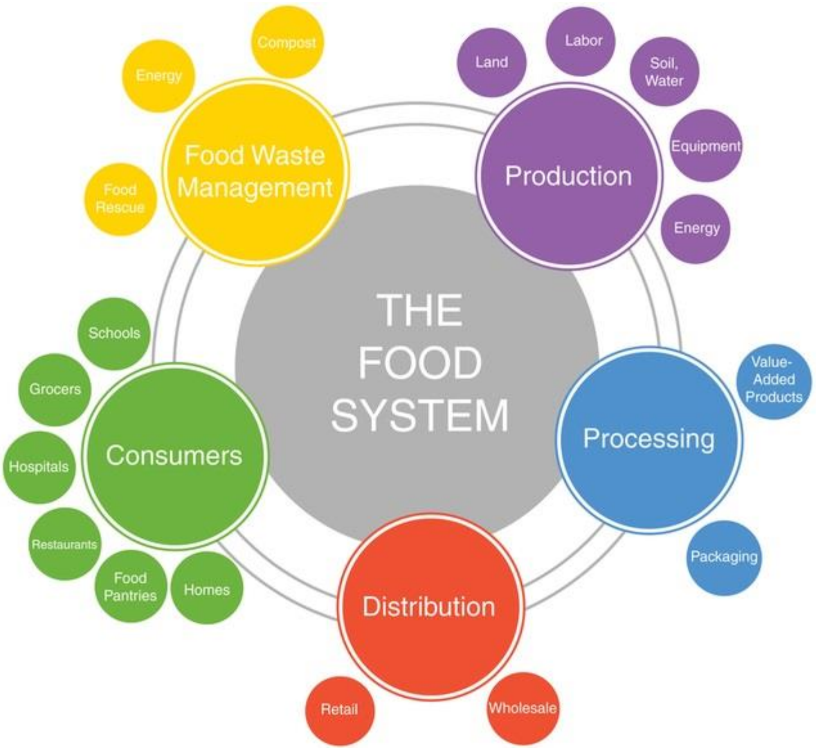
Geo-information Science and Remote Sensing

Thesis Report GIRS-2018-21

The Food Supply Chain and Public Transport Usages

Can the public transport network be used to optimize the food supply chain?

Koen R. Veenenbos



May 2018

WAGENINGEN UNIVERSITY & RESEARCH

The Food Supply Chain and Public Transport Usages

"Can the public transport network be used to optimize the food supply chain?"

Koen R. Veenenbos

Registration number: 951001856090

<u>Supervisors:</u>

Ron van Lammeren
Dirk Wascher

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre,

The Netherlands.

May 2018
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2018-21
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgement

This report was written for the Master Thesis Geo-Information Science. A part of the data was delivered by The Leibniz Centre for Agricultural Landscape Research (ZALF). Therefore I want to thank Dr. Ingo Zasada for his cooperation.
I would like to thank my first supervisor, Ron van Lammeren for his supervision and feedback throughout the entire thesis period, and my second supervisor, Dirk Wascher, for his input and feedback on the report. Furthermore, I want to thank Aldo Bergsma for helping me out with some ArcGIS problems and John Stuiver for the sessions needed to create a complete and correct network for the network analysis.

Abstract:

The world population is growing at a fast pace. The United Nations predict that the world population will be raised to 8.6 billion in 2030 and even 9.8 billion in 2050. With this enormous growth food problems will appear more often. Providing enough food at the right time and on the right location is one of the most difficult tasks in a food system.

This thesis focusses on improving the food supply chain using current public transport networks. The five largest supermarket chains (Aldi, Lidl, Netto, Rewe, and Edeka) in the area around Berlin were used as study case. The gathered data was all open source and came from Open Street Map and governmental institutions. First, the food demand per region was calculated by using the average food demand per person per day multiplied by the number of inhabitants of an assessment test area. Second, this regional food demand was spread out over the supermarkets located in corresponding area by using the number of square metres of the supermarkets. With an extended network analysis, the delivery routes from distribution centre to supermarket were determined. This resulted in the number of rides needed per supermarket chain. To deliver all the food, while only using regular trucks, 508 rides from the distribution centres must be made.

Afterwards, the same strategy was carried out, but this time bus routes were used as delivery routes instead of normal roads. Also, the capacity of the delivery vehicle was changed, because the public transport vehicle has two purposes (person and food transport). In total 78% of the distribution centres and 72% of the supermarkets is found to be reachable while using bus routes. Because the lower capacity of the delivery by bus, it will take 1321 rides to deliver all the supplies. Nevertheless, these results show that almost 70% of the food can be delivered by using public transport routes. Future studies should be focussing on the design of the buses. The Vehicles need adaptation in order to optimally transport both freight and persons.

Overall, this thesis is a first step in investigating the possibilities of using the current public transport network to improve the food supply system. It shows that the public transport network has great potential in completing these two tasks.

Table of Contents

1. Introduction

This thesis is linked to the FOODMETRES project. FOODMETRES has been funded by the European Union (FOODMETRES, 2015) and continues to drive further research by means of several follow-up projects. The project has developed a string of new assessment methods for food planning and food chain typologies addressing innovation and sustainability at the level of metropolitan regions (Wascher, Kneafsey, Pintar, & Piorr, 2015). Within the food chain innovation, the FOODMETRES researchers identified three distinct food supply chains: Short Regional Food Chains, Long Regional Food Chains and Long Global Food Chains. This research focuses on the short regional food chain which includes regionally produced food.
Next to their description of the different food chains, the researchers used six metropolitan regions (Berlin, Ljubljana, London, Milan, Nairobi, Rotterdam) as case studies. All six metropolitan regions have their own characteristics. The study area around Nairobi, the only not European region, is perfect to investigate the growth of a city. This raises questions, such as, what is the effect on the food prices and the food security. Another study area is the metropolitan region Rotterdam. This area is especially interesting for research on the short food supply chain, because of the large area with greenhouses.
Using one of these regions is interesting for this thesis, because there is more (geo-)data available. Institutions from the FOODMETRES project, such as The Leibniz Centre for Agricultural Landscape Research (ZALF), are willing to help providing data and advising in food research.
Furthermore, this chapter will give general information about the topic, followed by the problem definition and the research questions. The second chapter describes the methodology used in the thesis. The explanation is divided in per research question. Successively the results, conclusions, discussion and recommendations are written. These are followed by the references and last the appendix.

1.1 Food Systems

Providing enough food at the right time and on the right location is a major challenge. For this purpose, extensive logistic networks have been established over the past decades. These networks are key to the food supply chains (Dani, n.d.) and part of food systems at large (Mundler & Criner, 2016). These chains are connecting the farmers and horticulturalists, food manufacturers, suppliers, purchasing companies, supermarkets, shoppers and the consumers with each other (Planbureau voor de Leefomgeving, 2014). Using the example of waste management, Figure 1 illustrates the different components of the food system and ways of closing the circles with the aim to decrease negative impacts such as food waste, pollution or CO2-emissions.

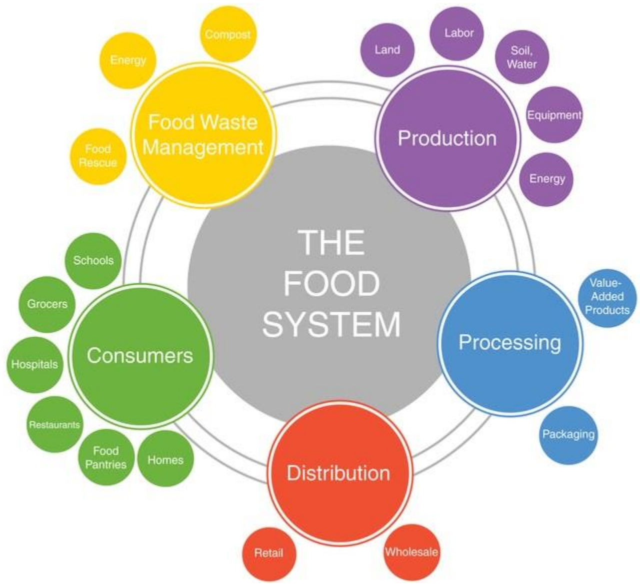


Figure 1. Closed food chain (Pon, 2012).

Key to the functioning of food systems are highly developed food supply chains linking food processing, distribution, and delivery to provide enough food at the right time on the right location. Pon's figure, Figure 1, shows an ideal situation of a food system.
The earlier mentioned farmers and horticulturalists form the basis of the food supply chain. At their farms, plantations and greenhouses the journey of the products starts. A common term for these food production locations is 'foodsheds'. Walter Hedden was one of the first authors who defined this term in 1929 (Butler & Butler, 2013). After Hedden, Arthur Getz reintroduced the term in 1991, when he argued that the entire world is our foodshed (Butler & Butler, 2013; Kloppenburg, Hendrickson, & Stevenson, 1996). Later the definition: *"A foodshed represents the land's resource that supports food production for a region or community: it is the geographic origin of a food supply."* was given by Blum-Evitts in her research how to conduct a regional foodshed assessment. (Blum-Evitts, 2009).

Additional to the term 'foodsheds' there is a term for the distance that the food travels from the start to end of the supply chain. These travelling distances are named 'food miles' by professor Timothy Land (Mundler & Criner, 2016). The definition given by the Cambridge dictionary is: "*Food miles are the distance between the place where food is grown or made and the place where it is eaten.*" (Cambridge, n.d.). Other ways of referring to the process are 'farm to fork' or 'field to table'.
In a previous study preformed in the US in 2001, researchers calculated the 'average distance' products travel to a market in Chicago (Pirog, Van Pelt, Enshayan, & Cook, 2001). For their calculations they used a 'Weighted Average Source Distance' (WASD). They find that the average distance of a product was 1245 miles in 1981. Afterwards they calculated the food miles for the market in 1998. The outcome showed that the average travel distance was grown with 22% to 1518 miles. Next to the market in Chicago they also did a research with data of three institutional markets in Iowa, which only sold local food. Here the average distance the food travels was 44.6 miles, which shows the large difference between local food and products gathered from all over the world. In Britain a similar study was performed in 2002. Researchers calculated the average distance that food travels to end up on a British plate is about 1500 miles.
In the past most of the food was produced locally, but over the years the prices of food production, processing and transport declined, which made it possible to import more food from further away. This eventually resulted in higher food miles (Mundler & Criner, 2016).

The whole journey from farmer to consumer can be seen as a large hourglass. Different visualisations have been made to show this phenomenon. Figure 2 for example shows a part of the supply chain of the United Kingdom, visualised by Tania Hurt-Newton. Though only schematic, the illustration shows the shape of the hourglass clearly: in the UK around 7000 suppliers deliver to four supermarket concepts with a market share of around 76%. Those supermarket concepts serve 25 million households (Nicholson & Young, 2012).

Figure 2. Hourglass of the UK food supply chain (Nicholson & Young, 2012).

Another more extended example of the hourglass is the visualisation of the food supply chain of the Netherlands. The Environmental Assessment Agency of the Netherlands published an infographic about the food chain (see Figure 3) in 2014. The 65000 farmers sell their products to the 6500 food manufactories. Then they trade with the 1500 suppliers. Next the 5 purchasing companies buy the goods and sell it to the 25 different supermarket concepts, which distribute it to their 4400 supermarkets. Lastly the 7 million shoppers do the groceries for the 16.7 million inhabitants of the Netherlands (Planbureau voor de Leefomgeving, 2014). The shape of the hourglasses gives the purchasing companies of the supermarket concepts a very strong position in deciding which prices should be used in their retails.
Besides showing the position of power, the shape of the hourglass also gives an indication where problems are most likely to occur when we look at the movement of the goods. Because all the products are moved through only a small number of distribution concepts, it is important that the logistic plans are organized

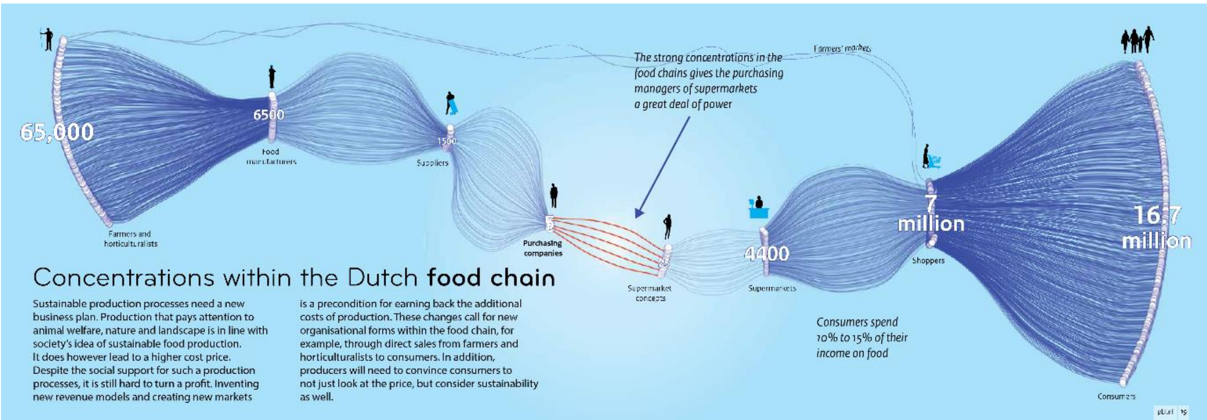


Figure 3. Example of the Dutch food supply chain (Planbureau voor de Leefomgeving, 2014).

well. These plans optimize the chosen locations of supermarkets, distribution centres and the optimal delivery route between them. When using a perfect distribution network, the impacts on the environment, such as CO2-emissions from the delivery trucks, stay at a minimum. Also, the influence on the traffic is minimalized because less delivery trucks are needed.

1.2 Problem Definition

The main problem underlying this research is the enormous pressure that arises on food supply chains. The pressure is caused by the still growing world population. By 2050, the world population has grown up to 9 billion inhabitants (King et al., 2017). The United Nations predict a world population of 8.6 billion for 2030 and even 9.8 billion for 2050 (United Nations, 2017). The population growth will lead to a 70% increase in food demand. All the extra food must be produced, distributed, consumed and recycled, which results in supply chains that are stretched to their maximum. If no improvements on the chains are made they could eventually collapse, which results in a food system that is unable to provide food on the right location on the right time.

Moreover, although the current system has proven to be capable of supplying all our food, it does so at the cost of environmental and social impacts which are externalized and not part of the product price. For example, more roads are constructed, or existing roads are broadened, while other transport modes are under-used. One of these under-used transport modes consist of the vehicles of the public transport. Therefore, this thesis will investigate the potential of using the existing public transport network to extend the food supply chain. The focus of the thesis will be on the connection between the distribution centres and the supermarkets. Especially in what way the public transport can be used for this connection to fulfil a double function? In Figure 1 this is mainly the red circles (distribution), but also the connection between the red and green (consumers) circles.

Public transport has been chosen as additional network, because of the low occupancy rate in some cities. The city of Helsinki showed in their annual report the under-used transport modes, see Figure 4. These number demonstrates the possibilities of using the public transport for more than only passenger transport. The occupancy rate is the average number of persons in a vehicle (car, bus, tram, metro, train) (European Environment Agency, 2008). The peak hours in the morning and afternoon are not usable for freight movement. The transport vehicles are just to full and extra freight cannot be added. This would only create new problems, such as extra stopping time for (un)loading. Nevertheless, besides the peak hours, large parts of the day the occupancy rate is very low (European Environment Agency, 2010). Especially the difference in occupancy rates of cities and provinces is large. In the cities centres, the transport vehicles are most of the time reasonably filled, but on the countryside or just outside the centre the public transport is often almost or completely empty (Vuurstaek et al., 2017). This has been mentioned in the news in different countries. In the Dutch newspaper NRC (Duursma, 2016) an article was written about the problem of empty buses in provinces and the vision on public transport of the government. The government declared in their vision that changes are needed. In another article in the American newspaper Statesman, Ben Wear stated that the number of daily bus users in Austin has dropped 20% from 2012 to 2016 (Ben Wear, 2016).

Next to the articles in the newspapers, the European Environment Agency (EEA) did a research in 2010. They compared the occupancy rate of public transport for several European countries. The outcomes of this research show large differences between countries (see Appendix 1). For example, the train occupancy rate of Poland was around 70% and declining, and in Norway the rate was only 32% and relatively steady (European Environment Agency, 2010).
Besides train occupancy rates they also compared the rates of long distance buses (see Appendix 2). They compared only the long distances buses because most of the data about short buses are not publicly available. The reason is the growing privatising of the bus services. Data of the long-distance buses shows that countries, such as Denmark, Spain and Italy, had a constant occupancy rate around 25%. The Czech Republic had with percentages between 60 and 70, the highest occupancy rate, but they also showed a declining trend.

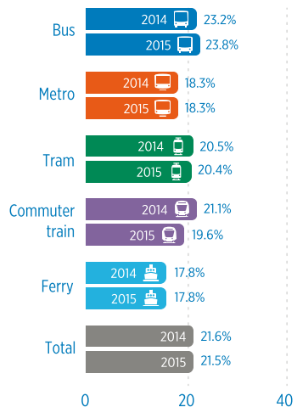


Figure 4. Occupancy rate Helsinki by type of transport (HSL Helsinki Region Transport, 2015).

More recently the city of Helsinki published the annual report 2015 about transport in the Helsinki region (HSL Helsinki Region Transport, 2015). In this report, among other things, they showed the occupancy rate of the different transport types in 2014 and 2015. Unfortunately, not all the annual reports, from the last years, contain these occupancy rates. This makes it difficult to determine a specific trend. What can be said is that all the different transport types have an occupancy rate between 17% and 24% (see Figure 4). The total percentage of both the years was around the 21%, which means that almost 80% of the public transport was empty. These numbers show the possibility of using the public transport network for other purposes.

The second reason for using the public transport as additional network to the food supply chain is the very consistent schedule that the transport vehicles have. Whether it is a train, tram, metro or bus all of them are following a frequent timetable. A research, executed in 2012 by the European Commission, showed that only five European countries score 90% or lower for the punctuality of local and regional trains (see Appendix 3). This includes all the trains that arrive with less than 5 minutes delay. The other 17 countries, which were publicised in the graph, had a punctuality of 92% up till 99% (Latvia).

The third benefit of using the existing public transport for food supply is that is relieves the need of lorries on the normal road network. Currently, most of the freights are transported over the road network (European Union, 2016). The food supply chain uses this road network for the distribution of the food, but the roads are very crowded. In his master thesis, Georges Fuchs examined the modal split of inland freight transports. He showed that 75% of inland freight movement was done on roads in 2003 (Fuchs, 2010). The remaining 25% is split in numbers around 18% by railways and 7% by inland waterways. The complete table with the data can be found in Appendix 4. When we compare this data with the published data from Eurostat (see Appendix 5), we see that these shares remained constant in the past years (Eurostat Statistics Explained, 2017). This means that, if all the extra food transport will be done by lorries that use the normal road network, this could create extra traffic jams. Besides the traffic jams, the lorries produce a large amount of CO_2 (Cefic, Care-Responsible, & ECTA, 2011) that is being expelled into the atmosphere.

By integrating food supply and public transport in the same network, the number of extra transport vehicles can be limited and with it the amount of negative impacts resulting from system inefficiency.

This study will focus only on movement by road. Other food transport modes such as movement by waterways, train or even by air are not taken into a count in this research. Waterways can be very useful delivery routes. In the Dutch cities Amsterdam and Utrecht this is already done for more than 10 years (Michiel Hegenaar, 2012). A boat delivers food and drinks to the restaurants and other boat takes care of the waste from the day before. This rules out the need for lorries on the small street in the city centre. However, due to the lack of waterways, not all cities are suitable for using water as supply routes. Therefore, while waterways can be an interesting transport mode for some cities, we decided to focus on the more general transport via road.

There are multiple ways to adapt the public transport vehicles to its new goal. First, it might be possible to adapt the design of buses. A newly designed should have enough seats for passengers and gets a special part where goods can be loaded. In the peak hours the part for the freight can be used as pitches for passengers. Outside the peak hours this part will be closed and used to place carts with food. A disadvantage of this method is that new vehicles need to be designed and build. Replacing old buses for new ones will cost a lot of time, organisation and money. Moreover, preparing the buses for peak hour and loading them with food supplies will be time-consuming. A benefit of this technique is that the vehicles can be designed as efficient as possible because they are specially build for this purpose.

A second option for combining freight transport and passenger movement, is by using trailers. These trailers are prepared at the distribution centres, which means that they only have to be attached to the buses. The advantage of this system is that it should not take too much time to attach and detach the trailer. This will save a lot of unloading time in comparisons with the first option, because only the trailer needs to be detached. The unloading of the trailer itself can be done afterwards and the empty trailer can be picked up by a new bus.

A disadvantage of this method is the that the size of the bus with trailer is very large, which is unhandy in crowded cities with small streets.
In Figure 5 an example of a bus with a trailer can be seen. These kinds of buses are already used in different cities around the world. The trailer can be used for transporting food, while using the first part for person transport.

Figure 5. Example of a bus with trailer ('Midi train MAN Göppel bus', n.d.).

1.3 Study Area

To model the possibilities of the current public transport via GIS, a case study had to be performed. Here-fore a study area had to be selected. For this thesis the state Berlin was chosen as study area. The state Berlin is one of the sixteen states of Germany and is surrounded by the state Brandenburg. Together they form the capitol region Berlin-Brandenburg. The state Berlin has more than 3.5 million inhabitants (Statistical Office of the Federation and States, 2018) and is still growing. Figure 6 shows the population development of Berlin-Brandenburg.
What stands out is the population growth in the areas around Berlin and the decline in almost all the other provinces of Brandenburg. This is a universal trend confirmed by the United Nations in their report: World Urbanization Prospects (Department of Economic and Social Affairs - United Nations, 2014). One of the key facts in the report, published in 2014, is that more than 50 per cent of the world's population is living in urban areas. Compared to 1950, when only 30 per cent of the populations was living in urban areas, this is enormous growth. The United Nations expect that in 2050 around 66 per cent of world's population is living in urban areas. Due to the process of urbanization, there will be a higher demand of food in the cities and consequently, a demand for a more efficient food supply system. The state Berlin thus is a great example of the global trend of growing cities and therefore chosen as study area.

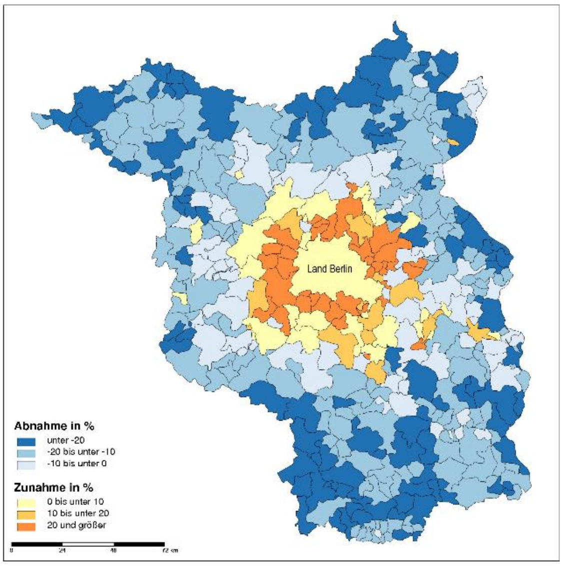


Figure 6. Population Development of Berlin-Brandenburg (State Office for Building and Transport, 2018).

Another reason to choose Berlin as study area, is that the state has been used as one of the six metropolitans in the FOODMETRES project. Before work on the thesis started, it seemed like this should gave an advantage in collecting data and background information. Unfortunately, this advantage was very small. Although the researchers of the FOODMETRES project were willing to help, to explain how they worked and what kind of data they used, most of the information was not applicable for the thesis.
As mentioned before the focus of this project is on the state Berlin, but because most of the distribution centres are located just outside the actual state Berlin we choose to expand the study area ten kilometres further than the A10, the Berliner Ring. This ensures that all the distribution centres close to the city are taken into a count for this project.
Furthermore, Germany offers a lot of open data which makes it easier to collect the needed information for the thesis. Also, the Government and their organisations are willing to provide data to researchers and students if it is used for science.

1.4 Research Objective and Research Questions

The main goal of this study is to estimate what the additional capacity of public transport to distribution in the food supply chain, when taking into a count speed, distance, capacities, and locations of start and end points. The objective of this research is:

"Improvement of the food supply chain using the current public transport networks."

To realize this objective the following three research questions, which structure this thesis, must be answered.

For the whole project the five largest supermarket chains are taken as case objects. This means that only Aldi, Lidl, Netto, Edeka, and Rewe supermarkets and distribution centres are used in this research.

1. **Where are the distribution centres and supermarkets located in a metropolitan region and how are they connected?**

   This research question, created to understand the existing food supply network, is a combination of two questions. The locations of the distribution centres must be known to have starting points for the supply routes and to give insight in the number of products to transport. The locations of the supermarkets (retail centres) present end points. A prepared network, based on OpenStreetMap data, will be used to calculate routes from distribution centres to retail centres by using the parameters of transport vehicles like lorries.

2. **How is the public transport network organized in the city Berlin, what is the occupancy rate of the public transport in this metropole and how can the distribution centres and supermarkets be connected by using the public transport network?**

   This second question is created to gain knowledge about the public transport network. It also exists of multiple sub questions.
   The first part is formulated to get an overview of the current public transport network of the study area. This includes the different kinds of transport vehicles such as, train, metro, tram and bus. For all these types the stops and travel routes are collected.
   The second part of the questions will give an indication if the public transport is even useable for freight movement.
   The last part studies how to connect the distribution centres and supermarkets while using the existing public transport network. This can be done by connecting to the closest bus stop (Point to point) or connect to the bus line (Euclidean distance).

3. **What is the additional capacity when the public transport network is used as extra supply network?**

   The last question has been formulated to get the actual difference of the transport capacity in tonnage between the two supply chains. The first chain only uses the current road network and the second supply chain uses the bus network. With these two scenarios the difference between the types of transport and the additional capacity can be determined.

2. Methodology

This chapter of the report will describe the methods that are used during the research to achieve the main objective: "Improvement of the food supply chain using the current public transport networks."

Firstly, a flow chart shows the steps of the research, then the data structure and, process of data gathering is explained. Furthermore, the three questions are discussed in different paragraphs. Additionally, in Appendix 7, 8 and 9 an overview of the toolboxes, toolsets and models created for this thesis is shown.

2.1 Flowchart

The flowchart in Figure 7 is used as a basic structure for this study. The yellow blocks contain the data sources and the method of gathering the data. This means that most of the data is collected from OpenStreetMap. Here-fore the two products GeoFabrik and Overpass Turbo are used. Furthermore, extra data, about food demand and borders, are gathered from governmental organisations. After gathering the different datasets, they are all clipped to the study area.
To create a ArcGIS feature for the study area a python script has been used. The script has been created by Aldo Bergsma, from the Wageningen University and Research, and was used in the course GeoTools. The python script is placed in the toolset and then dragged to a model. Opening the *CreateStudyAreaScript* tool gives you the option to fill in the xmin, ymin, xmax and ymax. These coordinates are the left lower and right upper corner points of your square. Also, the coordinate type must be selected. For the Berlin area the coordinate type WGS_1984_UTM_Zone_33N has been chosen.
Then two processes are created. The grey process involves finding the answers on the first research question, about the freight movement by road. And the blue process involves answering the second research question, involving the public transport. The red steps select and edit building data for the two processes. The green steps calculate the food demand per district for both the processes. The last step, shown by a dark grey block/symbol, is about comparing the movement by public transport and road to answer research question three.

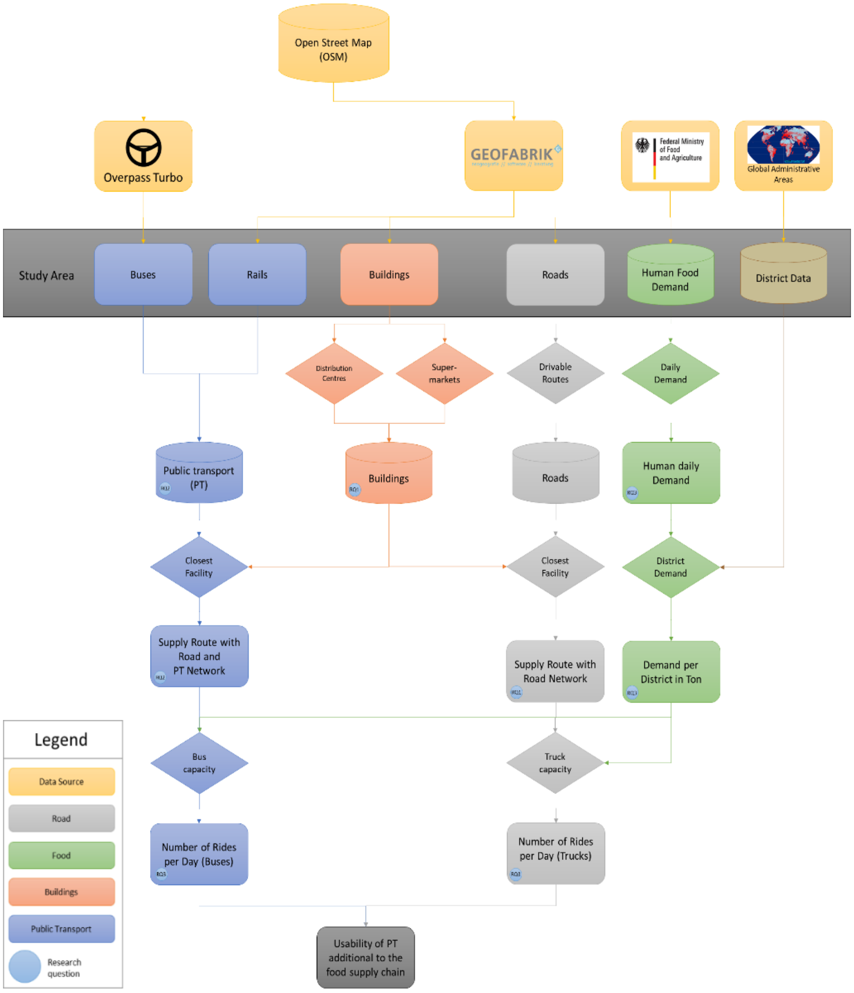


Figure 7. Flow chart describing the methodology.

2.2 Pre-processing Steps

In this paragraph the pre-processing steps of the thesis are discussed. This contains a description of the data storage structure, an explanation of the chosen the study area and clarification of the data gathering method. These steps are carried out before the first steps of the flow chart, shown in Figure 7.

2.2.1 Data Storage Structure

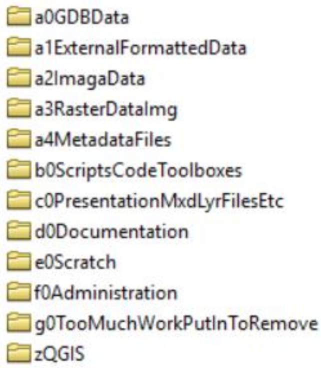


Figure 8. Data Storage Structure.

The data storage structure is based on the structure, which was used in the course GeoTools offered by the Wageningen University. Ten folders are created, all with their own purpose, see Figure 8. The first five folders, a0 until a4, are made for all different kinds of data. Folder *a0GDBData* contains three file geodatabases; one for the input and result data, a second for the working data and the third database is for all the temporary data. The second folder, *a1ExternalFormattedData*, contains all the raw data which cannot be placed in the file geodatabase in the a0 folder. For example, OSM files, shapefiles and excels sheets can be found in this folder. The third and fourth folder, *a2ImagaData* and *a4RasterDataImg*, are not filled with data. For this thesis topic no image or raster data is needed, but it is a standard to create them. *a4MetadataFiles* is created to hold the location of the metafiles describing the data.

The sixth folder, *b0ScriptsCodeToolboxes*, is created to store the toolboxes, toolsets, models and scripts. In Appendix 7, 8 and 9 a complete overview of the unfolded toolboxes for pre-processing, research question one and research question two can be found. This gives an overview of the steps that should be taken when working on the project.

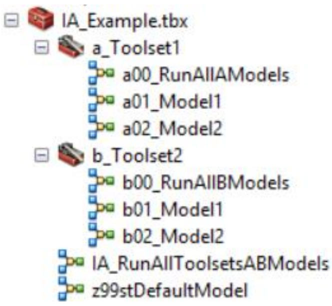


Figure 9. Example toolbox.

The *b0* folder has also its own structure. It contains four toolboxes, each with a separate task. In Figure 9 an example toolbox with the structure of the actual toolboxes of the thesis can be seen. It is used to explain the structure.

The example toolbox includes two toolsets, each with their own models. The first model in each toolset is called **00_RunAll*Models*, herein the stars are filled with the letter of the toolset. This model is made to run all the models of the toolset in a row. For example, model *a00_RunAllAModels* contains the models *a01_Model1* and *a02_Model2*. This is also done for the second toolset with model *b00_RunAllBModels*. The running order of the models is based on the order of importing, which means that the models should be imported in the correct order to the *a00* and *b00* models.

Additionally, each toolbox includes, besides the toolsets, also two models. The first model has the same function as the *00* models in the toolsets. This model, in the example toolbox model *IA_RunAllToolsetsABModels*, is used to run the *00* models of all the toolsets in a row. This means that all the models in the different toolsets are executed one after another.

The last model present in each toolbox is the model called *z99stDefaultModel*. As the name suggest it is default model containing all the correct settings and parameters. When a new model must be made, this model is copied to correct location and renamed. This ensures that all the models have the same settings.

The next folder, *c0PresentationMxdLyrFilesEtc*, contains the mxd files, which are used to save presentation files. Also, layerfiles for presenting specific layers are saved on this location.

Furthermore, the *d0Documention* folder is created for extra documentation about the models. Followed by a *e0Scratch* folder as location for scratch data and a *f0Administration* folder for administrative documents, such as user rights.

Besides these standard ten folders additionally a folder, *g0TooMuchworkPutinToRemove*, for models that took a lot of time to build, but eventually seemed unnecessary, and the folder *z0QIS* for QGIS data are added.

2.2.2 Data Gathering

Before starting on the methods for the research questions, the data is gathered. The gathered data is imported and pre-processed with the models in toolbox **IA_FoodSupplyChainPreprocessing**, see Appendix 7.

There are two different kinds of data used in the analysis. Firstly, the spatial data, which mostly comes from OpenStreetMap (OSM). The OSM data are downloaded from the Geofabrik website. Geofabrik is an organization who delivers prefixed dataset of the whole world (Geofabrik, 2017). Their datasets are created per continent, country, region and sometimes also sub regions.

2.2.2.1 Location data

For this thesis, location data of the provinces Berlin and Brandenburg are downloaded. The download contains 18 different shapefiles. One of the files is a layer with the buildings, which contains the descriptions of the buildings. This could for example be greenhouse, supermarket, hotel, residential and so on. The descriptions determine which buildings are used as locations for storing and selling food. Additional spatial information comes from the freight centrum of Berlin (Güterverkehrszentrum – GVZ Berlin). This information consists of the names and locations of the companies who are located at the three large logistic areas around Berlin. Figure 10 shows the three large logistic areas which are located around Berlin. Also, the webpages of the five supermarket chains, which are used in this analysis, have information about the location of their distribution centres. The distribution centres which were not available in the OSM data, are drawn or edited in OSM. The next time data from Geofabrik are downloaded these centres are also in the dataset.

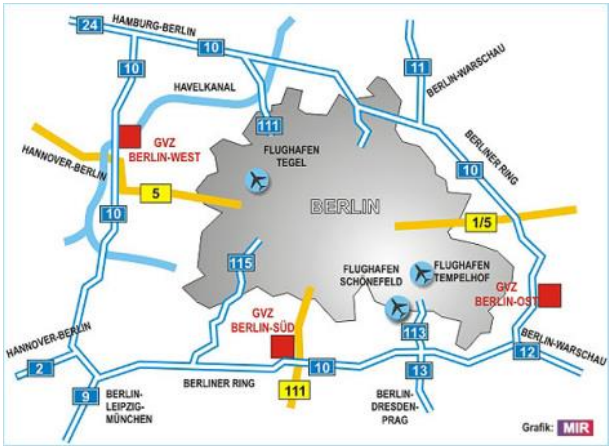


Figure 10. Logistic areas located around Berlin (Ministry of Infrastructure and Regional Planning, 2012).

With the knowledge of the two mentioned datasets and the webpages it is possible to select the distribution centres and supermarkets. This shows that there are a lot of subjects that can be studied with the available open data, if you know where to find the required data.

2.2.2.2 Road network

Next to the location data of the supermarkets and distributions data, information about the road network is required to determine the routes. For this another OSM layer is used. This layer is part of the earlier download data form Geofabrik. It contains all the roads divided in different categories, such as motorways, bicycle paths, footways, primary, secondary and so on. To use the correct roads, which are useable for freight transport, a precise selection is carried out. The selection is made in ArcMap and removes the bike

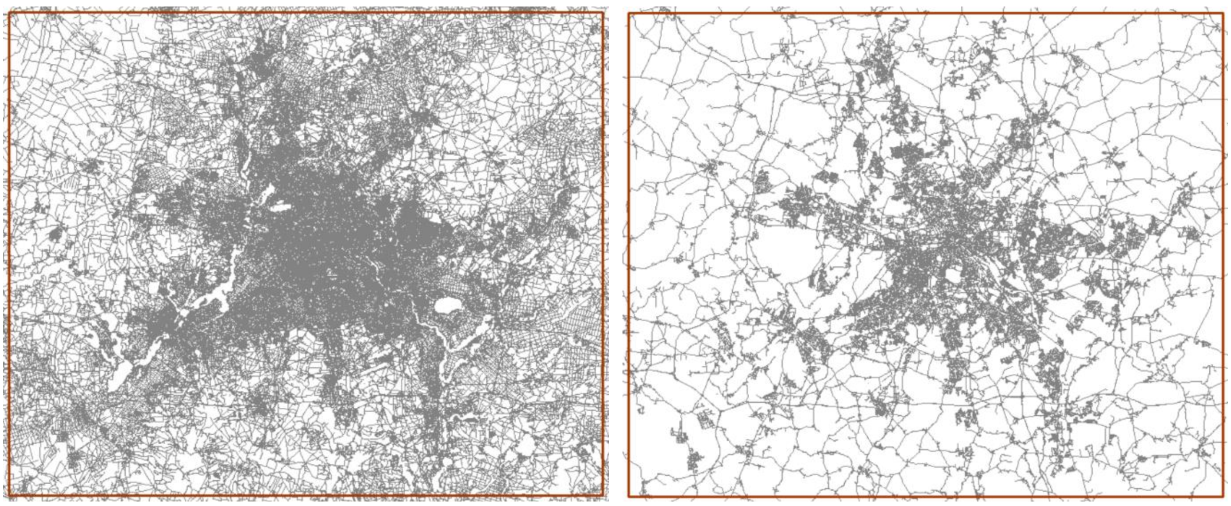


Figure 11. Complete OSM road network (left) and remaining roads after the selection (right).

lines, footpaths and other none driveable roads for large distribution vehicles. By using ModelBuilder a new model is developed which creates the selection and generates a new layer. The select tool is used to make the actual selection. The input for the tool is the OSM file with all the different types of roads. The expression selects only the roads which are usable for freight movement. (see Figure 11). After the selection is made the other part of the pre-processing can start.

As mentioned before the road data comes from OpenStreetMap. This means that it is freely available and easy to edit. Unfortunately, open source data has also some disadvantages. For example, everybody can alter the data of OpenStreetMap. This is not immediately a problem, but it also means that a lot of different styles are used when the data is created. Some persons are very precise in drawing and adding all the meta data and others are less secure. These issues do occur in the used road network data. Therefore, the models *c01*, *c02* and *c03* are made to delete, correct and add data. *c01stCopyAndAddFieldsToRoads* is created to delete overlaying polylines and to add the real geometry to the line segments. Deleting overlapping lines is done by first splitting all the lines in small segments by using the tool: *"Split Line at Point"*. The points for the split are create with the *"Feature Vertices To points"* tool. Afterwards the new output is than intersected with itself. This results in all the overlapping features. An extra field, with the name *"ToDelete"* is added to these features. The value is for all the features one. Now the overlapping features are recognizable and the *"ToDelete"* column can be joined to the early selected dataset. After the join all the features, with Null in the *"ToDelete"* column, are selected and saved as new dataset. As last part of the road pre-processing some unnecessary columns are deleted and the real geometry is calculated with the tool *"Add Geometry Attributes"*. This new geometry is necessary for the time that will be calculated later.

In *c02stEditMaxSpeedToRoads*, the maximum speed per road segment is added. This not needed for all the segments, because some features already have a maximum speed in the attribute table. The calculation, which determines wat the maximum speed is per road type can be seen in Figure 12. This Python script is part of the *"Calculate Field*" tool. It checks if the attribute from the OSM database *("maxspeed")* null is. When this happens the most common speed from the type is taken as maximum speed. Otherwise the present speed from the OSM database is taken as maximum speed.

```
def calculatespeed(maxspeed, fclass):
 if (maxspeed == 0 and fclass == 'motorway'):
  return 120
 elif (maxspeed == 0 and fclass == 'motorway_link'):
  return 60
 elif (maxspeed == 0 and fclass == 'primary'):
  return 50
 elif (maxspeed == 0 and fclass == 'primary_link'):
  return 50
 elif (maxspeed == 0 and fclass == 'secondary'):
  return 50
 elif (maxspeed == 0 and fclass == 'secondary_link'):
  return 50
 elif (maxspeed == 0 and fclass == 'tertiary'):
  return 50
 elif (maxspeed == 0 and fclass == 'tertiary_link'):
  return 50
 elif (maxspeed == 0 and fclass == 'trunk'):
  return 50
 elif (maxspeed == 0 and fclass == 'trunk_link'):
  return 50
 elif (maxspeed == 0 and fclass == 'residential'):
  return 30
 elif (maxspeed == 0 and fclass == 'unclassified'):
  return 30
 else:
  return maxspeed
```

Figure 12. Determine the maximum speed per road type.

After the calculation every road feature will have a maximum speed. Now the time per segment can be calculated. Here for also an extra field must be made. When this is done, the *"Calculate Field"* tool is used to calculate the time per feature. The following formula is used to calculated: *([LENGTH_GEO] / ([MaxSpeedComplete] / 3.6)) / 60*. In the formula the LENGTH_GEO is the newly added distance of a line segment. The MaxspeedComplete is the maximum speed in kilometres per hour. Dividing the speed by 3.6 transforms it to metres per second. Then dividing, the result of the distance divided by the speed, by 60 is needed to get the values in seconds. After this last step the road network is ready to use.

2.2.2.3 Public transport network

Following the locations of the food supply chain segments and the road network, the public transport network is essential for research question two. This network includes the bus stops and lines. The OSM file that is downloaded from Geofabrik also includes two layers of transport. The first one is a point layer with all the stops of the different transport types. And the second layer is a polygon layer, which contains the shapes of the larger stops, stations and terminals.

These two datasets deliver all the stops of the public transport for the area Berlin-Brandenburg. Similarly, the routes for trains, subways, light rails and trams are directly available in the OSM layer railways. Unfortunnaly, only the bus stops are accessible in OSM. There is no layer that contains the routes of the buses. Here for another technique is available. The bus lines can be mapped by using a technique which usages OSM roads, the bus stops and General Transit Feed Specification (GTFS) (Cich et al., 2016).

Firstly, the bus stops should be selected form all the public transport stops by using a selection on the column "fclass". After the selection the selected features are saved as a new layer. The created layer will be used to generate the bus lines. For this the recently developed method (Vuurstaek et al., 2017), which uses GTFS files in combination with the road network from OSM is used. The GTFS dataset, for the study area of this thesis, is an open data file and is offered by the Transport Association Berlin-Brandenburg (VBB).

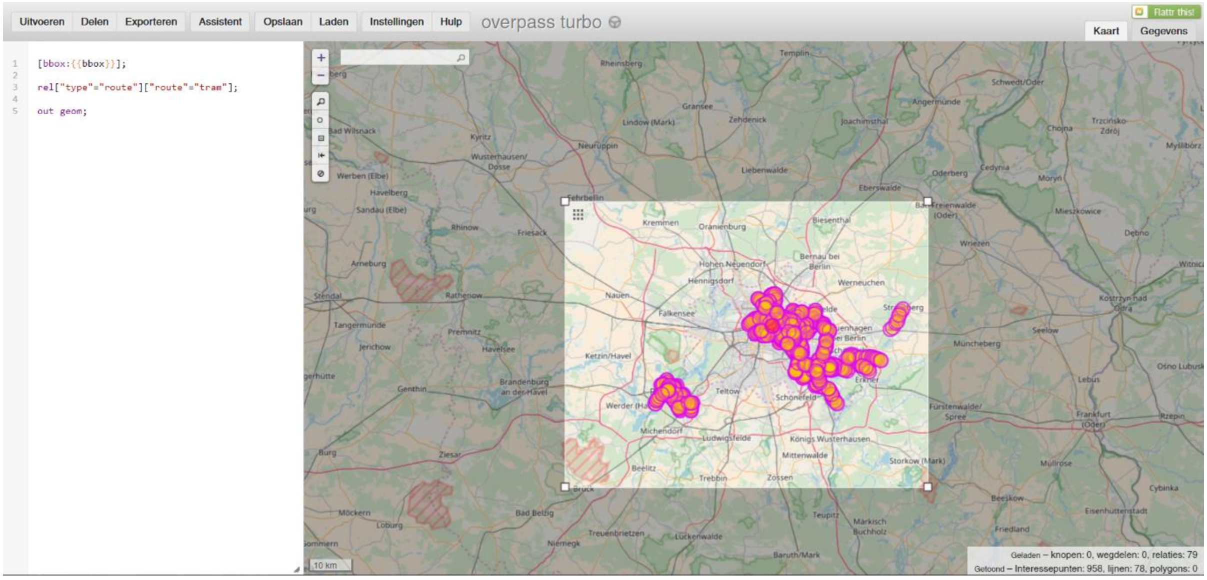


Figure 13. Example of Overpass turbo to select tram lines and stops in a selected area.

The dataset exists of nine text files. They contain different elements, such as the locations of the stops, routes, shapes and multiple others. Together they form the public transport network of the area Berlin-Brandenburg. The text file with the shapes of the bus lines can be loaded in QGIS. By using the steps explained in a manual to convert GTFS to shapefile with QGIS (Steven, 2016), a shapefile can be created. There has been chosen to do these steps in QGIS, because it is in this case easier and quicker than ArcGIS. The new shapefile now contains the bus lines for the study area, unfortunnaly not all the lines are following the road network. Because of the problem with the GTFS conversion there has been chosen to use the Overpass-turbo website to extract transport data. It also uses less intermediate steps which makes the possibility of making mistakes smaller. Overpass (Martin Raifer, n.d.) is an online tool created to extract data from OSM. By writing a small code a selection can be made over the whole OSM dataset. Figure 13 shows that three rows of code are needed to select the tram stops and lines in the selected area. The same kind of selection is made for the bus stops and lines. When the selection is made the selected parts can be downloaded as kml file. To transfer the file to a usable format QGIS is a good programme which is easily to use. It handles a lot of different file types of which kml is one. The kml download can be dragged from the saved location into QGIS. The kml file contains a layer with lines and other with points. By right clicking on the layers the *"save as"* option will pop-up. Saving them both as separate shapefiles leaves the opportunity to use the lines and points separately.

The shapefiles with the bus lines and stops are imported with the *b05stImportBusRoutesToDatabaseAndMakeSelection* model. Also, a spatial selection is made based on the borders of the study area. Both the lines and points are saved in working database.

As mentioned earlier in de paragraph about the road network, it is needed to delete overlapping lines. For the bus lines this is done in model *c05stCopyAndAddFieldsToBusRoutes*. First all the lines are dissolve with the *"Dissolve"* tool. Thereafter the output of the dissolve tool is used as input for the *"Multipart To Singlepart"* tool. And as last step the new length is calculated with the *"Add Geometry Attributes"* tool.

```
def calculatespeed(MaxSpeedComplete):
 if (MaxSpeedComplete is None):
  return 50
 else:
  return (MaxSpeedComplete)
```

Figure 14. Determine the maximum speed for bus lines.

After the new length is added the second model for processing the bus lines can be started. *c06stEditMaxSpeedToBusRoutes* is adding the maximum speed to all the bus lines, which have None as attribute value in the table. This is done with the script in Figure 14. It checks if the maximum speed is present otherwise it fills the table with a maximum speed of 50 kilometres per hour, which is the most common speed.

As last step for the bus lines the time per segment must be calculated. Here for the same model can be used as for the road network, only the names are chanced. After this step the bus lines are prepared and ready to use in the rest of the analysis.

2.2.2.4 Non-spatial data

Besides the spatial data, also the population numbers per region for Berlin and the average food consuming quantity per inhabitant are collected. The Leibniz Centre for Agricultural Landscape Research (ZALF) delivers the population numbers. These numbers are sent in an Excel file and needed to be transformed to an ArcGIS table. For the transformation the tool *"Excel to Table"* is used. To make the data in the new table spatial it must be joined to the districts of Berlin. The chosen districts are the so called *"Prognoseräume"* districts. This specific district division contains 60 areas with an average number of inhabitants between 55.000 and 60.000 (Government Berlin, n.d.). The join between the table and the districts is made based on the names of the districts. When the join is made the number of inhabitants is connected to the districts, which means that the data spatial is.

With the known number of inhabitants per districts, the food consumption per district can be calculated. Here for the average food consumption per person per day is multiplied by the number of inhabitants per districts (Models in toolset ***e_FoodDemandDistrictLevel***). This results in a total amount of food consumption per day per area. Dr. Ingo Zasada, from the Leibniz Centre for Agricultural Landscape Research (ZALF), who also worked on the FOODMETRES project, recommended to use the data from the Federal Ministry of Food and Agriculture. The ZALF published a report in 2008 with the average food demand per day for men and woman per state in Germany, see Appendix 6 (Federal Ministry of Food and Agriculture, 2008). The ministry carried out 15371 interviews on inhabitants spread over the country. These interviews formed the base for the national consumption statistics.

The last non-spatial data which are needed, are the capacities of transport by truck or public transport. These capacities determine the number of rides needed to deliver the required amount of food.

The capacity of the trucks is determined on 30 tons. This number is based on the maximum weight of a truck in most European countries, which is set on 40 tons (Butcher, 2009; THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 2015). Germany has the same maximum weight of 40 tons for trucks with a length between 16.50 and 18.75 metres (Neelevat, n.d.). If the weight of a truck, which lies around 10 tons (Kampen, 2003), is subtracted from the maximum total weight the capacity of a truck lays around 30 tons.

The public transport capacity is more difficult to determine because there is not such a solution for transporting persons and goods. Fortunately, there is a vehicle that has the possibility to be the solution, the Midi train MAN Göppel bus, see Figure 5. If the trailer part of this "train" will be transformed to a trailer which can contain food, it could be used as delivery vehicle. The trailer is easy to detach and can be pick up by another bus. The capacity of the trailer lies around 7 tons (Göppel Bus GmbH, n.d.), which is much less than a normal truck, but as additional delivery vehicle it could save many truckloads into the city.

2.3 Research Questions

In this part of the report the methodology used during the thesis is discussed. All the important steps form the model are explained. Some phases are skipped when they are already described earlier in the report.

2.3.1 Research Question One

With the data gathered, work on the first research question: *"Where are the distribution centres located in a metropolitan region and how are they connected to the supermarkets?"* can start. The question is defined to get knowledge of the metropolitan region that has been chosen as study area. All the models, needed for answering research question one, are saved in the toolbox **IA_FoodSupplyChainRQ1**.

2.3.1.1 D_FoodLocations

The first toolset, ***d_FoodLocations*** unfolded in Appendix 8, contains the models which are needed to divide the distribution centres and supermarkets in the five chosen chains (Aldi, Lidl, Netto, Edeka and Rewe).

Model *d01* selects and divides Aldi buildings, see Figure 15. First a *"Select"* tool selects all the buildings from a chain, for example Aldi. The input data

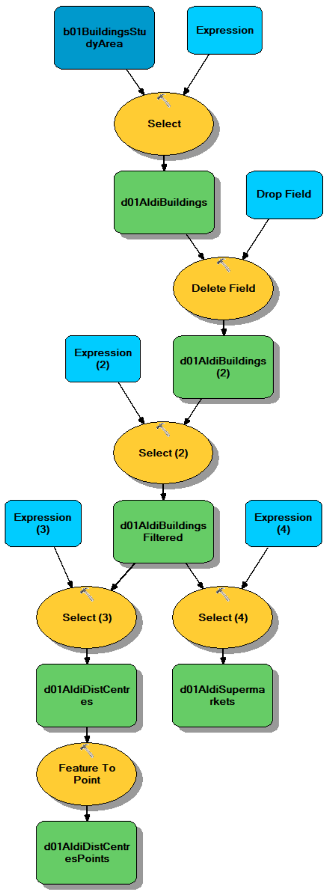


Figure 15. Model to select and divide Aldi buildings.

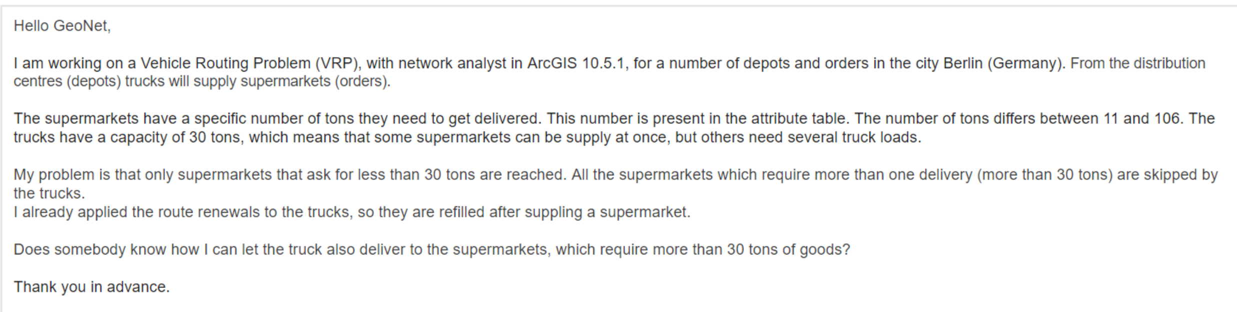
for this selecting is the dataset with all the buildings from the study area, which is prepared in model *b01stImportBuildingsShpToDatabaseAndMakeSelection*. The expression used for the selection, *"name LIKE '%ALDI%' OR "name LIKE '%Aldi%'"*, searches for the word Aldi in column with the names, but because of the percentage sign it does not matter what is in front or behind the word Aldi. After the selection the redundant data is deleted with the *"Delete field"* tool. As third step a filter takes out all the wrong Aldi selections. Wrong selections are already former or closed supermarket and parking lots. The fourth step split the distribution centres and the supermarkets in two different feature classes. This step makes, as well as the other two, use of the LIKE function in the expression.
This model is also made for the other four supermarket chains. There is deliberately chosen not to merge the five models, because with the separate models the opportunity to rerun only the data of one supermarket chain stays available.
After the five models have split the building data into separate features model *d06stSupermarketsSize* can be runt. It is created to add the size in square metres to all the buildings. This size is needed later to calculate the food demand per supermarket. Firstly, all the supermarkets get the supermarket name in a newly added column. This is done to make them al similar and easier to select. Secondly a field is added for the square size of the supermarket and the size is calculate with the python expression *!shape.area@squaremeters!*. Afterwards all the supermarkets are joined and converted to points. This new point feature class will be used later, in model *i03*, to define which supermarkets close enough are to the bus lines.
The last step from toolset *d* is model *d07stMergeDistributionCenters*. As the name suggests this model is made for merging the five distribution centres. The merged feature class will be used later, in model *i02*, to determine which distribution centres are close enough to the bus lines.

2.3.1.2 E_FoodDemandDistrictLevel

The second toolset ***e_FoodDemandDistricLevel*** is already mentioned in paragraph 2.1.3.4 Non-spatial data. The output of the last model in this toolset is the demand in tons per *"prognoseraume"* district.
Afterwards it seemed better to place this toolset in the pre-processing toolbox.

2.3.1.3 F_FoodDemandPerSupermarket

After completing the models in toolsets *d* and *e* it is time to continue with toolset ***f_FoodDemandPerSupermarket***. This toolset creates in multiple models the demand per supermarket.
The prepared *"prognoseraume"* areas are joint with the supermarket points. Furthermore, with the *"Summary Statistics"* tool the square metres per supermarket are summed. Then an extra field is added and with the calculation *"[ProgRegionTonFoodDemand] / [SUM_AreaSquareMeters]"* the food demand per square metre is calculate. Another extra field is added for the food demand per supermarket and calculate with the formula *"[AreaSquareMeters] * [TonPerSquareMeter]"*. When the supermarkets have their food demand, based on the size of the store, the polygons are converted to points and split in five layers, based on supermarket chain. As last step all the unnecessary fields are deleted.


Hello GeoNet,

I am working on a Vehicle Routing Problem (VRP), with network analyst in ArcGIS 10.5.1, for a number of depots and orders in the city Berlin (Germany). From the distribution centres (depots) trucks will supply supermarkets (orders).

The supermarkets have a specific number of tons they need to get delivered. This number is present in the attribute table. The number of tons differs between 11 and 106. The trucks have a capacity of 30 tons, which means that some supermarkets can be supply at once, but others need several truck loads.

My problem is that only supermarkets that ask for less than 30 tons are reached. All the supermarkets which require more than one delivery (more than 30 tons) are skipped by the trucks.
I already applied the route renewals to the trucks, so they are refilled after suppling a supermarket.

Does somebody know how I can let the truck also deliver to the supermarkets, which require more than 30 tons of goods?

Thank you in advance.

Figure 16. Question posted on https://community.esri.com (https://community.esri.com/thread/208929-vrp-multiple-deliveries-to-one-order)

Unfortunately, a problem was discovered during the creation of the routes between distribution centres and supermarkets. The supermarkets with a higher demand, in tons of food, than the maximum capacity of the transport vehicle, were skipped and would not get any food delivered. After trying everything by changing all the possible parameters, it still did not work. As last hope a question was posted on the ESRI community webpage, see Figure 16.

An employee of ESRI answered with a conformation that it was not possible to do this. He came with an idea of a work around. His idea meant that the individual supermarkets were split in multiple points with a maximum demand of 30 ton, the maximum capacity of a truck. For example, a supermarket with a demand of 74 ton will be split in two points of 30 ton and one with the remaining 14 ton. This method creates three points on the same location. A route will be calculated for each individual point. After deliveries to the points with a demand of 30 ton, the truck is empty, and it will return to the distribution centre to be refilled. The point with a demand of 14 ton leaves 16 ton in the truck, which will be used to supply other supermarkets.

Model *f02stAldiDivideSize* makes sure that all the supermarkets are split into points with a plurality of 30 ton or the remaining part. This is done with a simple selection statement. The expression, see Figure 17, determines the different groups that are needed for model *f03*.
The same method is used for the other four supermarket chains in models *f05*, *f07*, *f09*, and *f11*. Only the number of groups in the expressions changes depending on the highest demand of the supermarkets. If the supermarket with the highest demand needs a supply of 290 ton. There will be two extra expression rules. The first for the demand between 240 and 270 ton and second the second for the demand between 270 and 300 ton. This makes sure that the supermarket with a demand of 290 ton is covered by second added the expression.

	Expression
1	DemandTon <= 30
2	DemandTon > 30 AND DemandTon <= 60
3	DemandTon > 60 AND DemandTon <= 90
4	DemandTon > 90 AND DemandTon <= 120
5	DemandTon > 120 AND DemandTon <= 150
6	DemandTon > 150 AND DemandTon <= 180
7	DemandTon > 180 AND DemandTon <= 210
8	DemandTon > 210 AND DemandTon <= 240

Figure 17. Expression used to divide Aldi supermarkets in groups of 30.

The next model, *f03stAldiCreateSplitSuper*, creates new points for the divided supermarkets. For example, the supermarket with 74 ton gets the points called *"f03AldiDemand30And60Base30"*, *"f03AldiDemand60And90Base60"* and *"f03AldiDemand60And90Remain"*. The first point, Base 30, stands for the first 30 ton of the total demand. The second point, Base 60, for the second 30 ton of the total demand. And the last point, Remain, will complete the remainders of the total demand, in this case it will be 14 ton. the *"Copy Feature"* tool creates all the new feature classes, which are needed per supermarket. Next a field is added and filled with the number of tons it represents, this could be 30 for all the so called *"Base"* layers and the demand ton minus 30 for the remaining part.
The same procedure is followed for adding the service time to the attribute table. First a field is added and second the service time is calculated. This is done by filling the *"Base"* fields with 60 minutes and the remain field with two times the new demand, calculated in the previous step.
Afterwards all the feature classes are merged into one feature class, which contains all the supermarket points. This process is also carried out for the other four supermarkets in the models *f05*, *f07*, *f09*, and *f11*. With the output of this last step the toolset *f* is finished and all the input data for route calculation is made.

2.3.1.4 G_NetworksDatasetRoads

The next toolset, ***g_NetworkDatasetRoads***, contains the model that create the feature dataset for the road network, which is used to define the delivery routes between distribution centres and supermarkets.
The first step of the model *g01stRoadNetworkStudyArea* creates a feature dataset. This dataset is needed as location to build the network. With the tool *"Create Feature Dataset"* a dataset, with the name *g01NetworkStudyArea*, is created in the database *GDBProject2WorkingData*. Next the road data can be imported to the database. This is done by coping the earlier created *c03RoadTime*, with the *"Copy Features"* tool, to the database. The other steps of building the network must be performed manually.
With the previous model the feature dataset is created and filled with the needed road data. The next step is building the network dataset. For this the follow workflow is created:

- Right click on the feature dataset *g01stRoadNetworkStudyArea*, click *"New"* and select *"Network Dataset"*. This opens the network building wizard.
- Keep the name and choose *"10.1"* as version of the network. Click *"Next"*.
- Select the feature class *g01RoadStudyArea* and click *"Next"*.
- Select *"No"* as answer on the question: 'Do you want to model turns in this network?' and click *"Next"*.

- Click *"Connectivity"* and change the connectivity policy of the *g01RoadStudy* class to *"Any vertex"*. Click *"OK"* and *"Next"*.
- Select "None" as answer on the question: 'How would you like to model the elevation of your network features?' and click *"Next"*.
- In the right column click *"Evaluators"* this opens a new screen. Select attribute *"Length"* at the top. Change for From-To and To-From the Shape in column Value to *LENGTH_GEO*. Click "Apply" and "OK". The pop-upped screen closes. click *"Next"*.
- Type as travel mode *"Truck"* and click on the *"plus sign"* next to it. Change the Type into *"Truck"* and U-turns at Junctions into *"Allowed Only at Intersections and Dead Ends"*. Click *"Next"*.
- Select *"No"* as answer on the question: 'Do you want to establish driving settings for this network dataset?' and Click *"Next"*.
- *"Check"* the square in front of Build Service Area Index and Click *"Next"*.
- Click *"Finish"*.

After finishing the steps above the network dataset is build and can be used for calculating service areas and delivery routes.

2.3.1.5 H_CreateRoutesRoad

Now that the network dataset is ready to use the delivery routes can be calculated. Here for the toolset ***h_CreateRoutesRoad*** contains three different models.
The first model, *h01stMakeServiceAreas*, see Figure 18, creates per distribution centre a service area. This is done with the tool *"Make Service Area Layer"*. The tool only creates a starting layer and additional information must be added. As input data the in the previous toolset created network dataset is selected and five output names, for each supermarket chain one, are given to create five layers.
There is chosen to base the services area on driving time from the distribution centre. For this some parameters need to be set on specific values. The Impedance Attribute must be *"Minutes"*, Travel From or To Facility must be *"TRAVEL_FROM"* and Default Break Values must be *"60"*. These three parameters determine that the trucks can reach as far as a 60 minutes' drive.
After creating the service area layer the locations of the distribution centres must be add to the layer. This is done with the tool *"Add Locations"*. The input for this tool are five feature classes made in model *d01* with the distribution centres as points. Important is to set the search tolerance on *"250 metres"*. This is done because most of the distribution centre points are not direct to the network. One reason for this is the transformation from the distribution centres from polygons to points. The points are placed in the middle of the polygons, which is much further away than the outlines of the polygons. Also, the buildings are located on private terrain, which is not present as road in the network. With the search tolerance of 250 metres the distribution centres are connected to the road.

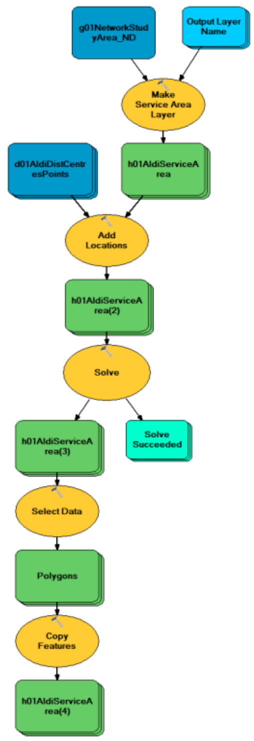


Figure 18. Model to create Service Areas.

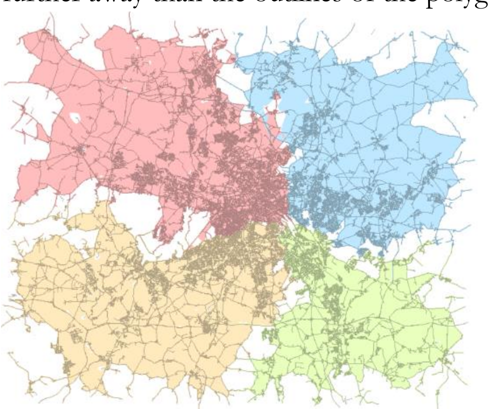


Figure 19. Service Areas of four Aldi distribution centres.

Subsequently, the service areas can be calculated with the *"Solve"* tool. Because these calculations are done with layer it is important to select the data with the *"Select Data"* tool and copy it to the working data database. The *"Select Data"* tool must be inserted in the model by clicking in the model *"Insert"*, *"Model Tools Only"* and then *"Select Data"*. Because it is a model only tool it cannot be found in the toolboxes or search beam. Figure 19 shows the output of the distribution centres of Aldi in combination with the overlaying road network.

The second model of the *h* toolset is the *h02stAldiTruckRidesPerServiceArea* model. This model calculates the number of truck rides needed per service area.

It starts with joining the services areas and the supermarkets. This is needed to calculate the total demand per service area. With a *"Spatial Join"* the supermarkets are all joined to the area where they are inside of.
Then an extra field, for the number of rides, is added to the attribute table. It is filled by using *"Calculate Field"* with the expression *"([DemandTon]+15) / %Truckload%"*. Hereby *"%Truckload%"* refers to a variable. This variable is inserted by clicking *"Insert"* at the top of the model and then clicking *"Create Variable"*. The reason for a variable is that it makes it easier to change if the truckload changes, but the same expression would work with the number 30 instead of %Truckload%. The 15 ton is added to the demand to round up, because it is half the truckload the it only influences the lower demand. For example, a demand of 40 means at least two truckloads, but will be rounded off to one truckload without adding the 15 extra tons. On the other hand, there is no influence for the supermarkets with a demand higher than 50% of a truck load. For example, a demand of 55 ton would be round up to two truck rides. With the 15 ton extra, the demand becomes 70 tons, which will be round off to two rides.
The next spatial join connects the distribution centres to the service areas. This delivers the result of this model, for every service area the total demand, number rides and name of the distribution centre in the service area.
The models *h04*, *h06*, *h08* and *h10* contain the same process for the other four supermarket chains.

The third model, *h03stAldiMakeRoadRoute*, creates the routes that can be taken by the trucks to deliver food from the distribution centres to supermarkets.
There are different options to calculate the shortest routes form point a to point b. The first option is to use the closest facility. This is an easy to use route solver, also available in ArcGIS. As the name suggest it searches for the closest facility using the shortest route, but it cannot visit multiple points after one another before going back to the distribution centre. That is the reason that it is not suitable for this problem.
For thesis a vehicle routing problem, also called traveling salesman, solution is needed. There is a tool in the network analyst toolbox in ArcGIS to create a vehicle routing problem. This is also the first step of model *h03*. The tool *"Make Vehicle Routing Problem Layer"* creates a basic layer where additional information can be added. The network dataset created in model *g01* is used as input dataset. This means that the routes are calculated over the *g01* network.
Afterwards extra data is added to problem layer. First the distribution centres are added with the tool *"Add Locations"*. Indicated in the tool is the function of the added data. The distribution centres are added as *"Depots"*. Also, a search tolerance is set on *"200 metres"*. This ensures that the distribution centres are connected to the road network. Even if they are not on the road, but in 200 metres of a road.
The next dataset which is added to the problem layer are the supermarkets. They are added with the same tool and the search tolerance stays on 200 metres, but this time they are added as *"Orders"* instead of depots.
For the next dataset a prepared feature class is made. To calculate the routes a layer is needed which contains the characteristics of the delivery vehicle. These are put in to the problem layer under the name *"Route"*. The first characteristic of the vehicles are the start and end of the route, in this case that is one of the distribution centre and the second is the capacity of 30 ton.

To refill the truck when they are empty, also route renewals are added to the routing problem layer. These route renewals are made specific per route. This means that every truck has his own route renewals. The properties of the route renewals determine to which truck they belong.
The last dataset which is added to the problem layer is the route zonal layer. This layer contains per distribution centre a service area. this are the areas made in model *h02*. There is chosen to make hard zones of the service areas. Because of this the distribution centres will delivery only to the supermarket in their service area.

With all the data added to the vehicle routing problem it can now be solved. This is done with the *"Solve"* tool. The input of the tool is the last step op the routing problem. The solve tool calculates the shortest routes, the heuristics used in this process are based on a tabu search metaheuristic and are proprietary (ESRI, 2017), between the distribution centres and the supermarkets. Sometimes multiple supermarkets can be affected by one truck before it reaches the maximum capacity it could carry. This saves extra rides between the distribution centres and supermarkets.
As last step of the *h03* model, via the *"Select Data"* tool the route layer is selected and copied to the database.
The models *h05*, *h07*, *h09* and *h11* contain the same procedure four the other four supermarket chains.
When all the models of this toolbox are carried out the first research question is answered.

2.3.2 Research Question Two

The second research question: *"How is the public transport network organized in the city Berlin and what is the occupancy rate of the public transport in this metropole?"* is created to get an overview of the existing situation of the public transport network. The models needed to gather answers for the second research question are collected in the toolbox **IA_FoodSupplyChainRQ2**.

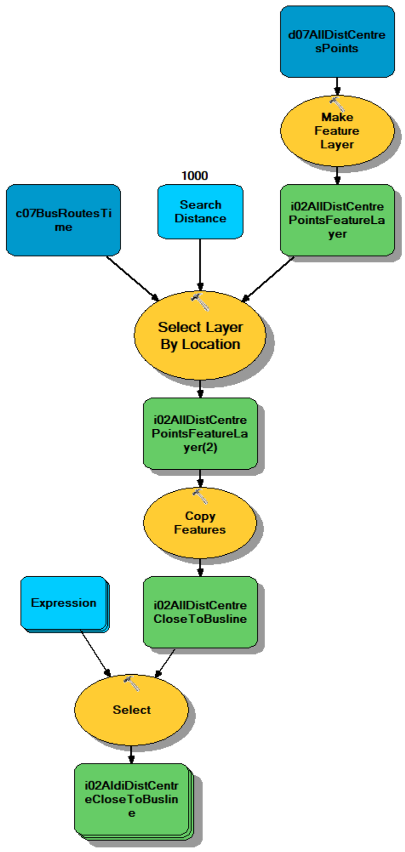


Figure 20. Model for selection distribution centres by location.

2.3.2.1 I_SelectSuitableSupersAndCentres

The first toolset in the toolbox contains the models which are needed to create the data for the vehicle routing problem with a bus network.

Model *i01stMergeSupermarkets*, see Figure 20, generates one feature class of the five datasets with the points per supermarket chain. This output file will be used later in the toolbox.

The next model, *i02stDistCentreCloseToBusline*, checks which distribution centres close enough are to be used in the analysis. It starts with converting the feature class with all distribution centre points to a layer file. Here for it uses the "Make Feature Layer" tool. This layer file is needed for the next tool because it only works with layer data.

The next tool is called *"Select Layer By Location"*. It uses an input feature layer, relationship, a selection layer and a search distance. The input layer is the feature layer with the distribution point, the output of the previous tool. The selection layer is the feature layer with the bus routes. This is the same layer that is the basis for the network that is built in the next toolset. The relationship between the two layer determines what kind of selection will be made. In this selection the question is which distribution centres are close by the bus lines. This means that an *"Intersection"* between the two layers is needed. Here for the search tolerance is set to *"1000 metres"*. The output of the selection includes all the distribution centres in a radius of 1000 metres of a bus line.

The result is copied and divided in the five supermarket chains by a selection based on the name.

Model *i03stSuperCloseToBusline* is made for the same purpose as *i02*, but instead of searching for distribution centres it looks this time for supermarkets. Only the search tolerance is set on *"100 metres"* instead of 1000.

After collecting all the useful supermarkets, they needed to be split, just like the supermarkets which were supplied by trucks. This is done in the models *i04stAldiBusDivideSize* and *i05stAldiBusCreateSplitSuper*. This time the capacity of the bus is only 7 tons, which means that the features needed to be split in a lot of extra features. For example, supermarket with a demand of 63 ton needs nine supply rides from a public transport supply unit. This results in nine points on the same locations. Eventually, Aldi had for example 244 point which needed a delivery. Compared to the 110 points Aldi had in the road network analysis this is much more.

The models *i06* to *i13* are processing the same procedure for Lidl, Netto, Edeka and Rewe.

2.3.2.2 J_NetworkDatasetBus

After preparing the input data, also the bus network must be prepared. In model *j01stBusNetworkStudyArea* is a feature dataset created to store the bus lines.

After the bus lines are added to the feature dataset, the topology of the bus line features must be checked. This must be done, because a lot bus stops are drawn as small loose lines next to the actual network. For checking the topology, the following procedure is used:

- Right click on the feature dataset *j01BusNetworkStudyArea*, click *"New"* and *"Topology"*.
- A wizard will open, click *"Next"*.
- Choose the defaults for the name and cluster tolerance and click *"Next"*.
- Check the bus line feature, *j01BusRoutesStudyArea*, as class to participate in the topology and press *"Next"*.
- Do not change the number of ranks, click *"Next"*.

- In the right column click *"Add Rule"*, select *"Must Not Have Dangles"* as rule and click *"OK"*. The pop-up closes and click *"Next"*.
- Click *"Finish"*
- Answer *"Yes"* on the question 'Would you like to validate it now?'.
- Drag the new topology layer into the screen.
- Start editing the *"j01BusRoutesStudyArea"* layer.
- Open the *Error Inspector* and select the *"Must Not Have Dangles"* error.
- Right click on the selected errors and click *"Select Features"*.
- Open the attribute table of the *"j01BusRoutesStudyArea"*.
- Open the *"Select by Attribute"* and create a selection on the selected features with the expression *"LENGTH_GEO < 50"*. This filters most of the bus stops which are located next to the actual bus network.
- Delete the selected features.

After correcting the wrong topology, the network dataset can be made. The procedure of producing a network dataset is explained in paragraph 2.2.1.4, but with the bus information.

2.3.2.3 K_CreateRoutesBus

Then the routes between the distribution centres and supermarkets over the bus routes can be calculated. For this the same models are used as in the toolbox *h* in paragraph 2.3.1.5. The models contain the same tools, but this time the bus data is used.

2.3.2.4 Occupancy Rate

The last part of the second research question is specified on the occupancy rate of the public transport. The original idea of the research was to also use time and occupancy rate of the public transport as part of the route analysis, but that proofed to be too difficult. Occupancy rate data of the public transport is hard to find. There is information free available in the Mobility of the City brochure of Berlin (Senate Department for Urban Development and the Environment of the State of Berlin, 2013). The Berliner Verkehrsbetriebe (BVG), the public transport company of Berlin, presents some numbers in their annual report. Table 1 shows the occupancy rate of the public transport network of Berlin. There is no specification in time or in the different types of transportation. The growth in the occupancy rate in Berlin has come to a standstill, which means that the remaining space potentially can be used for freight transport. If more specific data can be found difference can be made between the types of transport and the moment of the day. For this research occupancy rates are not included.

Adding occupancy rates would improve the results of the research. It could show when, where and how much space is left in the public transport vehicles. This shows if it is possible to use those transport vehicles for food distribution.

Table 1. Occupancy rate public transport Berlin (Berliner Verkehrsbetriebe, 2016).

Occupancy rate			
Year	2014	2015	2016
Percentage	17.8	18.2	18.2

2.3.3 Research Question Three

The last research question: *"What is the saving in tonnage transported by road, when the public transport network is used as additional supply network?"* has been formulated to notice what the effects are when the supply routes are changed. This research question is answered in the results and conclusions. Here the comparison between the tonnage by road and by public transport are given per supermarket chain.

3. Results

This chapter of the report shows the results of the thesis. It gives comparisons between the routes when using trucks and routes when the public transport (bus lines) are chosen. The results are given in numbers and maps showing the actual routes.

In the figure below a comparison is made between the data used for the network analysis. The maps, Figure 21 and Figure 22, show the distribution centres and supermarkets per supermarket chain. The columns on the left show the data for the road network, which answers a part of research question one. The columns on the right display the data used for the bus network.

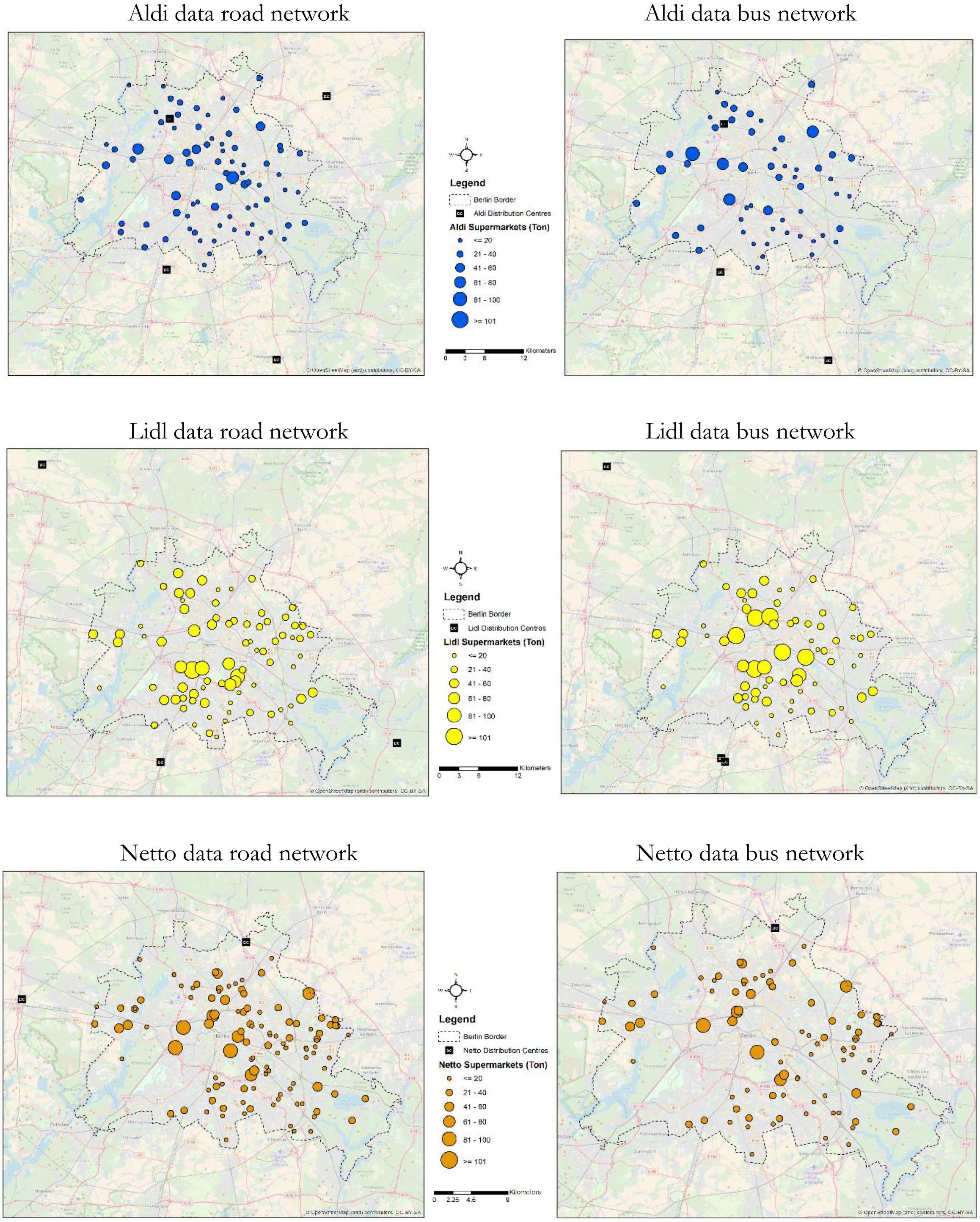


Figure 21. Location data Aldi, Lidl, and Netto used in the network analysis

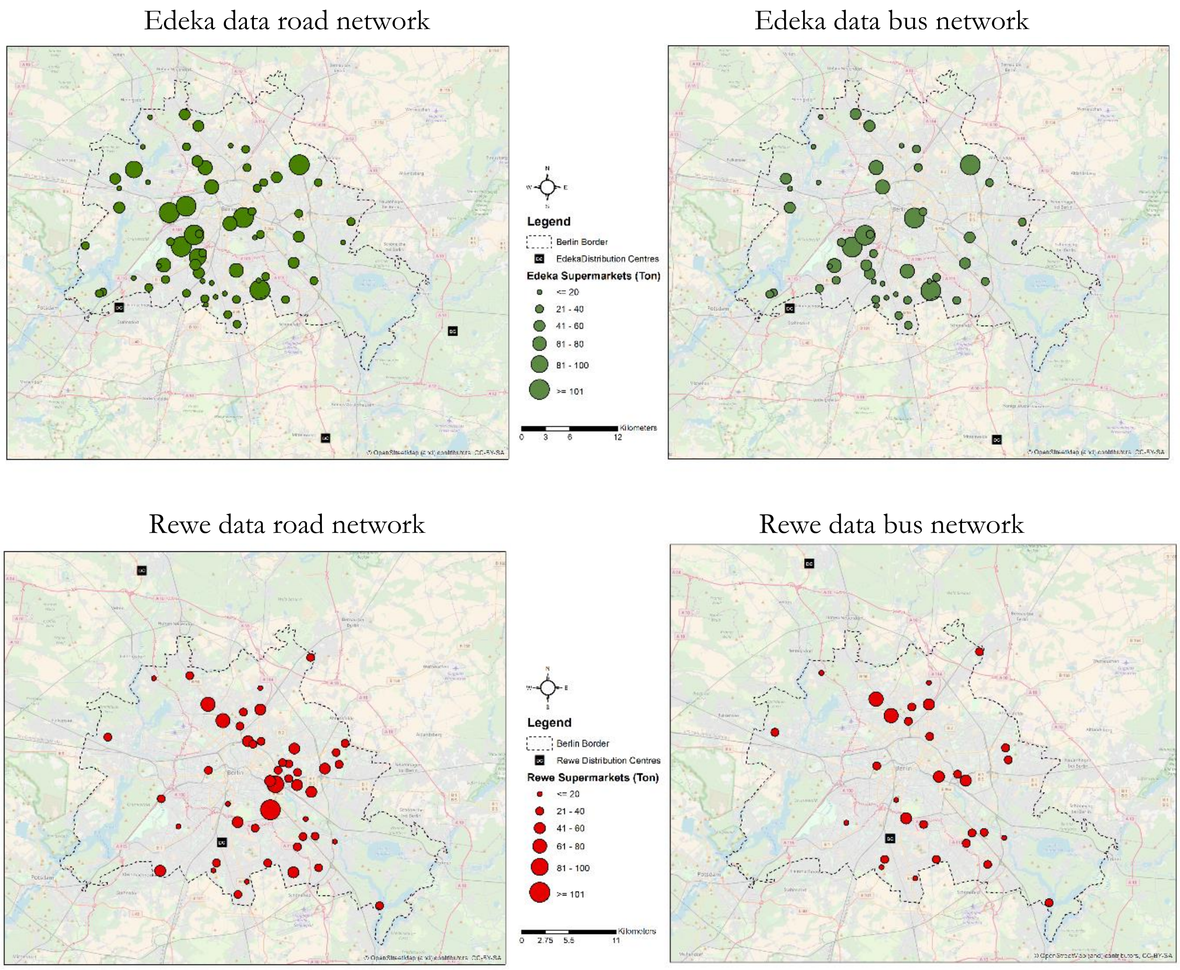


Figure 22. Location data Edeka and Rewe used in the network analysis

Because the difference between both situations is not always clear in the maps, the numbers of distribution centres and supermarkets are shown in Table 2 below for each supermarket chain.

Table 2. Numbers of distribution centres and supermarkets used per network type

		Truck		**Bus**	
		Number of Dist Centres	Number of Supermarkets	Number of Dist. Centres	Number of Supermarkets
Supermarket chain	Aldi	4	80	3	56
	Lidl	3	89	3	68
	Netto	2	114	1	85
	Edeka	3	65	2	49
	Rewe	2	48	2	30
	Total	14	396	11	288

Because not all the distribution centres and supermarkets are connected to the bus lines there are less locations to use as departure points and destinations. The next table, Table 3, gives an overview of the percentage of the usable locations (distribution centres and supermarkets) for the bus line analysis.

Table 3. Overview of the percentage usable locations for the bus line analysis

		Distribution Centres	Supermarkets
Supermarket chain	Aldi	75%	70%
	Lidl	100%	76%
	Netto	50%	75%
	Edeka	67%	75%
	Rewe	100%	63%
	Average	78%	72%

The next maps below, Figure 23 and Figure 24, show the delivery routes from distribution centre to supermarkets. The left column displays the routes when the road network is used, this is a part of research question one. The right column shows the routes which are created using the bus lines, which partly answers research question two.

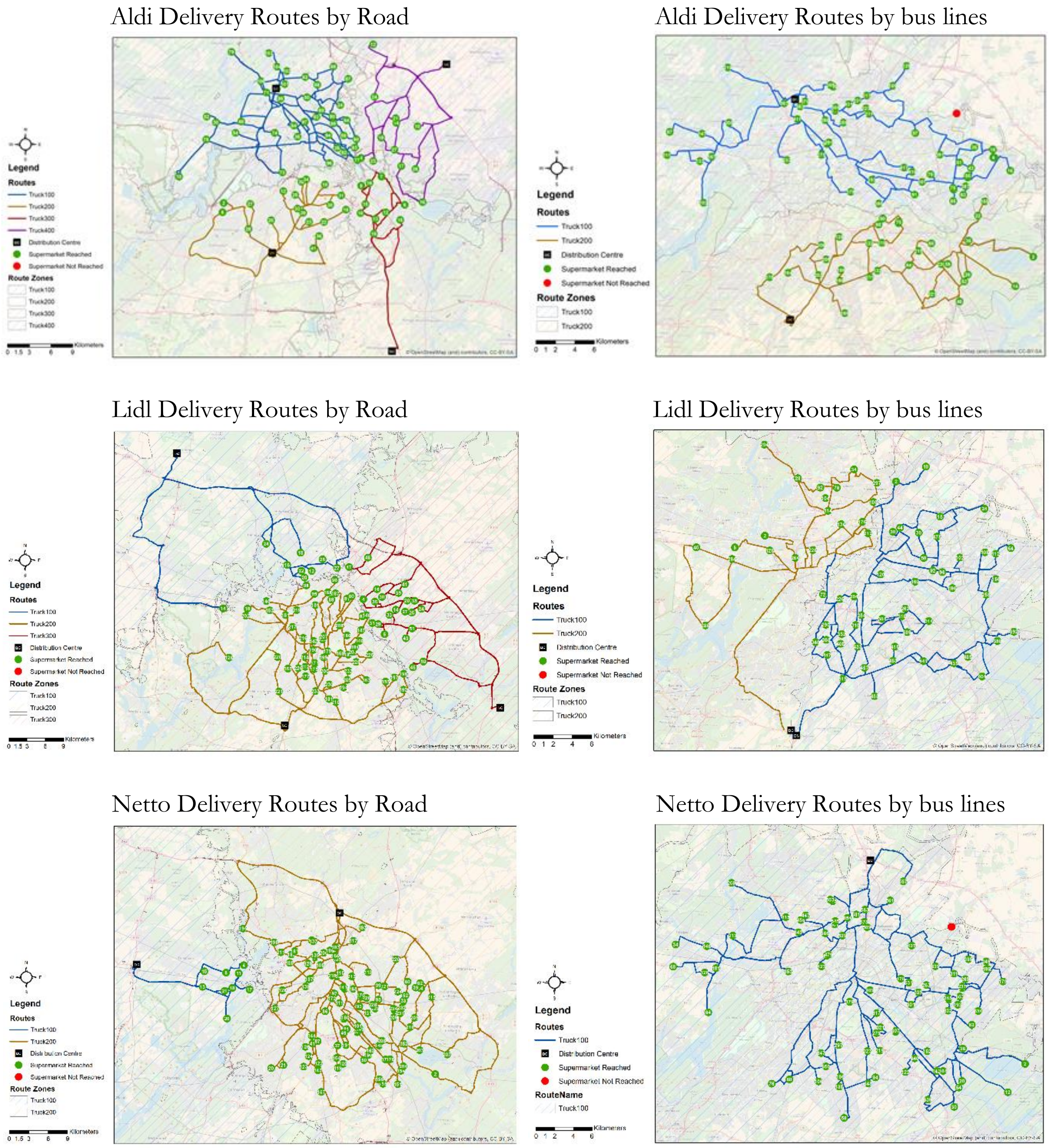


Figure 23. Overview of the delivery routes of Aldi, Lidl, and Netto by road or bus line

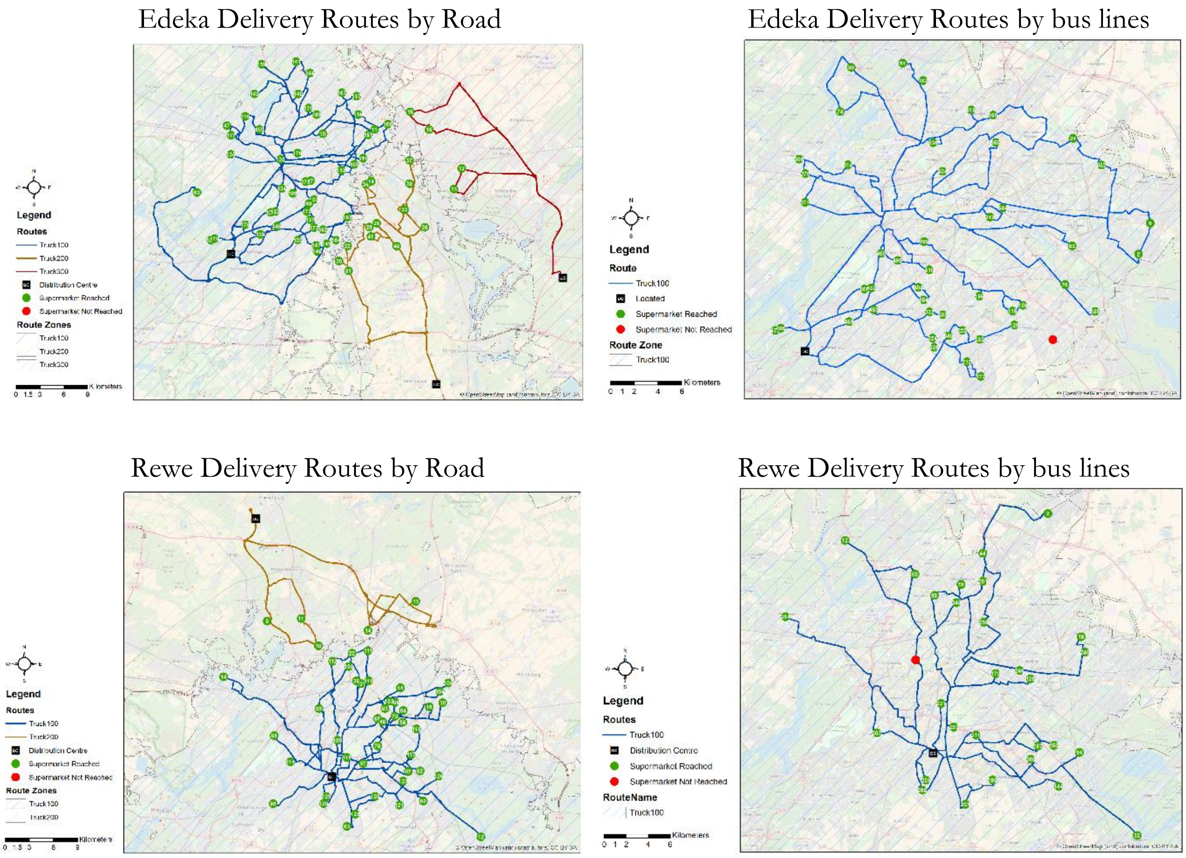


Figure 24. Overview of the delivery routes of Edeka and Rewe by road or bus line

The table below, Table 4, shows per supermarket chain the results of the route calculations. The routes are visualised in Figure 23 and Figure 24. The table displays the demand in tons that could be delivered by truck or bus. Also, the number of rides needed to deliver al the demanded supplies is presented in the table. The last column shows the percentage of the demand that could be delivered by using the bus lines, which answers research question three. These percentages are also the additional capacity when the public transport is used as food distribution network.

Table 4. Overview of the results of the truck and bus analysis

		Truck		Bus		Percentage demand that can be delivered by bus
		Demand	Number of Rides	Demand	Number of Rides	
Supermarket chain	Aldi	2251	84	1470	212	65%
	Lidl	3884	147	2740	395	71%
	Netto	2924	111	2233	315	76%
	Edeka	2951	102	2165	268	73%
	Rewe	1747	64	942	131	54%
	Total	13757	508	9550	1321	69%

4. Conclusions

In this chapter the results from the previous chapter will be explained and conclusions will be summarized.

The first conclusion is that on average more than 70% of the supermarkets can be reached using the bus network. This percentage is the supermarket data for the second research question. To reach 100% the bus network must be extended. Some locations in Berlin just do not have bus lines. That is also the reason that an average of 78% of the distribution centres can be used in route analysis using the bus network. This gives an insight in how the bus network is organised and which areas are not reachable with public transport, needed to answer research question two; "How is the public transport organized in the city Berlin".

Besides these unused locations, still four locations cannot be reached because of a technical error in the route calculation. The reason for this are the bus stops, drawn as loose lines next to roads, in the network. The supermarkets are connected to these lines because they are closer to the supermarket than the road. When the supermarkets are connected to the loose lines they cannot connect to the rest of the network. For that reason, the supermarkets are not reachable for the distribution vehicles.
These loose lines are mostly selected and deleted with the dangles method described in paragraph 2.3.2.2, but some lines can only be selected manually.

The second conclusion can be drawn with the data from Table 2 and Table 3 in the previous chapter. Table 2 and Table 3 show the numbers and percentages of usable distribution centres and supermarkets for the bus lines to be used in the analysis. 22% of the distribution centres are not close enough to bus lines. The percentages in the table show that all the distribution centres of Lidl and Rewe are close enough to a bus line to be used in the analysis. On the other hand, only one of the two Edeka distribution centres is reachable by bus. The average percentage of usable distribution centres lays around 78%, which means that Lidl and Rewe have a higher number than the average.

Table 3 shows that Lidl has the highest percentage of usable locations with 76% of the supermarkets, while the average percentage of supermarkets which are reachable by a bus line is 72%. Lidl (76%), Netto (75%), and Edeka (75%) have almost the same percentage of usable supermarkets.
In general, supermarkets should be reachable by bus for their consumers. This explains why the supermarkets show such high comparable percentages. They all want to be at least as reachable as the other chains, otherwise they could lose customers.

With the data from Table 4 also a few conclusions can be drawn. This table displays the food demand and the number of rides needed to deliver these goods, shown per supermarket chain.
The last row displays that total the demand that can be delivered by bus lines is on average 69%. This shows a high potential; the existing network can be very valuable for the delivery of goods. It answers research question three; "What is the additional capacity when the public transport network is used as extra supply network?".
It is still a challenging task with a lot of different parameters that must be calculated correctly, but this study gives an inside view of the possibilities. More than 50% of the food now delivered by individual truck rides can also be delivered in a multi-modal-integrated fashion saving energy and road movements. In case of Lidl, Edeka, and Netto this is even more than 70%. Netto shows with wisely chosen locations for supermarkets, that 76% of the distribution can be carried out with the use of public transport.

5. Discussion

In this chapter the findings of this study are discussed.

- First, using open data for a study is convenient, but its quality can be debatable. For this thesis almost all the data comes from OpenStreetMap. This makes it vulnerable for data faults. For example, the road or bus data can be drawn wrongly causing gaps in the network. Sometimes these faults are not filtered, which was the case with the small bus lines in the bus network. The topology of the data is checked and still this kind of error appear. For that reason, it can be said that the network is not 100% accurate.

- The second discussion point is the average food demand per person per day. The data, provided by the ZALF institute, is specific for Berlin, which gave us the opportunity to work with realistic data. However, it was hard to choose which numbers from the table (Appendix 6) should be taken as goods that are bought in a supermarket. For that reason, we chose to create an average food demand, from the numbers for men and women, and use all the available data. The average demand per person is set on 3956 grams a day (see paragraph 2.2.2.4). This includes the advised two litres of any kind of drinks. However, if people drink normal tap water their total demand is lower than estimated. Nevertheless, we chose to stick to average 3956 grams, because it was beyond the scope of this study to determine different minima and maxima. When another demand per person per day is chosen the results of this thesis will be different, because the amount of food that must be delivered to the supermarkets changes.

- Another point to discuss is the assumption that inhabitants of a "prognoseräume" district, assessment test areas, will do their groceries in the same district as where they live. This assumption gives a rough idea how many people go to supermarkets in a certain area. There is no open or governmental dataset with this information, which makes this assumption a logical step.

- Also, the total demand of a district is determined by the number of inhabitants times the average demand per person. To divide the food demand of a districts over the supermarkets, the size of the supermarkets is used. The total demand is divided by the size of the district to create a demand per square metre. Then the size of the supermarkets, in square metres, times the demand of one square metre gives the total demand per supermarket. The assumption is made that larger supermarkets sell more and have a higher demand. The relationship between large and small supermarkets is considered linear. With this method the demand of every supermarket can be determined based on only the size of the building.
  However, tinier supermarkets might attract a larger number of customers than expected, because they are location on busy locations.

- The fifth discussion point is the search distances of the models, which check if supermarkets and distribution centres are close enough to road or bus lines.
  - First, the search distance for the supermarkets and distribution centres to the roads is set on 200 metres. The reason is that most of the supermarkets and distribution centres are not exactly on the roads. Most of them have a parking lot or driveway in front of the building. And the points created from the original polygons of the buildings are placed in the centre of the polygon. This makes them located even further away for the roads. By setting the search range on 200 metres this displacement is solved.
  - Second, the distance from distribution centres to bus lines is regulated in our model. At the moment, the current bus lines are completely suitable for the purpose of food supply. Normally, if a bus line is 250 metres away from a distribution centre, it would not be used in this network analysis. But because the purpose of this research is; 'to find the possibilities of the bus system as additional distribution network', the search distance is set larger, on one kilometre. Now more distribution centres are used in the analysis. By doing so, we considered that bus routes can be changed a little in the future, in order to let them pass by distribution centres. Not the whole bus network has to be changed, only the small part close to the distribution centre.

- As earlier mentioned the average percentage of usable distribution centres lays around 78%, which means that Lidl and Rewe have a higher number than the average. This is not in line with the location of supermarkets: supermarket numbers show a different distribution. For example, the supermarket chain Rewe has an accessibility of 100% for the distribution centres, but 63% of the supermarkets is close enough to a bus line to be used in the analysis.

- Compared to the required number of truck rides the number of rides needed to deliver the goods are much higher for buses than for trucks. Obviously, the reason for this large difference is the capacity of a bus, 7 ton, in comparison with a truck, 30 ton. But if the delivery using the bus network can take care of suppling a large part of all supermarkets, it will save unneeded truck rides through the centres of large cities. Also, the environmental impact of the food supply chain will increase.

6. Recommendations

Making use of the bus network to substitute existing truck traffic is worth to be further explored. This thesis shows the possibilities of improving the food system by using existing routes. Because it uses these existing networks, the food system can be extended while no extra environmental impacts will occur. It can even replace delivery trucks which will make the food system even more sustainable. Hereby it is important to use the most environmentally friendly buses. The MAN Göppel bus is as an example how this bus could look. Today more environmental friendly alternatives are made. The use of electric buses should be considered. They can be charged at the bus stops and stations or at the drop and pick up locations of the trailers.

Introducing new transport modes will also raise all lot of new questions. For example, how should the public transport vehicles look like? Is it possible to design a vehicle that combines passenger transport and food transport? Where should the (un)loading platforms be located? What is the influence on the existing bus routes? Does the food transport has a negative effect on the public transport punctuality? What is the public opinion, vacant for such a solution?

When the idea, of using public transport as food distribution system, is going to be used, research similar to this thesis should be performed with more data of the public transport organisations and governmental institutions. Due to the more precise data it will give an even more accurate result.

Also, more precise data of the shopping locations of inhabitants in the city can be introduced. This will improve knowledge about what the demand per supermarket is. As mentioned in the discussion now the inhabitants must do their groceries in the assessment test area where they live.

Figure 25. Concept of the Robomart (Robomart, 2018a).

Besides the public transport, autonomous vehicles are seen as one of the most innovative solutions for food transport. In January 2018 an article was published about the concept 'Robomart' (Robomart, 2018; Shieber, 2018). The start-up uses a completely autonomous car, see Figure 25, to deliver food at specific locations. The costumer can follow the 'Robomart' on a map. They can click on the closest vehicle and select the products they want. It will drive to their location and the costumers can take their food out of the car.

Also, the car company Toyota came with a flexible electric automated car called: 'E-Palette', see Figure 26 (Toyota, 2018). In their examples they also use a drivable supermarket as idea. This indicates that not only upcoming businesses think about solutions on food distribution, but also the common and large companies do look at such innovations.

Figure 26. Toyota E-Palette Concept (Toyota, 2018).

Overall, this thesis gives an insight in the possibilities of using the public transport network as additional food supply network. It also raises a lot of new questions, which must be examined in future research.

References

Ben Wear. (2016). Capital Metro is hemorrhaging bus riders. *Statesman*. Retrieved from http://www.mystatesman.com/news/local/wear-pondering-cap-metro-ridership-plunge/QhYtceFqRQgpwrtG0BOCwJ/

Berliner Verkehrsbetriebe. (2016). *Geschäftsbericht 2016*. Berlin.

Blum-Evitts, S. (2009). *Designing a Foodshed Assessment Model: Guidance for Local and Regional Planners in Understanding Local Farm Capacity in Comparison to Local Food Needs Blum-evitts, Shemariah, "Designing a Foodshed Assessment Model: Guidance for Local and Regional Pla*. University of Massachusetts Amherst. Retrieved from http://scholarworks.umass.edu/theses%5Cnhttp://scholarworks.umass.edu/theses

Butcher, L. (2009). Lorry sizes and weights. *Business and Transport*.

Butler, M., & Butler, M. (2013). PDXScholar Analyzing the Foodshed : Toward a More Comprehensive Foodshed Analysis Follow this and additional works at : http://pdxscholar.library.pdx.edu/geog_masterpapers Toward a More Comprehensive Foodshed Analysis.

Cambridge. (n.d.). food miles Meaning in the Cambridge English Dictionary. Retrieved 11 October 2017, from http://dictionary.cambridge.org/dictionary/english/food-miles

Cefic, Care-Responsible, & ECTA. (2011). *Guidelines for Measuring and Managing CO 2 Emission from Freight Transport Operations* (Vol. march). Retrieved from https://www.ecta.com/resources/Documents/Best Practices Guidelines/guideline_for_measuring_and_managing_co2.pdf

Cich, G., Vuurstaek, J., Knapen, L., Yasar, A. U. H., Bellemans, T., & Janssens, D. (2016). Data Preparation to Simulate Public Transport in Micro-Simulations Using OSM and GTFS. In *Procedia Computer Science* (Vol. 83, pp. 50–57). https://doi.org/10.1016/j.procs.2016.04.098

Dani, S. (n.d.). FOOD SUPPLY CHAIN MANAGEMENT AND LOGISTICS. Kogan Page. Retrieved from https://www.koganpage.com/media/project_kp/document/food-supply-chain-management-and-logistics-sample-chapter.pdf

Department of Economic and Social Affairs - United Nations. (2014). *World Urbanization Prospects The 2014 Revision Highlights*. New Yotk. Retrieved from https://esa.un.org/unpd/wup/Publications/Files/WUP2014-Highlights.pdf

Duursma, M. (2016). Rijk en vervoerders willen ander openbaar vervoer. *NRC*. Retrieved from https://www.nrc.nl/nieuws/2016/12/14/rijk-en-vervoerders-willen-ander-openbaar-vervoer-5794730-a1536691

ESRI. (2017). Algorithms used by the ArcGIS Network Analyst extension. Retrieved 24 April 2018, from http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/algorithms-used-by-network-analyst.htm

European Environment Agency. (2008). *Occupancy Rates*. Copenhagen. Retrieved from https://www.eea.europa.eu/publications/ENVISSUENo12/page029.html

European Environment Agency. (2010). *Occupancy rates of passenger vehicles. Indicator Assessment | Data and maps*. Copenhagen. Retrieved from https://www.eea.europa.eu/data-and-maps/indicators/occupancy-rates-of-passenger-vehicles/occupancy-rates-of-passenger-vehicles-1

European Union. (2016). *Energy, transport and environment indicators. Eurostat Pocketbooks on Environment and energy* (2016th ed.). Luxemburg. https://doi.org/DOI 10.2785/33652

Eurostat Statistics Explained. (2017). Freight transport in the EU-28 modal split of inland transport modes. Retrieved 13 September 2017, from http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Freight_transport_in_the_EU-28_modal_split_of_inland_transport_modes_(%25_of_total_tonne-kilometres).png

Federal Ministry of Food and Agriculture. (2008). *National Consumption Study II*. Retrieved from https://www.bmel.de/SharedDocs/Downloads/Ernaehrung/NVS_ErgebnisberichtTeil2.pdf?__blob=publicationFile

FOODMETRES. (2015). The FP7 project 'Food Planning and Innovation for Sustainable Metropolitan Regions' (FOODMETRES). Retrieved 2 October 2017, from http://www.foodmetres.eu/

Fuchs, G. (2010). *Modal Split in European Freight Transport*. Lund University. Retrieved from http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1910255&fileOId=1910256

Geofabrik. (2017). GEOFABRIK // The Company. Retrieved 17 October 2017, from http://www.geofabrik.de/geofabrik/geofabrik.html

Göppel Bus GmbH. (n.d.). *Product information Göppel MIDI und MAXI TRAIN*. Augsburg. Retrieved from http://www.verkehrsverband-hochtaunus.de/media/Aktuelles+_+Verkehr/Buszug/Goeppel+Trains_produktinfo_ausführlich.pdf

Government Berlin. (n.d.). Der Bezirk und seine Räume - Berlin.de. Retrieved 13 March 2018, from https://www.berlin.de/ba-lichtenberg/auf-einen-blick/buergerservice/gemeinwesen/artikel.299604.php

HSL Helsinki Region Transport. (2015). *Annual Report 2015*. Helsinki. Retrieved from https://www.hsl.fi/sites/default/files/uploads/hsl_vuosikertomus_en_2015_sivut.pdf

Kampen, L. T. B. van. (2003). *Het ledig gewicht van motorvoertuigen*. Leidschendam.

King, T., Cole, M., Farber, J. M., Eisenbrand, G., Zabaras, D., Fox, E. M., & Hill, J. P. (2017). Food safety for food security: Relationship between global megatrends and developments in food safety. *Trends in Food Science & Technology Journal*, *68*, 15. Retrieved from http://ac.els-cdn.com/S0924224416303703/1-s2.0-S0924224416303703-main.pdf?_tid=950e6174-97b0-11e7-96c6-00000aab0f02&acdnat=1505217213_03006fcc10e682bfbbdaf0dfff61b577

Kloppenburg, J., Hendrickson, J., & Stevenson, G. W. (1996). Coming in to the foodshed. *Agriculture and Human Values*, *13*(3), 33–42. https://doi.org/10.1007/BF01538225

Martin Raifer. (n.d.). Overpass turbo. Retrieved 24 October 2017, from http://overpass-turbo.eu/

Michiel Hegenaar. (2012). Afval- en bierboten tegen de files - NRC. Retrieved 22 November 2017, from https://www.nrc.nl/nieuws/2012/10/20/afval-en-bierboten-tegen-de-files-1166847-a254653

Mundler, P., & Criner, G. (2016). Encyclopedia of Food and Health. *Encyclopedia of Food and Health*, 77–82. https://doi.org/10.1016/B978-0-12-384947-2.00325-1

Neelevat. (n.d.). Transport Duitsland. Retrieved 13 November 2017, from https://www.neelevat.nl/transport-duitsland#toggle-id-8

Nicholson, C., & Young, B. (2012). *The Relationship between Supermarkets and Suppliers: What are the Implications for Consumers?* Retrieved from http://www.europe-economics.com/publications/the_relationship_between_supermarkets_and_suppliers.pdf

Pirog, R., Van Pelt, T., Enshayan, K., & Cook, E. (2001). Food , Fuel , and Freeways : An Iowa perspective on how far food travels, fuel usage, and greenhouse gas emissions. *Leopold Center for Sustainable Agriculture*, (June), 37. Retrieved from http://ngfn.org/resources/ngfn-database/knowledge/food_mil.pdf

Planbureau voor de Leefomgeving. (2014). *Concentrations within the Dutch food chain [Infographic]*. Den Haag. Retrieved from http://www.pbl.nl/en/infographic/concentrations-within-the-dutch-food-chain

Robomart. (2018). Robomart. Retrieved 13 April 2018, from https://robomart.co/

Senate Department for Urban Development and the Environment of the State of Berlin. (2013). *Berlin Traffic in figures*. Berlin. Retrieved from https://www.berlin.de/senuvk/verkehr/politik_planung/zahlen_fakten/download/Mobility_en_komplett.pdf

Shieber, J. (2018). Robomart is the latest startup to try and unseat the local convenience store. Retrieved 13 April 2018, from https://techcrunch.com/2018/01/05/robomart-is-the-latest-startup-to-try-and-unseat-the-local-convenience-store/

Statistical Office of the Federation and States. (2018). Area and population. Retrieved 12 April 2018, from https://www.statistik-bw.de/Statistik-Portal/de_jb01_jahrtab1.asp

Steven. (2016). Converting a transit agency's GTFS to shapefile and GeoJSON with QGIS – Steven Can Plan. Retrieved 24 October 2017, from http://www.stevencanplan.com/2016/02/converting-a-transit-agencys-gtfs-to-shapefile-and-geojson-with-qgis/

THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. (2015). DIRECTIVE (EU) 2015/719. *Official Journal of the European Union*, *115*, 10. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L0719&from=EN

Toyota. (2018). Toyota Launches New Mobility Ecosystem and Concept Vehicle at 2018 CES® | CORPORATE | TOYOTA Global Newsroom. Retrieved 13 April 2018, from https://newsroom.toyota.co.jp/en/corporate/20546438.html

United Nations. (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*. https://doi.org/10.1017/CBO9781107415324.004

Vuurstaek, J., Cich, G., Knapen, L., Yasar, A. U. H., Bellemans, T., & Janssens, D. (2017). GTFS Bus Stop Mapping to the OSM Network. In *Procedia Computer Science* (Vol. 109, pp. 50–58). https://doi.org/10.1016/j.procs.2017.05.294

Wascher, D., Kneafsey, M., Pintar, M., & Piorr, A. (2015). *Food Planning and Innovation for Sustainable Metropolitan Regions – Synthesis Report.* Wageningen.

Appendices

Appendix 1: Occupancy rate of trains per European country

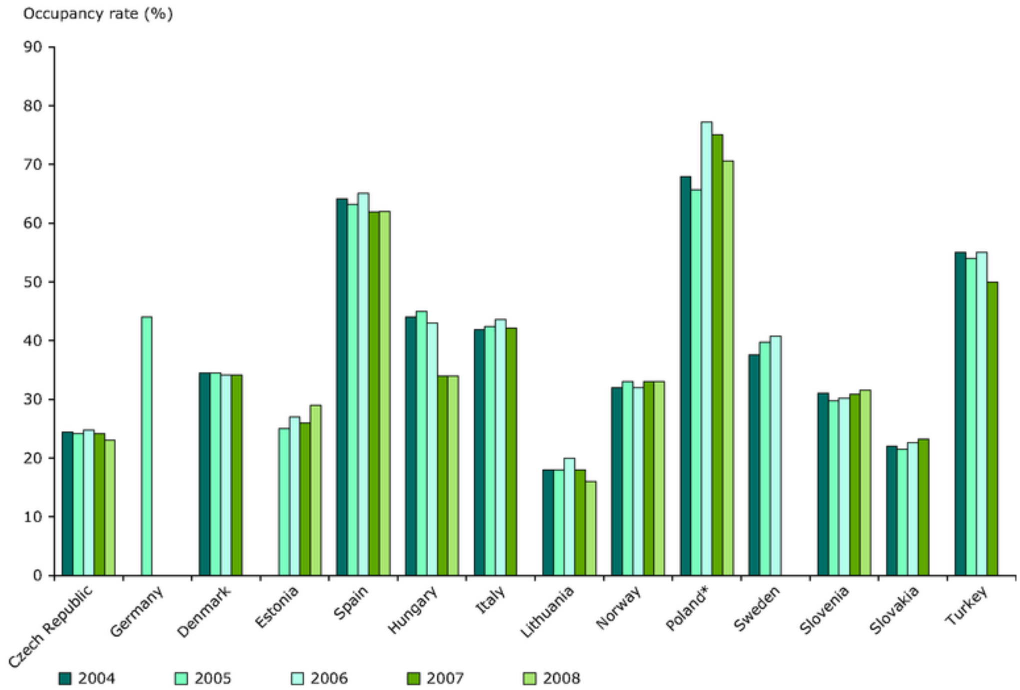


Figure 27. Occupancy rate of trains (European Environment Agency, 2010).

Appendix 2: Occupancy rate of long buses per European country

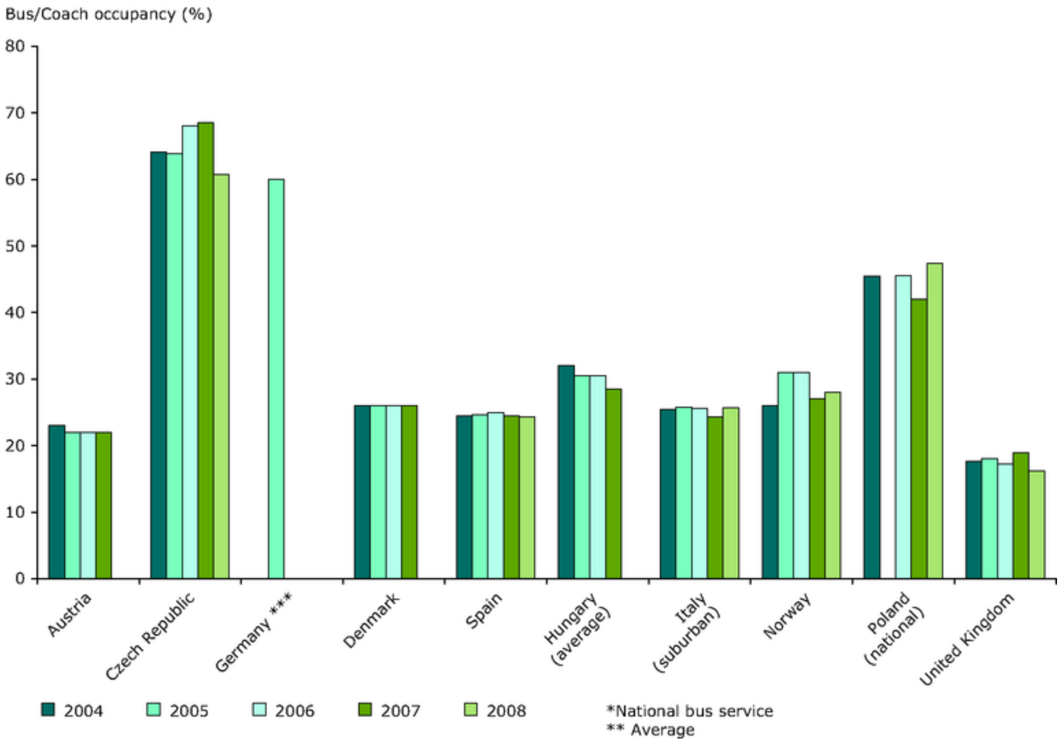


Figure 28. Occupancy rate of long distance buses (European Environment Agency, 2010).

Appendix 3: Punctuality of trains in Europe.

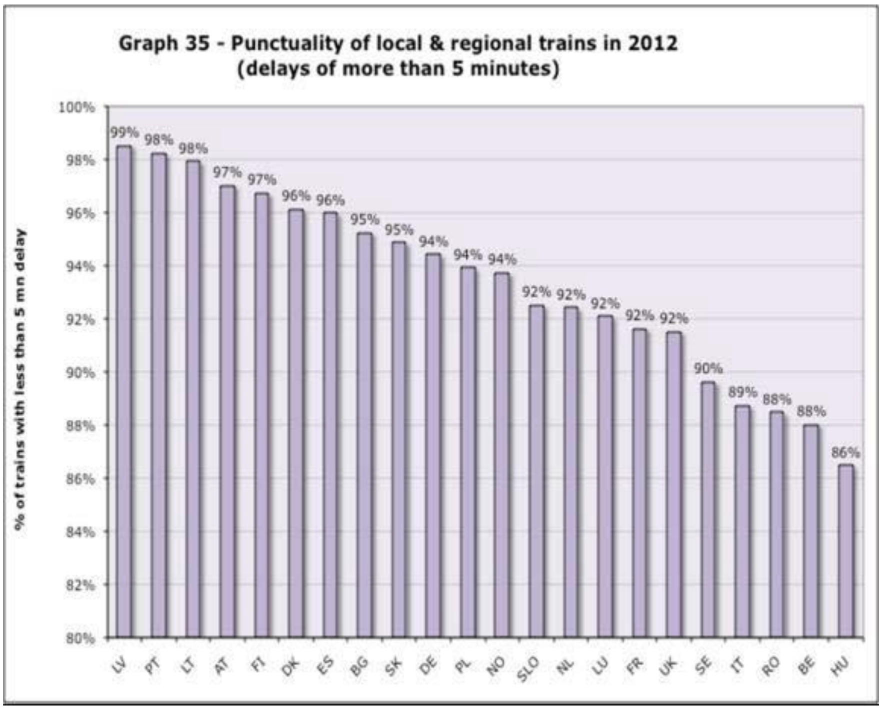


Figure 29. Punctuality of local and regional train in Europe. (European Commission, 2014).

Appendix 4: Modal split of the inland freight transport of the EU members in 2003

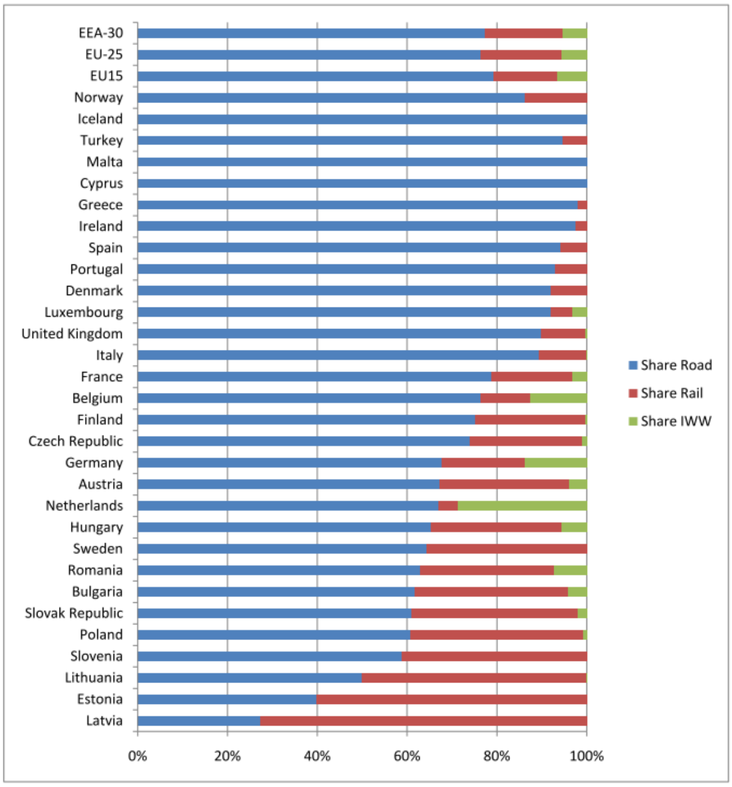


Figure 30. Inland freight transport per European country in 2003 (Fuchs, 2010).

Appendix 5: Modal split of inland freight transports for 2010-2015

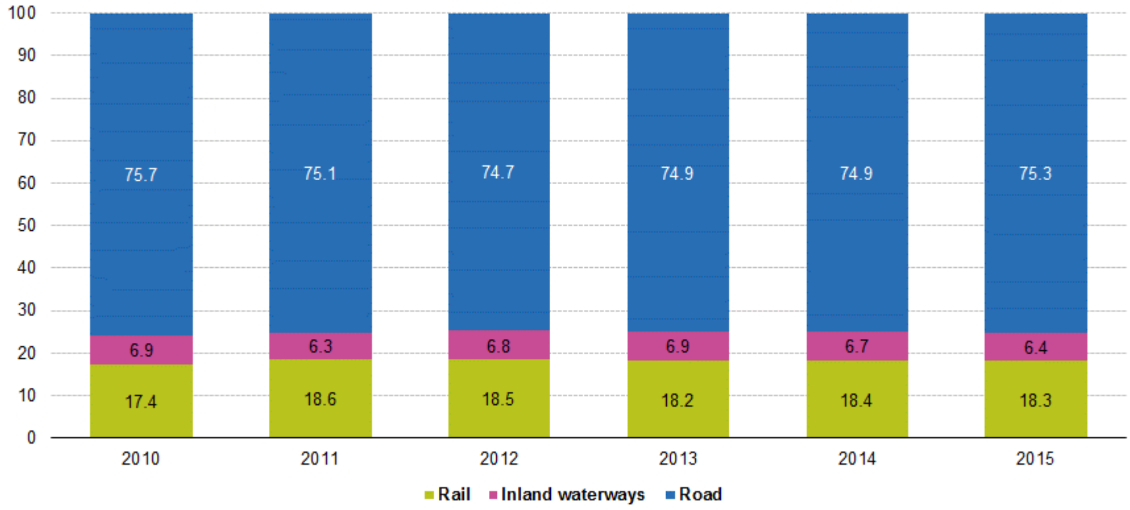
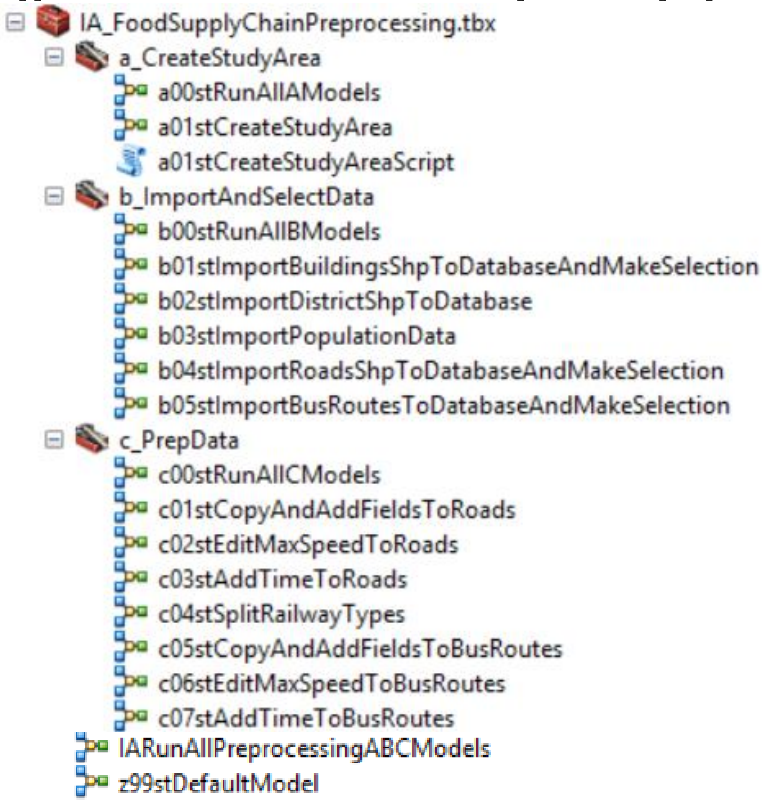


Figure 31. Inland freight transport for EU-28 (Eurostat Statistics Explained, 2017).

Appendix 6: Average food consumption Berlin

	Berlin				
	Men		Woman		Men/Woman
	Average	standard error	Average	standard error	Average
Bread	189	8,48	146	4,93	167.5
Bakery	46	3,52	33	2,00	39.5
Cereals / Products	38	3.06	32	2.04	35
Dishes based on cereals / products	53	4.13	45	2.42	49
Vegetables (or juice), mushrooms, legumes	124	6,32	149	6,87	136.5
- vegetables	118	6,14	141	6,72	-
Dishes based on vegetables	101	6,75	117	6,29	109
Potato products	77	4.02	57	2.43	67
Dishes based on potatoes	3	0,47	5	0,70	4
Fruits / products (or juice)	219	12,82	267	11,32	243
- Fruit	210	12.54	259	11.25	-
Fats and oils	37	2,00	23	0,97	30
Milk / products and cheeses	273	19,47	253	18,04	263
- Milk / mixed drinks	151	17.93	114	16.31	-
- Milk products	72	6,20	97	6,90	-
- cheese and quark	49	3,39	43	2,04	-
Dishes based on milk / products	19	2,38	14	1,44	16.5
Eggs	17	1.54	12	0.66	14.5
Eggs based on eggs	4	0,90	5	0,66	4.5
Meat / meat products and sausages	96	4,26	50	2,55	73
- Meat	44	2,90	23	1,50	-
- Meat products and sausages	52	2,90	27	1,59	-
Meat based dishes	52	3.9	26	1.55	39
Fish / Products and Crustaceans	18	1,74	14	0,94	16
Dishes based on Fish / Crustaceans	16	1,89	11	1,21	13.5
Soups and stews	109	7,13	73	4,90	91
Sweets	51	3,11	46	2,37	48.5
- Sweets	23	2,12	24	1,94	-
					-
Snack products	7	1,11	7	1,14	7
Water	1053	51.74	1083	47.78	1068
Coffee and tea (green / black)	605	33,84	588	26,62	596.5
Herbal and fruit tea	181	21.96	319	25.59	250
Fruit juices / nectars	271	24.15	213	16.51	242
Lemonades	170	24.59	89	16.4	129.5
Beer	260	28.65	40	6.75	150
Wine, sparkling wine	52	6,24	44	4,06	48
Spirits	4	1,16	1	0,45	2.5
Other (eg alcopops, alcoholic cocktails)	2	1,10	3	0,62	2.5
Total food consumption per day					**3956**

Appendix 7: Overview of the models and scripts used for pre-processing the data.



Appendix 8: Toolbox for research question one

- IA_FoodSupplyChainRQ1.tbx
  - d_FoodLocations
    - d00stRunAllDModels
    - d01stAldi
    - d02stLidl
    - d03stNetto
    - d04stEdeka
    - d05stRewe
    - d06stSupermarketsSize
    - d07stMergeDistributionCentres
  - e_FoodDemandDistrictLevel
    - e00stRunAllEModels
    - e01stBerlinPlanungsraumePopulation
    - e02stBerlinBezirksregionenPopulation
    - e03stBerlinPrognoseraumePopulation
    - e04stFoodDemandPerDistrict
  - f_FoodDemandPerSupermarket
    - f00stRunAllFModels
    - f01stTonPerSupermarket
    - f02stAldiDivideSize
    - f03stAldiCreateSplitSuper
    - f04stLidlDivideSize
    - f05stLidlCreateSplitSuper
    - f06stNettoDivideSize
    - f07stNettoCreateSplitSuper
    - f08stEdekaDivideSize
    - f09stEdekaCreateSplitSuper
    - f10stReweAldiDivideSize
    - f11stReweCreateSplitSuper
  - g_NetworkDatasetsRoads
    - g00stRunAllGModels
    - g01stRoadNetworkStudyArea
  - h_CreateRoutesRoad
    - h00stRunAllHModelsExceptH01
    - h01stMakeServiceAreas
    - h02stAldiTruckRidesPerServiceArea
    - h03stAldiMakeRoadRoute
    - h04stLidlTruckRidesPerServiceArea
    - h05stLidlMakeRoadRoute
    - h06stNettoTruckRidesPerServiceArea
    - h07stNettoMakeRoadRoute
    - h08stEdekaTruckRidesPerServiceArea
    - h09stEdekaMakeRoadRoute
    - h10stReweTruckRidesPerServiceArea
    - h11stReweMakeRoadRoute
  - IARunAllResearchQuestionOneDEFGHModels
  - z99stDefaultModel

Appendix 9: Toolbox for research question two

- IA_FoodSupplyChainRQ2.tbx
  - i_SelectSutibleSupersAndCentres
    - i00stRunAllIModels
    - i01stMergeSupermarkets
    - i02stDistCentreCloseToBusLine
    - i03stSuperCloseToBusline
    - i04stAldiBusDivideSize
    - i05stAldiBusCreateSplitSuper
    - i06stLidlBusDivideSize
    - i07stLidlBusCreateSplitSuper
    - i08stNettoBusDivideSize
    - i09stNettoBusCreateSplitSuper
    - i10stEdekaBusDivideSize
    - i11stEdekaBusCreateSplitSuper_TO_DO
    - i12stReweBusDivideSize
    - i13stReweBusCreateSplitSuper
  - j_NetworkDatasetBus
    - j00stRunAllJModels
    - j01stBusNetwerkStudyArea
  - k_CreateRoutesBus
    - k00stRunAllKModels
    - k01stMakeServiceAreasBus
    - k02stAldiBusRidesPerServiceArea
    - k03stAldiMakeBusRoute
    - k04stLidlBusRidesPerServiceArea
    - k05stLidlMakeBusRoute
    - k06stNettoBusRidesPerServiceArea
    - k07stNettoMakeBusRoute
    - k08stEdekaBusRidesPerServiceArea
    - k09stEdekaMakeBusRoute_TO_DO
    - k10stReweBusRidesPerServiceArea
    - k11stReweMakeBusRoute
  - IARunAllResearchQuestionTwoIJKModels
  - z99stDefaultModel

Appendix 10: Table of Content of the ZIP file

1. Report (PDF)
2. Midterm (PPT)
3. Final presentation (PPT)
4. Used datasets and created models (ArcGIS Structure)
5. Literature (Endnote xml)