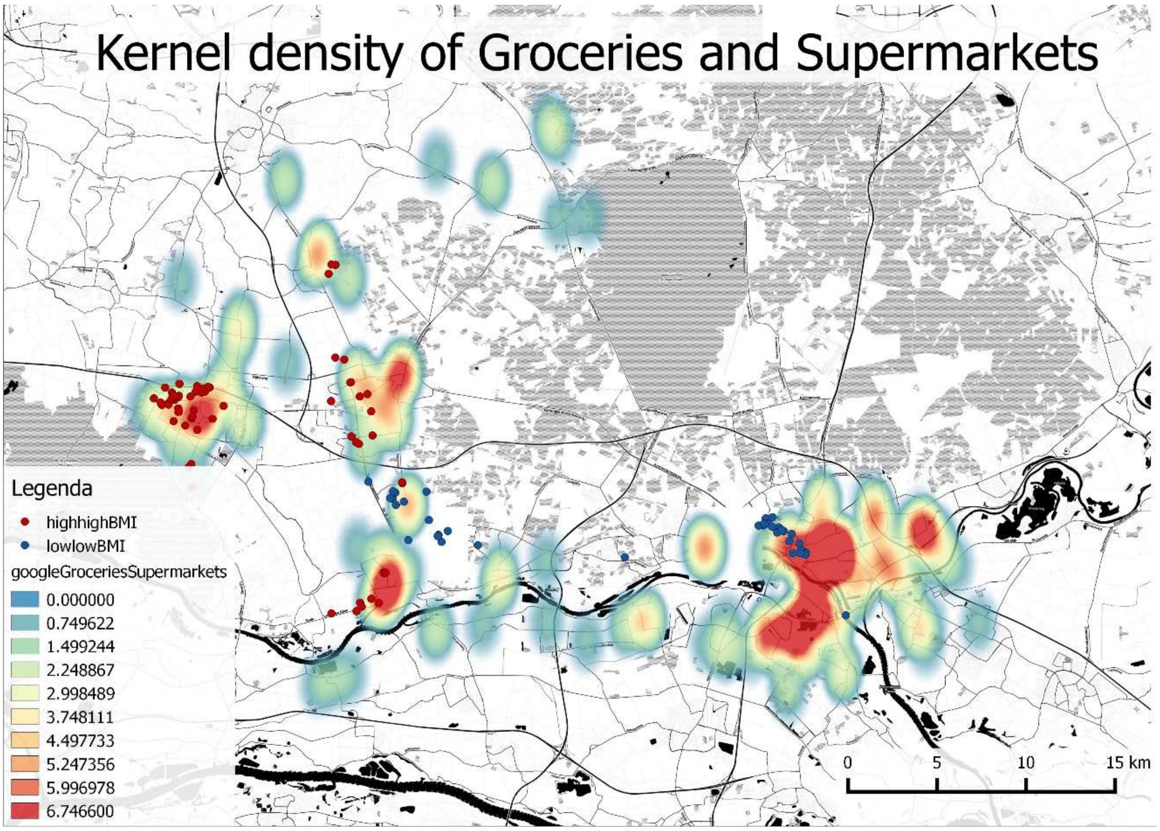

Figure 25 Kernel density of groceries and supermarkets

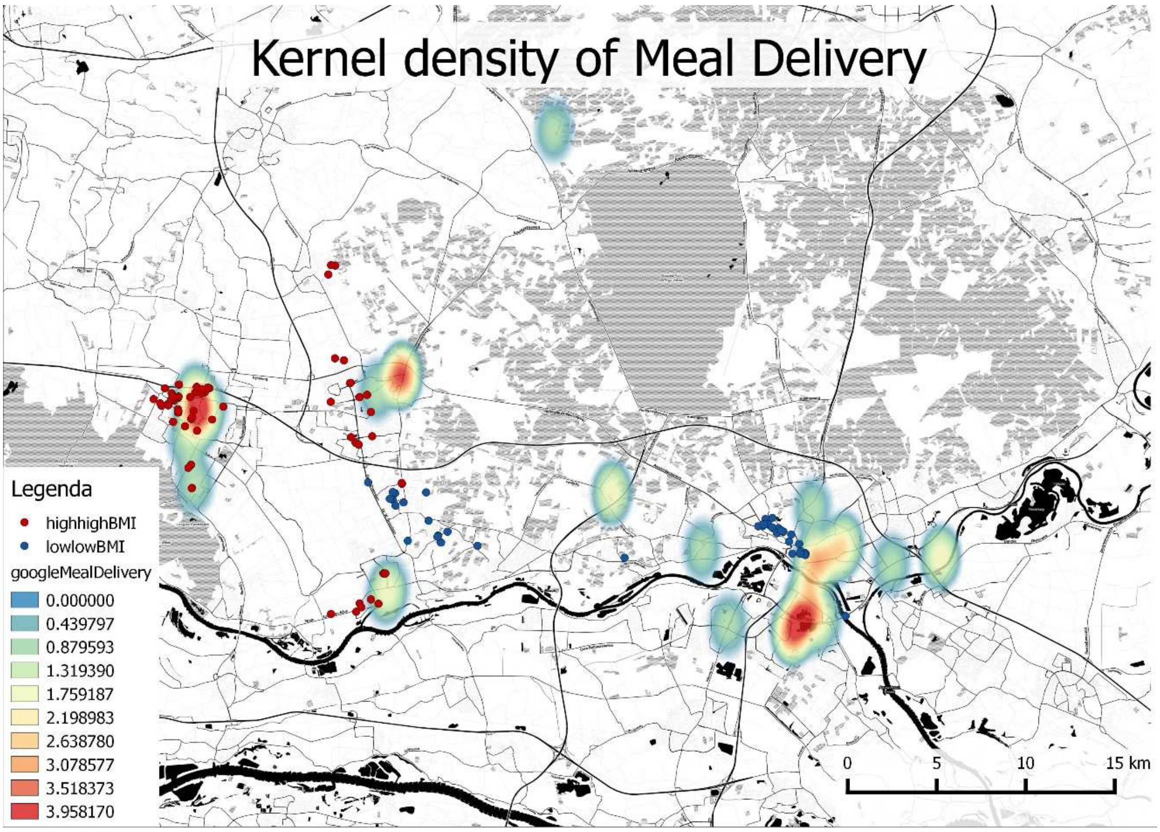

Figure 26 Kernel density of meal delivery

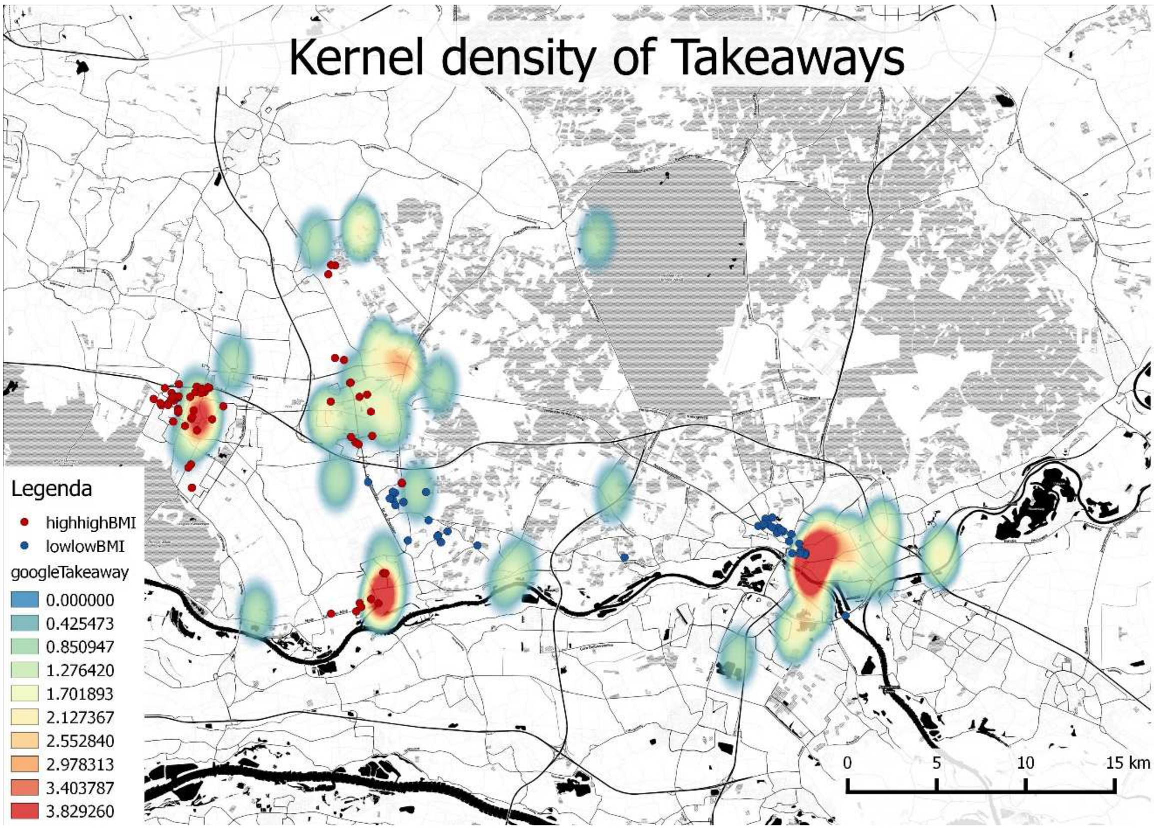

Figure 27 Kernel density of takeaways

4.4.3 Descriptive statistics

Descriptive statistics of variables used in this study have been calculated. They are displayed in Tables 16-19. Different variables contained different number of missing values. Therefore, each of the tables below contains number of rows taken in the account (n).

CBS data has showed that, in general, mean density of restaurants in the research area is higher than mean density of supermarkets. Mean distance (calculated by using all participants distance) to supermarkets (obtained from CBS) was longer than the mean distance to restaurants.

Table 16 Descriptive statistics of CBS data, n=1940

Variable (km)	mean	st dev	min	median	max
Mean distance to restaurant	858	484,7	100	800	3300
Mean distance to supermarket	794	510,6	200	600	4600
Mean density of restaurant	28	23,559	0,4	26,7	125,5
Mean density of supermarket	8	3,8724	0	9,6	24,4

Proximity to closest food retailer calculated as Euclidean distance (from individual's locations to the closest food retailer) was lowest in the case of restaurants and highest in case of takeaways.

Table 17 *Descriptive statistics of proximity (in meters) to closest retailer measured as Euclidean distance and network distance (*), n=1953*

proximity to nearest:	source**	mean	st dev	min	median	max
bakery	Google	640,6 *830,7**	496,8 *624,2**	6,2 *1,1**	533,9 *702,9**	4919 *5768,7**
café	Google	638,14 *830,2**	425,49 *523,3**	10,88 *2,3**	564,48 *755,8**	3050,87 *3553,2**
bar	Google	1545,9 *1936,6**	883,2 *1049,2**	15,4 *16,6**	1437,0 *1832,5**	4980,4 *5920,4**
grocery or supermarket	Google	525,6 *688,2**	414,3 *519,3**	3,23 *0,5**	433,49 *578,3**	3869,48 *4408,2**
restaurant	Google	359,26 *392,44**	264,19 *275,76**	3,68 *0,04**	303,49 *339,73**	2177,24 *2247,16**
takeaway	Google	1033,7 *1284,2**	723,5 *824,2**	7,2 *1,2**	913,9 *1183,6**	5357 *6303,6**
retailer	Google	288,42 *480,26**	211,93 *332,63**	3,23 *0,04**	241,69 *422,15**	1776,86 *2869,52**
restaurant	eet.nu	369,81 *506,14**	272 *365,27**	1,55 *0,02**	302,22 *426,91**	1990,79 *3040,19**
snack bar	eet.nu	519,26 *698,9**	365,93 *474,4**	14,97 *5,8**	460,44 *627,4**	3377,90 *4424,5**

**source of the locations

Density of retailers was the highest in case of restaurants, and lowest in case of convenience stores and bars.

Table 18 Descriptive statistics of density of retailers within the CBS neighbourhood, n=1935

Density of:	source*	mean	st dev	min	median	max
bakeries	Google	0,6656	1,0545	0	0	6
cafés	Google	0,9273	2,0234	0	0	19
bars	Google	0,07476	0,28004	0	0	2
convenience stores	Google	0,03379	0,19702	0	0	2
groceries and supermarkets	Google	1,2366	1,5977	0	1	9
meal delivery centres	Google	0,2033	0,5052	0	0	4
restaurants	Google	3,217	5,151	0	2	49
takeaway	Google	0,2355	0,5393	0	0	3
restaurants	eet.nu	3,401	5,960	0	1	56
snack bars	eet.nu	0,8725	1,2814	0	0	6

*source of the locations

Table 19 Descriptive statistics of retailers density within network buffer, n=1953

Density of:	source*	mean	st dev	min	median	max
bakeries within 800m	Google	1,1767	1,4379	0	1	9
bakeries within 1600m	Google	3,8792	3,1790	0	3	25
cafés within 800m	Google	2,545	6,417	0	1	79
cafés within 1600m	Google	10,008	15,869	0	5	130
bars within 800m	Google	0,254	1,1028	0	0	14
bars within 1600m	Google	0,9913	2,5377	0	0	23
groceries and supermarkets within 800m	Google	2,3845	2,7859	0	1	26
groceries and supermarkets within 1600m	Google	7,199	6,266	0	7	52
meal delivery centres within 800m	Google	0,5218	1,0518	0	0	6
meal delivery centres within 1600m	Google	1,7808	2,0603	0	1	8
restaurants within 800m	Google	7,35	12,79	0	3	124
restaurants within 1600m	Google	27,549	30,004	0	15	214
takeaway within 800m	Google	0,579	1,2251	0	0	10
takeaway within 1600m	Google	2,3318	2,5244	0	1	16
restaurants within 800m	eet.nu	8,059	14,547	0	3	136
restaurants within 1600m	eet.nu	29,942	32,738	0	17	228
snack bars within 800m	eet.nu	1,6728	2.0148	0	1	16
snack bars within 1600m	eet.nu	6,040	4,611	0	5	35

*source of the locations

4.4.4 Multiple regression analysis

Euclidean distance proximity

The first group of variables investigated in regression analysis as possible factors influencing the diet were distances to closest retailer, measured as Euclidean distance. A model was run for each proximity (to closest retailer) separately and included additional variables: sex, education level and income. Proximity, density and additional variables were calculated/extracted per individual diet holder's location (not for central locations of the clusters/neighbourhoods). Results of mentioned models are presented in Tables 20 and 21 and in Appendix V.

Results suggest that distance to the **closest café** may influence BMI (coef. 0,0005; p-value 0,04): people living closer to cafes have higher BMI. Although this model explained only 5,97% of the variability of the BMI data around its mean. The second association was found between CBS **mean distance to restaurants within 3 km from the neighbourhood** centre and kcal intake (with 1km increase in mean proximity kcal intake increased by ~83 kcal). However, this model has a low adjusted R-squared (0,61%). Finally **distance to closest food retailer** was negatively associated with kcal intake (coef. -2,726, p-value 0,0359) which means that the longer the distance to the place where food can be purchased, the lower the kcal intake. This model's adjusted R-squared was 0,2%.

No significant associations between DHD-index and food environment were found, so the DHD results table has been excluded from this chapter. It is included in Appendix V.

Table 20 Multiple regression on BMI as dependent variable

distance to the closest:	**BMI** incl. sex, age, education and income n = 1868				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurant (in km)*	-0,0016	0,9936	5,76%	5,50%	5,15%
supermarket (in km)*	0,3136	0,1078	5,89%	5,63%	5,28%
retailer**	0,0001	0,8097	5,00%	4,75%	4,41%
bakeries**	0,0002	0,3026	5,81%	5,56%	5,20%
café**	**0,0005**	**0,0423**	**5,97%**	**5,71%**	**5,34%**
bar**	0,0002	0,0580	5,94%	5,69%	5,30%
convenience store**	0,0000	0,6155	5,77%	5,52%	5,15%
groceries or supermarket**	0,0003	0,2005	5,84%	5,59%	5,22%
meal delivery centre**	0,0001	0,2195	5,83%	5,58%	5,19%
restaurant**	0,0004	0,2542	5,82%	5,57%	5,21%
takeaway**	-0,0001	0,2906	5,81%	5,56%	5,21%
restaurant***	-0,0001	0,7561	5,76%	5,51%	5,15%
snack bar***	0,0000	0,9814	5,76%	5,50%	5,14%

Data source: * CBS, ** Google, ***eet.nu, significant results are bold

Table 21 Multiple regression on kcal intake

distance to the closest:	**kcal intake** *incl. age* *n = 1570*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurant*	**82,9200**	**0,0086**	**0,74%**	**0,61%**	**0,33%**
supermarket*	-19,4700	0,5242	0,33%	0,20%	0,00%
retailer**	**-2,726**	**0,0359**	**0,32%**	**0,20%**	**0,00%**
bakery**	-0,0173	0,5800	0,32%	0,19%	0,00%
café**	0,0252	0,4857	0,33%	0,21%	0,00%
bar**	0,0063	0,7157	0,31%	0,18%	0,00%
convenience store**	-0,0129	0,0681	0,51%	0,39%	0,14%
grocery or supermarket**	-0,0720	0,0562	0,53%	0,41%	0,15%
meal delivery centre**	-0,0140	0,2933	0,37%	0,24%	0,00%
restaurant**	-0,0399	0,4839	0,33%	0,21%	0,00%
takeaway**	0,0149	0,4704	0,33%	0,21%	0,00%
restaurant***	0,0488	0,3883	0,35%	0,22%	0,00%
snack bar***	-0,0614	0,1583	0,43%	0,30%	0,05%

Data source: * CBS, ** Google, ***eet.nu, significant results are bold

Network proximity

By a regression analysis, investigated proximity (measured as network distance) on BMI, kcal intake and DHD. The results of this step (concerning BMI and kcal intake as dependent variables) are presented in Tables 22 and 23. The results from models using DHD as dependent variable have been moved to Appendix V.

Table 22 Multiple regression on BMI as dependent variable

distance to the closest:	**BMI** incl. sex, age, education and income n = 1868				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
retailer*	0,0004	0,1154	5,08%	4,88%	4,57%
bakery*	0,0002	0,2909	5,81%	5,56%	5,20%
cafés*	0,0003	0,0675	5,93%	5,67%	5,30%
bar*	**0,0002**	**0,0491**	**5,59%**	**5,70%**	**5,32%**
convenience store*	0,0005	0,0994	5,89%	5,64%	5,28%
grocery or supermarket*	0,0002	0,2169	5,83%	5,58%	5,21%
meal delivery centre*	0,0002	0,5010	5,78%	5,53%	5,17%
restaurant*	-0,0001	0,3957	5,79%	5,54%	5,19%
takeaway*	0,0000	0,9531	5,76%	5,50%	5,15%
snack bar**	-0,0000	0,8764	5,00%	4,75%	4,40%
restaurant**	0,0001	0,7698	5,76%	5,51%	5,15%

Data source: * Google, **eet.nu, significant results are bold

Results suggest that BMI may be associated with the distance to the closest **bar** (coef. 0,0002, p-value 0,0491). Adjusted R-squared for this model was 5,70%. In case of kcal intake, one association has been found. Distance to the closest **grocery or supermarket** was negatively associated with kcal intake (coef. -0,0639, p-value 0,0338). In other words, the closer you live to a supermarket or grocery store, the more kcal you consume. However, this model's adjusted R-squared was only 0,46% which is really low.

Table 23 Multiple regression on kcal intake as dependent variable

distance to the closest:	**kcal intake** *incl. age* *n = 1570*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
retailers*	- 0,0352	0,4366	0,32%	0,20%	0,00%
bakeries*	-0,0201	0,4182	0,34%	0,33%	0,00%
cafés*	0,0167	0,5671	0,32%	0,20%	0,00%
bars*	0,0079	0,5897	0,32%	0,19%	0,00%
convenience store*	-0,0378	0,4029	0,35%	0,22%	0,00%
groceries or supermarkets*	**-0,0639**	**0,0338**	**0,59%**	**0,46%**	**0,20%**
meal delivery centres*	-0,0612	0,2588	0,38%	0,26%	0,00%
restaurants*	0,0185	0,3071	0,37%	0,24%	0,00%
takeaway*	0,0302	0,4797	0,33%	0,21%	0,00%
snack bar**	-0,0386	0,2551	0,37%	0,24%	0,00%
restaurant**	-0,0364	0,2838	0,37%	0,25%	0,00%

*Data source: * Google, **eet.nu, significant results are bold*

CBS neighbourhood density

Neighbourhood densities of food retailer were investigated next. The results of models are displayed in Tables 24 and 25. Models with DHD as a dependent variable again did not find any significant predictors (results moved to Appendix V). Models using BMI as a dependent variable have found a negative association with density of **bakeries** (coef. -0,1905, p-value 0,0347). This model has low adjusted R-squared: 5,73%. Kcal intake was negatively associated with densities of **cafes** (coef. -23,8900, p-value 0,001), **restaurants** (coef. -7,5650, p-value 0,0023) and **food retailers** in general (coef. -4,566, p-value 0,0037). This means that the more restaurants/cafes/food retailers in a neighbourhood you have, the more kcal you consume. However these models had really low adjusted R-squared (between 0,69% and 0,87%).

Table 24 Multiple regression on BMI as dependent variable

density of:	**BMI** incl. sex, age, education and income n = 1868				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurants*	-0,0038	0,3537	5,80%	5,55%	5,20%
supermarkets*	-0,0259	0,3100	5,81%	5,56%	5,19%
retailers**	0,0021	0,8410	5,46%	5,51%	4,73%
bakeries**	**-0,1905**	**0,0347**	**5,98%**	**5,73%**	**5,37%**
cafés**	-0,0429	0,3571	5,80%	5,55%	5,21%
bars**	-0,1482	0,6536	5,77%	5,51%	5,15%
convenience stores**	-0,0639	0,8920	5,76%	5,50%	5,16%
groceries or supermarkets**	-0,0254	0,6635	5,77%	5,51%	5,15%
meal delivery centres**	-0,1111	0,5482	5,77%	5,52%	5,18%
restaurants**	-0,0176	0,3370	5,80%	5,55%	5,21%
takeaway**	0,1542	0,3837	5,80%	5,54%	5,18%
restaurants***	-0,0095	0,5486	5,77%	5,52%	5,18%
snack bars***	0,0347	0,6335	5,77%	5,52%	5,13%

*Data source: * CBS, ** Google, ***eet.nu, significant results are bold*

Table 25 Multiple regression on kcal intake as dependent variable

density of:	**kcal intake** *incl. age* *n = 1570*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurants*	-0,8476	0,1938	0,41%	0,28%	0,02%
supermarkets*	-2,1310	0,5910	0,32%	0,19%	0,00%
retailers**	**-4,566**	**0,0037**	**0,82%**	**0,69%**	**0,43%**
bakeries**	-20,7700	0,1585	0,43%	0,30%	0,05%
cafés**	**-23,8900**	**0,0010**	**0,99%**	**0,87%**	**0,61%**
bars**	-75,5800	0,1689	0,42%	0,29%	0,02%
convenience stores**	-11,9500	0,8754	0,30%	0,18%	0,00%
groceries or supermarkets**	-17,0100	0,0763	0,50%	0,37%	0,13%
meal delivery centres**	-20,0000	0,5042	0,33%	0,20%	0,00%
restaurants**	**-8,1590**	**0,0045**	**0,81%**	**0,69%**	**0,42%**
takeaway**	-50,8100	0,0745	0,50%	0,38%	0,12%
restaurants***	**-7,5650**	**0,0023**	**0,89%**	**0,76%**	**0,50%**
snack bars***	-21,3900	0,0713	0,51%	0,38%	0,14%

*Data source: * CBS, ** Google, ***eet.nu, significant results are bold*

Network buffer density

The last stage of regression analysis included network buffer density of food retailers (within 800m and 1600m distance from diet holder's home) as possible factors influencing BMI/DHD/kcal intake. In this group, only models using kcal intake as a dependent variable have found some associations. The results of these models are showed in Table 26. Results of models investigating relation between retailer density and BMI/DHD have been moved to Appendix V.

Kcal intake was significantly related with density of: **cafes** (coef. = -6,62 in case of 800m buffer), **bars** (coef. = -35,7900 in case of 800m buffer), **groceries and supermarkets** (-4,8490 in case of 1600m buffer), **restaurants** (-3,1860 in case of 800m buffer) and **takeaways** (-24,3900, in case of 800m buffer).

Table 26 Multiple regression on kcal intake as dependent variable

density of:	**kcal intake** *incl. age* *n = 1570*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurants*	-0,8476	0,1938	0,41%	0,28%	0,02%
supermarkets*	-2,1310	0,5910	0,32%	0,19%	0,00%
retailers 1600m**	**-0,6558**	**0,0127**	**0,71%**	**0,57%**	**0,32%**
retailers 800m**	**-1,9184**	**0,0028**	**0,89%**	**0,76%**	**0,52%**
bakeries 1600m**	-5,7570	0,2336	0,39%	0,26%	0,02%
bakeries 800m**	-19,8800	0,0621	0,52%	0,40%	0,14%
cafés 1600m**	**-2,5289**	**0,0086**	**0,74%**	**0,61%**	**0,40%**
cafés 800m**	**-6,6180**	**0,0034**	**0,85%**	**0,72%**	**0,52%**
bars 1600m**	**-13,5650**	**0,0249**	**0,62%**	**0,49%**	**0,29%**
bars 800m**	**-35,7900**	**0,0060**	**0,78%**	**0,65%**	**0,45%**
groceries or supermarkets 1600m**	**-4,8490**	**0,0466**	**0,55%**	**0,43%**	**0,19%**
groceries or supermarkets 800m**	-6,4470	0,2341	0,39%	0,26%	0,02%
meal delivery centres 1600m**	-1,9180	0,7950	0,31%	0,18%	0,00%
meal delivery centres 800m**	-9,7100	0,5049	0,33%	0,20%	0,00%
restaurants 1600m**	**-1,0258**	**0,0438**	**0,56%**	**0,43%**	**0,20%**
restaurants 800m**	**-3,1860**	**0,0063**	**0,77%**	**0,65%**	**0,43%**
takeaway 1600m**	**-14,5240**	**0,0162**	**0,67%**	**0,54%**	**0,30%**
takeaway 800m**	**-24,3900**	**0,0468**	**0,55%**	**0,43%**	**0,19%**
restaurants 1600m***	**-0,9965**	**0,0325**	**0,59%**	**0,47%**	**0,23%**
restaurants 800m***	**-2,9390**	**0,0041**	**0,82%**	**0,70%**	**0,48%**
snack bars 1600m***	-3,7020	0,2608	0,38%	0,25%	0,01%
snack bars 800m***	-12,4450	0,0952	0,48%	0,35%	0,11%

*Data source: * CBS, ** Google, ***eet.nu, significant results are bold*

4.4.5 Summary of spatial relationship analysis

Regression models conducted in this analysis investigated each of the food retailers' variable in separate models. These models have found some associations, but the adjusted R-squared were very low (0,20% - 5,73%), which means that models poorly predicted BMI, kcal intake and DHD values. None of the models have found an association between DHD and the food environment variable. In case of BMI, influencing factors were:

- Euclidean distance to closest café,
- Network distance to closest bar,
- Neighbourhood density of bakeries.

In case of kcal intake, more factors have been found. They were:

- Euclidean distance to closest restaurant,
- Euclidean distance to closest food retailer,
- Network distance to closest supermarket or grocery,
- Neighbourhood density of food retailers,
- Neighbourhood density of cafes,
- Neighbourhood density of restaurants,
- Network buffer (800m and 1600m) density of food retailers,
- Network buffer (800m and 1600m) density of cafes,
- Network buffer (800m and 1600m) density of bars,
- Network buffer (1600m) density of groceries or supermarkets,
- Network buffer (800m and 1600m) density of restaurants,
- Network buffer (800m and 1600m) density of take away.

Most often factors were associated with cafes and restaurants, but also with food retailers in general. However, low adjusted R-squared obtained in these models suggest that these models are bad in predicting, and therefore not reliable. Thus, it is suggested that more regression models, using more food retailers' variables at once should be conducted. Hopefully they can result in higher adjusted R-squared values. Eventually, other variables proven to be influential should be included in this type of analysis. This can lead to building a model that will predict the kcal intake, DHD or BMI much better than models used in this study. More about this can be found in subchapter 5.3 Recommendations.

5 Conclusions, discussion and recommendations

5.1 Introduction

This chapter is meant to summarize this study. Firstly, it describes the answers to research questions. Secondly, it compares the results of this study with results obtained by other studies (finding similarities and differences). Comparison of the results leads to conclusions. This chapter also presents limitations and strengths of this study. Finally, recommendations for future research are presented (5.4).

5.2 Conclusions

This study explored the research regarding the relationships between the food environment and diet. The summary of the results, ordered by research questions is presented below.

RQ1: How to spatially express human diets?

Diet/weight variables investigated in previous studies included food (as specific products, e.g. fruits and vegetables, etc.) or nutrients (e.g. calcium, proteins, etc.) intake and diet indices like BMI. These variables have been studied in many configurations. Most often fruit, vegetable intake and BMI were investigated. This study has used kcal intake, DHD and BMI.

Spatial representation of diet variables has been expressed in a few ways. Some of the studies have been using exact locations of diet holders (e.g. Larsen et al., 2015) while other studies used polygons with aggregated locations, resulting in a count of them (per neighbourhood) (e.g. Clark et al., 2014). In the case of the second scenario, neighbourhoods were represented by using: census areas, postcode areas or grid. Aggregation resulted in polygons containing a count of the points (representing diet holders within a polygon/cell) or mean value of variable per polygon. Diet holder's location has been expressed as locations of their homes or schools.

This study used exact locations of diet holders. It investigated if different diet variables are spatially clustered. The existence of clusters determined which locations should be further investigated and compared.

RQ2: How to spatially express the food environment?

Similarly to diet variables, food environment was spatially expressed as point locations and as polygons with a count of retailers per neighbourhood. Retailer types included varied between studies. Some researchers used a specific type of retailer (e.g. Fraser & Edwards, (2010) have used fast food restaurants) while others used all places where food can be purchased (e.g. Larsen et al. (2015)). This study has used all food retailers that could be located via open data sources. These retailers were used separately (using different retailer types) and as a one group (all retailer types).

GIS methods used to describe local food environment included measuring density of retailers and proximity to the closest retailer. They were calculated, using different techniques. The density of retailers (per diet holder) was calculated using: circular buffer, network buffer, kernel density, spatial clustering. Proximity to the closest retailer was calculated as: Euclidean distance, network distance, population weighted distance, modelling travel time. Mentioned methods expressed food environment per study participant.

Interestingly, the spatial patterns of different retailer types looked similar. It was expected, because normally city centres are more retailer-dense.

RQ3: What approach can be used to study the impact of the food environment on diet patterns?

This study has used various methods in order to investigate if there is a relation between a diet/weight status and the local food environment in the Netherlands.

The comparisons of different data sources and methods were done in order to create a methodology which can be used again in similar studies. It is difficult to select the best method for this type of study because the results were mostly insignificant. The choice of methods used in this study was mostly dictated by approaches found in the literature. However, there were some exceptions. To my knowledge, the methods comparing local food environment with low and high BMI clusters were developed in this study and therefore, not used before.

In case of proximity network distance, two methods have been used (Euclidean distance and network distance). Network distance seems to be better because it is a more realistic representation of the distance. Both proximities (Euclidean distance and network distance) were used in final analysis (separately). Comparison of the results of both approaches did show differences. The regression model using Euclidean distance to the closest food retailer found association with BMI and cafés, while the regression model using network distance found associations between BMI and bars. Comparison of low and high BMI clusters also found differences between mentioned proximity methods. The values of mean Euclidean distance to the closest retailer and network distance the closest retailer (calculated for both, low and high BMI clusters), often greatly differed. The difference was often more than 200 meters (e.g. in the case of mean distance to the closest bar). The network distance values were always higher than Euclidean distances, because they were calculated by using street network instead of straight line between points. Thus, the differences in values were expected. However, it was interesting to see if these differences will be also present in results of analysis (that used both Eucludean and network distance as input). The results showed that the selection of proximity method can greatly change the results of analysis (see the beginning of this paragraph).

Calculation of retailer density included 3 methods: network buffer, kernel density and spatial clustering. In comparing network buffer and spatial clustering (to predefined neighbourhoods), network buffer seems to be better. This method was better because every participant was located in the centre of its service area (which was not always the case while using CBS neighbourhood). Besides this, service areas had similar sizes. In the case of CBS neighbourhoods (which were better for visualization purposes), sizes of the polygons differed, which made analysis less uniform and probably less reliable.

Methods which are recommended for next studies in the Netherlands include proximity to closest retailer calculated as network distance, and density of neighbourhood retailers calculated as number of retailers within selected network buffer. These variables can be used to describe local food environment of clusters of high and low values (e.g. BMI, DHD or kcal intake) located by Moran's I.

In the case of data, this type of research in the Netherlands is still lacking accurate food retailer data, so no recommendation is given at this point.

RQ4: Does a Dutch food environment influence dietary patterns?

Results have shown that some of the characteristics of the local food environment like retailer density or proximity may influence diet. However, in most of the cases these characteristics (if proved to be significant) were only weakly correlated with the diet/weight. Cluster analysis gave more promising results. It helped to locate clusters of high and low values and compare the average local food environments of them (by using mean values of proximity and density of local food retailers).

Multivariate Linear Regression results indicated that the higher the neighbourhood density of the cafes, the lower the kcal intake (-23,89 kcal per additional café). Also, a similar association with neighbourhood restaurants density was found, which proved a high density to be correlated with less kcal intake. However, the adjusted R-squared in both cases was very low. It is expected that results may improve by modifying the regression model. More promising results were obtained by using cluster analysis. It was found there that high and low BMI cluster clearly differ in their spatial relation to food retailers. Therefore it may be an influencing BMI of people. Interestingly, people from high BMI (obese people) cluster had on average ~40 restaurants less within 3km from their home than people from low BMI cluster. Besides that, obese people had ~24 cafés, ~24 restaurants and ~6 grocery stores or supermarkets less within 1600 meters from home than normal weight people. Finally, the average person from the low BMI cluster had ~10 restaurants more within 800 meters from home than person from high BMI cluster.

Limitations

This study encountered some limitations. Firstly, the diet measurements were self-reported, which means they might be not reliable. It was already found by Larsen et al. (2015), Spence et al. (2009), Morland & Evenson (2009) that self-reported measurements were often underestimated.

Secondly, the open source data on food retailers was not perfect. Some places included did not exist anymore, whereas other existing places were not included. This definitely interfered with the final results. It is unknown how strongly.

Thirdly, WAVER data (introduced in Chapter 3.3 Diet data) have missed the values of kcal intake or DHD. Because some of the participants were missing this data, they were excluded from the analysis. It would be beneficial if they do not miss this data, what would result in bigger analysis sample. Additionally, it is possible that NQplus sample was not representative enough.

Finally, buffer sizes could have been chosen incorrectly because so far no buffer choice approach has been created yet. Therefore, it is unknown what buffers are best in this type of study..

Potential implications of these limitations are unknown, but it is suspected that they are playing a role in the correctness of the study results. Therefore it is recommended to minimize them with the next study investigating the same topic.

Strengths

The strength of this study is the fact that BMI was not self-reported as was often a case in previous studies. Self-reported data was proven to be less reliable because people often under-report height and over-report weight (Spence et al., 2009; Morland & Evenson, 2009). Another strength was a literature review and investigation on possible retailer data sources which may be used in the Netherlands.

Such studies were not conducted yet in the Netherlands. Thus this is the first study investigating a Dutch case, which also makes it important.

Summary

This study investigated what was done and what still can be done in the area of research on relation between diet and food environment. It succeeded in reviewing studies on this topic and drawing conclusions from it. These conclusions have been used in creating a proper methodology, which was used in a Dutch case study. The mentioned methodology can be also used in future research in the Netherlands.

5.3 Discussion

Answers to research questions were carefully compared with what was found in reviewed studies. The results of this comparison are presented in this subchapter. Additionally, the major conclusions and critical evaluation of this study have been added to this chapter.

Data

Data used in previous researches varied between studies. This study used predefined diet data (NQplus) and food retailers data downloaded from open sources. Previously conducted studies used predefined (collected in another study) diet data as well. Only ~22% of the investigated studies collected diet data themselves, mostly via questionnaires. In the case of food retailer data, most of the studies have used governmental data from municipalities. Governmental data seems to be more reliable in terms of accuracy than open source data. Therefore, studies that used them may get more accurate results than this study (which used open sources).

Investigated studies often used a specific group of people (e.g. children, pregnant women). This study sample included various age groups and both genders.

Research methods

Multiple methods have been used in other studies' analysis. The approach of using density and proximity was more common. Similarly to this study, proximity to closest retailer was calculated as Euclidean distance or by network distance. In the case of density of neighbouring retailers, investigated studies used predefined neighbourhoods (e.g. census track, postal code neighbourhoods) or service areas (constructed as network or circular buffers around diet holder's locations). This study has used a few of these methods (Euclidean distance, network distance, network buffer as service area and predefined CBS neighbourhoods).

An important part of this study was cluster analysis with Moran's I tool. Surprisingly, this method was not used before in the same way. Moran's I was previously used by Kloog et al. (2009) and Zenk et al. (2005). They used it in order to check spatial clustering of residuals obtained in Ordinary Least Squares regression. Frank et al. (2006) used Moran's I to check the degree of food outlet locations. In the case of this study, Moran's I was used to identify clusters of people with high and low values concerning diet or weight. After identifying clusters, their local food environment was characterized by using mean of proximity and density of retailers.

Most of the studies investigated point/point and point-polygon relations. The first one included relation between locations of study participants and the locations to her/his closest food retailer (proximity between them). The second one did investigate the relation between the location of study participant's location (together with his/his diet) and the density of food retailers within the service area of investigated participant (expressed as buffer or CBS neighbourhood). This study used exact location of diet holder. Most of the investigated studies shared this approach. However, some of them snapped the location of study participant to the centre of the neighbourhood where the study participant lived. This approach was used by Pearce et al. (2007) – where centroids of meshblocks were used.

Results

Investigated studies found various relationships between diet and local food environment. Retailers which have proven to be influential were **supermarkets, restaurants**, **fast food restaurants**, **groceries**, **convenience stores** and **supercentres**. In this study most retailers from these groups (supermarkets, restaurants, fast food restaurants, groceries, convenience stores) have been investigated as well. Sometimes the associations which were found were the same, sometimes very different. For example, Block et al. (2011) have found that for every 1km increase in driving distance to the closest grocery store, BMI decreased by 0.06 units. My study has found a similar relationship: people from low BMI clusters live in areas of lower grocery and supermarket density than high BMI clusters. However, there were also studies that have found the opposite of my results. Gibson et al. (2011) have found that the more restaurants, the higher the BMI, while my study found that people from low BMI clusters live in places with higher density of restaurants than people from high BMI groups (the same relationship has been found by Mehta & Chang (2008) and Inagami et al. (2009)). It is an interesting result because you expect that these people are eating in restaurants more often and therefore they have higher BMI. In fact, they might visit restaurants more often, but the food composition may be quite different compared with the ones at fast food places. There is also a possibility that the closeness to restaurants does not influence their visits in the restaurants and they visit them with the same frequency as people living further from them. We cannot know for sure what is behind these results, so it might be interesting to investigate it further.

Fast food restaurants or snack bars were not found significant in this study while Block et al. (2011) have found association between their density and BMI. Grocers were not analysed as a separate category, therefore the results found by Smiths et al. (2013) cannot be compared with the results from this study. Distance to convenience stores prove to be associated with BMI, as it was found by Berge et al. (2014). However this was found only in cluster analysis, not in regression. Finally, supercentres were not investigated in this study, therefore they cannot be compared.

Interestingly, people from the normal weight cluster lived in areas of higher density of restaurants and cafés, but their average distance to the closest meal delivery, convenience store, takeaway, grocery store or supermarket was higher than in case of obese people. These results are surprising because the kernel densities of different retailer types showed that the density patterns are similar in case of all retailers. Thus the density of the restaurants will be highest in similar locations like the highest density of groceries and supermarkets. The results suggest that the further away to one group of retailers you live and the more retailers from second group you have in your neighbourhood, the less likely you are to became obese. Taking into account fact that density patterns are similar for all retailers, it is expected that the highest density of retailers from both groups will be located in similar places. To investigate these relations further, another service areas sizes should be used. It is interesting to see if the results will be similar while using different buffer size to calculate the retailers density.

This research is important because it helps to understand how local environment (in this case food environment) influences our diet choices, as well as how to investigate this relation. Researchers and practitioners can use this knowledge in order to study this topic more extensively and/or change policies concerning food environment.

Taking into account the limitations and strengths of this study, it can be concluded that it could have been better. The value of it lay especially in creating proper methodology which may be used in future studies. Downsides include diet data sample size and the accuracy of retailer data. It leads to conclusion that the value of the results is questionable. More accurate food retailer data and bigger sample size of diet data would considerably increase the quality and therefore the value of the results.

Our diet is influenced by many factors. It is still being investigated how complex these relationships are. We do not know yet if and how strong the relationship between diet and local food environment is, but this study results suggest that the density of retailers in your local food environment is indeed influential. The investigation of the Dutch case confirmed that. Therefore, it is crucial to continue the research on these topics in order to benefit from it in the future. The findings of this study can be helpful in creating a healthier environment or in preventing obesity or lowering obesity rates.

5.4 Recommendations

This study also resulted in recommendations for future research. They are gathered together in this section.

Firstly, it is recommended to collect data from as many sources as possible and compare them with each other. This will help in the selection of the most accurate dataset for analysis. Because the more accurate the data is, the more reliable the results of the study are, which improve results. Therefore, the initial focus should be placed on comparing available data sources in order to find the most reliable one.

This study struggled with Modifiable Areal Unit Problem (MAUP). It concerns the selection of aggregation level used for aggregating food retailers per neighbourhood. It was mentioned in previous studies (Cerin et al., 2011) that the selection of a proper unit is difficult. Therefore it is suggested to try different units (census areas, postcode areas, or grid) because using other sizes may give different results.

Another spatial units used were buffer zones (service areas). It also created problems because no approach for selecting proper buffer size exists. Thus the decision on buffer size should be made carefully.

Obtained data can be analysed in many ways. Methods used in this study favour spatial statistics like Moran's I. Advanced analysis like Ordinary Least Square regression and Geographically Weighted Regression are highly recommended, as well.

It is suggested to validate the relationships that are found. It can be done by making a questionnaire and giving it to people who are supposed to be influenced by their local environment. For example, if the relationship between BMI and number of restaurants has been found, the people from "influence group" should be asked if they visit their neighbouring restaurants and how often they are doing it. By analysing the answers, we can find out if the relationships are real.

Including all main locations (house + work place/school) in the analysis can also be helpful. None from the investigated studies included them both at once. They were only investigated separately.

Future studies could also use additional information. The knowledge where people do their shopping would be helpful. It can be obtained via questionnaires or via GPS tracking (e.g. by a mobile App). Studies may use mobile applications to track and save the movement of the respondent, which will help to investigate which retailers were visited. So far, only one from the investigated studies (Shearer et al., 2014) used GPS trackers in the investigation. The results were compared with home-based measurements and the conclusion was that traditional home-based approaches overestimate the importance of the neighbourhood food environment. It was also concluded that these measures provide only modest evidence of linkages between the food environment beyond the residential neighbourhood boundary and dietary intake. GPS tracking measures were more accurate than other GIS methods because the exact shops encountered by study participants were identified. Additionally, the research was not limited to the neighbourhood. Summarizing, use of GPS tracking can be a beneficial tool in a study on diet and food environment relation.

Investigating the purchasing behaviour by collecting the shopping bills of the study participants is also worth considering. It can help to investigate what shops the people are visiting and what they are buying. The information about the locations of the places where they are buying food would help to answer the question about if people are purchasing food in their neighbourhood or somewhere else (e.g. on their way back from work). Additionally, it could help to investigate if and how the choice of the shop affects the food choices. This method may partly replace GPS tracking because it would reveal the location of shops visited by study participants. The problem is that it would be time consuming (collecting bills and typing in their content).

Because this is a first study analysing this problem in the Netherlands, it should be repeated (in a different city, possibly with bigger group of people) to have a better insight into the problem. If similar associations will be found, our understanding of the relation between local food environment and diet/weigh will be better.

The results of this study suggest that the spatial dimensions of the local food environment have an impact on diet. However, we cannot assess yet the importance of this impact. This study is a first start, and may support future research because it investigated where and how to obtain geo-data and the best approaches to process these.

6 References

Algert, S. J., Agrawal, A., and Lewis, D. S. (2006). Disparities in access to fresh produce in low-income neighborhoods in Los Angeles. *Am J Prev Med* **30**, 365-370.

Apparicio, P., Cloutier, M. S., and Shearmur, R. (2007). The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. *Int J Health Geogr* **6**, 4.

Banwell, C., Hinde, S., Dixon, J., and Sibthorpe, B. (2005). Reflections on expert consensus: A case study of the social trends contributing to obesity. *European Journal of Public Health* **15**, 564-568.

Berge, J. M., Wall, M., Larson, N., Forsyth, A., Bauer, K. W., and Neumark-Sztainer, D. (2014). Youth dietary intake and weight status: healthful neighborhood food environments enhance the protective role of supportive family home environments. *Health Place* **26**, 69-77.

Block, J. P., Christakis, N. A., O'Malley, A. J., and Subramanian, S. V. (2011). Proximity to food establishments and body mass index in the Framingham Heart Study offspring cohort over 30 years. *Am J Epidemiol* **174**, 1108-1114.

Bodor, J. N., Rose, D., Farley, T. A., Swalm, C., and Scott, S. K. (2008). Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutr* **11**, 413-420.

Boone-Heinonen, J., Gordon-Larsen, P., Kiefe, C. I., Shikany, J. M., Lewis, C. E., and Popkin, B. M. (2011). Fast food restaurants and food stores: longitudinal associations with diet in young to middle-aged adults: the CARDIA study. *Arch Intern Med* **171**, 1162-1170.

Booth, S. L., Sallis, J. F., Ritenbaugh, C., Hill, J. O., Birch, L. L., Frank, L. D., Glanz, K., Himmelgreen, D. A., Mudd, M., Popkin, B. A., Rickard, K. A., Jeor, S. S., and Hays, N. P. (2001). Environmental and Societal Factors Affect Food Choice and Physical Activity: Rationale, Influences, and Leverage Points. *Nutrition Reviews* **59**.

Botto, L. D., Mulinare, J., and Erickson, J. D. (2000). Occurrence of congenital heart defects in relation to maternal multivitamin use. *American Journal of Epidemiology* **151**, 878-884.

Buck, C., Börnhorst, C., Pohlabeln, H., Huybrechts, I., Pala, V., Reisch, L., and Pigeot, I. (2013). Clustering of unhealthy food around German schools and its influence on dietary behavior in school children: a pilot study. *International Journal of Behavioral Nutrition and Physical Activity* **10**.

Burgoine, T., Alvanides, S., and Lake, A. A. (2011). Assessing the obesogenic environment of North East England. *Health Place* **17**, 738-747.

Canada, H. (2013). Measuring the food environment in Canada. *Health Canada*.

Casagrande, S. S., Franco, M., Gittelsohn, J., Zonderman, A. B., Evans, M. K., Fanelli Kuczmarski, M., and Gary-Webb, T. L. (2011). Healthy food availability and the association with BMI in Baltimore, Maryland. *Public Health Nutr* **14**, 1001-1007.

Caspi, C. E., Sorensen, G., Subramanian, S. V., and Kawachi, I. (2012). The local food environment and diet: a systematic review. *Health Place* **18**, 1172-1187.

Cerin, E., Frank, L. D., Sallis, J. F., Saelens, B. E., Conway, T. L., Chapman, J. E., and Glanz, K. (2011). From neighborhood design and food options to residents' weight status. *Appetite* **56**, 693-703.

Cetateanu, A., and Jones, A. (2014). Understanding the relationship between food environments, deprivation and childhood overweight and obesity: evidence from a cross sectional England-wide study. *Health Place* **27**, 68-76.

Charreire, H., Casey, R., Salze, P., Simon, C., Chaix, B., Banos, A., Badariotti, D., Weber, C., and Oppert, J. M. (2010). Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutr* **13**, 1773-1785.

Clark, E. M., Quigg, R., Wong, J. E., Richards, R., Black, K. E., and Skidmore, P. M. (2014). Is the food environment surrounding schools associated with the diet quality of adolescents in Otago, New Zealand? *Health Place* **30**, 78-85.

Clarke, G. E., H.; Guy, C. (2002). Deriving Indicators of Access to Food Retail Provision in British Cities: Studies of Cardiff, Leeds and Bradford.

Cummins, S., and Macintyre, S. (1999). The location of food stores in urban areas: a case study in Glasgow. *British Food Journal* **101**, 545-553.

Donkin, A. J. M., Dowler, E. A., Stevenson, S. J., and Turner, S. A. (1999). Mapping access to food at a local level. *British Food Journal* **101**, 554-564.

Frank, L. G., Karen; McCarron, Margaret,; Sallis, James; Sealens, Brian; Chapman, James (2006). The Spatial Distribution of Food Outlet Type and Quality around Schools in Differing Built Environment and Demographic Contexts. *Berkeley Planning Journal* **19**.

Fraser, L. K., Clarke, G. P., Cade, J. E., and Edwards, K. L. (2012). Fast food and obesity: a spatial analysis in a large United Kingdom population of children aged 13-15. *Am J Prev Med* **42**, e77-e85.

Fraser, L. K., and Edwards, K. L. (2010). The association between the geography of fast food outlets and childhood obesity rates in Leeds, UK. *Health Place* **16**, 1124-1128.

Gibson, D. M. (2011). The Neighborhood Food Environment and Adult Weight Status: Estimates From Longitudinal Data. *American Journal of Public Health*.

Glanz, K., Sallis, J. F., Saelens, B. E., and Frank, L. D. (2005). Healthy nutrition environments: Concepts and measures. *American Journal of Health Promotion* **19**, 330-333.

Gustafson, A. A., Sharkey, J., Samuel-Hodge, C. D., Jones-Smith, J., Folds, M. C., Cai, J., and Ammerman, A. S. (2011). Perceived and objective measures of the food store environment and the association with weight and diet among low-income women in North Carolina. *Public Health Nutr* **14**, 1032-1038.

Inagami, S., Cohen, D. A., Brown, A. F., and Asch, S. M. (2009). Body mass index, neighborhood fast food and restaurant concentration, and car ownership. *J Urban Health* **86**, 683-695.

Inagami, S., Cohen, D. A., Finch, B. K., and Asch, S. M. (2006). You are where you shop: grocery store locations, weight, and neighborhoods. *Am J Prev Med* **31**, 10-17.

Jago, R., Baranowski, T., Baranowski, J. C., Cullen, K. W., and Thompson, D. (2007). Distance to food stores & adolescent male fruit and vegetable consumption: mediation effects. *Int J Behav Nutr Phys Act* **4**, 35.

Janevic, T., Borrell, L. N., Savitz, D. A., Herring, A. H., and Rundle, A. (2010). Neighbourhood food environment and gestational diabetes in New York City. *Paediatric and Perinatal Epidemiology* **24**, 249-254.

Jennings, A., Welch, A., Jones, A. P., Harrison, F., Bentham, G., van Sluijs, E. M., Griffin, S. J., and Cassidy, A. (2011). Local food outlets, weight status, and dietary intake: associations in children aged 9-10 years. *Am J Prev Med* **40**, 405-410.

Kloog, I., Haim, A., and Portnov, B. A. (2009). Using kernel density function as an urban analysis tool: Investigating the association between nightlight exposure and the incidence of breast cancer in Haifa, Israel. *Computers, Environment and Urban Systems* **33**, 55-63.

Kosti, R. I., and Panagiotakos, D. B. (2006). The epidemic of obesity in children and adolescents in the world. *Cent Eur J Public Health* **14**, 151-159.

Laraia, B. A., Siega-Riz, A. M., Kaufman, J. S., and Jones, S. J. (2004). Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Prev Med* **39**, 869-875.

Larsen, K., Cook, B., Stone, M. R., and Faulkner, G. E. (2015). Food access and children's BMI in Toronto, Ontario: assessing how the food environment relates to overweight and obesity. *Int J Public Health* **60**, 69-77.

Laska, M. N., Graham, D. J., Moe, S. G., and Van Riper, D. (2010). Young adult eating and food-purchasing patterns food store location and residential proximity. *Am J Prev Med* **39**, 464-467.

Latham, J., and Moffat, T. (2007). Determinants of variation in food cost and availability in two socioeconomically contrasting neighbourhoods of Hamilton, Ontario, Canada. *Health Place* **13**, 273-287.

LeDoux, T. F., and Vojnovic, I. (2014). Examining the role between the residential neighborhood food environment and diet among low-income households in Detroit, Michigan. *Applied Geography* **55**, 9-18.

Lee, L. v. (2014). The Dutch Healthy Diet index: Development, Evaluation, and Application.

McKinnon, R. A., Reedy, J., Morrissette, M. A., Lytle, L. A., and Yaroch, A. L. (2009). Measures of the Food Environment. A Compilation of the Literature, 1990-2007. *American Journal of Preventive Medicine* **36**, S124-S133.

Mehta, N. K., and Chang, V. W. (2008). Weight status and restaurant availability a multilevel analysis. *Am J Prev Med* **34**, 127-133.

Michimi, A., and Wimberly, M. C. (2010). Associations of supermarket accessibility with obesity and fruit and vegetable consumption in the conterminous United States. *Int J Health Geogr* **9**, 49.

Moore, L. V., Diez Roux, A. V., and Brines, S. (2008). Comparing Perception-Based and Geographic Information System (GIS)-based characterizations of the local food environment. *J Urban Health* **85**, 206-216.

Morland, K., Diez Roux, A. V., and Wing, S. (2002). The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study.

Morland, K., Diez Roux, A. V., and Wing, S. (2006). Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. *Am J Prev Med* **30**, 333-339.

Morland, K. B., and Evenson, K. R. (2009). Obesity prevalence and the local food environment. *Health Place* **15**, 491-495.

Murakami, K., Sasaki, S., Takahashi, Y., Uenishi, K., Japan Dietetic Students' Study for, N., and Biomarkers, G. (2009). Neighborhood food store availability in relation to food intake in young Japanese women. *Nutrition* **25**, 640-646.

Pearce, J., Blakely, T., Witten, K., and Bartie, P. (2007). Neighborhood deprivation and access to fast-food retailing: a national study. *Am J Prev Med* **32**, 375-382.

Pearce, J., Hiscock, R., Blakely, T., and Witten, K. (2008). The contextual effects of neighbourhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. *J Epidemiol Community Health* **62**, 198-201.

Pearce, J., Witten, K., and Bartie, P. (2006). Neighbourhoods and health: a GIS approach to measuring community resource accessibility. *J Epidemiol Community Health* **60**, 389-395.

Pearson, T., Russell, J., Campbell, M. J., and Barker, M. E. (2005). Do 'food deserts' influence fruit and vegetable consumption?--A cross-sectional study. *Appetite* **45**, 195-197.

Penchansky, R., and Thomas, J. W. (1981). The concept of access. Definition and relationship to consumer satisfaction. *Medical Care* **19**, 127-140.

Seliske, L. M., Pickett, W., Boyce, W. F., and Janssen, I. (2009). Association between the food retail environment surrounding schools and overweight in Canadian youth. *Public Health Nutr* **12**, 1384-1391.

Shearer, C., Rainham, D., Blanchard, C., Dummer, T., Lyons, R., and Kirk, S. (2014). Measuring food availability and accessibility among adolescents: Moving beyond the neighbourhood boundary. *Soc Sci Med*.

Smith, D., Cummins, S., Clark, C., and Stansfeld, S. (2013). Does the local food environment around schools affect diet? Longitudinal associations in adolescents attending secondary schools in East London. *BMC Public Health* **13**, 70.

Spence, J. C., Cutumisu, N., Edwards, J., Raine, K. D., and Smoyer-Tomic, K. (2009). Relation between local food environments and obesity among adults. *BMC Public Health* **9**, 192.

Sturm, R., and Datar, A. (2005). Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health* **119**, 1059-68.

Swinburn, B. A., Metcalf, P. A., and Ley, S. J. (2001). Long-term (5-year) effects of a reduced-fat diet intervention in individuals with glucose intolerance. *Diabetes Care* **24**, 619-624.

Taubes, G. (1997). New study says low-fat diet can lower blood pressure. *Science* **276**, 350.

Timperio, A., Ball, K., Roberts, R., Campbell, K., Andrianopoulos, N., and Crawford, D. (2008). Children's fruit and vegetable intake: Associations with the neighbourhood food environment. *Preventive Medicine* **46**, 331-335.

Willett, W. C., and Trichopoulos, D. (1996). Nutrition and cancer: a summary of the evidence. *Cancer Causes Control* **7**, 178-180.

Yen, P. K. (1998). Stopping heart disease with diet. *Geriatric nursing (New York, N.Y.)* **19**, 50-51.

Zenk, S. N., Schulz, A. J., Israel, B. A., James, S. A., Bao, S., and Wilson, M. L. (2005). Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *Am J Public Health* **95**, 660-667.

7 Appendix I: Contents of DVD

- Scripts: obtaining google data, obtaining eet.nu data, changing addresses into coordinates
- Shapefiles:
  - Google food retailers,
  - eet.nu restaurants and fast foods,
  - OSM food retailers,
  - CBS neighbourhoods.
- Maps
- Presentations (midterm, final)
- Report (.docx, .pdf)

8 Appendix II: Results from reviewed studies

Study	Main Findings
Laraia et al., 2004	• significantly lower mean DQI-P score for women living greater than 4 miles from a supermarket; • a significant decreasing trend in mean DQI-P with increasing distance from a convenience store; • location and proximity of food retailers are significantly associated with overall composite dietary score of pregnant women
Pearson et al., 2005	• distance to nearest supermarket and potential difficulties with grocery shopping were not significantly associated with either fruit or vegetable consumption
Sturm and Datar, 2005	• the number of fast-food restaurants in the neighbourhood was significantly associated with BMI gain
Inagami et al., 2006	• those who own cars and travel farther to their grocery stores also have higher BMI; • individual-level demographic characteristics were associated with variability in BMI; • college education was associated with lower BMI; the better predictor of BMI was not the individual's specific choice of the grocery but the location of where the average resident shopped
Morland et al., 2006	• the presence of supermarkets was associated with a lower prevalence of obesity and overweight; • the presence of convenience stores was associated with a higher prevalence of obesity and overweight; the presence of grocery stores was positively associated with the prevalence of overweight, obesity, diabetes and

	hypertension; • the presence of convenience stores was also associated with an increased prevalence of overweight, obesity and hypertension
Jago et al.,2007	• distance to fast food restaurants were associated with fruit and vegetable consumption among male adolescents; • distance to the nearest small food store was positively associated with high fat vegetable consumption; • distance to the nearest fast food restaurant was negatively associated with high fat vegetables consumption; • fruit and vegetable consumption was inversely associated with access to small stores
Bodor et al.,2008	• greater fresh vegetable availability within 100m of a residence was a positive predictor of vegetable intake; • having a small food store within this same distance was a marginal predictor of fruit consumption; access to a small food store within 100m of the residence was marginally associated with an increased fruit intake; • no association was found between intake and access to supermarkets, which differs with prior research in this area.
Pearce et al., 2008	• the consumption of the recommended daily intake of fruit was not associated with living in a neighbourhood with better access to supermarkets or convenience stores; • access to supermarkets was not related to vegetable intake; • individuals in the quartile of neighbourhoods with the best access to convenience stores had 25% lower odds of eating the recommended vegetable intake compared to individuals in the base category (worst access)
Mehta et al., 2008	• fast-food restaurant density and a higher ratio of fast-food to full-service restaurants were associated with higher individual-level weight status; • a higher density of full-service restaurants was associated with lower weight status
Inagami et al., 2009	• a high concentration of local restaurants is associated with BMI; car owners have higher BMIs than non-car owners; • individuals who do not own cars and reside in areas with a high concentration of fast food outlets have higher BMIs than non-car owners who live in areas with no fast food outlets, higher restaurant density is associated with higher BMI among local residents; • the local fast food environment has a stronger association with BMI for local residents who do not have access to cars.

Morland and Evenson, 2009	• the prevalence of obesity was lower in areas that had supermarkets and higher in area with small grocery stores or fast food restaurants
Murakami et al., 2009	• neighbourhood store availability for confectioneries and bread was significantly positively associated with the intake of confectioneries and bread
Seliske et al., 2009	• none of the individual food retailers was associated with an increased likelihood of overweight; • at 1 km, students attending schools with at least one food retailer had a lower relative odds of overweigh; • at 5 km, students attending schools with the highest exposure to the total food retailer index had a lower relative odds of overweight compared with students attending schools with no exposure; • exposure to various types of food retailers in school neighbourhoods was not associated with an increased likelihood of overweight in Canadian school-aged youth
Spence et al., 2009	• the lower the ratio of fast-food restaurants and convenience stores to grocery stores and produce vendors near people's homes, the lower the odds of being obese; • RFEI within 800m of the home was negatively associated with obesity prevalence
Fraser and Edwards, 2010	• a higher density of fast food outlets was significantly associated with the child being obese (or overweight/obese); • there is also a significant association between fast food outlet density and areas of higher deprivation.
Janevic et al., 2010	• no association between food environment measures and gestational diabetes was found; • a significant association between healthy food outlets and a crude measure of obesity but not gestational diabetes was found; • association between the lack on healthy food outlets and pre-pregnancy weight >200lbs was found, while there was no association for unhealthy food outlets
Michimi and Wimberly, 2010	• in nonmetropolitan areas; distance to supermarket had no associations with obesity or F/V consumption for all supermarket size categories; • obesity prevalence increased and F/V consumption reased with increasing distance to supermarket in metropolitan areas, but not in nonmetropolitan areas.

Block et al., 2011	• each 1-km increase in distance to the closest fast-food restaurant was associated with a 0.11-unit decrease in BMI; • the authors did not find a consistent relation between access to fast-food restaurants and individual BMI
Boone-Heinonen et al.,2011	• fast food consumption was related to fast food availability in low-income respondents; • greater supermarket availability was generally unrelated to diet quality and fruit and vegetable intake and relationships between grocery store availability
Burgoine et al.,2011	• decreased residential density (increased obesogenicity) is associated with a generally increased risk of overweight, although with high significance only in areas of the very lowest residential density; • food availability was not found to be significantly associated with BMI (overweight or obesity)
Casagrande et al., 2011	• high availability of healthy foods was associated with significantly higher BMI compared with individuals living in neighbourhoods with low availability of healthy food after adjustment for demographic variables; • there was a positive association between the availability of healthy food and higher BMI among individuals living in predominantly white neighbourhoods
Cerin et al.,2011	• the healthful-food availability score for grocery stores was associated with walking for eating purposes in women but not in men; • residents living within 1 km from convenience stores showed positive associations of walking for eating purposes with number of grocery stores and retail density.
Gustafson et al., 2011	• individuals who lived in census tracts with a convenience store and a supercentre had higher odds of perceiving their neighbourhood high in availability of healthy foods than individuals with no store; • individuals with a supercentre in their census tract weighed more than individuals without one; those who lived in a census tract with a supercentre and a convenience store consumed fewer servings of fruits and vegetables
Jennings et al., 2011	• Availability of BMI-healthy outlets in neighborhoods was associated with lower body weight; neighborhood availability of BMI-unhealthy outlets was inversely associated with body weight; • unhealthy food intake was also associated with availability of BMI-unhealthy food outlets; features of the built environment relating to food purchasing opportunities are

	correlated with weight status in children.
Gibson, 2011	• for residents of urban areas, the neighborhood density of small grocery stores was positively and significantly related to obesity and BMI.; • for individuals who moved from a rural area to an urban area over a 2-year period, changes in neighborhood supermarket density, small grocery store density, and full-service restaurant density were significantly related to the change in BMI over that period
Fraser et al., 2012	• the consumption of fast food was associated with a higher BMI SD score; higher body fat percentage; and increased odds of being obese; • the relationship between the accessibility of outlets and consumption did vary over space, with some areas (more rural areas) showing that increased accessibility was associated with consumption, whereas in some urban areas increased accessibility was associated with lack of consumption
Buck et al., 2013	• food stores and fast food restaurants do not significantly cluster around schools; • the consumption of junk food in young children is not influenced by spatial availability of unhealthy food
Smith et al., 2013	• There were significant positive relationships between the distances travelled to grocers and healthy diet scores though effects were very small.; • significant negative relationships between proximity to takeaways and unhealthy diet scores also resulted in small parameter estimates; • no statistically significant relationships between the count of food outlets and diet scores; healthy diet scores are positively correlated with the minimum distance to grocery stores, within both 400 and 800 metre buffers; • unhealthy diet scores are negatively correlated with the median distance to takeaways within 400 m, and the minimum distances to grocers within 800 m and takeaways at both distances

Berge et al., 2014	• having a convenience store close was significantly associated with higher BMI z-score in adolescent girls and having a riskier neighborhood score was associated with higher BMI z-score for adolescent boys; • having a fast food restaurant close was significantly associated with higher fast food consumption in adolescent boys
Cetateanu and Jones, 2014	• a positive association between the density of unhealthy food outlets in a neighbourhood and the prevalence of overweight and obesity in children; • the prevalence of overweight and obesity was positively associated with deprivation, with a negative association with professional employment for all outcomes; • for the older children there remained a statistically significant positive trend between overweight and obesity and obesity and obesity and the number of both 'fast food' and 'other unhealthy' outlets
Clark et al., 2014	• both distance to and density of food outlets were associated with dietary quality in adolescents, a high density of certain food outlets such as cafes, restaurants, supermarkets and takeaways around schools, was associated with a higher DQI score in boys, every 100m increase in distance to the nearest food outlet of any type was associated with a decrease in DQI score for girls only, showing that having less access to local food outlets had a small negative effect on diet quality
Shearer et al., 2014	• there were no associations between home-based measures of availability and accessibility and dietary intake and only one for GPS-based measures, with greater distance to convenience stores associated with greater fruit and vegetable consumption
Larsen et al., 2015	• Living in an area with a higher density of healthy food outlets and in close proximity to a supermarket decreased the odds of being overweight or obese; • distance to the closest supermarket was significantly related to the odds of being overweight or obese, while the density was not significant; • as distance to the nearest supermarket increases, so too does the corresponding odds ratio. an extra kilometre increases the odds of being overweight or obese by nearly 1.5 times; • living in a neighbourhood with a higher density of healthy food retailers, lowers the odds of being overweight and obese, while proximity does not appear to be important; • living in an area with a higher density of healthy food

	outlets and in close proximity to a supermarket decreases the odds of being overweight or obese, independent of income or gender, while unhealthy food outlets do not appear to relate to body weight.

9 Appendix III: Indices explained

DHD index (Dutch Healthy Diet index):

The DHD index includes components on physical activity, vegetables, fruit, dietary fibre, saturated fatty acids, trans fatty acids, consumption occasions with acidic drinks and foods, sodium, and alcohol. Scores for each component range between 0 (no adherence) and 10 (complete adherence) points. The DHD-index was inversely associated with energy intake and positively associated with most micronutrient intakes when adjusted for energy intake. Range: 0-90

BMI (Body Mass Index) = mass in kg/(height in meters)^2

BMI	Weight Status
Below 18.5	Underweight
18.5 – 24.9	Normal
25.0 – 29.9	Overweight
30.0 and Above	Obese

DQI (Diet Quality Index): DQI scores encompass dietary variety, adequacy, moderation, and balance. Scores range from 0 to 100 with higher scores reflecting better diet quality.

DQI-P (Diet Quality Index for Pregnant women): The DQI-P was based on eight dietary characteristics - percentage of recommended servings per day of grains, vegetables, and fruits, percentage RDA of folate and iron, AI of calcium, percentage of calories from fat, and meal pattern score.

10 Appendix IV: Evaluation of validation points

Table 27 Validation Google Food Retailers

FID	placename	address	correct	incorrect
52	Vreemde Streken Eetwinkel	1e Kloostersteeg 3	X	
48	Restaurant O Mundo	5 Mei Plein 1	X	
85	Koekoekpizzapannenkoek	5 Mei Plein 13	X	
26	De Overkant B.V.	Bevrijdingsstraat 38	X	
67	Zeezicht	Bevrijdingsstraat 38		X
2	Limburgia Wageningen	Bevrijdingsstraat 48	X	
22	Jojo's Café	Bevrijdingsstraat 7	X	
72	Sphinx Pizzeria-Steakhouse	Bevrijdingsstraat 9	X	
78	Cafetaria Het Stekkie	Gerdesstraat 2	X	
55	Eeterij H'eerlijk	Heerenstraat 47	X	
18	Café Het Gat van Wageningen	Herenstraat 31	X	
28	Cafe De Zaaier	Herenstraat 33	X	
53	Eetcafe H41	Herenstraat 41	X	
65	Ali Baba	Herenstraat 43	X	
3	Bagels & Beans	Hoogstraat 10	X	
54	Restaurant 't Carillon	Hoogstraat 12	X	
46	Brasserie De Blije Boedha	Hoogstraat 13	X	
0	Brood- en Banketbakkerij van Voorthuizen	Hoogstraat 14	X	
16	De Vlaamsche Reus B.V.	Hoogstraat 21	X	
7	J.S. Cuisine De Keuken	Hoogstraat 5	X	
8	Kruidvat	Hoogstraat 65	X	
1	Bakker Bart	Hoogstraat 76	X	
69	Grand-café Suisse	Hoogstraat 88	X	

42	Thais restaurant My Asia Wageningen	Hoogstraat 9	X	
40	Subway	Hoogstraat 93	X	
47	Toko Radjawali	Junusstraat 19	X	
15	Columbus	Junusstraat 27-A	X	
31	Morning Tales (by Greek Food Tales)	Junusstraat 45	X	
76	Pizzeria Grillroom Cleopatra	Junusstraat 8	X	
74	Cafetaria 't Passantje	Kapelstraat	X	
23	Café 't Centrum 2011 Wageningen	Kapelstraat 2	X	
24	Villa Bloem espressobar	Kapelstraat 2-A	X	
86	Da Martini restaurant	Kapelstraat 3	X	
19	Side-Walk	Kapelstraat 9		X
49	Hof Van Wageningen Hotel En Congrescentrum	Lawickse Allee 9	X	
64	EatCetera	Lawickse Allee 9	X	
83	Jaap Venendaal Groep	Lawickse Allee 9	X	
35	Cafe de Korenbeurs	Markt 11-13	X	
14	Café De Tijd	Markt 12	X	
34	The Doctor	Markt 14	X	
57	Colors World Food	Markt 15	X	
13	Brasserie de Stad	Markt 16		X
66	Eetcafe Buurman & Buurman	Markt 18	X	
58	Steakhouse	Markt 2A	X	
32	Sportsbar De Malle Molen	Markt 2-H	X	
68	Turks Eethuis Ilayda	Markt 4	X	
17	eetcafe De Kater	Markt 8	X	
63	Restaurant Drinks and Bites	Markt 9	X	
37	Sixpack B.V.	Markt 9	X	
61	Het Oude Pakhuis	Molenstraat 4	X	

20	Café Loburg	Molenstraat 6	X	
51	Sa Lolla	Molenstraat 6	X	
36	Stichting Cultureel Café Wageningen	Nieuwstraat 12		X
84	Restaurant Ivory	Poststraat 8		X
29	Café XL	Riemsdijkstraat 6		X
21	Sjop '86 B.V.	Rouwenhofstraat 1-A	X	
6	Toko Indrani	Salverdaplein 2-A	X	
11	Maller-Chou V.O.F.	Salverdaplein 4		X
56	Restaurant Toledo	Schoolstraat 15	X	
30	Poolcafe Infinity	Stadsbrink 12	X	
60	Chinees-Mongools Restaurant King's Garden B.V.	Stadsbrink 1-M	x	
39	Domino's Pizza Wageningen	Stadsbrink 34	X	
4	Albert Heijn	Stadsbrink 375	X	
12	AH Stadsbrink	Stadsbrink 375	X	
5	Lidl	Stadsbrink 4-10	X	
9	Foladi Groenten & Fruit	Stadsbrink 443	X	
33	Cafetaria Eetsalon De Dubbeldekker	Stadsbrink 551	X	
25	Lunchcafe de Serre	Stationsstraat 7	X	
10	Paul en Maartje	Stationsstraat 70	X	
71	Jeruzalem	Veerstraat 5	X	
27	CafeDaniels	Vijzelstraat 10	X	
70	Cafe-Carre Eten & Drinken B.V.	Vijzelstraat 2	X	

Existing places not included in Google points:

1. Zoetwaren
2. HEMA
3. Flavours
4. Heroes of Taste
5. Zuivelhoekje

6. Lazuur
7. Watami
8. De Hoek
9. Penny Lane
10. Taste
11. Ijssalon Antonio
12. Le Perron
13. Shashima Palace Lounge
14. Tante uit Marokko
15. Florissant
16. Ijssalon Cicuto
17. De Urker Visspecialist
18. Zamzam

Table 28 Validation of Eet.nu points

FID	name	street	correct	incorrect
5	Vreemde Streken	1e Kloostersteeg 3	X	
16	O Mundo	5 Mei Plein 1	X	
33	Koekoekpizzapannenkoek	5 Mei Plein 13	X	
19	IJssalon Antonio	Bevrijdingsstraat 48	X	
30	Sphinx	Bevrijdingsstraat 9	X	
38	Cafetaria Het Stekkie	Gerdesstraat 2	X	
6	Eetcafé H 41	Herenstraat 41	X	
22	Ali Baba	Herenstraat 43	X	
2	Eeterij H'eerlijk	Herenstraat 47	X	
32	Bagels & Beans	Hoogstraat 10	X	
15	't Carillon	Hoogstraat 12	X	
14	Croissanterie Pepain	Hoogstraat 13	X	
12	Hema Lunchroom	Hoogstraat 59	X	
3	Broodje Bram	Hoogstraat 72		X
9	Bakker Bart	Hoogstraat 76	X	
31	Grand Café Suisse	Hoogstraat 88	X	
40	My Asia	Hoogstraat 9	X	

36	Subway	Hoogstraat 93	X	
20	Toko Radjawali	Junusstraat 19	X	
25	Cleopatra	Junusstraat 8	X	
37	Cafetaria Passantje	Kapelstraat 12	X	
26	Da Martini	Kapelstraat 3	X	
23	Bij de Buuren	Markt 11		X
11	De Tijd	Markt 12	X	
21	Colors World Food	Markt 15	X	
10	Brasserie de Stad	Markt 16		X
4	Eetcafe Buurman & Buurman	Markt 18	X	
17	't Steakhouse	Markt 2a	X	
42	Ilayda	Markt 4	X	
8	Eetcafé de Kater	Markt 8	X	
24	Drinks & Bites	Markt 9	X	
0	Het Oude Pakhuis	Molenstraat 4	X	
28	Sa Lolla	Molenstraat 6	X	
41	Toledo	Schoolstraat 15	X	
18	IJssalon Cicuto	Schuylensteeg 5	X	
34	De Hoek	Stadsbrink 1h	X	
1	King's Garden	Stadsbrink 1m	X	
29	Watami	Stadsbrink 2	X	
35	Domino's Pizza	Stadsbrink 34	X	
39	Snackpoint De Dubbeldekker	Stadsbrink 551	x	
13	De Serre	Stationsstraat 7	X	
27	Jeruzalem	Veer 5	X	
7	Café-Carré Eten & Drinken	Vijzelstraat 2	X	

Existing places not included in eet.nu points:

1. Morning Tales
2. Flavours
3. Florissant

11 Appendix V: Multiple regression analysis (no significant) results

Euclidean distance proximity

	DHD *incl. sex, age, moderate activity per day* *n = 1344*				
distance to:	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurant*	-0,9495	0,0938	4,21%	3,92%	3,49%
supermarket*	-0,2274	0,6827	4,02%	3,73%	3,33%
bakeries	0,0003	0,6312	4,02%	3,74%	3,33%
cafés	-0,0001	0,9137	4,01%	3,72%	3,31%
bars	0,0000	0,9381	4,01%	3,72%	3,29%
convenience store	0,0001	0,6730	4,02%	3,73%	3,30%
groceries or supermarkets	0,0003	0,5934	4,03%	3,74%	3,33%
meal delivery centres	0,0001	0,5200	4,04%	3,75%	3,32%
restaurants	0,0000	0,9756	4,02%	3,72%	3,27%
takeaway	0,0003	0,3970	4,06%	3,77%	3,34%
restaurant**	-0,0010	0,2850	4,09%	3,80%	3,36%
snack bars**	0,0008	0,2898	4,09%	3,80%	3,32%
* CBS					
** eet.nu					

Network distance proximity

	DHD *incl. sex, age, moderate activity per day* *n = 1344*				
distance to:	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurant*	-0,8761	0,1001	4,21%	4,01%	3,65%
supermarket*	-0,4559	0,3719	4,08%	3,88%	3,55%
bakeries	0,0002	0,6100	4,04%	3,84%	3,50%
cafés	-0,0002	0,7044	4,04%	3,83%	3,49%
bars	0,0000	0,9422	4,03%	3,82%	3,47%
retailers	-0,0002	0,7850	4,03%	3,83%	3,46%
groceries or supermarkets	0,0004	0,4010	4,08%	3,87%	3,53%
restaurants	0,0003	0,7075	4,04%	3,83%	3,47%
takeaway	0,0002	0,5773	4,05%	3,84%	3,49%
restaurant**	-0,0008	0,2611	4,11%	3,91%	3,54%
snack bars**	0,0005	0,3344	4,09%	3,88%	3,48%
* CBS					
** eet.nu					

Neighbourhood density

density of:	DHD *incl. sex, age and income per citizen* *n = 1344*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurant*	**-0,0036**	0,7382	**4,03%**	3,83%	3,43%
supermarket*	0,0352	0,5943	4,05%	3,84%	3,48%
bakeries	-0,1276	0,5977	4,05%	3,84%	3,48%
cafés	-0,0235	0,8420	4,03%	3,82%	3,46%
bars	-0,2112	0,8145	4,03%	3,82%	3,47%
convenience store	0,0400	0,9750	4,03%	3,82%	3,44%
groceries or supermarkets	-0,1223	0,4284	4,07%	3,86%	3,48%
meal delivery centres	-0,7579	0,1291	4,19%	3,98%	3,63%
restaurants	-0,0165	0,7232	4,04%	3,83%	3,47%
takeaway	-0,6751	0,1530	4,17%	3,96%	3,58%
restaurant**	-0,0144	0,7203	4,04%	3,83%	3,47%
snack bars**	-0,3858	0,0535	4,28%	4,08%	3,73%
* CBS					
** eet.nu					

Network buffer density

density of:	BMI incl. sex, age, education and income n = 1868				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurants*	-0,0038	0,3537	5,80%	5,55%	5,20%
supermarkets*	-0,0259	0,3100	5,81%	5,56%	5,19%
bakeries 1600m	-0,0446	0,1525	5,86%	5,61%	5,26%
bakeries 800m	-0,0355	0,5935	5,77%	5,52%	5,15%
cafés 1600m	-0,0104	0,1022	5,89%	5,64%	5,31%
cafés 800m	-0,0147	0,3774	5,80%	5,54%	5,18%
bars 1600m	-0,0769	0,0548	5,94%	5,69%	5,37%
bars 800m	-0,1827	0,0643	5,93%	5,68%	5,34%
groceries or supermarkets 1600m	-0,0231	0,1488	5,86%	5,61%	5,26%
groceries or supermarkets 800m	-0,0042	0,9053	5,76%	5,50%	5,13%
meal delivery centres 1600m	-0,0352	0,4619	5,78%	5,53%	5,17%
meal delivery centres 800m	-0,0154	0,8651	5,76%	5,51%	5,14%
restaurants 1600m	-0,0057	0,0797	5,91%	5,66%	5,32%
restaurants 800m	-0,0060	0,4459	5,79%	5,53%	5,17%
takeaway 1600m	-0,0081	0,8363	5,76%	5,51%	5,15%
takeaway 800m	0,0225	0,7800	5,76%	5,51%	5,14%
restaurant 1600m**	-0,0048	0,1105	5,89%	5,63%	5,29%
restaurant** 800m	-0,0044	0,5190	5,78%	5,52%	5,16%
snack bars 1600m**	-0,0207	0,3360	5,80%	5,55%	5,19%
snack bars** 800m	0,0113	0,8171	5,76%	5,51%	5,12%

density of:	DHD *incl. sex and age* *n = 1344*				
	coef	p-value	R-sq	R-sq (adj)	R-sq (pred)
restaurants*	-0,0036	0,7382	4,03%	3,83%	3,43%
supermarkets*	0,0352	0,5943	4,05%	3,84%	3,48%
bakeries 1600m	-0,0735	0,3620	4,08%	3,88%	3,47%
bakeries 800m	-0,1082	0,5461	4,05%	3,85%	3,46%
cafés 1600m	-0,0085	0,5939	4,05%	3,84%	3,39%
cafés 800m	-0,0205	0,6111	4,04%	3,84%	3,36%
bars 1600m	-0,0728	0,4644	4,06%	3,86%	3,40%
bars 800m	-0,1582	0,5014	4,06%	3,85%	3,30%
groceries or supermarkets 1600m	0,0073	0,8572	4,03%	3,82%	3,41%
groceries or supermarkets 800m	-0,0045	0,9612	4,03%	3,82%	3,43%
meal delivery centres 1600m	0,0819	0,5190	4,06%	3,85%	3,49%
meal delivery centres 800m	-0,1372	0,5880	4,05%	3,84%	3,47%
restaurants 1600m	-0,0007	0,9364	4,03%	3,82%	3,40%
restaurants 800m	-0,0110	0,5940	4,05%	3,84%	3,40%
takeaway 1600m	-0,0244	0,8102	4,03%	3,82%	3,43%
takeaway 800m	-0,1664	0,4377	4,07%	3,86%	3,43%
restaurant 1600m**	-0,0002	0,9834	4,03%	3,82%	3,41%
restaurant** 800m	-0,0057	0,7561	4,03%	3,83%	3,39%
snack bars 1600m**	-0,0046	0,9343	4,03%	3,82%	3,44%
snack bars** 800m	-0,0959	0,4567	4,07%	3,86%	3,48%