MASTER

# Investigating traffic flows entering cities based on openly available data and a geographic information system
## The case of Eindhoven

Hulleman, Hugo F.

*Award date:*
2023

# Investigating traffic flows entering cities based on openly available data and a geographic information system

## The case of Eindhoven

**Master Thesis:**
7Z45M0 - Graduation project Urban Systems & Real Estate (45 ec)

**Author:**
H.F. (Hugo) Hulleman – 1244239

**Date of the final Colloquium:**
5 July 2023

**Graduation Supervision Committee:**
Committee chair: Prof. Dr. T.A. (Theo) Arentze
1st supervisor: Dr. Ing. P.J.H.J. (Peter) van der Waerden
2nd supervisor: Ir. V. (Valeria) Caiati

*Statement public availability*
*This Master thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Integrity.*

(Page intentionally left blank)

# Table of Contents

# List of Figures

# List of Tables

# Abstract

This report presents the development of an easy-to-use method based on openly available data to examine the traffic flow at entry points of a city. The study aims to address the limited knowledge regarding the number and characteristics of (car) trips using different entry points of a city by proposing a method consisting of five steps: filtering, mapping, routing, visualizing, and validating. The data set used for the method relies on national travel surveys, which are openly available in various countries. A case study conducted on the city of Eindhoven, the Netherlands, is used to assess the method's performance. While some challenges were encountered, the method can provide valuable insights into entry point usage relatively easily. Recommendations for future studies include developing a user-friendly tool and addressing data aggregation. The research enhances understanding of entry points and offers valuable information for policymaking and infrastructure planning. The ease of use should enable various parties to use the method.


**Keywords:** entry points of cities, traffic flows, policymaking, GIS, Eindhoven

(Page intentionally left blank)

## Summary

The demand for car-based transportation increases due to the growing trend of travelling longer distances and the expansion of car ownership. The increase in demand puts pressure on the existing transportation supply. Furthermore, the growth of the urban population and the trend towards denser cities specifically affect the transportation supply within cities. However, expanding transportation infrastructure within cities may not be desirable or can be difficult due to limited space caused by existing buildings and protected green areas. Congestion problems often occur at entry points of cities, where higher capacity roads generally are connected to the limited infrastructure capacity within the cities. Therefore, this study focuses on traffic flows entering a city via entry points, aiming to provide valuable insights into traffic volume and characteristics and support the implementation of effective traffic flow policies by developing a new method.

To improve the understanding of the factors influencing transport demand and to get inside into the existing models, a literature study is conducted. Transport demand is derived from the need to fulfil activities at different destinations. The type of activity or travel purpose often determines a trip's spatial and temporal characteristics. These characteristics include, for example, origin location, destination location, the day of the week and the hour of the day. Cities generally experience higher transport demand than rural or suburban areas due to the concentration of activity locations and population density.

Transportation infrastructure provides the transportation supply. Specific infrastructure is needed to accommodate different modes of transportation. The transportation mode choice depends on several factors, including availability, cost, time, comfort, safety, and the activities planned during the trip. Multiple modes can be used for one trip, but locations to transfer need to be provided for these types of trips. Multiple routes usually exist to get to an activity location. The route choice is often based on travel time. However, other factors such as distance, directness, habit, speed, road types, and information provided by navigation services can also influence the choice.

Supply-demand mismatches can lead to congestion, delays, and bottlenecks. Bottlenecks occur at locations where multiple traffic streams converge, or capacity decreases, causing congestion problems. Addressing mismatches can be challenging to solve as the relationship between supply and demand in transportation is complex. Increasing supply, like by adding lanes to roads or increasing the frequency of public transportation, can lead to induced demand, where increased accessibility stimulates more trips than before the intervention.

Transportation demand models play a crucial role in gaining insight into supply-demand mismatches. Usually, origin-destination (O-D) matrices are determined to model the demand. O-D matrices can be determined for demand prediction and to reveal past or current demand. The four-step and activity-based models are examples of prediction models. Although these models do not provide detailed insights into entry points in an easily accessible manner, their characteristics can be leveraged to develop a method that does. The method for revealing demand depends on the type of available input data. Three data types have been identified, active solicitation, traffic counts and big data. Each type has its own positive and negative aspects. Data types can be combined in a study to make use of the positive aspects and reduce the influence of the negative aspects on the findings.

To address the research gap and develop an easy-to-use and openly accessible method for examining traffic flow at entry points by policy makers, a five-step method was designed; filtering, mapping, routing, visualizing and validating. The first step involves filtering the necessary data from a national travel survey. In this first step, the data has been split based on trip purpose, day of the week, and time of the day to get more detailed insights into the characteristics of the trips. In the mapping step, the

data is geo-coded, and the origin, entry, and destination points are determined. Next, routes between the origin and destination points are generated in the routing step, allowing for the calculation of the number of trips entering the city at the entry points. The results are then visualized to illustrate spatial differences in transportation demand and to communicate the findings visually. Finally, the methodology is validated to ensure accuracy by comparing the findings to traffic count data. The validation is intended to serve three purposes, to assess the potential biases of the survey and if the data is not overly affected by them, to check the representativeness of the weighting used in the method, and to check if assumptions made about the route choice reflect real-life. A large part of the process could be automated by using scripts. The automation is intended to make future use of the method easier.

A case study based on Eindhoven, the Netherlands, was conducted to trial the method. The developed method could be entirely performed. However, some software limitations were found during the routing and visualization steps which made the method less easy to use. Furthermore, due to a limited amount of data and the software limitations, origin and destination areas had to be aggregated into larger areas, and travel time needed to be aggregated into two categories. As a result, the insights obtained were less detailed than initially intended. The method’s usability may also be affected by the unique characteristics of the transportation network, as well as the geographical and political features of cities. Consequently, several manual actions like defining the city borders and the destination points need to be performed when using the method, impacting the ease of use. Additionally, finding adequate data for validation came with difficulties.

The developed method provides valuable insights into the traffic flow at entry points of cities, enabling an improved view of the situation at entry points. However, limitations such as the spatial and temporal aggregation, the software problems and the difficulties of finding adequate data for validation must be acknowledged when applying the method and interpreting the results. Recommendations for future research include developing user-friendly tools or interfaces to help make the needed decisions, thus easing the use of the method for policy-makers. Furthermore, exploring alternative software solutions could help to overcome the issues found. The method can contribute to more effective transportation planning and policy implementation by providing unique insights into the entry points of cities.

# 1. Introduction

## 1.1 Context

Transportation is an essential aspect of the development and welfare of cities (Herranz-Loncan, 2007; Knaap & Oosterhaven, 2011; Rietveld, 1994). Historically, the primary mode of transportation was walking. Therefore, people could not live far away from their daily jobs and travel distances needed to be low (Bairoch, 1988). Cities needed to be dense to make it possible to participate in urban life. Distances between activity locations and housing could increase when new transportation technologies arose (Harris, 2004). Especially when cars became widely accessible to people as the costs decreased and income rose, it became possible to live in rural areas and still access the amenities of the city within a reasonable time (Harris, 2004). Cars also made living in one city and working in another easier. Commuting longer distances became normalized, and thus more and longer trips were made (Davis & Boundy, 2019; Harris, 2004; Waller et al., 2021). A comparison between the population and vehicle growth in the United States is shown in Figure 1.

The increase in demand for transportation due to the mentioned changes needs to be facilitated by infrastructure. Expanding infrastructure is relatively easy outside the build-up areas; however, it is much more difficult within cities due to existing buildings and protected green space. On the other hand, it is important to consider that people need transportation to participate in activities. Most of these activities are located within cities and not around the highways. The location of activities means that people need to exit the high-capacity roads to enter the lower-capacity city street network when travelling from outside the city to activity locations inside the city.

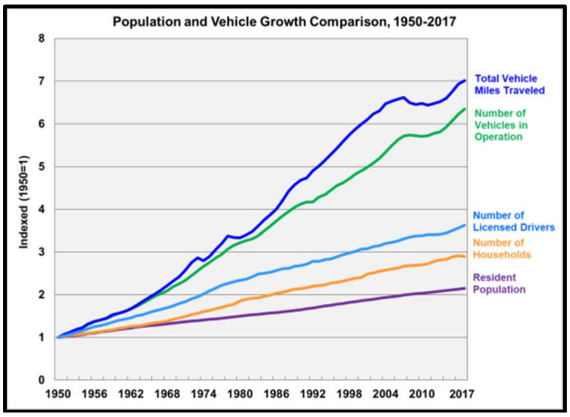



*Figure 1 Population and vehicle growth comparison in the USA, 1950-2017 (Davis & Boundy, 2019)*

People enter the city's street network at a specific location. In this study, the location where people enter is called the 'entry point', a term also used in other contexts to describe, for instance, where people enter a city centre by Borgers and Timmermans (1986). At entry points of cities, congestion problems often occur due to the high demand and the limited space to expand infrastructure. Therefore, this study will focus on traffic flows entering a city via entry points.

## 1.2 Problem analysis and research question

Combining an increased urban population, car-ownership, car-use, travel distances, and interurban travel tendencies with increasingly dense cities causes many problems in cities worldwide (OECD, 2010; Ritchie & Roser, 2018; Waller et al., 2021). These problems are diverse and numerous. However, in this section, the focus is on the issue of space-use due to car-based transportation, as this issue can clearly illustrate the concept of entry points and the importance of studying them.

Cities have limited space, and cars require a significant amount of room per person during movement and parking when not in use (Gössling, 2020; Nieuwenhuijsen & Khreis, 2016). Cities attract many trips. The high number of trips on a city's limited road network causes congestion, noise, pollution, and unsafety. Because the road network outside cities can relatively easily be expanded and the network within cities cannot, difficulties arise where the two networks meet. Problems occur when people leave the wide highways entering the city to travel to their final destination. A clear example of this problem can be observed at locations referred to as bottlenecks.  Where people enter the city, the capacity of roads generally decreases compared to the capacity of highways (Rodrigue, 2020; Wada et al., 2021). A schematic overview of a bottleneck at an entry point of a city is shown in Figure 2. In addition to the problem of bottlenecks, when people arrive at their destination with a car, the car still uses space in the form of parking. The need for car parking limits the available space for other, more beneficial land uses like housing, parks, or shops (Gössling, 2020).

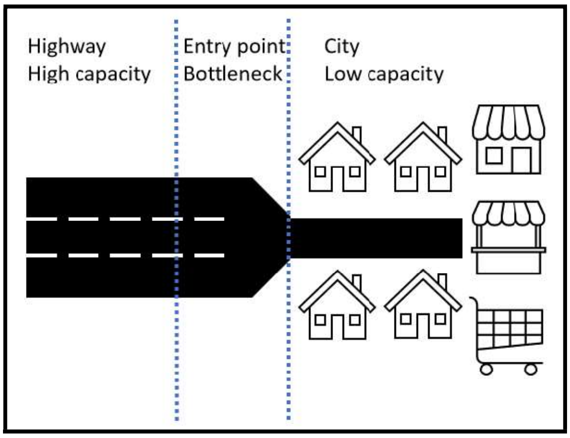



*Figure 2 Schematic view of a bottleneck at an entry point of a city*

It is widely acknowledged that cars are valuable transportation tools, but they also cause many problems. Therefore, cities worldwide are implementing policies, adding attractive alternative modes by changing or adding infrastructure and improving road networks to increase traffic flow, capacity, and safety (Gössling, 2020; Nieuwenhuijsen & Khreis, 2016). Many approaches adopted by cities focus on traffic within cities. However, cities often attract many people from outside the city. A possible method to address trips from outside cities involves constructing Park-and-Ride facilities on the city's outskirts (Karamychev & van Reeven, 2011). The use of Park-and-Ride facilities retains the benefit of cars, being able to travel from remote and scattered areas to cities while reducing the use of space where it is limited by employing more space-efficient vehicles. Park-and-Ride facilities are built where space is more abundant, and land is cheaper than in the inner city they serve. Several factors can influence the effectiveness of Park-and-Ride implementations in reducing the number of cars entering a city (Bos et al., 2003; Mingardo, 2013; Tian et al., 2017; Zijlstra et al., 2015). According to Zijlstra et al. (2015) a critical factor is the location. Therefore, it is crucial to study their usefulness at different locations. Gathering information about potential users of different entry points is therefore essential.

Traffic counts are often used as a simple and fast solution to gather insights into the number of cars on a road. However, traffic count data lacks crucial information for determining the effectiveness of specific interventions, such as information about the origin, destination and purpose of trips. For example, in a scenario where a high trip-count on a route entering a city prompts the proposal of a Park-and-Ride facility, it may be the case that most trips originate relatively close to the city, which would probably mean that improving bike infrastructure and/or local bus routes would probably provide a more effective solution. Similar scenarios can be thought of when it comes to the usefulness of knowing the destination and purpose of trips.

Methods do exist to retrieve information about the origin and destination from traffic counts, for example, in the study from Bell (1983) and Bera & Rao (2011), further discussed in section 2.7.2. However, these methods are often highly technical and will likely cost a substantial amount of time and money to implement, forming a barrier to further research. The fact that the methods are highly technical also brings that the research generally needs to be outsourced to research or consultancy firms, which typically comes with extra costs and may lead to a reduction of the influence of the policymakers over what is researched (Grijzen, 2010). Similar issues can be seen for many other methods for gaining insights into traffic flow.

The development of a new method could potentially solve part of the issues stated above. This method should be usable by policy makers and potential other parties outside of the academic field to gain initial insights into traffic flows at entry points of the city, including the origin, destination and purpose of trips. Low or zero monetary costs should be a priority to make the method accessible to use without the need for budgetary commitments. The method's output should provide a first indication of where problems like congestion and excessive land-use might occur to a person with a rudimentary knowledge of the city's transportation network. The method should also be able to be used to check whether a proposed solution to a problem makes sense at a particular location and where further research might be useful. Ultimately, the goal is to create an easy-to-use method based on openly accessible data to get initial insights into traffic flows at entry points of cities.

More detailed knowledge about people entering a city can enable more effective implementation of several policy options, including building Park-and-Ride facilities. Specifically, it is important to identify the number of people using different entry points, their origins, the purpose of travel, transport method, and arrival timing to get a broad view of the users of entry points. As far as aware, there is limited knowledge about the number and characteristics of trips entering a city and how to examine these trips in more detail. This knowledge gap leads to the research question of this report:

*How can an easy-to-use and openly available method be developed that makes it possible to examine the traffic flow at entry points of a city?*

The sub-questions are:

1. What are the externalities of car trips entering cities, and how are they caused?
2. What is the gap in knowledge in the literature when it comes to examining traffic entering cities?
3. What are the steps that need to be taken to be able to examine incoming traffic flows, and how can they be performed?
4. How and to what extent can the method be automated to improve ease of use?
5. How well does the method perform when applied to a case study in terms of the reliability of the output and in terms of accessibility, and ease of use?

The research aims to develop an easy-to-use and openly available method to determine traffic flow entering a city using publicly available travel data and widely available software. The focus is on the entry points of a city. The developed method seeks to reveal information about the number of users of different entry points beyond the findings of traffic counts, namely, the origin and destination location (aggregated), the motive for travel, when the trip took place and the route people likely travelled along. The findings should be presented to be easily communicated and used for developing policies.

## 1.3 Relevance

### 1.3.1 Societal relevance

Congestion is a widespread problem, and significant investments are being made to address this issue (Rijksoverheid, 2021; U.S. Department of Transportation, 2021). Knowledge about many distinct aspects of a city is needed to effectively invest in alleviating traffic. Incoming traffic is a significant factor in congestion. Therefore it is essential to expand the understanding of this topic. The use of open data and widely available software makes the method available to many parties. Unlike many existing methods, the method developed in this study can be used on a low budget and thus performed without a significant commitment. The use of open data reduces not only the costs but also the time it takes to perform the method, as waiting times on third parties to provide the data are minimized because no exact data requests are needed. Cities and municipalities can utilize the method as a supporting tool to substantiate policy decisions and plan future interventions.

To give a specific example, a new P+R location in Eindhoven, the Netherlands, is currently not used to its total capacity (Burg, 2021). The relatively newly opened facility in the South of the city, Genneper

Parken at the Aalsterweg, is, on average, used by 25 cars per day (according to Burg (2021) in the ED), whereas the maximum capacity equals 641 spaces. Significant investments were put into this multi-level modern facility. Several cities around the Netherlands and the rest of the world also invested in P+R locations, with mixed success stories (Dijk et al., 2013; Meek et al., 2008; Mingardo, 2013; Zijlstra et al., 2015). For example, Groningen, the Netherlands, built five separate P+R locations to lower the number of car trips into the city centre (Groningen Bereikbaar, 2021). Since opening, the most popular two P+R locations have been expanded to meet the demand. Other notable Dutch examples are in Amsterdam, Rotterdam, Maastricht and ‘s Hertogenbosch (Kennisplatform CROW, 2015). Building Park-and-Ride facilities can be expensive, land-intensive, and time-consuming (Kennisplatform CROW, 2015). Therefore, getting more insight into a Park-and-Ride facility's potential at specific locations is crucial. A logical place for a Park-and-Ride facility is at the entry point of a city, making it a relevant place to investigate the traffic flow (Zijlstra et al.,2015).

### 1.3.2 Scientific relevance

The scientific relevance of the developed method is multi-fold. Firstly, it contributes to existing traffic models by focusing specifically on the entry points of cities. While traffic flow overall has been extensively studied, the entry points of cities have received little attention. By addressing this gap, the method provides a deeper understanding of the dynamics of traffic circulation in relation to the entry points of cities.

Secondly, the method draws upon the concept of entry points from the context of shopping centres, as discussed by Borgers & Timmermans (1986). This comparison allows for a new perspective in traffic flow research and modelling. By transferring the concept of entry points to the study of city traffic, a focus is put on a specific and essential part of trips which has not been studied comprehensively.

Lastly, the method opens ways for comparative studies between different cities. By examining the structure and user characteristics of entry points in various urban areas, researchers can compare and analyse the impact of city layouts, road networks, and other factors on the formation of bottlenecks. This comparative analysis can provide valuable insights into how the design and organization of cities influence transportation patterns and help identify best practices or potential areas for improvement in urban planning and management.

Overall, the scientific relevance of the developed method lies in its contribution to traffic modelling, its novel perspective on city entry points, and its potential for comparative studies to enhance understanding and inform decision-making in the field of transportation.

## 1.4 Research design

The research design depicted in Figure 3 consists of five phases, of which some are split into several sub-phases. In the first phase, a literature study was performed to determine the knowledge gap the proposed method would try to fill. The research question and the objective are then determined based on the knowledge gap. In the next step, the method has been developed to answer the research question based on findings in the literature study. In this phase, the software programs SPSS, QGis and Excel are used. Other programs with the same functionality can be used to perform the method if these are not available to the user. The developed method is applied in a case study. Finally, conclusions are written based on the developed method and case study.

The method uses open data from the Netherlands. Therefore, it should be able to be used for any Dutch city. It may be possible to perform the method on cities outside of the Netherlands if the relevant data input is available, but this is not within the scope of the research.

## 1.5 Readers guide

The first chapter introduces the research question and problem and gives context to the question and problem. It also provides the structure of the research. In chapter 2, a literature background study is performed. The literature study concludes by answering the first two sub-questions. Chapter 3 introduces the proposed method. The method consists of five steps which are discussed in detail. For each step, it is explained how and why the step is performed, answering sub-question three. It is described to what extent the steps are automated, answering sub-question four. In the fourth chapter, the proposed method is applied to the city of Eindhoven, the Netherlands, as a case study, answering the fifth sub-question. By performing the case study, the developed method can be assessed on the reliability of the findings, the ease of use and the accessibility. In the chapter's conclusions, special attention is given to how the findings can be used to implement Park-and-Ride locations. The report finalizes with a conclusion and reflection on the performance of the proposed method and the strengths and possible limitations.

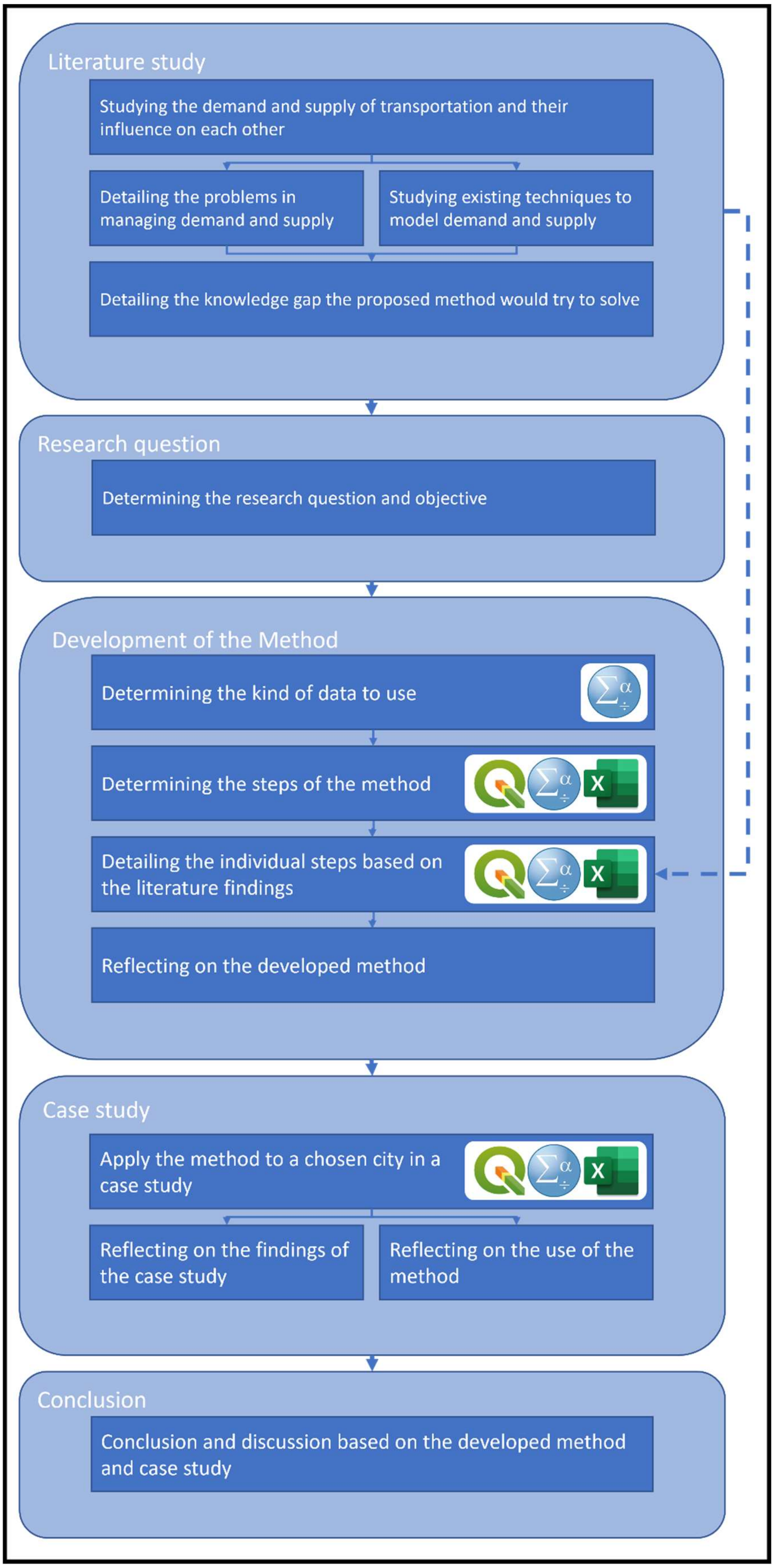


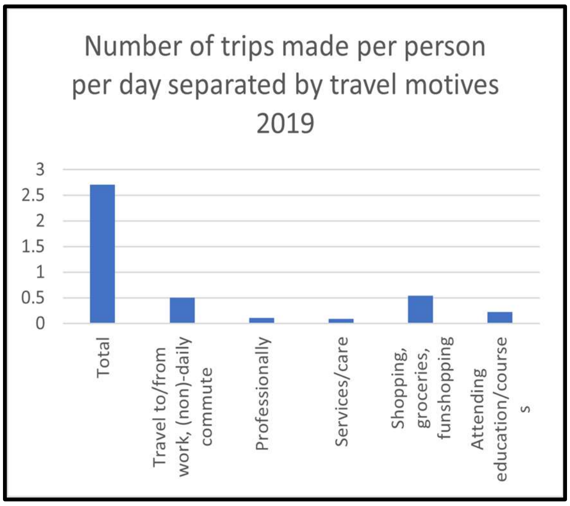

*Figure 3 The research design with an indication of the software used in specific steps*

# 2. Literature

Before proposing a new method to study incoming traffic flows at entry points, it is essential to understand why and when trips are made, what modes and infrastructure are available to make a trip, and how the trips are conducted. The first part of the literature study focusses on answering these four basic questions (sections 2.1, 2.2, 2.3, 2.4). Why and when trips are made is explained as part of the demand for transportation. The supply side explains the question of what infrastructural elements are available. Based on the demand and the supply, choices are made on how to conduct a trip. The subsequent sections give insight into how supply and demand influence each other and what problems occur when demand and supply mismatch (sections 2.5 and 2.6). The last part of the literature study focuses on existing traffic flow models. Problems in the transport network should be prevented as much as possible, and a solution must be found if they occur to keep an efficient transport network. Transportation models are created to get insights into prevention and finding solutions. There are many kinds of models available already. However, as far as can be found, there are no methods to study specifically incoming traffic flows at the entry points of a city. The several existing approaches for modelling traffic are explained in section 2.7. The chapter will end with a description of what is missing in the literature regarding the research question presented in section 1.2.

## 2.1 Demand for transportation

Transport demand can be defined as the need for people to make trips over distances to satisfy different requirements (Ertman et al., 2016). Transport demand is an example of derived demand, which means that the demand for travel arises from the need to achieve a specific purpose or activity at the destination (Bates, 2007; Bucsky & Juhász, 2022; Meyer & ITE (Institute of Transportation Engineers), 2016; Rodrigue, 2020). Thus, activities are the reason why trips are made. These activities may occur daily, weekly, or occasionally (e.g., commuting, visiting family and friends, and vacation). Therefore, trip timing depends on the activity type being carried out. As activities are the reason for travel, they may also be described as motives or purposes for travel. Examples of different travel motives and the distribution in the distance and number of trips of these motives can be seen in Figure 4. Activities differ when and where they are conducted; therefore, the overall derived demand for trips differs over time and in space (Bates, 2007; Ertman et al., 2016; Ke et al., 2017; Shen et al., 2020). Differences in demand over time and space can be called spatial-temporal heterogeneity (Shen et al., 2020). Spatial-temporal heterogeneity includes daily peak hours, differences in demand between the days of the week, seasonal factors, and differences in the urban landscape and the available transport modes (Ivanchev et al., 2015; Meyer & ITE, 2016; Shen et al., 2020). Figure 5 shows the temporal heterogeneity of trips on a typical day in the United States based on travel motives. Spatial-temporal heterogeneity can also be attributed to economic development level differences, lifestyle choices, the distribution of resources between regions and the layout of the built environment (Guan et al., 2020; Shen et al., 2020).



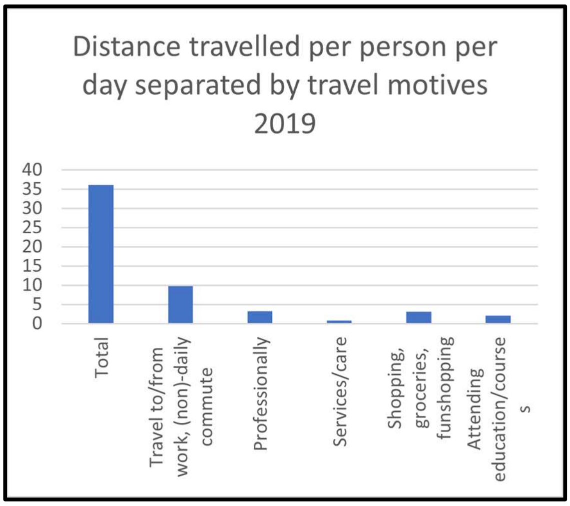




*Figure 4 Average number of trips and distance travelled per person per day in 2019 in the Netherlands separated by travel motives (CBS, 2022a)*

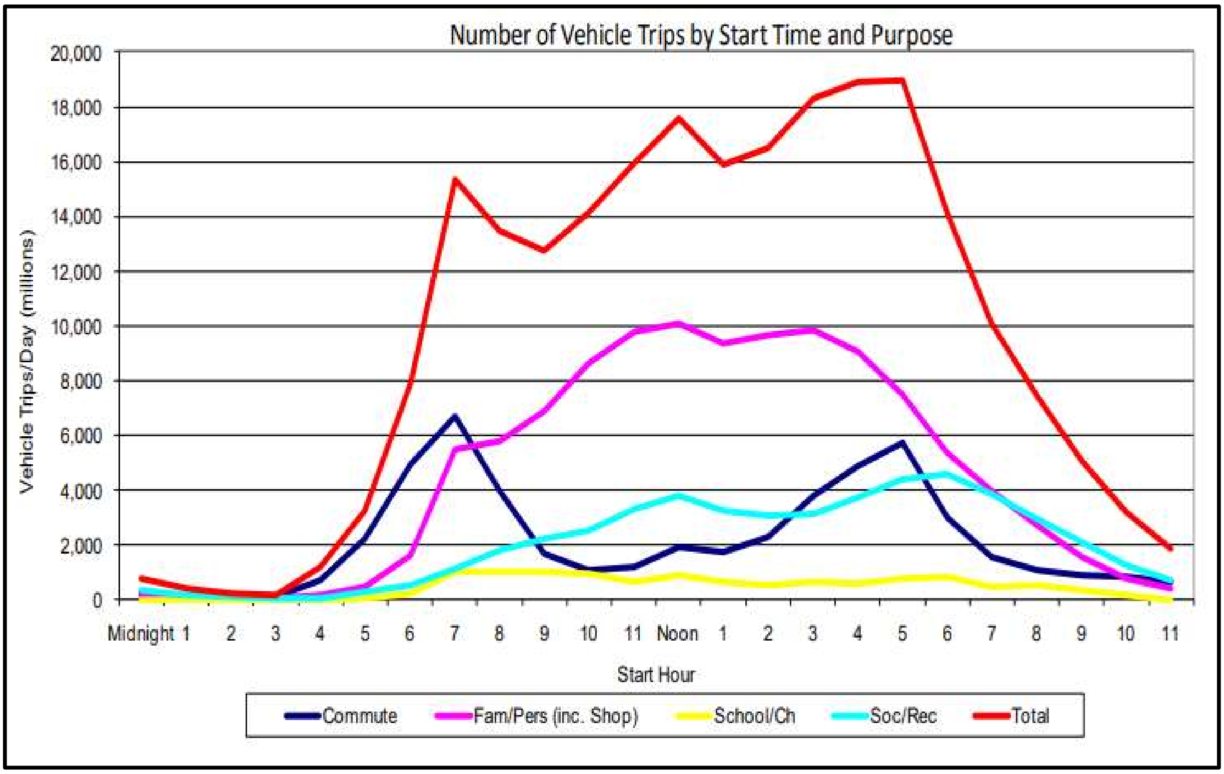



*Figure 5 Trip distribution during the day based on activities in the USA (Santos et al., 2011)*

Spatial heterogeneity is apparent when comparing cities with rural or suburban locations. In cities, more places exist where various activities can be conducted. Furthermore, location for specific activities like theatres and stadiums are often located in cities. The density of activity locations is higher than in rural or suburban areas (Cervero & Kockelman, 1997; Ewing et al., 2018; Swinney et al., 2018). Therefore, transportation demand is higher towards cities than towards rural or suburban locations. The higher population density also increases the trip demand inside towns because of the higher number of people who have a demand for trips (Ewing et al., 2018). The layout of a city has a considerable influence on transportation demand. The spatial distribution of land uses determines where demand is generated and where activities can be performed (Rodrigue, 2020). With a mix of housing, commercial, work and service buildings, distances between these locations tend to be shorter. A mix in land use means that activity locations can be reached with shorter trips and transport modes fit for short distances (e.g. walking, cycling). The land use mix thus influences the demand for transport (de Vos, 2015; Tao et al., 2020).

Temporal differences derive from the differences in starting and ending times of activities. Generally, work starts around 9 o'clock and ends at around 5 o'clock. Therefore, demand for commuting trips is high in the morning and afternoon, see Figure 5 (Santos et al., 2011). Differences between the day of the week can also be significant (Ivanchev et al., 2015). For example, fewer people work on weekends, commuting demand will be low, but more trips for other activities may be conducted.

## 2.2 Supply for transportation

There needs to be a supply of transportation to satisfy the demand for transportation. Transportation supply consists of infrastructure, vehicles, and (transit) services. Infrastructure provides the paths between the origin and destination of a trip. A plethora of modes of transportation can use these paths. Paths can be catered to one specific mode, like railways, or multiple modes, like roads. For the different modes to be used, additional amenities specific to the transportation modes must be at or near activity locations, e.g., train stations, bus stops, and car or bike parking. The supply of transportation encompasses everything that is needed to conduct a trip.

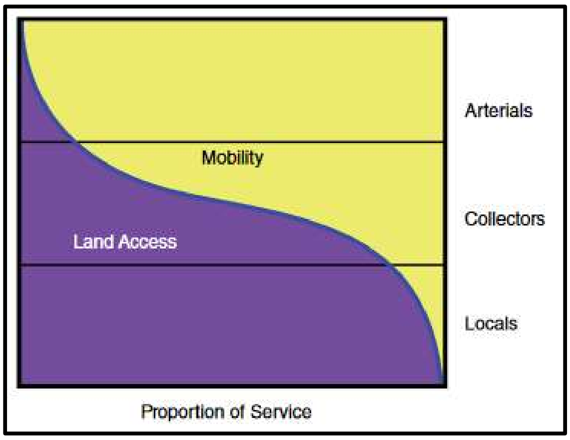



*Figure 6 Road categories and the functions (Meyer & ITE, 2016)*

Road and public transportation supply can be categorized into distinct types with specific functions (Meyer & ITE, 2016). Figure 6 shows an example of how roads can be categorized. Arterials, mostly highways and provincial roads, have mobility as their primary function. Arterials have just a few connecting roads and high speeds, resulting in a fast and efficient way to travel long distances (Meyer & ITE, 2016). Because of the higher travel speed, arterial roads are attractive even when they are not always the most direct or shortest route (Bovy & Stern, 1990; Meyer & ITE, 2016; Thomas & Tutert, 2015). To accommodate the resulting higher demand, arterials are designed with higher capacity than local roads. Collections are built to connect local roads to arterials and vice versa. Collectors are typically designed for medium speed and with more connections than arterials. Collectors connect to local streets. Local streets are where most origin and destination points are located. Public transportation can similarly be split into specific functions. Mainline railroads and long-distance busses serve as the arterials for public transport, whereas trams, metro, busses, and regional rail can be used as collectors and local services (Meyer & ITE, 2016).

## 2.3 Mode choice

When there is demand for a trip and a transport supply, a traveller needs to choose which transport mode to use. The mode choice is dependent on numerous factors. Critical considerations for this choice are the different modes' availability and costs, both real and perceived (Meyer & ITE, 2016). To choose a mode, the supply of infrastructure and service must first exist and be accessible to the user. If different transport modes are available and affordable, other aspects are considered to make a choice, e.g., time, habit, comfort, safety, distance, and other activities that will be conducted during a trip (Domarchi et al., 2008; Meyer & ITE, 2016). Household properties significantly influence the choice of automotive transport; car ownership and income are prominent factors (Shen et al., 2020). Figure

7 shows the distribution of usage of different modes by the average number of daily trips and the distance travelled per person in the Netherlands. The figure shows that the number of bike trips is almost the same as trips by car, whereas the distance covered is much smaller for bikes, indicating an influence of trip distance on the mode choice (CBS, 2022a). A person can use multiple modes during one trip, called intermodal transportation. Park-and-Ride is an example of intermodal transportation. Cars are less likely to be used in intermodal trips than trains or busses where stations or stops first need to be accessed by another mode (Pitsiava-Latinopoulou & Iordanopoulos, 2012).

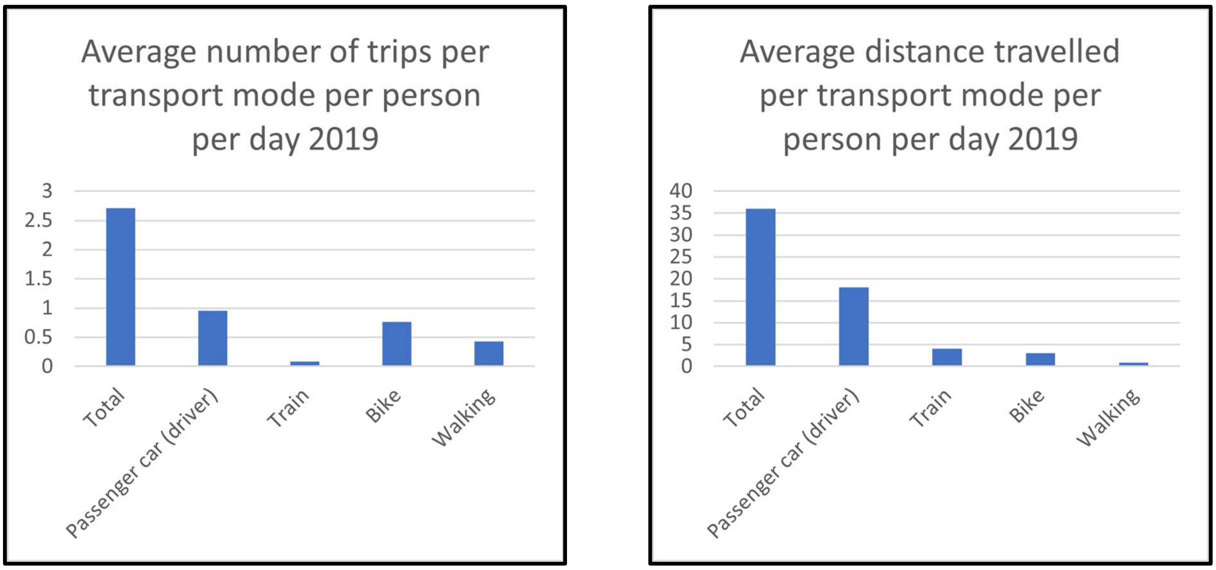



*Figure 7 Average number of trips and distance travelled per person per day in 2019 in the Netherlands separated by transport mode (CBS, 2022a)*

## 2.4 Route choice

Once a transport mode is selected, a traveller must decide on the route. Route choice depends on mode choice because different modes are linked to specific preferences and infrastructure features (Bovy & Stern, 1990). Travel time is often seen as one of the most crucial variables in choosing a route and is used to model route choice (Bovy & Stern, 1990; Meyer & ITE, 2016). Travel time is, however, not always used as the determining variable. Several variables can influence this choice, including shortest distance, directness, habit, speed, types of roads and information provided by navigation services (Bovy & Stern, 1990; Jafari et al., 2022; Stinson & Bhat, 2004; Thomas & Tutert, 2015; Wagner et al., 2021). A simple example of route options with different characteristics is shown in Figure 8. Two routes can be chosen from the origin to the destination, a direct route using local roads and an indirect route partly using a faster arterial highway. Although the second route may cover a longer distance, it might have a shorter overall travel time. The ultimate route choice depends on the traveller's preferences and the trip's circumstances(Thomas & Tutert, 2015).


*Figure 8 Schematic view of route options with different characteristics*

## 2.5 Supply influencing demand and vice versa

The demand for transportation can influence the supply for transportation and vice versa. Supply is often increased when demand is high along a piece of infrastructure. Lanes can be added to roads, and new routes can be built to complement the existing road network. For public transportation, the frequency of a line can be increased, or more lines and bus stops can be added. As an added effect of adding more infrastructure, demand can increase. A rise in demand due to increased supply is sometimes called induced demand (Hymel et al., 2010; Mokhtarian et al., 2002; Rodrigue, 2020). Induced demand can be a response to a cost reduction due to increased accessibility (Bucsky & Juhász, 2022; Rodrigue, 2020). This cost reduction includes the perceived costs of a trip, including time savings. Induced demand can happen for a specific mode, where trips would always have happened, but with another transport mode, due to spatial or temporal changes, or entirely new trip demand can be added (Mokhtarian et al., 2002; Rodrigue, 2020). Induced demand demonstrates how demand can influence supply, and its effects can be perceived as either positive or negative depending on the objective of an infrastructure project. When it comes to road widening projects, induced demand operates in opposition to the intended purpose of reducing traffic congestion. On the other hand, in case of improving bike infrastructure, resulting in an increased use of bicycles, generally aligns with the goals of such a project. In this case, induced demand is also described as “built it, and they will come” (Cervero et al., 2013; Félix et al., 2020).

## 2.6 Supply and demand mismatch

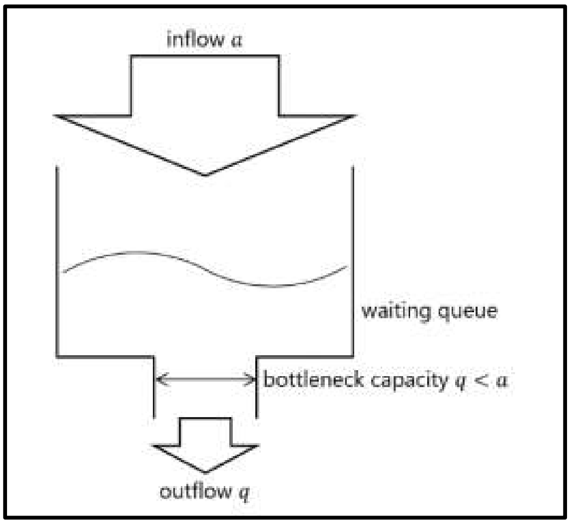



*Figure 9 Schematic view of a bottleneck (Wada et al., 2021)*

In several places, supply and demand mismatch. Mismatches can happen at specific times, like peak hour congestion or special events. An example of areas where transport demand exceeds supply regularly is called a bottleneck. Bottlenecks often occur at places where multiple streams of traffic come together (Wada et al., 2021). Bottlenecks may form at these locations as traffic streams come together at a place where the capacity is lower than the demand. Figure 9 illustrates a bottleneck; the inflow α can be one wide road or multiple roads coming together (Wada et al., 2021). Bottlenecks are of significant concern for cities as they lead to regular delays (Rodrigue, 2020). During peak hours and events, this can build up to problems affecting other parts of the system (Meyer & ITE, 2016). Roads at bottlenecks may be widened to more lanes to accommodate more traffic, but this may not always help because of the concept of induced demand explained above. Other options are to create or improve alternative modes of transport, like adding bike infrastructure, improving public transportation, or creating different routes.

Mismatched demand and supply do not always lead to delays or traffic jams. Some places, such as rural areas, may have a deficient transport supply for specific modes or in general. Basic requirements cannot be or are challenging to reach due to the lack of transport supply, leading to a difficult life. These places can be called transport deserts or, specifically for public transport, transit deserts (Jiao & Dillivan, 2013). Transit deserts create car-dependent areas, leading to significant problems (Jomehpour Chahar Aman & Smith-Colin, 2020; Rodrigue, 2020).

These are all examples of where demand outpaces supply, but supply can also outpace demand. Unnecessarily wide roads for low traffic levels or underused parking lots are examples of supply outpacing demand. Infrastructure construction and maintenance costs are generally high. Too much infrastructure at places with insufficient demand to justify the expenses leads to unnecessary costs and claims on other land uses, like greenery. In suburban areas in the United States, this has led to money problems, bankruptcies of counties and cities, and excessive land use(Burchell et al., 2005).

## 2.7 Modelling traffic demand

Transportation demand models try to minimise the problems due to supply and demand mismatches by studying or predicting the demand for transportation given a specific supply situation. In this way, the provided supply can be better fitted to the demand. To estimate which routes are used, it is necessary to know the origin and destination points of trips. For the estimation of the routes, an origin-destination (O-D) matrix is usually created (Abrahamsson, 1998; Cascetta & Russo, 1997; Egu & Bonnel, 2020). An O-D matrix contains information on the number of trips made between different zones in a region (Abrahamsson, 1998). Multiple O-D matrices can be created for the same study area to represent different characteristics, like travel purpose or transport mode (Cascetta & Russo, 1997). Several different methods exist to create an O-D matrix. O-D matrices can be determined for demand prediction and to reveal past or current demand. The four-step and activity-based models are examples of prediction models, which are discussed in section 2.7.1. Many methods exist to determine an O-D matrix for revealing past travel demand. The method for revealing demand depends on the type of available input data. Several input data options and methods are discussed in section 2.7.2. The primary method of performing route allocation is discussed in section 2.7.3. In section 2.7.4, model validation is discussed.

### 2.7.1 Predicting demand

Two major streams of demand models exist, the four-step model and the activity-based model. These methods are well-established in the literature (McNally, 2007; McNally & Rindt, 2007). The four-step emerged in the early 1950s as a trip-based method (Weiner, 2016). In the 1970s, the trip-based approach was re-considered as a method for travel forecasting as it was thought to be inadequate to accommodate the significant development occurring in urban, environmental, and energy policy (McNally & Rindt, 2007). At this time, the activity-based approach was extensively studied. Both are still in use. The four-step (2.7.1.1) and activity-based (2.7.1.2) models mainly differ in how the demand is predicted.

#### 2.7.1.1 Four-step model

The four-step model is primarily used for predicting future travel demand and the performance of the transportation supply. The model is used mainly on a regional or sub-regional scale (McNally, 2007). As the name states, the model consists of four steps: trip generation; trip distribution; mode choice; route choice, shown in Figure 10. The input for the trip generation step, named activity system in Figure 10, is typically represented by land-use, socio-economic or demographic data for several spatial units (McNally, 2007). The first step represents the demand for transportation. The next three steps are based on the transport system or transportation supply. The last step of the four-step model is route choice. A modal origin-destination matrix is loaded on the model network. An equilibrium is found based on the assumption that all paths are used to an equal resistance. After this process, the traffic flows in the used network are determined. With the flows known, it is possible to, for example, determine if the current infrastructure supply is still adequate if demand changes.

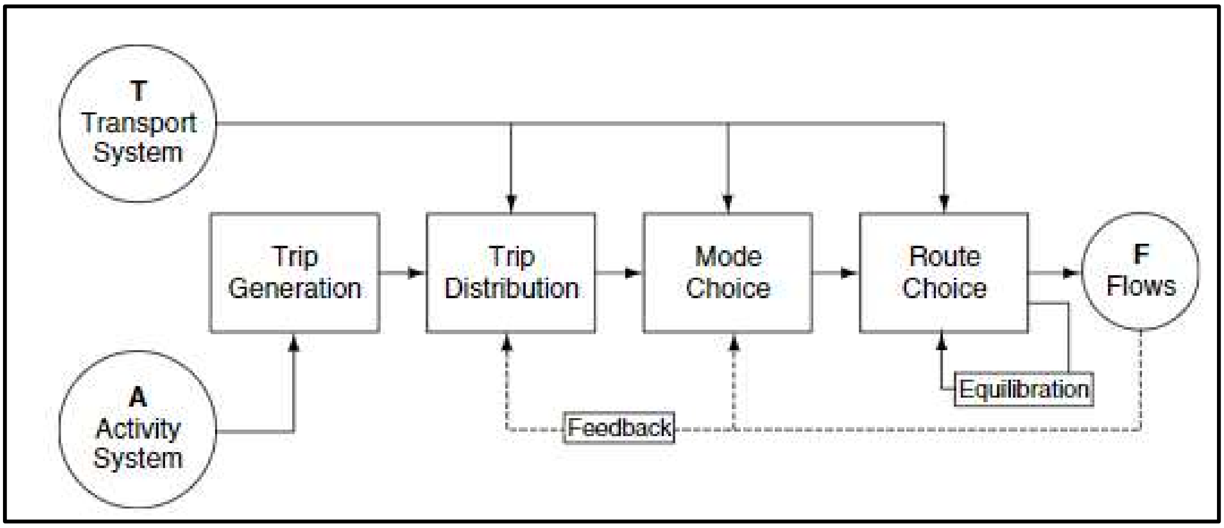



*Figure 10 The four-step model (McNally, 2007)*

Rasouli & Timmermans (2014) identify four major shortcomings with the four-step model: the models lack integrity, the assumption of independence between the four steps, the strong aggregate nature, and the lack of behavioural realism. The first two shortcomings mentioned are criticism of the linear way of working and the disconnect between the four steps. For example, the first step, where the demand is determined, is entirely independent of the transport system in the model (McNally & Rindt, 2007; Rasouli & Timmermans, 2014). Phenomena discussed in section 2.5, like induced demand, are therefore not considered in the four-step model. The strong aggregate nature of the four-step model refers to the aggregation of all origins and destinations to single points in space and the little differentiation between periods in time. Aggregation of spatial and temporal can lead to significant aggregation bias. In the trip generation step, an average of households in a particular area is taken, which is not representative of the actual households in the area, as people are fundamentally different and in different situations. Also, in the later steps, the aggregation is influential. Taking a central point as the origin and destination of a large area means that origins and destinations at the borders are likely not to be presented realistically. The fourth criticism from Rasouli & Timmermans (2014) is that the four-step models do not consider individual constraints, creating a lack of behavioural realism. For example, an individual may have limited monetary resources; the model may predict a destination outside this individual's budget.

#### 2.7.1.2 Activity-based model

The first steps in the development of the activity-based model were made by Chapin (1974) and Hägerstrand (1970). The approach was introduced as an alternative to the four-step models (Chu et al., 2012; McNally & Rindt, 2007; Rasouli & Timmermans, 2014). The activity-based approach for transportation modelling attempts to model the transportation system in much more detail than the four-step model, as it attempts to model the complexity of daily human behaviour (McNally & Rindt, 2007). The philosophy behind the activity-based approach is based on the fact that demand for transportation is derived from activity participation, as also discussed in section 2.1. Further characteristics of the activity-based approach are the use of sequences or patterns of behaviour as the unit of analysis, the influence of household and other social structures on the travel and activity behaviour, the interdependencies of spatial, temporal, transportation, and interpersonal constraints on activities and travel behaviour and the scheduling of activities in time and space (Chu et al., 2012; McNally & Rindt, 2007).

The high level of detail is cited as the most significant shortcoming of the activity-based approach. The approach is more complex than the four-step model, requiring more input data to create the model. McNally & Rindt (2007) ask whether this high level of detail is needed for the models' purposes. However, Rasouli & Timmermans (2014) argue that if the institutional constraints are not included, an activity-based model would not need much more data than a traditional four-step model. Furthermore, the calculations are much more complex than the four-step model (McNally & Rindt, 2007; Rasouli & Timmermans, 2014). The four-step and activity-based models have both positive and negative sides; therefore, it depends on the specific circumstances which model is most suitable for use.

### 2.7.2 Revealing demand

Instead of predicting travel demand, the use of revealing demand gives an indication of the actual past travel demand. There are different types of input data options for the use of this approach. The primary data types in the literature are active solicitation, traffic counts and big data. Depending on the input data, there are different methods to create an O-D matrix. These methods all come with positive and negative aspects, which are discussed below.

#### 2.7.2.1 Active solicitation

Traditionally, active solicitation like (household) travel surveys are used to determine O-D matrices (Egu & Bonnel, 2020). In these surveys, participants are directly asked about trip-related information like the origin, destination, purpose, transport mode and timing of trips. Additionally, personal information (and household information) is often included in the surveys. Because of the direct availability of crucial information, setting up an O-D matrix based on travel surveys is a relatively simple process. The findings of travel surveys can be particularly helpful for long-term transportation planning (Waller et al., 2021). However, several studies describe problems using (traditional) active solicitation survey methods. The first problem is that they rely on people's ability to precisely report their travel behaviour (Stopher & Greaves, 2007; Wolf et al., 2003). This problem could, however, be solved by the application of GPS (Wolf et al., 2001, 2004). Another significant criticism is that carrying out travel surveys is costly and generally on a small scale (Bagchi & White, 2005; Chen et al., 2016). Additionally, several studies have indicated a falling survey response rate (Bonnel, 2003; Groves, 2006; Stopher & Greaves, 2007). Although these studies are relatively old, a more recent study from Holtom et al. (2022) found a significant trend of increasing response rates on surveys in general. To summarise, active solicitation, like travel surveys, is a useful approach for collecting data to determine O-D matrices. However, producing the data comes with significant challenges that can lead to survey biases.

#### 2.7.2.2 Traffic counts

Another option to obtain O-D matrices is to use traffic count data. An early example of using traffic counts can be found in the study of Bell (1983). Traffic counts can give exact values on the number of vehicles using a specific stretch of road. The counting can be manually performed or automated, usually by induction loops. Automation allows the collection of large amounts of data with limited labour costs, although the metering infrastructure may be expensive to install and maintain (Waller et al., 2021; Zheng & Mike, 2012). The O-D matrices must be estimated because traffic count data does not directly include origin and destination information. For the estimation of O-D matrices, traffic counts of many locations are combined, where more locations mean a more reliable estimation (Bostanci et al., 2023). The estimation can be performed in numerous ways. Virtually all approaches use prior information on the O-D matrix, often in the form of an old matrix or a sample survey (Abrahamsson, 1998). The four main statistical approaches are Entropy Maximizing, Maximum Likelihood, Generalized Least Squares and Bayesian interference estimators (Abrahamsson, 1998; Bera

& Rao, 2011; Cascetta & Nguyen, 1988; Cascetta & Russo, 1997; Yang et al., 2015). Bera & Rao (2011) and Abrahamsson (1998) also discuss several other approaches to derive O-D matrices from traffic count data. The approaches are more technical than needed for active solicitation, but survey biases are avoided, and the data collection is generally easier and cheaper than traffic surveys.

#### 2.7.2.3 Big data

The use of big data for research purposes has risen in recent years as more data becomes available. Through smart devices, a large amount of data is collected, including travel data. Navigation data providers, social media platforms, mobility-as-a-service aggregators and fitness trackers can all be used to collect data on travel behaviour (Waller et al., 2021). These data sources and many others could potentially be used to derive O-D matrices. Examples of data already used in travel demand studies are smart cards, mobile network probes, social networking, automatic fare, and Bluetooth traffic monitoring data (Bagchi & White, 2005; Barcelö et al., 2010; Bonnel et al., 2018; Egu & Bonnel, 2020; Janzen et al., 2018; Waller et al., 2021; Yang et al., 2015). Difficulties in using big data arise when setting up an approach to use the data, as the collected data will be unlikely to cover the needed information to answer specific research questions (Waller et al., 2021). Additionally, the method of obtaining the data may also lead to unintended biases, for example, specific trips might be reported more often than others (Chen et al., 2016; Waller et al., 2021).

#### 2.7.2.4 Combining different data types

Because all data types have positive and negative aspects, many studies combine different data types. One way of combining data types is to derive the O-D matrix using one data type and validating the findings using another type of data, like in the studies from Bonnel et al. (2018), Bostanci et al. (2023) and Egu & Bonnel (2020). Another type of data can also be used for calibrating the findings, like in the studies from Bostanci et al. (2023) and Cascetta & Russo (1997). Combining two types of data can minimise the influence of issues like collection biases, thus improving the accuracy of studies.

### 2.7.3 Route assignment

Regardless of the model, each trip will ultimately have a determined origin and destination. The supplied transportation network links these origins and destinations together via a route. The origin and destination zones used to create an O-D matrix are often converted into centroids, representing the origin and destination points of the zones (Abrahamsson, 1998). One widely used method to determine the shortest path between origin and destination points is the Dijkstra algorithm, introduced in 1959 by Edsger W. Dijkstra (Dijkstra, 1959). The algorithm uses a network of nodes and paths, with each path having a cost connected to it (time, distance, monetary cost), as illustrated in Figure 8 (p. 22) in a simple form. Using the algorithm, the path of least resistance is determined. Different assignment methods may be used to determine the path of least resistance, like proportional assignment and equilibrium assignment. In proportional assignment, congestion is not taken into account. The 'all-or-nothing' method is an example of proportional assignment where all trips are assigned to the cost-minimising routes (Abrahamsson, 1998). For equilibrium assignment, congestion is taken into account. The cost of a link depends on the traffic volume assigned to the link. The assignment method tries to satisfy Wardrop's first equilibrium principle, where no traveller can achieve a lower travel cost by switching to another route (Wardrop, 1952). The Dijkstra algorithm may have a long computation time in large networks, like a national road network, especially when the equilibrium assignment method is used. To make the algorithm more efficient, the *'contraction hierarchies'* algorithm was introduced by Geisberger et al. (2008). It improves the calculation time by introducing (virtual) paths between major nodes, which skips all in between nodes. Calculations must be performed for every node to determine the lowest cost route. Because many nodes can be skipped when using the *'contraction hierarchies'* algorithm, a large number of calculations are not needed.

Figure 11 illustrates this with one road between the origin and destination. The road has five intersections (intersecting streets are not drawn). The cost to travel between each point is 1. To lower the number of calculations needed, node 1 is also directly connected via a virtual path with node 5; the cost is equal to the sum of the costs of the intermediate links. By using contraction hierarchies on a large scale, the number of calculations can be reduced significantly compared to the original algorithm.

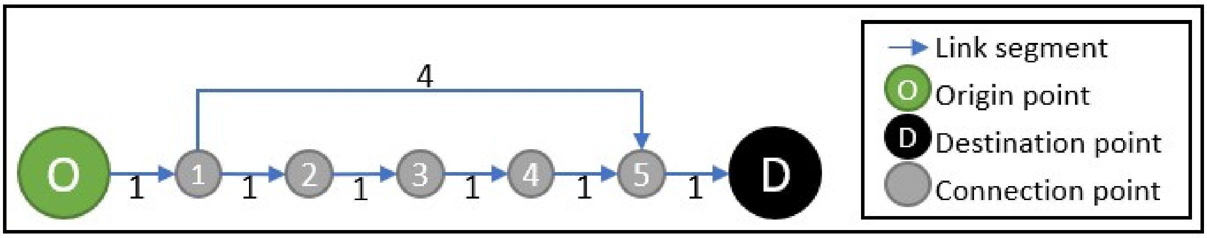



*Figure 11 Illustration of contraction hierarchies algorithm*

### 2.7.4 Validation

For the creation of any model, the last step in the process is often the validation of the model output (Aumann, 2007). Validation is critical to determine if the created model is a 'close enough' representation of the real world and can give a model more explanation power and legitimacy (Aumann, 2007; Jackson et al., 2000). The validity of a model can be tested in several ways. First, the validity of the operational and conceptual aspects of the model should be evaluated (Kerr & Goethel, 2014). Operational validation tests if the model output coincides with observed data, like in the study by Bostanci et al. (2023). The operational validation can be performed by making a statistical comparison between the output of the model and an independent data set (Kerr & Goethel, 2014). When performing the comparison, it needs to be checked if the independent data set truly represents the model's output and, if not, how the test outcome will be influenced (Rykiel, 1996). The conceptual aspects of a model can be validated by checking if the theory and assumptions on which the model is based are justifiable (Kerr & Goethel, 2014). Here it is tested if the choices made to develop the model are logical and follow the findings from the literature. A model's conceptual and operational aspects can also be assessed on the validity by the judgement of an expert. Experts in the field of study are asked to assess if the output and the concept are logical based on their knowledge and expertise (Ashcroft et al., 2016). Another part of the validation process of a model could be a sensitivity analysis. In this analysis, the effects of changing specific parameters on the overall output of the model are tested (Kerr & Goethel, 2014). By performing a validity assessment, the reliability of a model can be strengthened.

## 2.8 Conclusion

To summarize, the demand for transportation is derived from participating in activities. Cities have a high density of activity locations and thus attract many trips. Cars make it possible and convenient to travel from rural and scattered areas to activity locations in cities. Outside of cities, there is generally enough space to build enough infrastructural supply to match the demand for transportation. Within cities, space is limited because of the existing buildings and protected green space. Mismatches between supply and demand cause problems like congestion. The problems are prevalent at bottlenecks, often occurring at locations where the traffic enters the city. Models exist to predict the demand on a road network and the needed supply and traffic flow of a road network, like the four-step and activity-based models. However, these models are often difficult to create or access and do not focus on the entry points of cities. Revealing demand approaches can be used to examine the actual demand for transportation. Three data types have been identified, active solicitation, traffic counts and big data. Each type has its own positive and negative aspects. Data types can be combined in a study to make use of the positive aspects and reduce the influence of the negative aspects on the findings. To conclude, because of the problems at entry points caused by traffic entering a city, it would be beneficial to have insights into the characteristics of entering traffic. Although many methods exist to study traffic, as far as aware, no easy-to-use and openly accessible method exists which can create insights into traffic entering cities. For these reasons, it would be valuable to develop a method that can create the needed insights in an accessible manner.

(Page intentionally left blank)

# 3. Method

The developed method aims to make it possible to examine incoming traffic flows at entry points in an easy-to-use and openly accessible manner. The goal of the method is to generate initial insights into the traffic flow at entry points of cities, identify potential problems at entry points and help to identify the reasons behind problems. Additionally, the method aims to pinpoint specific entry points where interventions may be useful or most effective. The information generated from using the method should facilitate or substantiate transport policy choices taken by cities or decision-makers. A specific focus on entry points brings a new perspective to the existing literature. The method is to reveal which city entry points are used for interurban travel, how much, when, for what purpose, where the trips originate, and their destination points based on open data.

This chapter starts with explaining the method outline (section 3.1), discussing what type of data is used and introducing the steps of the developed method. The input data needed for the calculations is discussed in detail in section 3.2. In the following sections, the five steps of the method are described in detail, filtering (section 3.3), mapping (section 3.4), routing (section 3.5), visualizing (section 3.6), and validating (section 3.7). The chapter concludes with a reflection and summary of the developed method in section 3.8.

## 3.1 Method outline

An approach needs to be selected based on the goal of the method described above. Predicting and revealing demand are the two primary existing options for investigating traffic flow, as discussed in section 2.7. As the goal is to reveal information about the traffic flow, the developed method uses a revealing demand approach. Several options for input data have been discussed in section 2.7.2. To select a type of input data, the primary positive and negative aspects of each type are summarized. Traffic counts can accurately measure traffic flows at specific locations, but it is difficult to determine the origin-destination (O-D) matrix, and they do not directly contain information about the purpose of trips (Abrahamsson, 1998). Travel surveys do contain purpose, origin and destination information but are difficult to setup if not already available, and they may be prone to sampling biases and underreporting (Janzen et al., 2018; Stopher & Greaves, 2007; Waller et al., 2021; Wolf et al., 2003). Big data does provide many cases, but often vital information is missing, and collection biases are likely (Chen et al., 2016; Waller et al., 2021). The use of travel surveys would be beneficial because of the inclusion of origin, destination and purpose information. As national travel surveys are an openly available data source in several countries, and these surveys are suitable for reaching the research objective, one of the concerns of using travel surveys can be addressed. The issue of sampling bias and underreporting can be addressed by using a different data type as validation, as performed in the study by Bostanci et al. (2023).

It is chosen to base the method on a national travel survey, which is validated by another data set. Generally, a national travel survey covers a much more extensive range of trips than are relevant. Therefore, the method starts with filtering the necessary data from the larger set. The research concerns geographical information. Therefore, the data set needs to be geocoded to perform the analysis, which is performed in the mapping step using a geographical information system (GIS)-based program. With knowledge about where trips originate and their destination, routes can be created to connect these points. By generating routes, it is possible to find which entry points are used and how much. The results are then visualized to make communicating and interpreting the findings easy. Lastly, the findings are validated to strengthen the reliability of the output and to check for sampling bias and underreporting of trips. The method thus consists of five steps; filtering, mapping, routing, visualizing, and validating, illustrated in Figure 12.

The method's output is visualized in several maps, illustrating the number of trips passing the entry points of a city. Visualization is important for analysing the findings, and conveying the findings clearly to policy makers and other involved parties (Wallner et al., 2018). The maps show the spatial differences in transportation demand. Each map displays a part of the trips that entre a city based on sub-divisions determined by trip purpose, arrival period and trip day. The last two sub-divisions mentioned show the temporal differences in transportation demand. The method uses revealed and existing travel survey data and a map of the chosen country divided into subdivisions, e.g., postal code areas or municipalities. The method could be easily replicated in any city if the data is available.

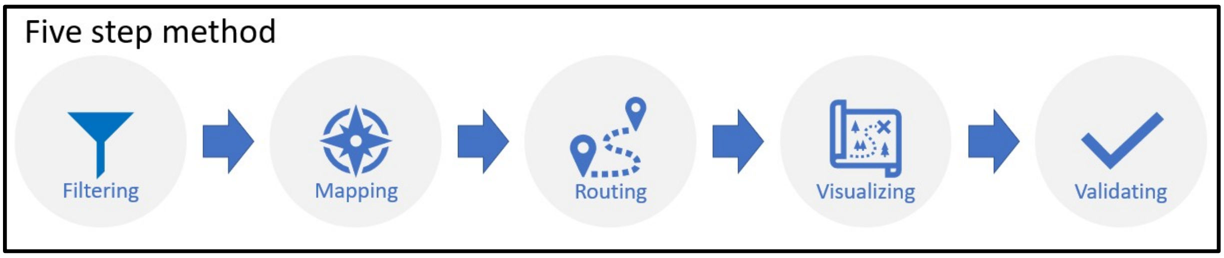



*Figure 12 Five steps of the method*

## 3.2 Input data

The data that is needed for the method is listed in Figure 13. First, a data set should be available to analyse 'city entering traffic' based on revealed data. National travel surveys are available in several countries (Ahern et al., 2013). Examples are the national travel survey of the Netherlands, UK, and Canada (CBS, 2022c; Department of Transportation UK, n.d.; Statistics Canada, n.d.). The developed method is based on the Dutch survey from CBS called ODiN (Onderweg in Nederland) (until 2017 called OVG/MON/OViN) (CBS, 2022c). This survey is conducted yearly among the Dutch population and asks them to keep a travel diary for one particular day of the year. The method described could also be applied to other travel-related data sets if it has similar information available. Essential data in the data set are the origin and destination location of trips, mode of travel, time of arrival, day of the week, travel purpose and travel distance. The survey will only cover a part of the population. If the data set also includes trip-based weights, it is possible to extrapolate the survey data to the entire population and year.

From the mapping step onwards, a map of subdivisions of the country of choice is used. These subdivisions include municipalities, counties, neighbourhoods, or postal code areas. The subdivisions used in the map also need to be available in the travel survey. As a Dutch example, the municipality map from the 'wijk- en buurtkaart' map set from CBS (2021a) can be used. Or the more detailed map of the Dutch postal code areas, also from CBS (2022a). For the routing step, a road network map is needed. OpenStreetMap is the primary provider of publicly available road network data (OpenStreetMap contributors, 2023). If available for the selected city, validation data is used in the validation step. What this could entail is described in section 3.7.

**Needed data and examples**

National travel survey
- Onderweg in Nederland (ODiN) (CBS, 2022b)
- National Travel Survey Canada (Statistics Canada, n.d.)
- National Travel Survey UK (Department of transportation UK, n.d.)

Maps of subdivisions for geocoding
- Map of municipalities
- Map of counties
- Map of postal code areas

Road network
- OpenStreetMap (OpenStreetMap contributors, 2023)

Data for validation
- Traffic count data
- Bluetooth traffic monitoring

*Figure 13 List of needed input data with examples*

## 3.3 Filtering

The first step of the method is filtering the necessary data from the travel-related data set. The necessary data in the context of the current study is explained in this chapter. Before an analytical study can be performed, a selection needs to be made of the data. Thus, the data required for this study is filtered from the complete data set to be analysed further. The filtering process is divided into three parts: selection, weighting, and subdivisions, as shown in Figure 14. Five selection criteria are discussed to find the necessary data points. Based on these criteria, a base map can be created in the visualization step of the method. A trip-based weight is added to extrapolate the data over the total population. In the last section, four subdivisions are made.

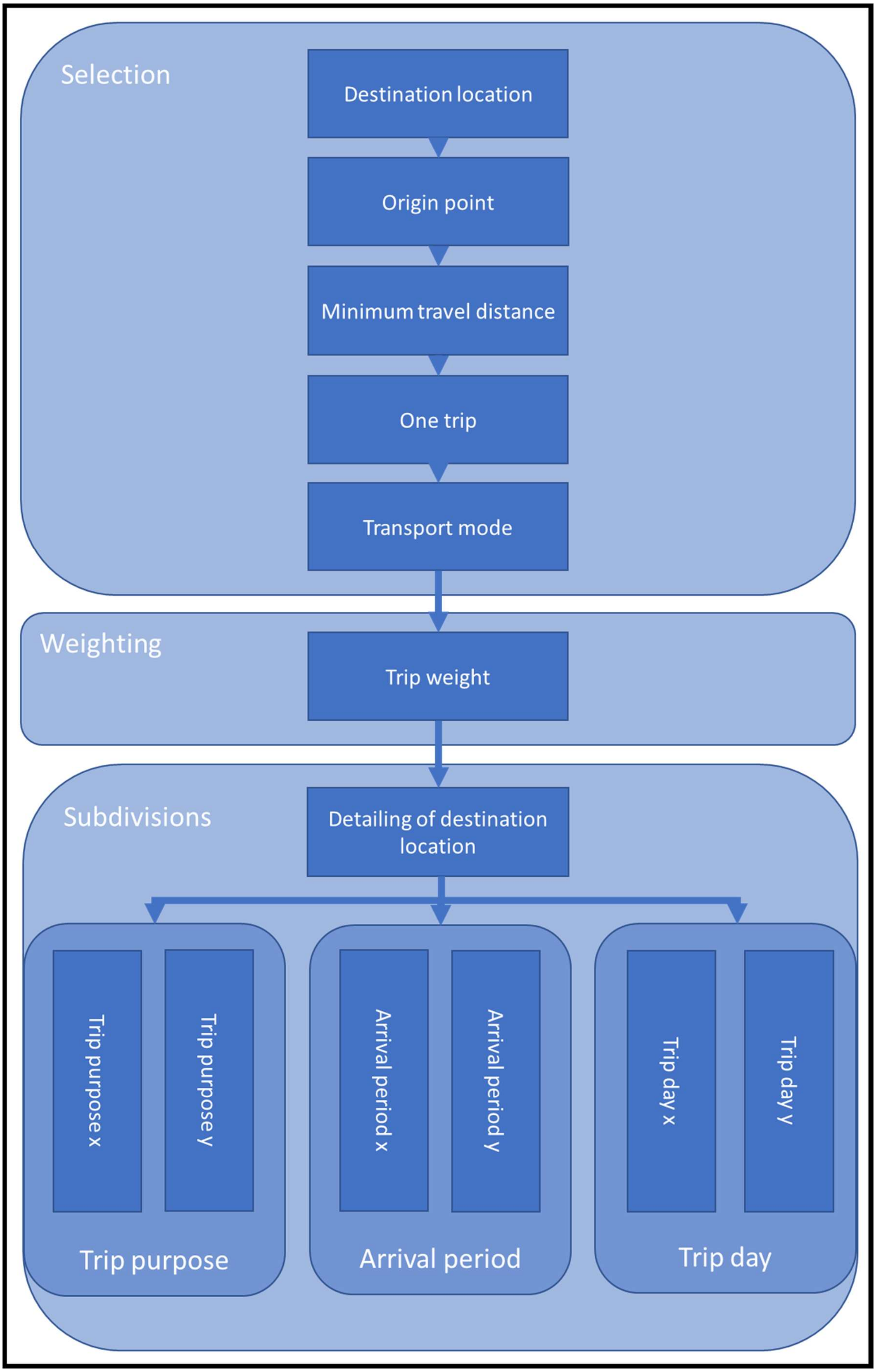



*Figure 14 Flowchart filtering step*

### 3.3.1 Selection

This phase will reduce the data set to the data points used to create the base map. A general and a study-specific formula will be given for each selection criterion. The general formula can be used when using other travel-based data sets.

#### 3.3.1.1 Destination location

The method aims to illustrate incoming traffic into a specifically chosen city. Therefore, the first step in the selection process is to select trips with their destination location inside the city. This selection process may be easy for some cities as municipality borders follow the city borders, but for others, municipal and city borders may differ. As an example, this difference is shown in Figure 15. The municipality of Groningen (the Netherlands) includes a suburban area, several villages, and a large rural area within its borders. The maximum level of detail of the ODiN database is 4-digit postal code areas; the differentiation can therefore be done on this level. How exactly the border of the city is determined depends on the chosen city and what kind of features it has. Examples of what can be used to form the border are geographical features like rivers, canals, lakes, mountains, green belts, and nature reserves. Large infrastructural features like ring roads, highways, and railroads could also be used. These border features can only be used if the postal code area borders also follow the selected feature. As an example, the postal code coloured red in Figure 15 covers a part of the city inside the highway ring road and a suburban and rural area on the other side of the highway. Therefore, in such cases, a choice should be made, exclude this area, include this area but acknowledge it when making conclusions based on the results, or include all the suburban areas, which makes the choice of border consistent. A choice must be made depending on the final goal of using this method. The rural areas and villages in separate postal code areas should always be excluded (orange in Figure 15) as these are clearly not part of the city.

The formula below should be filled in when a choice is made on which border of the destination area to take. In ODiN, the destination postal code areas are called *'AankPC'*. The *'AankPC'* variable is used in the filter to select the trips arriving in the chosen destination area, as shown in the formula below. To simplify the filter, a dummy variable is created with the chosen postal code areas indicated with a 1 and the other areas with a 0. The variable is called DestArea.

$$Area\ of\ destination = \text{"}Chosen\ postal\ code\ areas\text{"}\ (General)$$

$$(AankPC = postcode1\ OR\ AankPC = postcode2\ OR\ \ldots) \rightarrow (DestArea = 1)\ (Study\ specific)$$

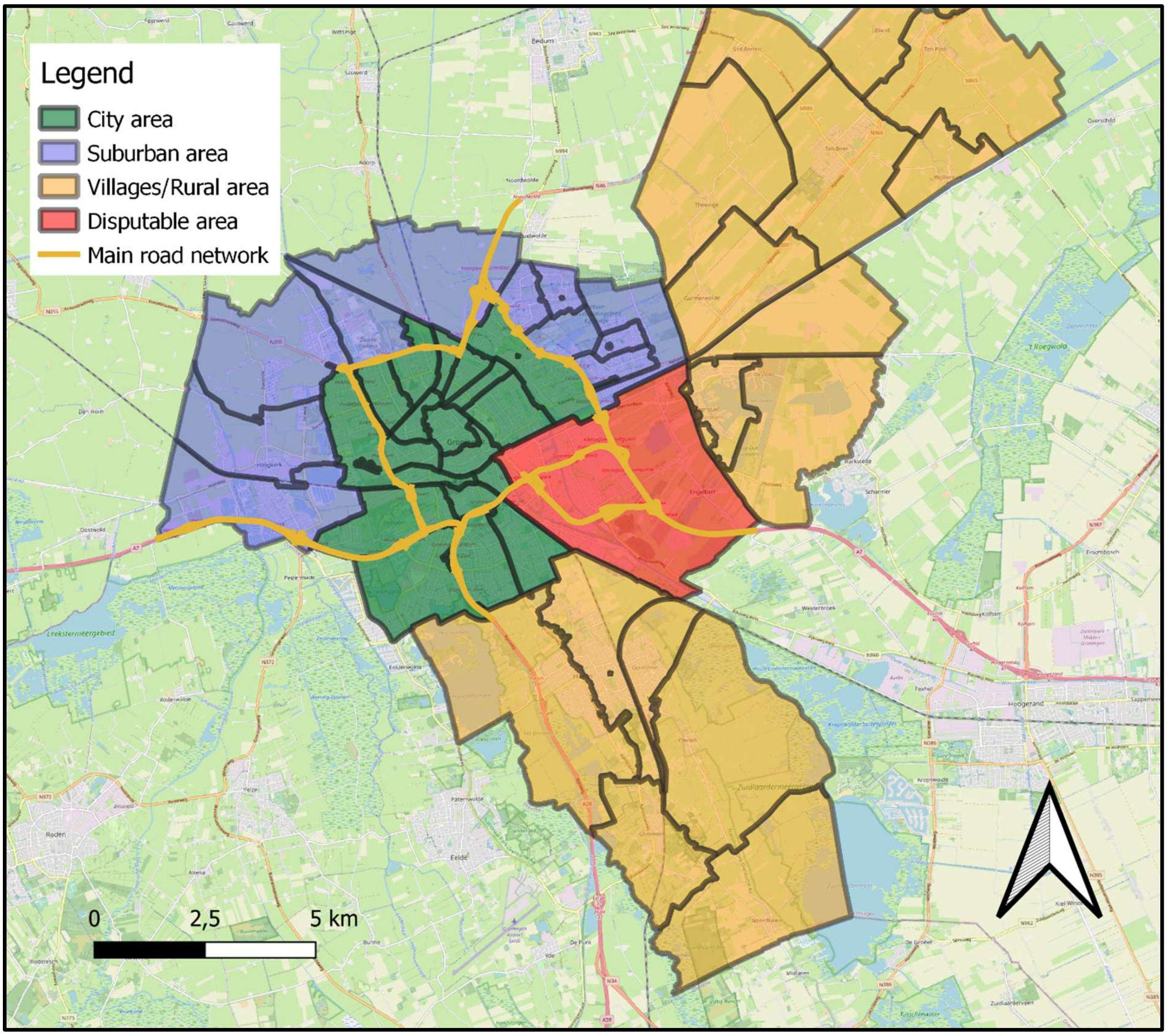



*Figure 15 The municipality of Groningen separated into three areas and a disputable area based on postal code areas and zone type*

### 3.3.1.2 Origin point

The goal of the method is to study incoming traffic into a city. Therefore, traffic originating inside the chosen city should be excluded from the data set. The area selected as the destination area is again used to define the city. To simplify the formula, the VertPC (origin postal code in ODiN) is recoded to a dummy variable where the chosen postal code areas for the city are 1 and the other values are 0. In this thesis, the variable will be called OriginArea.

$$Area\ of\ origin\ \neq \text{"}Chosen\ postal\ code\ areas\text{"}\ (General)$$

$$OriginArea = 1\ (Study\ specific)$$

### 3.3.1.3 Minimum travel distance

As seen in 3.2.1.1, determining the city borders can be arbitrary, mainly because the postal code borders do not always align with a logical physical city border. A minimum travel distance criterion is added to the selection process to create a less rigid border. This criterium ensures that the traffic included is incoming traffic, not traffic that originates just outside the chosen border. The exact distance that needs to be taken depends on the size of the selected city. To determine the minimum distance, take the distances between the centre or the central business district to the outer edges of the city and average out these distances. This average distance will be filled in the formula for x. The total trip distance is called *'AfstV'* in ODiN and is coded in hectometres.

$$Travel\ distance \geq x\ (General)$$

$$AfstV \geq x\ (Study\ specific)$$

#### 3.3.1.4 One trip

This selection criterium needs to be added because of the specific setup of the ODiN data set. The ODiN data points consist of individual trip legs. Tip legs are parts of an entire trip divided based on the transportation mode. The method will use the data regarding the trips as a whole. Each trip leg includes data for the personal, household, entire trip, and trip leg attributes. Therefore, only one trip leg is needed per trip. The *'Verpl'* variable in ODiN is a dummy variable which tells if a case is the first trip leg of a trip and can therefore be used to make use that every trip is only counted once. The following equation ensures only one trip leg remains, and thus every trip is only counted once. Because this criterium does not change depending on the chosen city or study specifics, the formula can already be filled in.

$$One\ trip = 1\ (General)$$

$$Verpl = 1\ (Study\ specific)$$

#### 3.3.1.5 Transport mode

A specific transport mode should be chosen depending on the ultimate goal of using this method. Because the choice of transport mode influences the route choice, it is important to select a mode (Jafari et al., 2022). Some modes use separate infrastructure, e.g., rail, separated bike routes, and highways. Besides the infrastructure, the type of route that should be generated also depends on what mode is used; the fastest or shortest route may make more sense (Jafari et al., 2022; Stinson & Bhat, 2004; Thomas & Tutert, 2015). For example, when the study's primary goal is to measure car flows or study where a new Park-and-Ride location should be placed, the mode of transport will be a car. Suppose the study asks for a scenario where every person would choose one particular mode of transport. In that case, no manipulation must be done in the data set regarding the transport mode, as all cases will be left in the data set. If a mode of transport is selected, an assumption is made that the primary transport mode is used for the whole trip. In particular, this is highly likely to be true for car trips, as discussed in section 2.3 (Pitsiava-Latinopoulou & Iordanopoulos, 2012). The primary transport mode, called 'Hvm' in ODiN, is thus used in the formula for the filter shown below.

$$Main\ transport\ mode = x\ (General)$$

$$Hvm = x\ (Study\ specific)$$

### 3.3.2 Weighting

All the unrelated cases are now excluded from the data set. At this point, a weight is added to the data set. The data set is based on a sample of the population. The insights generated by the method can be increased from the sample to the population level using a weight. In ODiN, three different weights are included, for households, per person and per trip. The weights show how many households, persons, or trips one case represents in the entire population (for details, see CBS (2021b)). The trip weight (FactorV in ODiN) is used as this method uses the trip data.

### 3.3.3 Subdivisions

From this point in the process, no more cases are excluded from the dataset. However, to get more detailed information, the data set will be subdivided into different categories. Based on the subdivisions, extra maps can be made which give more details on what kind of traffic uses the city's entry points. First, the destination area, at this point being the whole city, is subdivided into (sets of) postal code areas. In the last part of this section, three examples are given to subdivide the data; trip purpose, arrival period, and trip day because they align with a part of the objective for creating the method. These could be any other subdivisions if the data is available in the data set. The last three subdivisions are divided into two categories because creating more categories is likely to result in categories with not enough cases.

#### 3.3.3.1 Detailing of destination location

The city is subdivided into small destination areas to give more detailed information about the destination in the city. The extra details are important because trips might pass through different entry points depending on the exact destination location in a city. The destination areas should be as small as possible to reduce how much the trips are aggregated. For ODiN, this is to the level of postal code areas. However, as having more destination areas means more detail, it also requires more computing time and power; therefore, careful consideration must be made to decide the number of destination areas. If destination areas are aggregated, this should be done so that the choice of entry point is least affected. Therefore, much like how the city borders are divided in section 3.3.1.1, the destination areas are determined using different geographical features and important infrastructure. Higher-level roads, for example, ring roads, are another way to divide the destination areas as they could influence the route choice in a significant way, and thus also the entry location (Meyer & ITE, 2016). Disrupting waterways through a city is another way that route choice would be different depending on what side of the waterway the destination is located. The city centre or central business district should be considered one of these areas because, in most cities, this location will attract a large amount of traffic due to the density of activity locations areas (Cervero & Kockelman, 1997; Ewing et al., 2018). When the destination areas are determined, a new variable called PC is created based on the ODiN variable *'AankPC'* (destination postal code area). Each area gets a specific value. Below, an example function is given for one of the destination areas as a general and study-specific function. When the PC variable is created, the output is exported as a cross table with the municipality as rows and the PC variable as the columns. The syntax used for exporting is shown in Appendix 2.1.

$$Destination\ Area = \text{"}Postcode\ 1\text{"}\ OR\ Destination\ Area = \text{"}Postcode\ 2\text{"} \rightarrow PC = 1\ (General)$$

$$AankPC = \text{"}Postcode\ 1\text{"}\ OR\ AankPC = Postcode\ 2\text{"} \rightarrow PC = 1\ (Study\ specific)$$

#### 3.3.3.2 Trip purpose

The trip purpose is discussed in section 2.1 as the reason for the demand for transportation (Bates, 2007; Bucsky & Juhász, 2022; Meyer & ITE, 2016; Rodrigue, 2020). The trip purpose can also affect how transport policies are perceived and what the potential effects of policies are. As two examples, Andrew Kelly & Peter Clinch (2006) show a significant difference between business and non-business trips when parking costs are increased, and Azari et al. (2013) found the same difference regarding congestion pricing policies. Therefore, the dataset is subdivided based on the trip purpose. A dummy variable is made for a chosen trip purpose (e.g., work, '*Worktrip*' = 1) and the opposite of the selected purpose (e.g., non-work, '*Worktrip'* = 0). A trip purpose is determined based on the reason for conducting the study. The formula is added to the filter in the syntax shown in Appendix 2.1 to export two tables, one for trip purpose x and one for trip purpose y.

$$Trip\ purpose = x, Trip\ purpose = y\ (General)$$

$$Worktrip = 0, Worktrip = 1\ (Study\ specific)$$

#### 3.3.3.3 Arrival period

The following subdivision concerns the arrival times of trips in the chosen city. The arrival time is of interest because the usage may fluctuate depending on the time of day, adding pressure on the road network at specific times, as discussed in section 2.1 (Santos et al., 2011). The arrival time is used as the method focuses on incoming traffic into a city. The arrival time will be closer to when trips enter the city than the departure time. The arrival time included in ODiN is aggregated into two arrival periods to have enough cases in each set. These two periods of the day form a dummy variable. Examples of these periods are peak (*'Peak'* = 1) and off-peak (*'Peak'* = 0) hours or night and day trips. Again, two subsets of the data are exported for each destination area, one for arrival period x and one for arrival period y.

$$Arrival\ period = x, Arrival\ period = y\ (General)$$

$$Peak = \ 0, Peak = 1\ (Study\ specific)$$

#### 3.3.3.4 Trip day

Lastly, the data set is subdivided into which day a trip is made. Just like the time of day, the day of the week also influences traffic numbers, as explained in section 2.1 (Ivanchev et al., 2015). The most logical subdivision is to divide the dataset based on workdays and weekends, as this will likely give the most significant differences. However, other subdivisions may be of interest in specific circumstances. A dummy variable is made for the two sets of days, with as an example, workdays (*'Weekend'* = 0) and weekend days (*'Weekend'* = 1). Two subsets of the data are exported based on the chosen two groups of trip days.

$$Trip\ day = x, Trip\ day = y\ (General)$$

$$Weekend = 0, Weekend = 1\ (Study\ specific)$$

## 3.4 Mapping

After the filtering process, as detailed in section 3.3, is completed, several preparations must be made before starting the route allocation process. The exported sub-sections of the data set of all subdivisions are combined into one file. The resulting file is imported into the chosen geographical information system software (for this study QGis) and connected to geographical locations by joining it to an existing map (section 3.4.1) (QGIS Development Team, 2020). Secondly, origin points are created (section 3.4.2). Next, the destination points must be determined and mapped depending on the specificity of the destination area(s) (section 3.4.3). Finally, entry points are added to the map (section 3.4.4).

### 3.4.1 Preparing the data

The first step is to prepare the selected data from the filtering process for use in QGis. The data must be linked to geographical locations to be used in the geographical information system software. The data exported in section 3.3.3 includes the name of the origin municipalities and the number of trips originating from the municipalities. Using the municipalities' names, the exported data is linked to a map of the Dutch municipalities from the “wijk- en buurtkaart” map set from CBS (2021). How the linking of the map attributes and the data is performed can be found in Appendix 3.

### 3.4.2 Origin points

The software used for the routing step (discussed in section 3.5.1) has some constraints for the input layers. The input layers must be points, and the points must be within 350 meters of the road network accessible by the chosen transport mode; otherwise, the software will not be able to create a route. The origin municipalities will be converted into one point per municipality, which represents the origin of every trip originating in the municipality. To consistently convert the polygon-based map of the municipalities into points, the polygons are converted to their geometric centroid point. The software will move the origin points to the road network when the points are within 350 meters of the road network. If points are not within this buffer, the points need to be moved to the closest road using the Euclidean distance. Moving the point can be performed manually or using the *'snap geometries to layer'* tool in QGis by snapping the points to the road network from OpenStreetMap (OpenStreetMap contributors, 2023). An example is shown in Figure 16; here, the centroid is 700 meters away from the closest road and thus must be moved. Exceptional cases may occur where other solutions are more appropriate (see section 4.2.1 of the case study). By performing the described actions, a map is created with one point for each municipality, representing the origin point for all trips starting in the municipality.


*Figure 16 Example of moving a centroid to a location where the routing plugin works*

### 3.4.3 Destination points

The destination areas are the areas where the trips end. These areas are determined as described in section 3.3.3.1. The areas are mapped out using the postal code area map from CBS. The postal code areas that are combined in the destination areas are combined into one polygon. Because of the constraints mentioned in section 3.4.2, the created destination areas are converted into points by taking the centroids of the areas. These centroids are all moved to the closest road regardless of if they are within 350 meters of a road. These points have a significantly higher influence on the final output than the origin points, as the focus is on incoming traffic. It is, therefore, essential to place the points on the road network. Finally, the destination points are all separated into individual point layers, as this is needed in the routing step. The final product is several layers, each containing one unique destination point within the chosen city.

### 3.4.4 Entry points

The method's primary goal is to find the number of trips entering a city using a particular road. The number of trips will be counted at particular points, which in this thesis are called: entry points. On every road entering the city, an entry point is placed. 'The city' is defined as the area within the borders defined in section 3.3.1.1. The points should be located precisely on the road network used in the routing process. At some intersections, this may be a challenging process with several turning points; therefore, it is advised to use polygons or lines to cover the whole road instead of points. Instead of a line or polygon, adding the entry points after the routes are generated is also possible. Regardless of the chosen method, points should be used to visualize the final result and create a clear, easy-to-understand map. The output of this step is a layer consisting of the entry points for visualization and a layer with entry polygons or lines for the calculations in the routing step.

## 3.5 Routing

Routing is done using a route assignment algorithm (see section 2.7.3 for an explanation of route assignment algorithms). Route assignment is not a part of the QGis functionality; therefore, a plugin is used. The goal of traffic assignment is to connect the origin points selected from the data set with the determined destination points in the city via existing infrastructure.

### 3.5.1 QGis Plugin for Routing

To generate the routes, a plugin for QGis is needed. The choice for which plugin to use is based on six criteria: The software has to be still up-to-date, open-source, there needs to be an option to load a data set into the plugin in one go, it should be possible to use the traffic assignment within a script, the plugin should be able to perform many calculations at the same time, and the correct transport mode needs to be supported. Against these criteria, eight different plugins were tested. How the plugins scored on the criteria can be found in Appendix 1.

The ORS Tools plugin is the only plugin that fulfils these constraints, at least for the transport modes of car, bike, walking and using a wheelchair. It is possible to do the calculations for public transport using OpenTripPlanner if a person wants to use this method for public transport trips. However, it is not possible to automate the calculation.

The ORS Tools plugin algorithm is based on OpenStreetMap. It uses the contraction hierarchies principle to calculate the shortest distance or fastest time path between two points (Geisberger et al., 2008) (for an explanation of contraction hierarchies, see section 2.7.3).

### 3.5.2 Generating routes

Routes are generated from each origin point to each destination point using ORS Tools. The routes are created for each destination point individually, as this is needed to join the right data points to the correct routes. The routes are generated based on the assumption that people will choose the fastest route as this is more likely to be preferred than the shortest route by the population of the research, as discussed in section 2.4 (Bovy & Stern, 1990; Meyer & ITE, 2016). When using ORS Tools, all attributes from the input layers are excluded from the output. The attributes are needed, however, to know how many trips a specific route represents. An option in ORS Tools makes it possible to keep one attribute field that can later be used to join all needed attributes to the routes, solving the abovementioned problem. The municipality code attribute of the origin point is kept to be used for joining. The *‘All-by-all option’* must be chosen to create routes from all origin points to a destination point. The transport mode chosen in the selection process needs to be selected.

The data subsets exported from ODiN are joined to the routes using the municipality code included in the attribute table of the generated routes. Appendix 4 shows the script created to generate the routes automatically and join the routes to the ODiN data. The output of the route generation process is a set of line layers to every destination point and the number of trips separated in the subdivisions defined in section 3.3.3 as fields in the attribute table.

### 3.5.3 Counting entering trips

The number of trips passing a particular entry point can be extracted when the routes are generated. The extraction uses the entry points layer created in section 3.4.4 and the route generation step's final result. First, all generated route layers to every destination point must be merged into one layer. The number of trips each line represents per subcategory is summed up based on which entry point it crosses. This action is performed using the QGis tool: *'Join attributes by location (summary)'.* A script is written to automate the merging and summing process. The script can be found in Appendix 5.

### 3.5.4 Verification of the routing step

During the routing step, miscalculations may occur due to possible missing links in the underlying road network data set or possible errors when running the software. To ensure accuracy, it is essential to visually check the output and assess whether the generated routes make sense. It may be helpful to compare the routes with those of other geographic information system programs to ensure routes are correctly generated. When examining the routes, it is important to check if they primarily follow arterial roads. As for the fastest option, most of the trip is expected to be on highways (see section 2.2), which will be left only when close to the city. If the routes deviate from this expectation, the underlying road network map should be checked for any missing links or other errors. Additionally, the number of trips using an entry point can be verified. When summed, the entry numbers should be the same as the number of cases in the data set.

## 3.6 Visualizing

The goal of visualizing the output is to make it possible to visually communicate the differences in the use of the entry points, routes and where the trips originate. The visualization is done by creating a map in QGis for each subcategory determined in section 3.3.3 that conveys the revealed information to the viewer in one overview. The number of trips using a part of the road network and entry points is visualized by the thickness of the route lines. A thicker line means more users of a part of the network. The number of trips originating from a municipality is visualized using a colour scheme, with the darkest colour representing the highest number of trips. In sections 3.6.1 and 3.6.2, the needed preparations for the visualization are discussed. The prepared layers are stylized to be displayed on a map in section 3.6.3. Additionally, essential details are added to fully visualize the information generated by the proposed method, which is discussed in section 3.6.4.

### 3.6.1 Preparation of the routes

Before the information can be visualized, some preparations need to be made. The generated routes are stored as a layer consisting of continuous lines from origin points to destination points overlaying on top of each other. As routes come together from different origins, the total number of trips must be summed. To sum up the number of trips, the routes need to be cut into separate links at the intersections with other routes. The overlapping links are then summed together to find the total number of trips using a specific link. Figure 17 shows a schematic view of the output of splitting the routes into links and summing the number of trips per route link. The splitting process can be performed using the *'v.clean'* tool from the GRASS (Geographic Resources Analysis Support System) plugin using the *'break'* option (GRASS Development Team, 2016). The tool breaks a route at every intersection with another route. If the GRASS plugin causes issues because of too large an input, the *'line explode'* algorithm can be used. The algorithm takes a line and explodes it into a set of lines representing the segments of the original line. The explode algorithm is less prone to crashes but creates more segments than needed. It is, therefore, preferred to use the *'v.clean'* tool if possible. After the routes are separated into segments, the *'Join attributes by location (summary)'* algorithm used in section 3.5.3, is used again. The number of trips per subcategory is summed by location for each link. The output of this process is a line layer with route segments which includes the number of trips using that link per subcategory in the attribute table. A script is written to automate the process and can be found in Appendix 6.

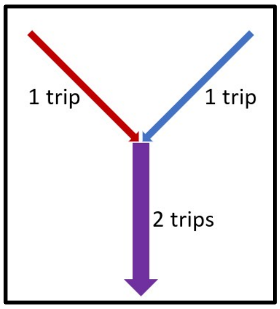



*Figure 17 Schematic view of the output of splitting and summing the route links*

### 3.6.2 Preparation of the origin areas

The origin areas are used to visualize where trips originate on the map. The data set created after the filtering step in section 3.3.3 is joined to the origin area map. The output of this step is a layer of origin areas whose attribute table contains information about the number of trips towards the chosen city for each subcategory.

### 3.6.3 Stylize the routes and origin areas

The number of trips a route segment represents can be visualized by linearly increasing the line's width. The width of the route segments, the number of classes and the range of widths used for the visualization depend on the number of cases within the output. An alternative to using line width would be 3D maps, where the number of trips is represented by the height of lines, as performed by Chapleau & Morency (2005). The goal of the visualization is to give as much detail in the maps without lines from different routes starting to overlap and still have a clear view of which roads are used.

The number of trips originating in an origin area is visualized using a sequential single-hue colour scheme to represent continuous data. Nine colour classes and a white class represent the data as this is the maximum number of colours that can be differentiated by colour-blind people, making the map accessible (Harrower & Brewer, 2003). The white class represents no trips originating from that area, and the darkest colour represents the highest number of trips. Text labels with the number of trips or 3D maps could be used as alternatives or additions to the visualization using colours if this would make the communication of the findings clearer.

The lines will be coloured orange, and the origin areas will be coloured with different hues of blue, as these colours are the best to differentiate between for the colour-blind (Katsnelson, 2021). The specific hues of blue are selected by using a website which outputs colour combinations for maps that can be distinguished by colour-blind people (Harrower & Brewer, 2003).

### 3.6.4 Map details

Several details are added to the maps after the main layers are stylized. The entry and destination points are added to the maps showing where trips are counted and where the routes terminate. This information is essential for the interpretation of the map. To finalize the visualization, the map is prepared for presentation by adding essential items, like a scale bar. An example of the output of the visualization step, developed by performing a case study, can be seen in Figure 37 (p. 61).

## 3.7 Validating

The last step in the proposed method is validating the output. Several options exist to validate the method's output, as discussed in section 2.7.4. In this chapter, the focus is put on one validation option; comparing the method’s output with an independent data set. In this process, it is tested to what extent the outcome of the proposed method mirrors the real world. This step is dependent on what kind of data exists and is available, and therefore strict steps to follow cannot be made. Different possibilities may exist for specific cities.

### 3.7.1 Options for validation

One option is to compare the outcome to other survey data, as Park et al. (2020) did in a study to improve forecasting travel demand. A problem with this approach would be that the comparison is again made against a population sample. Another option is to compare the number of trips passing an entry point, and actual traffic volume counts (Apronti & Ksaibati, 2018; Bostanci et al., 2023; Horváth et al., 2017). Many cities count traffic at specific locations using, for example, induction loops. A comparison can be made if this data is available for the studied city and multiple entry point locations. When using traffic counts, the difference in the kind of trips counted in the traffic counts and the kinds of trips that are included in the findings of the method need to be taken into account. Traffic counts count all vehicles, whereas the method's output does not include commercial trips, foreign, and local trips but does count car passengers as trips. The use of big data sources, as described in section 2.7.2.3, can also be an option for validation. If no data can be found, a last option would be to perform manual vehicle counts at entry locations. This last option makes it possible to gather information at the exact location of the determined entry points, but it is time-consuming, and errors could be made (Zheng & Mike, 2012).

### 3.7.2 Statistical tests

When relevant data for validation are available, statistical analysis must be performed to investigate if the method's output can predict the selected validation data. A simple regression analysis has to be performed to test how well of a predictor the method's output is of traffic counts. The test aims to determine how well the method's output fits real observations. In case of perfect fit, the test results will show that the data sets are equal and exhibit a $R^2$ value of 1. The prediction is expected to be imperfect because of the difference in measurements highlighted in section 3.7.1.

Before performing the test, it is essential to ensure that the assumptions for the suggested tests are satisfied. The assumptions of simple regression analysis are normality, homoscedasticity and linearity (Field, 2009; Statistics Solutions, n.d.). The assumption of normality refers that the residuals of the regression should follow a normal distribution. Normality can be tested by examining a normal Predicted Probability (P-P) plot. Homoscedasticity refers to if the residuals are equally distributed and can be checked by plotting the predicted values and the residuals on a scatterplot. Linearity refers to if the predictor variable has a linear relationship with the outcome variable. Linearity can be tested by plotting the predictor against the outcome variable.

The regression analysis can be performed if the data sets meet the assumptions. The regression model is expected to predict the validation data significantly well because both data sets measure the same phenomenon, namely the number of trips passing a specific road. The $R^2$ is not expected to be extremely close to 1 because of the likely difference in trips counted between the data sets. However, the method's output should explain a reasonable part of the variation.

### 3.7.3 Interpreting results

The significance value can tell if the model, overall, results in significantly good degree of prediction of the validation data (Field, 2009). By looking at the $R^{2,}$ the percentage of the variance in the validation data that the method's output explains can be extrapolated. An $R^2$ value close to one would indicate high prediction power, whereas a value close to zero shows low prediction power. When the $R^2$ value is low, the method's output and the validation data should be checked for errors. Furthermore, the assumptions made about the comparability of the two data sets need to be examined. The sample size will likely be low, therefore, the prediction power must be high to make reasonable claims that the model fits well. If the model does not fit well, particular care should be taken when using it to justify policy choices.

## 3.8 Discussion and Conclusion

The five steps of the method are discussed, answering the question of how a method would look like that makes it possible to examine the traffic flow from outside a city into a city. Large parts of the filtering, routing and visualizing steps are automated using Python and SPSS scripts, making it easy and fast to use the method. The mapping and validation steps have not been automated as they are specifically catered to the chosen city. All cities have a unique road layout; therefore, the mapping step consists of several manual steps, like the placement of the entry points. How to perform these manual tasks and how the locations need to be selected are described in detail in the method. How the validation is performed heavily depends on what validation data is available. Due to this dependency, it is chosen not to automate this step, and it thus needs to be performed manually. Several options for validation data are given, and a general approach is described to perform the step. A small summary is given in Figure 18, showing the automated and manual steps needed to use the method for other cities.

Several choices need to be made when performing the method, which depends on the specific layout of the city. These choices are mainly made in the filtering step, where the minimum distance of a trip is chosen, and the borders for the city are selected. The need to make these choices may make the method more complex. To help make reasonable decisions, criteria have been described in the method to help make the choices.

**Summary of the process**

| | | |
|---|---|---|
| 1. | Determine filter(s) | (manual) |
| | Export data based on filter (Appendix 2) | (automated) |
| 2. | Combine exported data and geocode to origin area (Appendix 3) | (automated) |
| | Determine origin points, destination points and entry points | (manual) |
| 3. | Generating routes (Appendix 4) | (automated) |
| | Determine number of entering trips (Appendix 5) | (automated) |
| 4. | Prepare visualizing the routes (Appendix 6) | (automated) |
| | Stylize routes and origin areas and map detailing | (manual) |
| 5. | Validation | (manual) |

*Figure 18 Summary of the automated and manual steps to perform the method*

(Page intentionally left blank)

# 4. Case study Eindhoven

A case study has been performed based on the city of Eindhoven, the Netherlands, to illustrate and trial the method described in chapter 3. Eindhoven has been selected for the case study because of its clear urban structure (see section 4.1), which helps to illustrate the method clearly. Moreover, the city has several traffic-related issues the municipality is working on solving (Gemeenteraad Eindhoven, 2013; van Hal et al., 2022). The issues in the city make it possible to show how the method's output can help with decision-making regarding transportation policy. In this chapter, first, some relevant background is given about Eindhoven. Second, the used data sets are introduced. Then the five steps of the method proposed in chapter 3 (see Figure 12, p. 32) will be implemented on Eindhoven. The chapter ends with a conclusion and discussion of the proposed method and the outcome of the case study.

## 4.1 Background of the City

Eindhoven is a highly radial city. The city has grown significantly in the car era, meaning there was space to build wide collector and arterial roads directly connecting the city centre with the highway ring at the city's North, West, and South (Oldenziel et al., 2016). An inner ring road encircles the inner city, see Figure 19. To the East, the city is bordered by a green belt, crossed at several locations by larger roads. A part of the Eindhoven municipality lies outside the highway ring, making it a separate part of the city. The postal code areas of the city roughly follow the arterial and ring roads.

The city has struggled to attract users to a newly built Park-and-Ride location at the Aalsterweg (see Figure 19) in the South of the city (Burg, 2021). The Park-and-Ride facility consists of a newly built parking garage with 641 spaces and a bus stop directly connected to the city centre and central station of Eindhoven, mostly using dedication bus lanes. The facility was built to reduce the number of cars that drive into the city to reduce traffic congestion and to offer a more sustainable alternative to driving (Gemeente Eindhoven, 2019). An unknown number of cars pass the facility when entering the city, but only a few make use of the facility. To showcase how the method could be used to study a specific location, a specific focus will be put on the Aalsterweg in the conclusion of this chapter. The goal is to give insights into the type of trips that pass the facility and to try to figure out possible reasons why the facility is not used to its potential based on the method’s output. Using the issue of the low user numbers of the Park-and-Ride makes it possible to illustrate how the method would be carried out and how the generated information can be used.

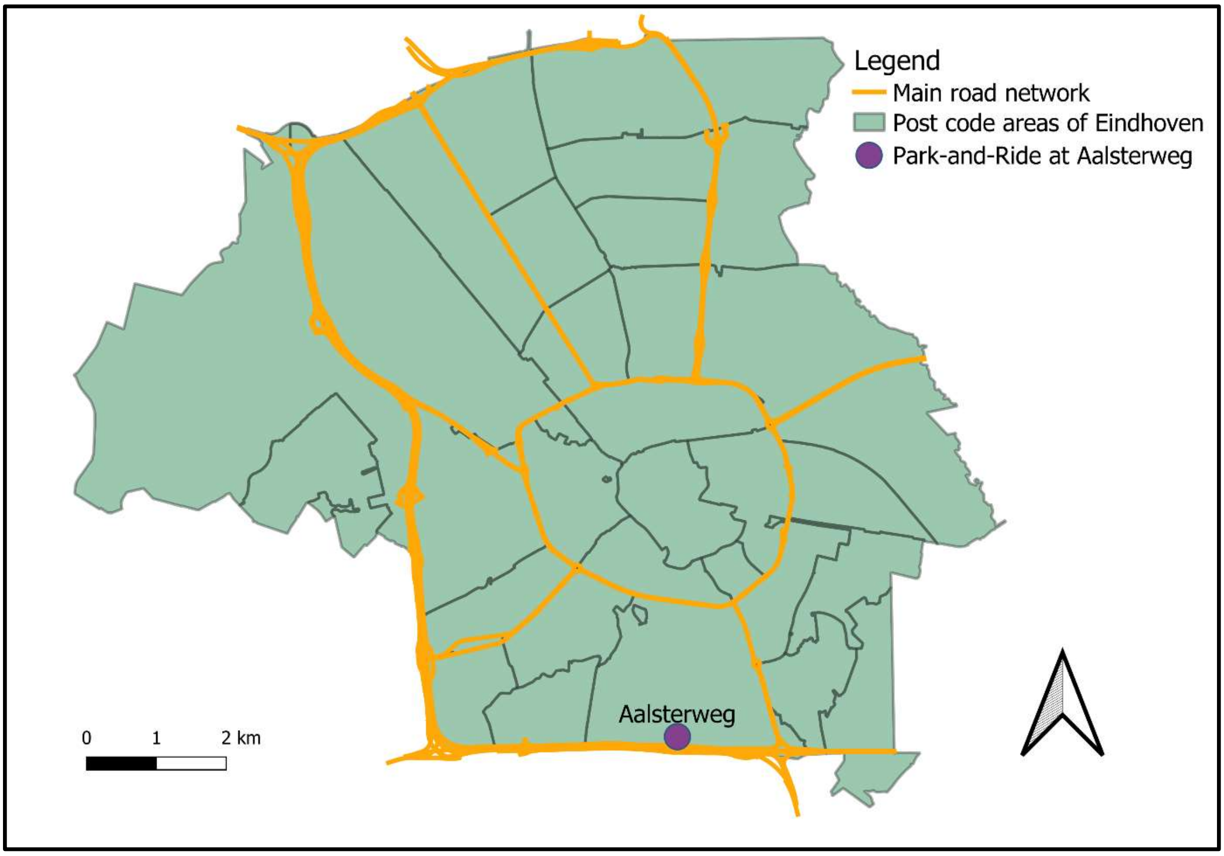



*Figure 19 Map of the layout and postal code areas of Eindhoven with the Park-and-Ride at the Aalsterweg indicated*

## 4.2 Data

A list of the data used in the case study can be seen in Figure 20. Four years of ODiN/OViN (2017, 2018, 2019 and 2020) are used to make the sample size large enough to produce reliable insights. It should be noted that 2020 had fewer trips than the other years because of Covid19 measures in the Netherlands that year. The four data sets have been combined using SPSS. The difference in coding between the years has been considered when combining the data sets. Several municipalities are merged to form bigger municipalities over the years. The changed municipalities were coded in their respective new municipalities on the first of January 2021. If a municipality was split over multiple municipalities, all cases have been added to the municipality with the highest population gain due to the merger. The “wijk- en buurtkaart 2021” municipality map from CBS (2021) is used to geocode the origin areas in section 4.4. The “numerieke postcode” map from CBS (2022a) is used for the destination areas determined in section 4.3.3. OpenStreetMap is used for the routing step and the visualization of the road network (OpenStreetMap contributors, 2023). Traffic count data is used in the validation step to be compared to the output of the case study (Dufec, 2021).

**Used data**

National travel surveys

- OViN 2017 (CBS, 2018b)
- ODiN 2018 (CBS, 2019)
- ODiN 2019 (CBS, 2020)
- ODiN 2020 (CBS, 2021b)

Maps for geocoding

- Wijk- en buurtkaart 2021 (CBS, 2021a)
- Numerieke postcode (PC4) (CBS, 2022a)

Road network

- OpenStreetMap (OpenStreetMap contributors, 2023)

Traffic count data

- Data set 1 (Dufec, 2021)
- Data set 2 (Dufec, 2021)

*Figure 20 List of data used for the case study Eindhoven*

## 4.3 Filtering

The first step in the proposed method is filtering out the relevant trips. The filtering is performed and exported to Microsoft Excel using the SPSS Syntax shown in Appendix 2.2. For the case study, the process shown in Figure 14 (p. 34) is filled in for Eindhoven, as shown in Figure 21.

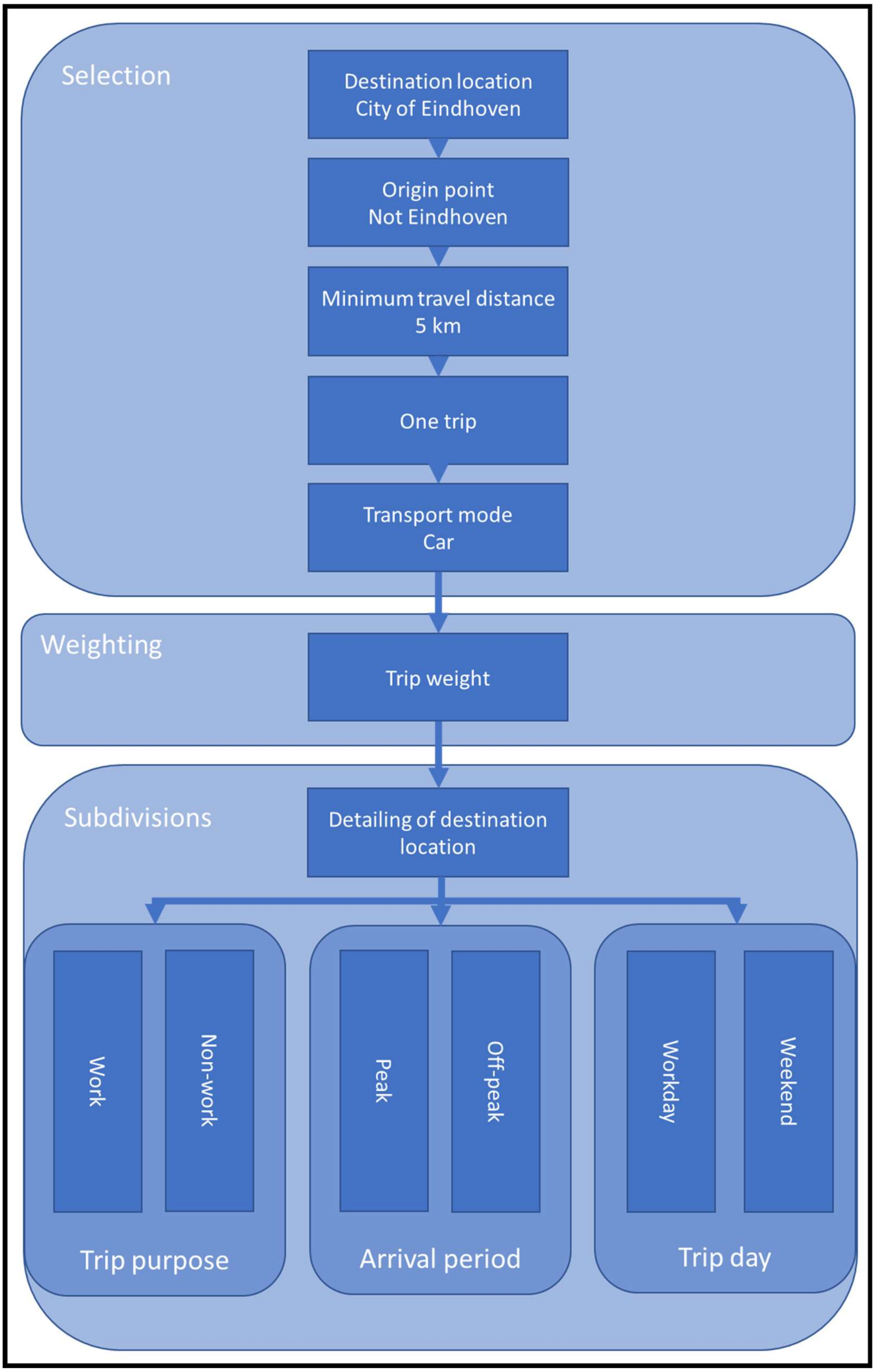



*Figure 21 Filtering step defined for the case study Eindhoven*

### 4.3.1 Selection

The relevant data is selected in the first stage of the filtering step. After each added selection criterion, the number of remaining cases is shown. In this stage, the total number of cases in the data set is reduced from 657,530 to 1,307.

#### 4.3.1.1 Destination location

First, the borders of the city are defined. Within the municipal boundaries of Eindhoven, no rural areas or separate villages, as described in section 3.3.1.1, are included, meaning that there are no areas in the municipality that are definitely not part of the city. The North and South of the municipality are bordered by the highway ring and the East by a green belt. These are logical borders of the city. In the West, however, the neighbourhood of Meerhoven and the airport are located on the other side of the highway ring. To consistently use clear city borders divided by infrastructure and nature, a choice is made to exclude areas outside the highway ring, as shown in Figure 22.

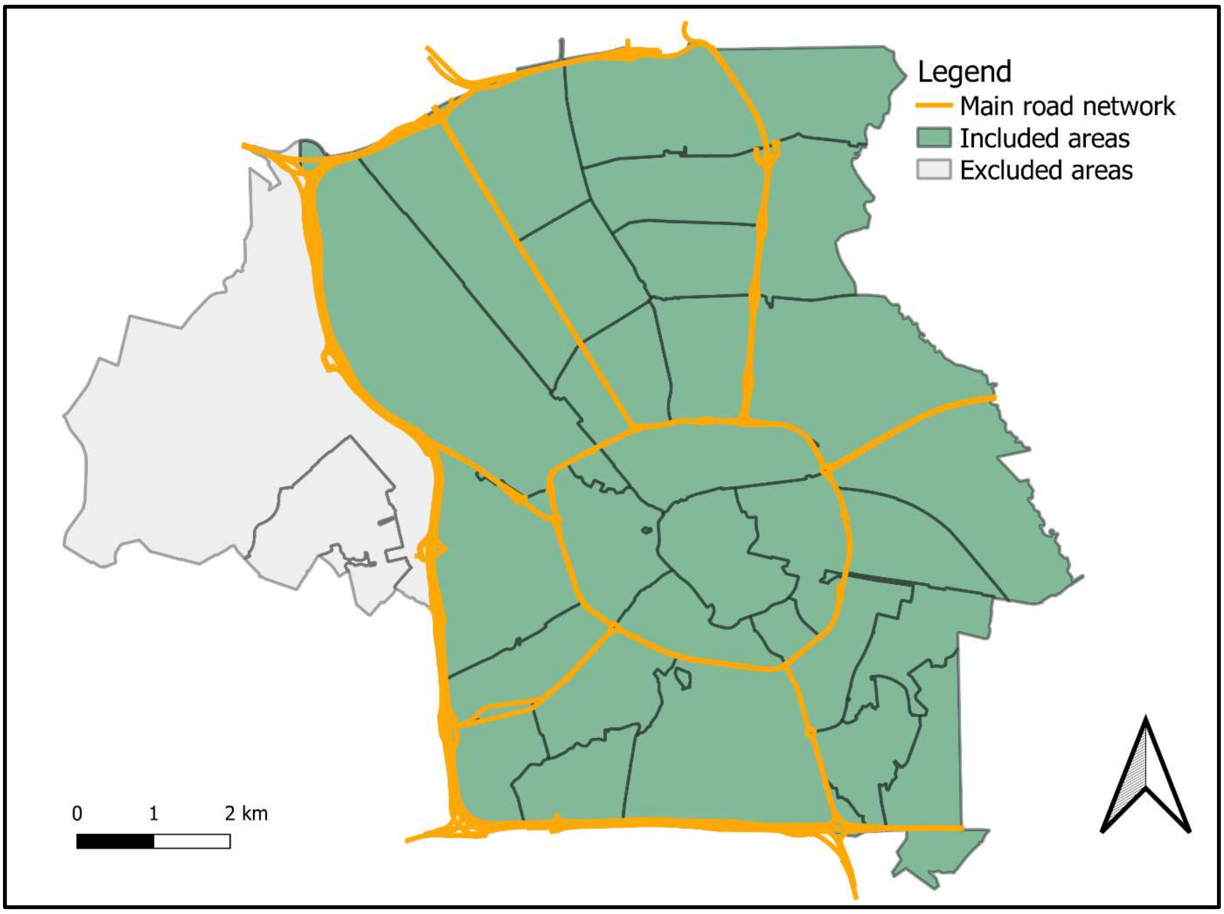



*Figure 22 Map of the selected postal code areas of Eindhoven municipality forming the destination area*

The destination (AankPC) variable of ODiN is recoded into the DestArea variable with the included postal code areas designated as 1 and all other postal codes set as 0. The formula used in the filter is, therefore: DestArea = 1. This selection significantly reduces the number of cases in the data set. A significant reduction is expected as the focus of the data set is changed from country-wide trips to specifically Eindhoven focussed trips. The process is shown in Figure 23.

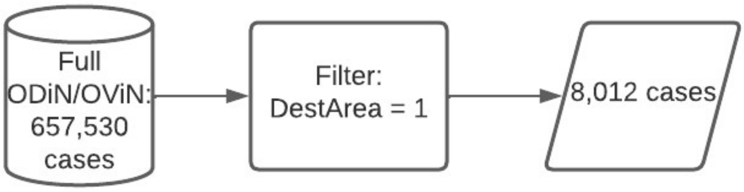



*Figure 23 Filter flowchart for the destination area (part 1 of selection)*

#### 4.3.1.2 Origin point

The origin area is the area outside of the city. Here, 'the city' is defined as the destination area determined in section 4.3.1.1. The OriginArea dummy variable is created where all cases originating in the city, originating in a different country and all missing or unknown are coded as 0, and the other cases are coded as 1. In Figure 24, the formula is filled in as a filter, significantly reducing the number of cases. The high reduction shows that many trips are made within the city borders.

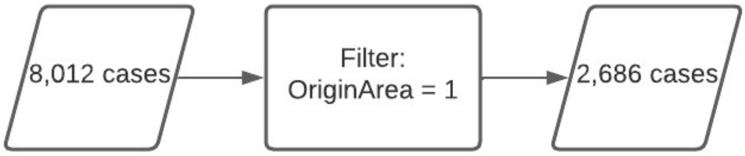



*Figure 24 Filter flowchart for the origin point (part 2 of selection)*

#### 4.3.1.3 Minimum travel distance

Because cities differ in size, the minimum travel distance criterion has not yet been assigned in the method. How this can be done is explained in section 3.3.1.3. The radius of the city of Eindhoven differs in the cardinal directions. To the south, east and west, the distance equals around 4 kilometres; to the north, the distance is 6.5 kilometres. As an estimated average, 5 kilometres has been chosen as the cut-off distance. All trips below 5 kilometres are filtered out. ODiN uses hectometres as the unit for distance; thus, the $x$ in the formula from section 3.3.1.3 is 50. Figure 25 shows a slight reduction in cases due to the filter. Most trips included were already longer than 5 kilometres.

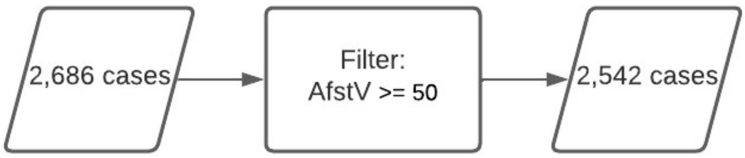



*Figure 25 Filter flowchart for the minimum travel distance (part 3 of selection)*

#### 4.3.1.4 One trip

The filter in Figure 26 makes sure each trip is counted only once. Seven hundred twenty-three cases were removed because trips were counted twice or more times as they were split into trip legs.

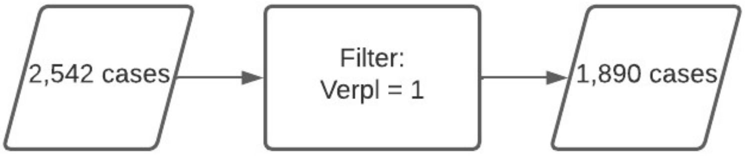



*Figure 26 Filter flowchart for selecting every trip once (part 4 of selection)*

#### 4.3.1.5 Transport mode

Because the problems described in the introduction and literature study relate to car traffic and the relevance of car traffic for transport policy-making, the car is selected as the studied transport mode.

OViN/ODiN differentiates between the primary transportation mode (mode used for the longest part of a trip) and the trip-leg transportation mode (mode used for one part of a trip). The primary transportation mode is used because the method is based on full trip data. It is assumed that the primary mode is used to perform the complete trip. This assumption should be valid as people using the car as the main mode of transportation for a trip are not likely to change mode, see section 2.3 (Pitsiava-Latinopoulou & Iordanopoulos, 2012). In the OViN/ODiN data set of the selected years, only 2% of trip legs with the main transport mode being a car also use other transport modes.

OViN/ODiN also differentiates between car driver and car passenger. These will both be selected as both are a manifestation of car transport demand and are of interest for policy-making.

A new variable has been added to the dataset called “Car” to simplify the selection process. “Car” = 1 when the main transport mode equals car driver or passenger, and “Car” = 0 when other modes are used. Figure 27 shows the number of remaining cases after applying this last selection criterium. 1307 trips (cases) remain in the data set.

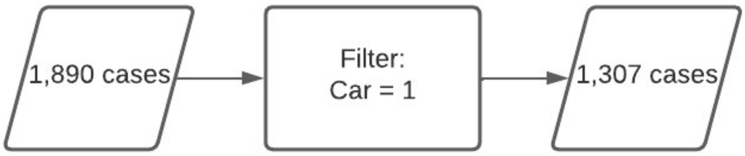



*Figure 27 Filter flowchart of the transport mode (part 5 of selection)*

### 4.3.2 Weighting

As described in section 3.3.2, the trip-based weight from OViN/ODiN is added to the data set. The output of this stage represents the number of trips towards Eindhoven in four years as OViN/ODiN data sets from four years are used as input. 153 million trips towards Eindhoven in a period of four years are thus used in the analysis, as shown in Figure 28.


*Figure 28 Flowchart for weighting*

## 4.3.3 Subdivisions

### 4.3.3.1 Detailing of destination location

The city is divided into 13 destination areas. Some postal code areas have been combined to create the destination areas. In this way, the needed computing power is reduced. The reduction is needed because of limitations to the available hardware when performing the case study. The aggregation of the areas has been done based on cardinal directions and important infrastructure, as discussed in section 3.3.3.1. In this way, the aggregation is expected to have a limited effect on the final output of the method, as the selected entry point is unlikely to be different in most cases. The destination areas are grouped into three districts: city centre; inner ring; outer ring, as seen in Figure 29. The numbering of the destination areas is done based on the districts. This numbering can later be used when interpreting the output. The correspondence between postal codes and destination areas is reported in Appendix 8. The areas are coded into the new PC variable, whose values are shown in Figure 29. A cross-tabulation is exported based on PC and the origin municipality to be used in the routing step using the syntax shown in Appendix 2.2. The cross table indicates the number of trips from each origin municipality to each destination area. Table 1 shows the total number of trips terminating in each destination area.

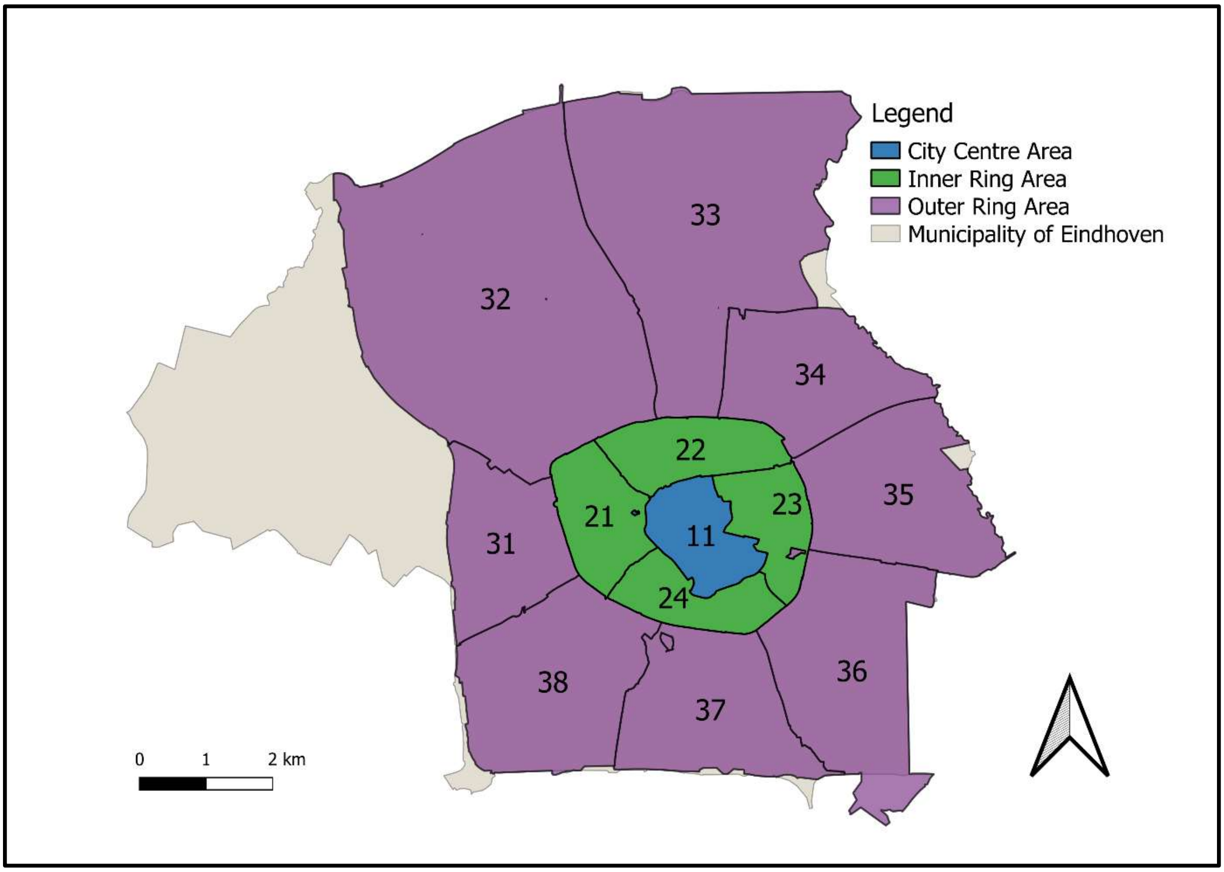



*Figure 29 The city of Eindhoven divided into destination areas based on postal code areas and the location within the city.*

*Table 1 Number of cases per destination area*

| ID | Destination area | Number of cases (weighted) |
|---|---|---|
| *11* | *City centre* | 18,414,009 |
| *21* | *Inner ring – Strijp* | 10,507,133 |
| *22* | *Inner ring – Woensel / TU/e* | 8,364,298 |
| *23* | *Inner ring – Tongelre / Stratum* | 8,889,428 |
| *24* | *Inner ring – Gestel / Stratum* | 7,741,156 |
| *31* | *Outer ring – Hurk / Het Ven* | 11,021,263 |
| *32* | *Outer ring – Strijp / Woensel* | 24,686,795 |
| *33* | *Outer ring – Woensel* | 27,956,227 |
| *34* | *Outer ring – Eckart* | 3,949,427 |
| *35* | *Outer ring – Tongelre* | 5,457,390 |
| *36* | *Outer ring – Stratum* | 6,464,556 |
| *37* | *Outer ring – Gennep / Stratum* | 6,802,395 |
| *38* | *Outer ring – Gestel* | 13,274,453 |
| **Total** | | **153,528,530** |

#### 4.3.3.2 Trip purpose

For policy-making, it is interesting to know for which purpose trips are made (Azari et al., 2013; Kelly & Clinch, 2006). In this additional criterion, all work trips and all non-work trips are selected. Work trips include the commute to a job or study location and business trips. These purposes are often combined in similar studies (Cools et al., 2010; Krizek, 2003). From a policy point of view, work and education institutions can have similar influences on how trips are made. One example is the government pushing these institutions to incentivise sustainable modes of transport.

A new variable is created called “Worktrip”. “Worktrip” = 1 when a trip is made to go from or to a work or study location or is business-related. “Worktrip” = 0 when the trip is made for other reasons. The output is exported using the syntax in Appendix 2.2, with the functions below added to the filter. This process results in two tables, one with the work trips and one with non-work trips. The process results in 66,168,163 work trips and 87,360,369 non-work trips, 43% and 57% of the total cases, respectively.

$$Worktrip = 1, Worktrip = 0 \text{ (= Non-Work)}$$

#### 4.3.3.3 Arrival period

The number of trips conducted is related to the time of day, as discussed in section 2.1 (Ke et al., 2017; Shen et al., 2020). Differences are expected to be the largest when comparing peak and off-peak hours. Morning peak hours are defined by CBS (2016b) as 07.00 until 09.00, and evening peak from 16.00 until 18.00 (CBS, 2016a). This study will allocate a trip to peak hours if the arrival time falls within the defined peak hours. Arrival time is chosen because of the focus on the last part of the trip, entering the city. A peak hour variable is added to the data set for the selection process. The results are 48,825,042 peak hour trips and 104,703,486 off-peak hour trips, meaning that 32% of cases are assigned to the peak hour subcategory and 68% to the off-peak hour. The following functions are added to the syntax of Appendix 2.2 to export the two tables.

$$Peak = 1, Peak = 0 \text{ (= Off-Peak)}$$

#### 4.3.3.4 Trip day

Like the time of day, the day of the week can also significantly affect how much a road is used; the most significant difference is suspected to be between a workday and a weekend day, see section 2.1 (Ivanchev et al., 2015). A new variable is created called “Weekend”. “Weekend” = 1 when a trip is made on a weekend day. “Weekend” = 0 when the trip is made on a workday. 25% of the cases are weekend trips, and 75% are workday trips, which are 37,923,514 and 115,605,005 cases, respectively. The following functions will be added to the syntax of Appendix 2.2 to export the two tables.

$$Weekend = 1, Weekend = 0 \text{ (= Workday)}$$

## 4.4 Mapping

### 4.4.1 Preparing the data

All exported data from the steps discussed in section 4.3.3 are combined into one file using the method shown in Appendix 3. The resulting data set is imported into QGis and joined to the municipality map from CBS (2021a).

### 4.4.2 Origin points

Using the municipality map from CBS, the origin points are created by taking the geometric centroids of each destination area. The centroids further than 350 meters from the road network are moved to the nearest car-accessible road. It was noticed that the ferries from the Frisian Islands municipalities were not part of the road network. The origin points are moved from the islands to the harbour where the ferries land, as shown in Figure 30. The resulting origin points for the whole of the Netherlands can be seen in Figure 31.

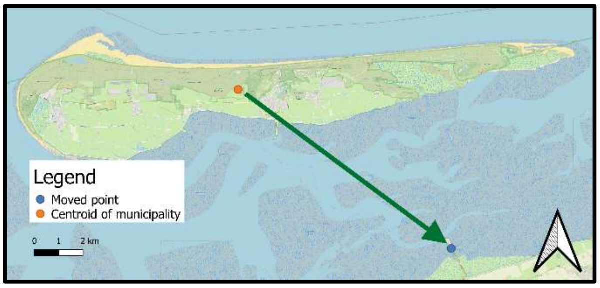



*Figure 30 Special example of moving a centroid to a location where the routing plugin works*

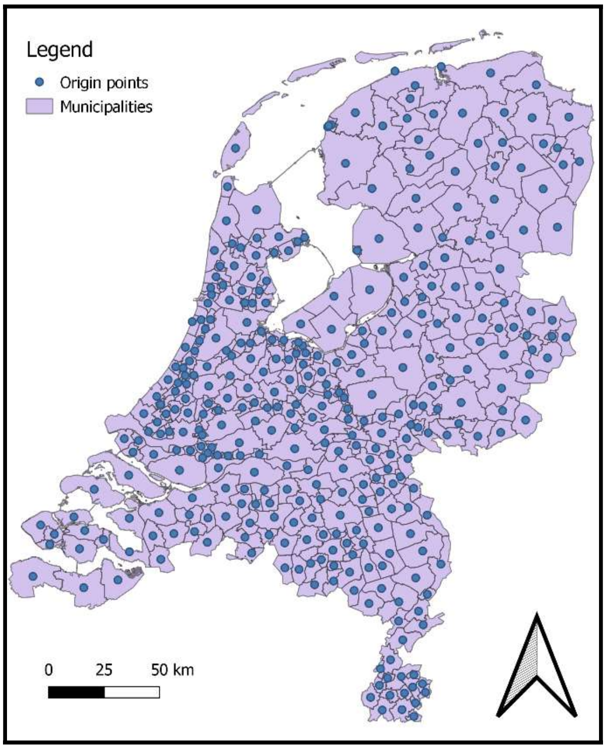



*Figure 31 Map of the spatially aggregated origin points*

### 4.4.3 Destination points

The destination areas shown in Figure 29 (p. 53) were used to determine the destination points in Eindhoven. The geographic centroids were calculated to form the destination points as done for the origin points. These points were then moved to the nearest car-accessible road. The destination points are given the same index as their destination areas. The points are all in separate layers within QGis, as this is needed in the routing step. The aggregated destination points can be seen in Figure 32.

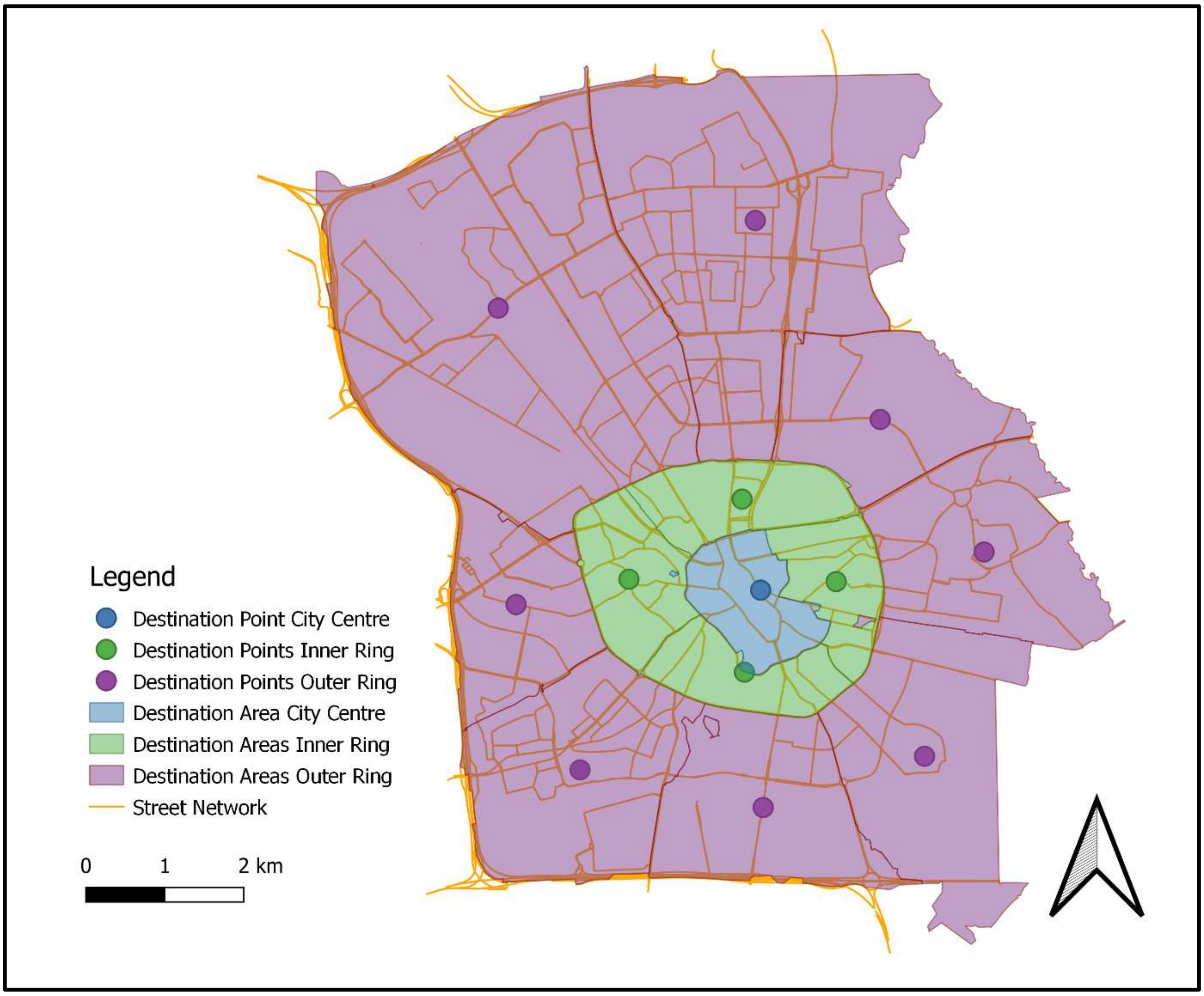



*Figure 32 Map of the spatially aggregated destination points*

### 4.4.4 Entry points

A measurement location must be chosen to calculate the number of trips entering the city using a specific road. The city borders defined in section 4.3.1.1 are used to determine the entry point location.

The highways are not included; the entry points are where traffic leaves the highway. A limited number of car-accessible roads cross the determined borders. On these roads, entry points are added. Specifically, the entry points are located before the first junction to capture all traffic using one entry point. A total of 24 points are selected, shown in Figure 34.

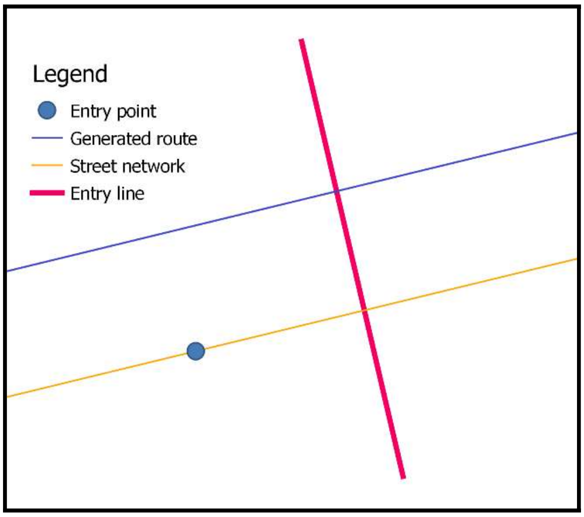



*Figure 33 Schematic view of the use of entry points versus lines*

The entry points are used to count the entering trips, as described in section 4.5.2. The points must be at the exact location where the routes will be generated to perform the counting step. The ORS tools plugin generates its routes slightly next to the street network that can be downloaded from OpenStreetMap (with the same projection). The points must be converted to lines or polygons covering the entire road to ensure the entry point layer will intersect with the routing lines. Figure 33 illustrates this solution. The orange line represents the OpenStreetMap road network, the blue line is a generated route (see section 4.5.1), and the pink line is the "entry line", which can be used instead of the entry point. The entry points, as shown in Figure 34, will be used in the visualization step (see section 4.6).

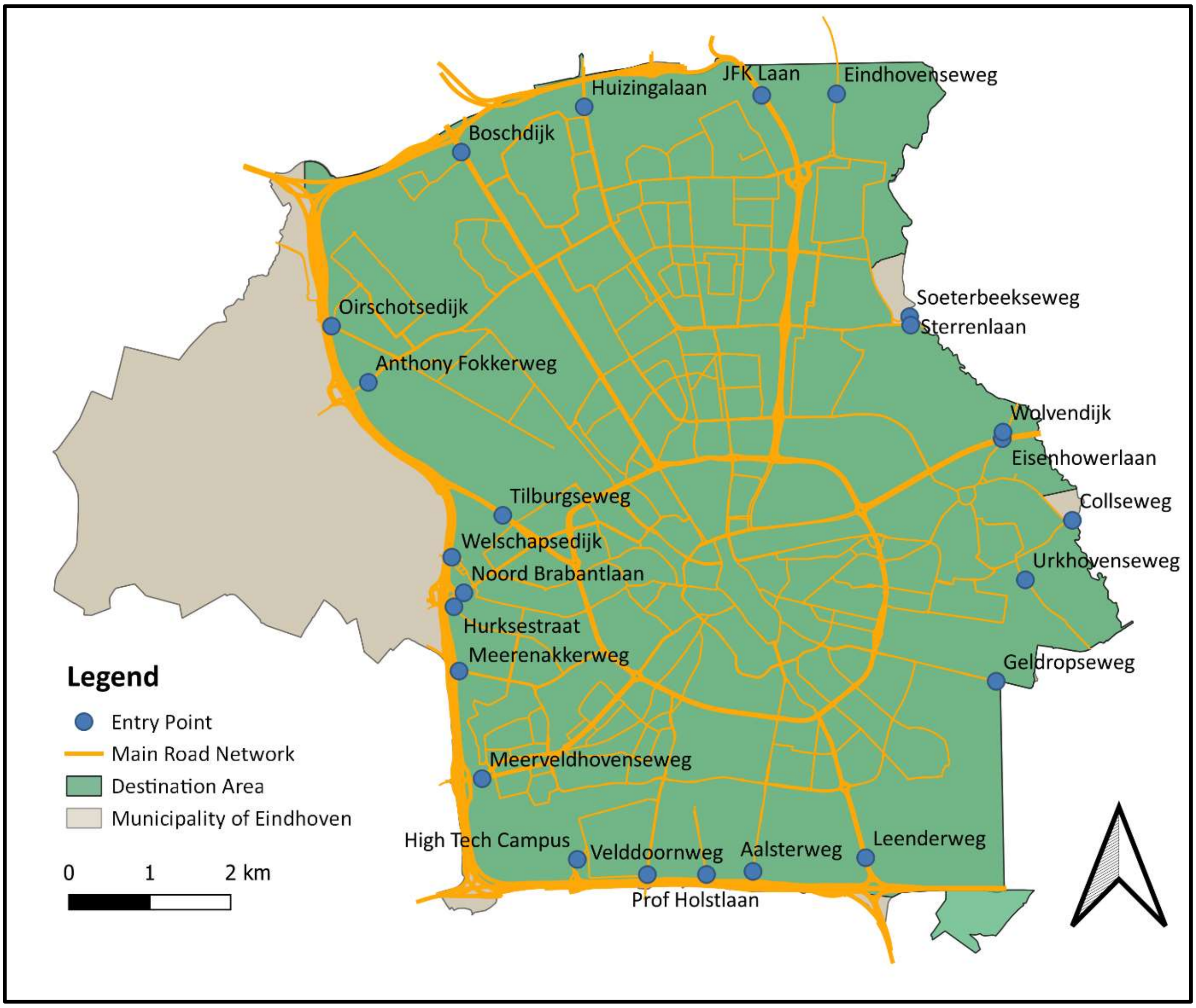



*Figure 34 The entry points for the case study of Eindhoven*

## 4.5 Routing

In the routing step, the routes are generated between the origin and destination points. After that, the number of trips passing the entry line will be counted. Finally, the output will be verified to check if all routes are generated and counted correctly.

### 4.5.1 Generating routes

The ORS Tools plugin for QGIS generates the routes, as explained in section 3.5. Using the script shown in Appendix 4, the origin points with more than 0 trips to a specific destination point in Eindhoven are extracted, after which the routes are generated. Because of limitations of the ORS Tools plugin (limited number of allowed requests per minute and day), the script is run 13 times for all the destination points once. The resulting map is shown in Figure 35.

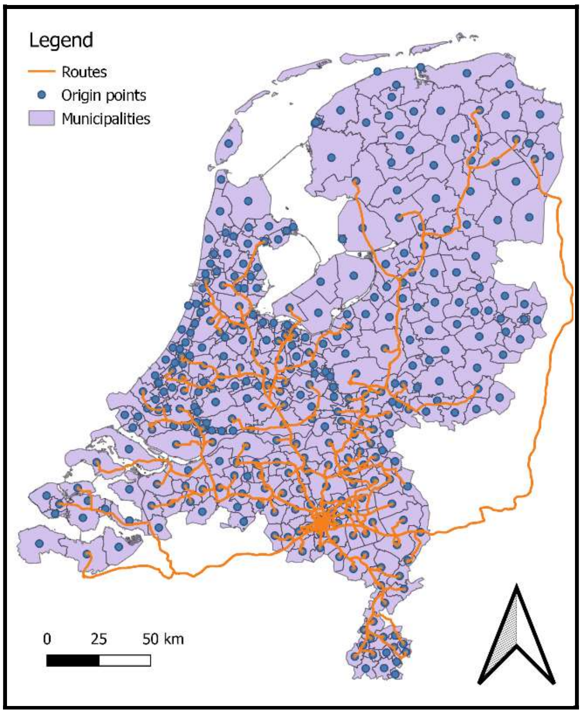



*Figure 35 Map of the generated routes for all trips to Eindhoven*

### 4.5.2 Counting entering trips

The trips passing the entry points (intersecting lines) are summed using the script in Appendix 6. The findings can be found in Appendix 9. The highest number of trips is counted at the entry point of the John F. Kennedylaan, as shown in Table 2. It makes sense that this road is the most used as it connects many of the trips originating North of Eindhoven directly to the city centre and ring road, and the road has a high-speed limit. The John F. Kennedylaan is shown in Figure 36, with arrows indicating the primary traffic flows towards the road. According to the model, ten minor entry points (ID 2, 7, 15, 18, 19, 20, 21, 22, 24, 25) are not used. In particular, many of these roads are dedicated to local access trips, which are not included because of the minimum travel distance criterium and due to the spatial aggregation of the destination and origin areas to centroids. For the roads with assigned trips, the exact values have been recalculated in percentages, showing the ratio of trips between the entry points (Table 2 and Appendix 10) and the ratio of trips per subcategory (Table 3). The percentages are used when interpreting the case study results (see section 4.8.1).

*Table 2 Percentage of the total number of trips per subcategory passing an entry point*

| ID | Entry Point | All Trips |
|---|---|---|
| *1* | *John F. Kennedylaan* | *28%* |
| *3* | *Boschdijk* | *4%* |
| *4* | *Antony Fokkerweg C* | *10%* |
| *5* | *Tilburgseweg* | *8%* |
| *6* | *Meerenakkerweg* | *5%* |
| *8* | *Aalsterweg* | *5%* |
| *9* | *Leenderweg* | *12%* |
| *10* | *N270* | *13%* |
| *11* | *Soesterbeek* | *2%* |
| *12* | *Geldropseweg* | *3%* |
| *13* | *Meerveldhovenseweg* | *3%* |
| *14* | *Hurksestraat* | *3%* |
| *16* | *Prof Holstlaan* | *2%* |
| *17* | *Noord Brabantlaan* | *1%* |
| *23* | *Urkhovenseweg* | *0%* |
| | **Total** | *100%* |

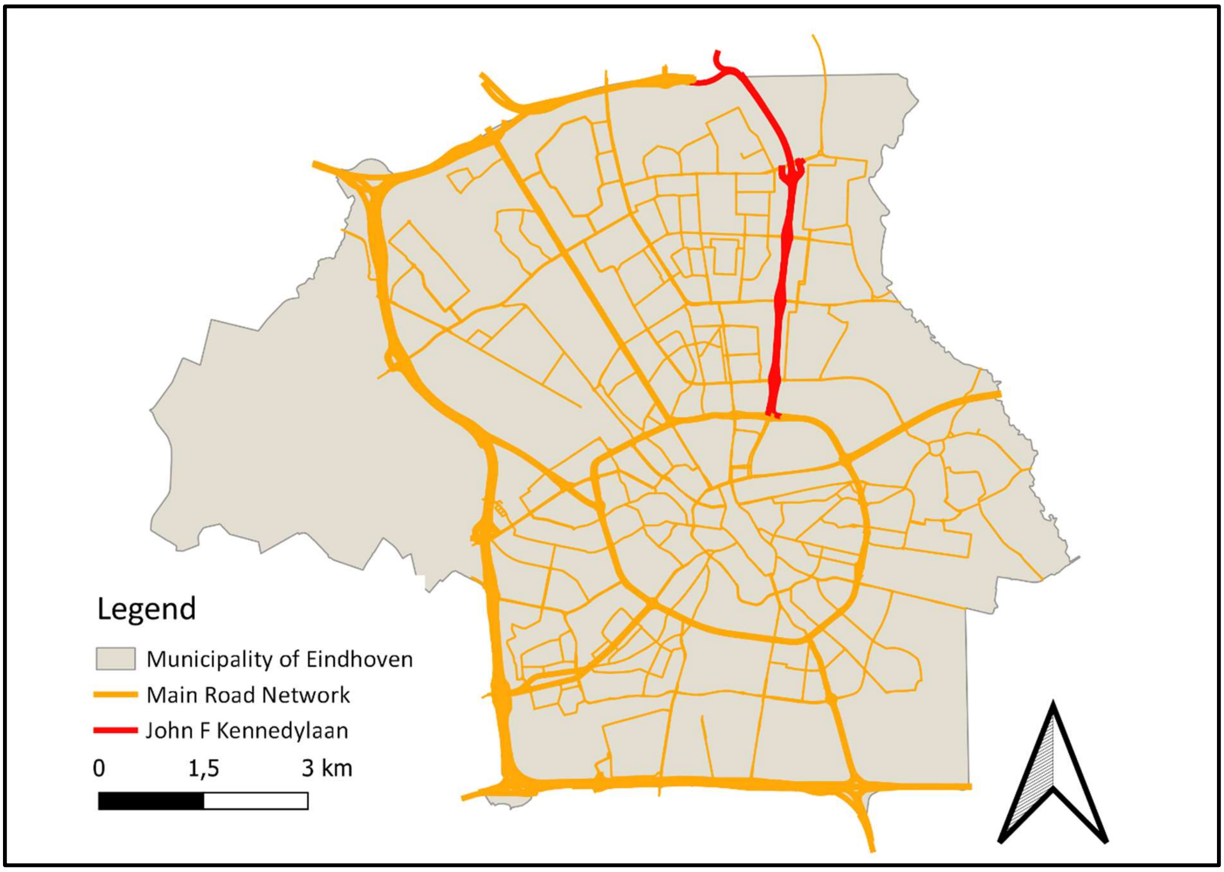



*Figure 36 The John F Kennedylaan as the main entry point for trips coming from the North*

*Table 3 Percentage of the total number of trips per entry point divided into sub-categories*

| ID | Entry Point | Work | Non-Work | Peak | Off-Peak | Weekend | Workday |
|---|---|---|---|---|---|---|---|
| *1* | *JFK laan* | 41% | 59% | 32% | 68% | 27% | 73% |
| *3* | *Boschdijk* | 34% | 66% | 31% | 69% | 22% | 78% |
| *4* | *Antony Fokkerweg C* | 54% | 46% | 36% | 64% | 21% | 79% |
| *5* | *Tilburgseweg* | 46% | 54% | 30% | 70% | 27% | 73% |
| *6* | *Meerenakkerweg* | 60% | 40% | 43% | 57% | 11% | 89% |
| *8* | *Aalsterweg* | 45% | 55% | 26% | 74% | 31% | 69% |
| *9* | *Leenderweg* | 58% | 42% | 38% | 62% | 23% | 77% |
| *10* | *N270* | 37% | 63% | 30% | 70% | 26% | 74% |
| *11* | *Soesterbeek* | 10% | 90% | 23% | 77% | 23% | 77% |
| *12* | *Geldropseweg* | 28% | 72% | 25% | 75% | 21% | 79% |
| *13* | *Meerveldhovenseweg* | 46% | 54% | 35% | 65% | 24% | 76% |
| *14* | *Hurksestraat* | 17% | 83% | 27% | 73% | 40% | 60% |
| *16* | *Prof Holstlaan* | 36% | 64% | 15% | 85% | 12% | 88% |
| *17* | *Noord Brabantlaan* | 5% | 95% | 12% | 88% | 39% | 61% |
| *23* | *Urkhovenseweg* | 26% | 74% | 24% | 76% | 26% | 74% |
| **Total** | | 43% | 57% | 32% | 68% | 25% | 75% |

### 4.5.3 Verification of the routing step

The routes created are visually checked based on the principles explained in section 3.5.4. The routes first go to high-level roads and stay on these roads until close to or within the city, as predicted based on the findings in the literature study (see section 2.2). The checked routes are visualized in the next step and shown in Figure 37 and Figure 38. Checking the total output against the total input shows that no cases are lost during the process. The total numbers per subcategory are also checked for lost cases. Also, no lost cases are found for the subcategories.

## 4.6 Visualizing

The output of the routing step is visualized using the method described in section 3.6. Using the script shown in Appendix 6, the route layers are prepared for the visualization. The preparation entails cutting the routes into individual links, summing the total number of trips per link, and linking the total number of trips generated per subcategory in each municipality to the municipality map. As discussed in section 3.6.1, the ‘v.clean’ tool from the GRASS plugin is ideally used for cutting up the routes as it only cuts where two routes intersect. However, the tool crashes when the amount of input data is too high. In the case study, the amount of input data is too high, therefore, the alternative ‘explode’ algorithm is used, as shown in the script in Appendix 6. The municipality map with the joined input data set, as performed in section 4.4.1, is used to visualise the origin areas. The visualization of the origin areas can be seen in the background of Figure 37. The total number of trips to all destination points from each origin area is used for the visualization.

The number of trips using a route link is visualized by the width of the lines. The routes are separated into ten classes for visualization. The width of the lines ranges from 0.3 mm to 4.0 mm. The number of trips originating in each origin area is visualized using colour. The data is classified into ten classes based on the Natural Breaks classification. The number of classes and the choice of colour is based on what is distinguishable by colour-blind people, see section 3.6.3 (Harrower & Brewer, 2003). The Natural Breaks classification is used as the data does not follow a clear linear or logarithmic line. Using the Natural Breaks classification, the differences between the number of trips are most apparent; however, this choice should be considered when analysing the model's output. An overview and a zoomed-in map are made for the complete data set and for each sub-category. The resulting maps for all trips towards Eindhoven are shown in Figure 37 and Figure 38. The maps of the different subcategories can be found in Appendix 12. All maps use the same classes to make comparisons possible. To analyse one specific entry point, it is also possible to make a map where only the routes using this entry point are visible together with the origin municipalities of these routes. An example of a map focussed on one entry point is given in Figure 39. The map shows the routes using the Aalsterweg as their entry point into the city. An analysis of all the results is provided in section 4.8.1 of the report.

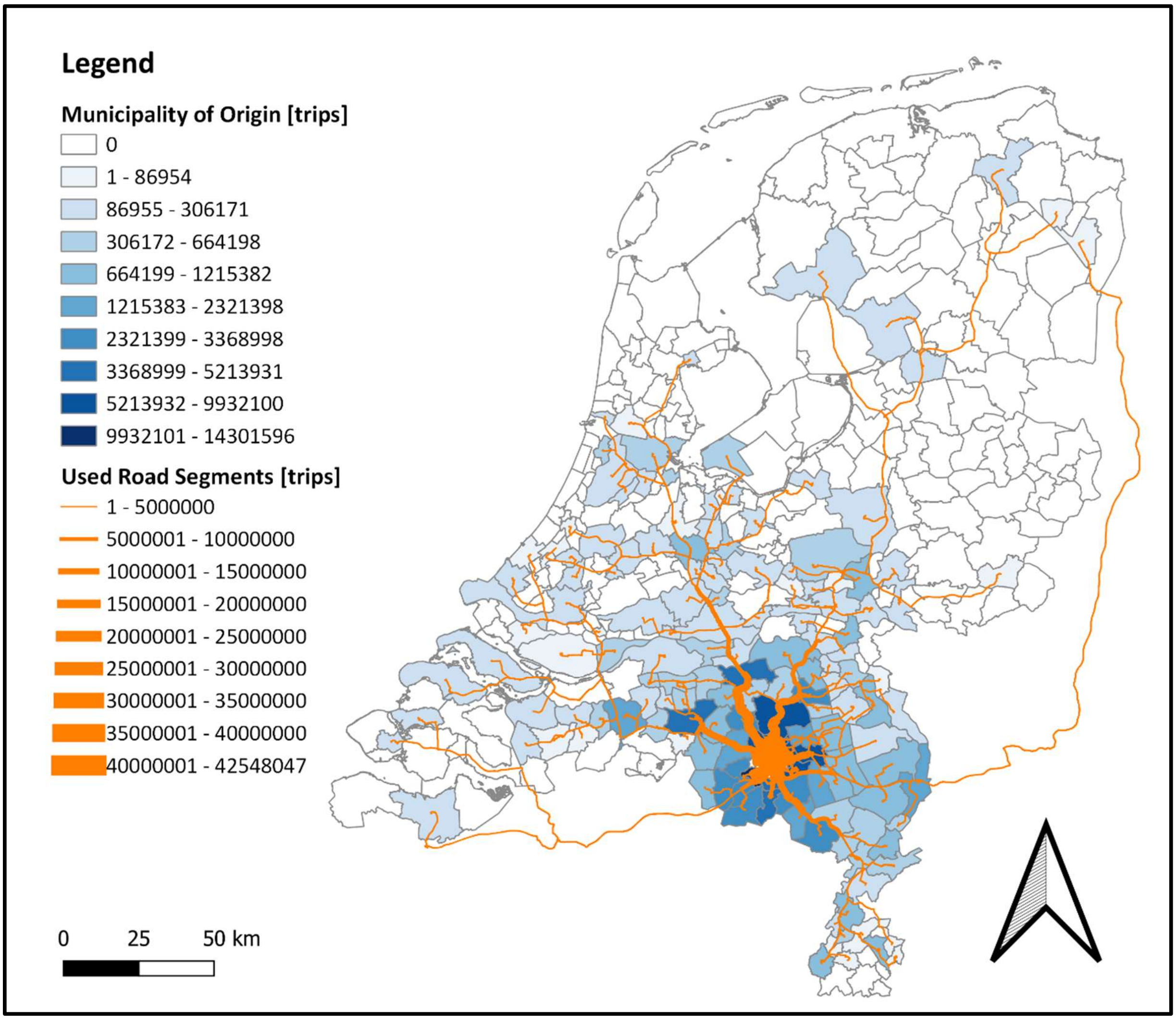



*Figure 37 Case study Eindhoven visualization, all trips towards Eindhoven 2017-2020, overview*

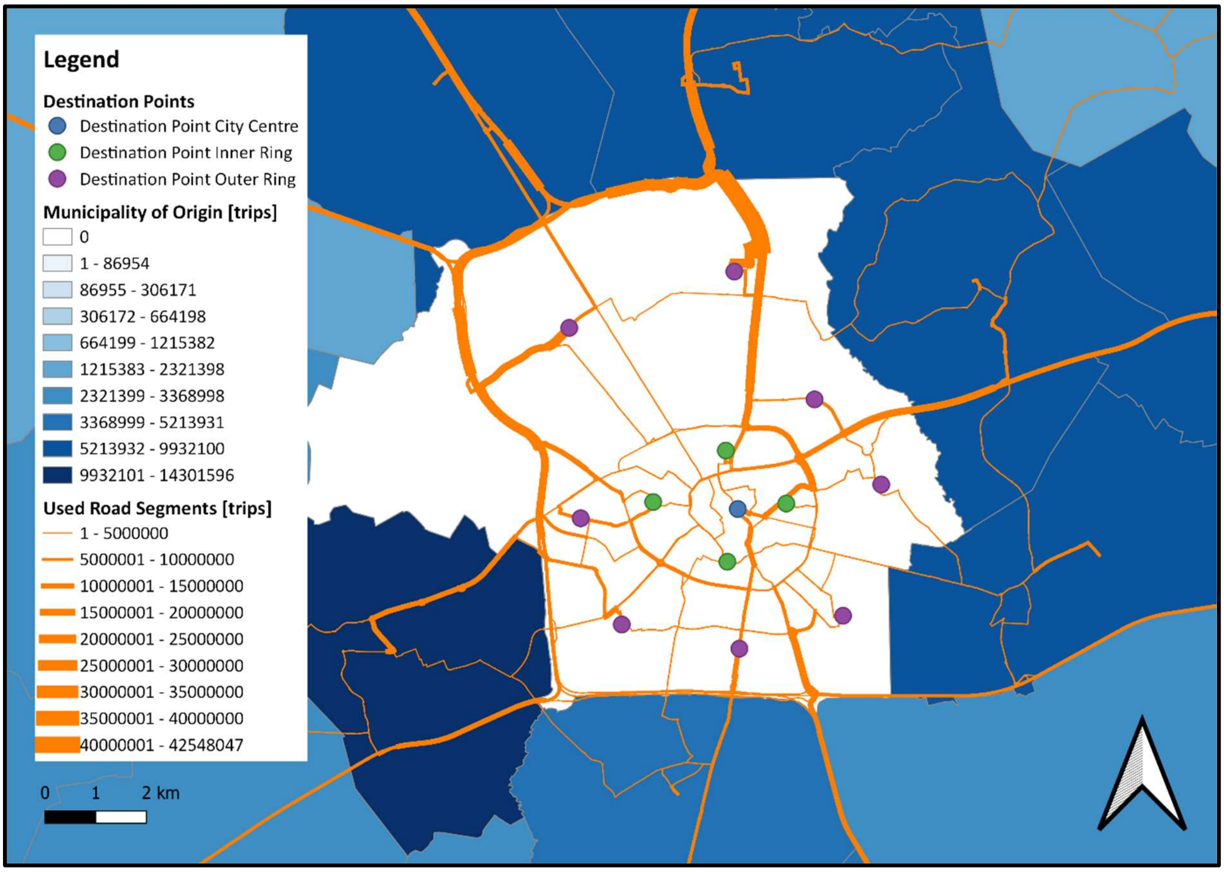



*Figure 38 Case study Eindhoven visualization, all trips towards Eindhoven 2017-2020, zoomed*

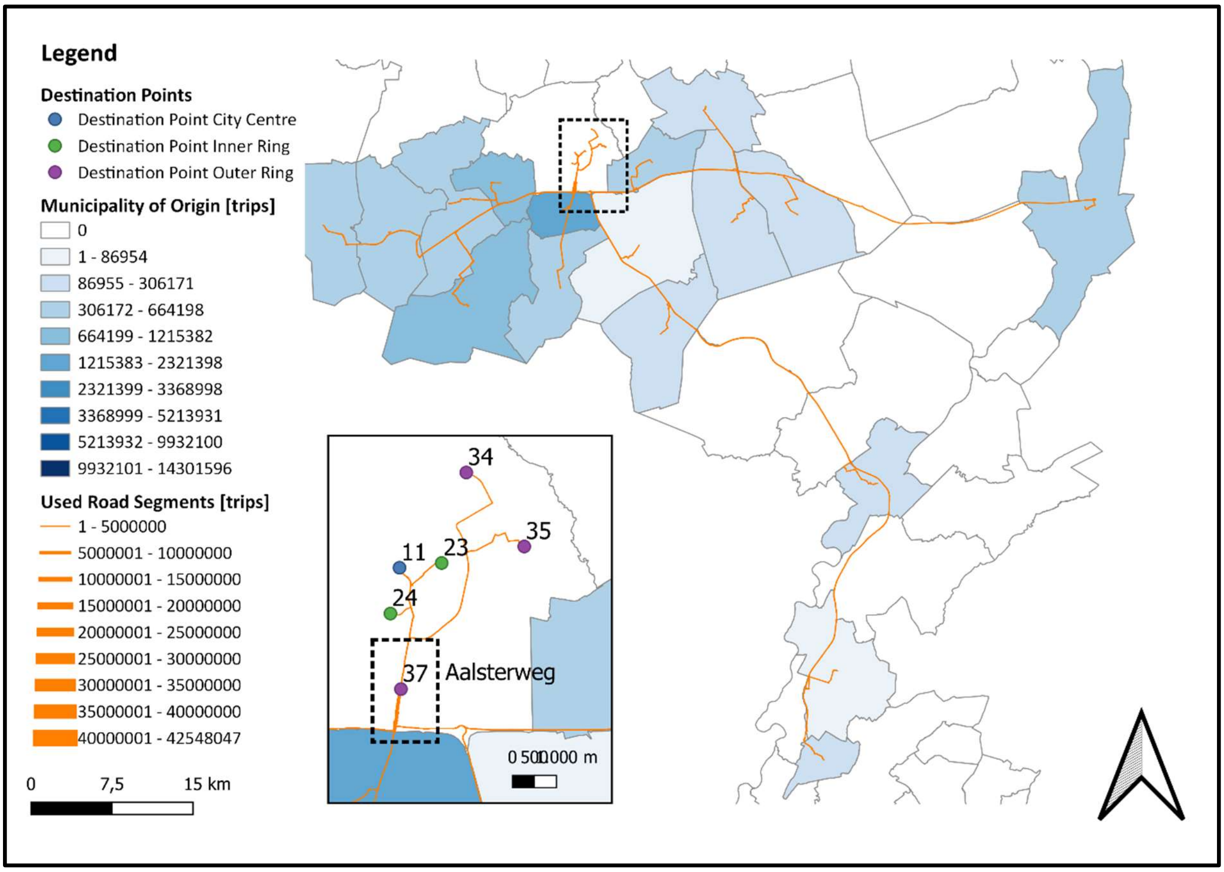



*Figure 39 Map of trips using the Aalsterweg, 2017-2020*

## 4.7 Validating

The last step of the method is validating the generated number of trips per entry point. In this case study, the output will be validated by comparing the number of trips per entry point with traffic count data by performing a correlation and regression analysis. In section 3.7.1, several options are given to perform this step.

### 4.7.1 Input data

One data set is found with observed traffic counts for all trips and trips made in peak hours (Dufec, 2021) (referred to as 'data set 1'). For workday traffic, two data sets are found (Dufec, 2021) ('data set 1' and 'data set 2'). Both data sets are used separately in the analysis. The analysis also uses the average of the data sets for workdays. Both data sets are measured in periods of several weeks and are on city level. Data set 1 is included data from 2016 until 2021, and data set 2 from 2011 until 2019. For the validation, the average number of vehicles per day from 2017 until 2020 is taken, if available, as this corresponds with the survey data used in the method. Data set 1 includes the exact location of where the traffic is counted. Some traffic count locations differ from the entry points selected in the case study. Figure 40 shows the traffic count location used as the validation points and the entry point they are combined with. Most validation points are much closer to the inner city than the used entry points; this needs to be considered when interpreting the results. Because of the locations of the validation points of the Tilburgseweg and Noord Brabantlaan, these two points are combined into one measurement, resulting in 9 measurements. It is unclear where the traffic count took place for data set 2. The total number of trips found in Appendix 9 is divided by 1,461 for all trips and peak trips, as this is the combined number of all days in 2017, 2018, 2019 and 2020. The total trips for the workday trips are divided by the number of workdays in the same period, which is 1,044. The input data for the statistical test is shown in Table 4. It is expected that the values of method output per entry point are different from the traffic count data because of several reasons, including the difference in measurement location, the inclusion of local, foreign and commercial trips in the traffic count data, the differences in measurement period and the inclusion of car passengers as separate trips in the method.

*Table 4 Data input for the validation step based on the case study findings and traffic counts from Dufec (2021)*

| | | **Method Output Trips per Day** | | | **Traffic Counts per Day** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **ID** | **Entry Point** | **All** | **Workday** | **Peak-Hour** | **All** | *Workday 1* | *Workday 2* | **Workday Average** | **Peak-Hour** |
| *1* | *JFK laan* | *29,369* | *30,168* | *9,309* | *32,249* | *35,239* | *39,673* | *37,456* | *9,644* |
| *3* | *Boschdijk* | *4,336* | *4,737* | *1,345* | *9,255* | *10,317* | *10,898* | *10,607* | *2,627* |
| *8* | *Aalsterweg* | *5,227* | *5,077* | *1,358* | *8,275* | *8,603* | *9,268* | *8,935* | *1,992* |
| *9* | *Leenderweg* | *12,212* | *13,200* | *4,630* | *12,980* | *13,858* | *14,847* | *14,352* | *3,721* |
| *10* | *N270* | *13,397* | *13,961* | *4,000* | *19,797* | *21,525* | *21,997* | *21,761* | *5,749* |
| *12* | *Geldropseweg* | *2,800* | *3,108* | *712* | *11,447* | *12,083* | *11,467* | *11,775* | *3,004* |
| *13* | *Meerveldhovenseweg* | *3,159* | *3,340* | *1,105* | *7,443* | *8,035* | *7,303* | *7,669* | *1,955* |
| *16* | *Prof Holstlaan* | *2,029* | *2,510* | *312* | *4,037* | *4,585* | *4,715* | *4,650* | *1,459* |
| *5+ 16* | *Noord-Brabantlaan en Tilburgseweg* | *10,037* | *10,042* | *2,747* | *16,908* | *18,655* | *20,105* | *19,380* | *5,377* |

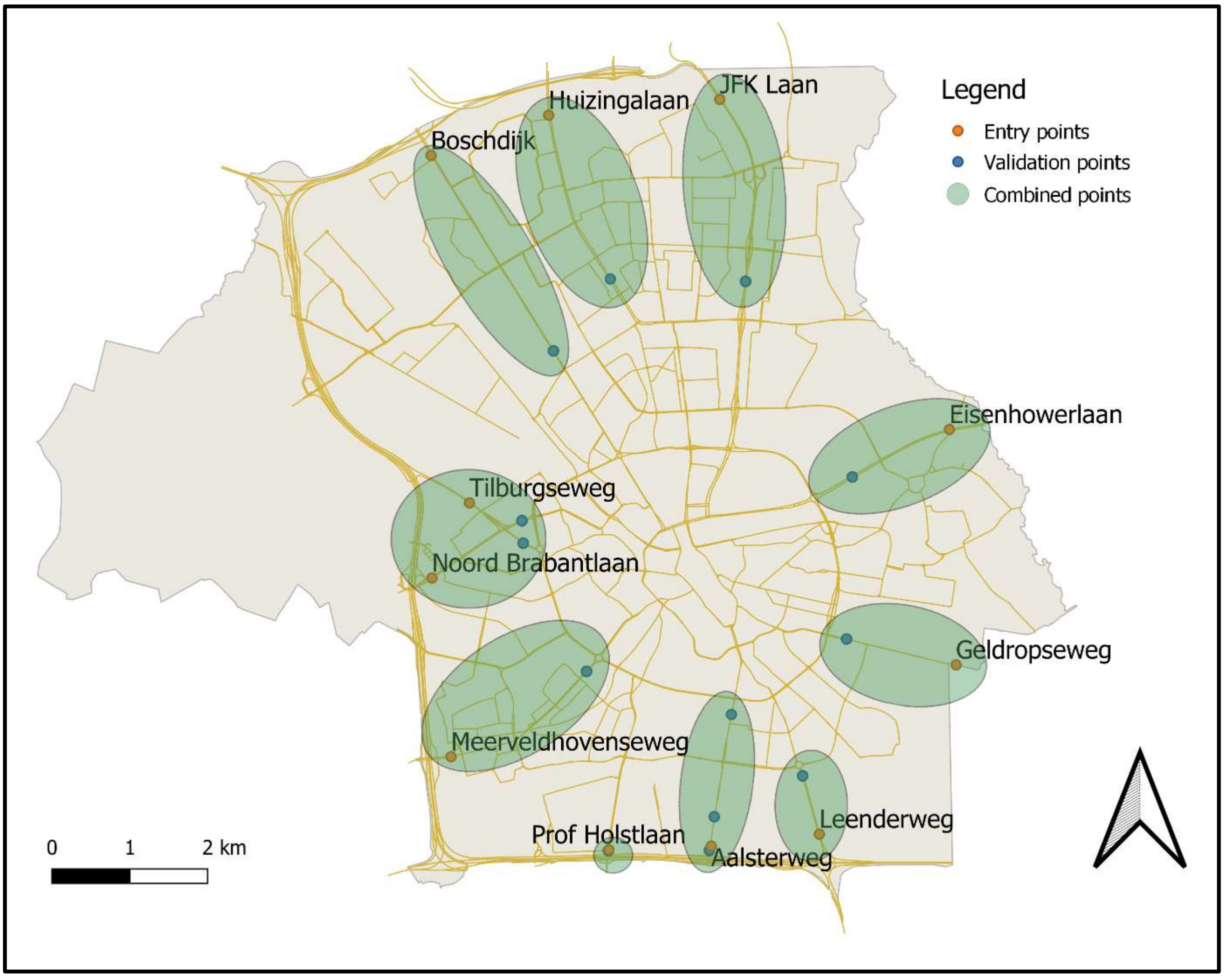



*Figure 40 The validation points and the entry points they represent*

## 4.7.2 Statistical test

A simple linear regression has been performed to test how well the method's output can predict the actual traffic counts at the different entry roads. In total, five regressions are performed, one for all trips, three for workday trips (data set 1, data set 2 and the average) and one for peak-hour trips. The simple linear regression analysis assumptions have been met for all five combinations. The analysis performed to test for the assumptions for all-trip regression is described next. The analysis for the other four combinations can be found in Appendix 11. Normality is tested by creating a Predicted Probability (P-P) plot, see Figure 41. The observations roughly follow the linear line, indicating that the residuals of the regression follow a normal distribution. Homoscedasticity is checked by plotting the predicted values against the residuals on a scatterplot, see Figure 42. No obvious pattern can be distinguished, but because of the low number of values, it is difficult to give a fully satisfying answer to the question of homoscedasticity. For now, it is assumed that the values are homoscedastic. This assumption will be discussed in the interpretation of the results. Linearity is tested by plotting the two data sets against each other, see Figure 43. The plot shows that the two data sets have a linear relationship, thus satisfying the last assumption.

The five regression analyses can be performed as the assumptions have been met. The output of the analyses can be found in Table 5. All regression analyses show very high explanation power and a significantly good degree of prediction.

*Table 5 Regression output case study Eindhoven*

| | Regression | | | ANOVA | |
|---|---|---|---|---|---|
| **Trip Type** | **R** | **$R^2$** | **Adjusted $R^2$** | **F** | **Sig.** |
| *All Trips* | *0.957* | *0.915* | *0.903* | *75.460* | *<0.001* |
| *Workday 1* | *0.952* | *0.906* | *0.893* | *67.698* | *<0.001* |
| *Workday 2* | *0.964* | *0.930* | *0.920* | *92.585* | *<0.001* |
| *Workday combined* | *0.959* | *0,920* | *0,909* | *80.512* | *<0.001* |
| *Peak-hour* | *0.927* | *0.859* | *0.839* | *42.766* | *<0.001* |

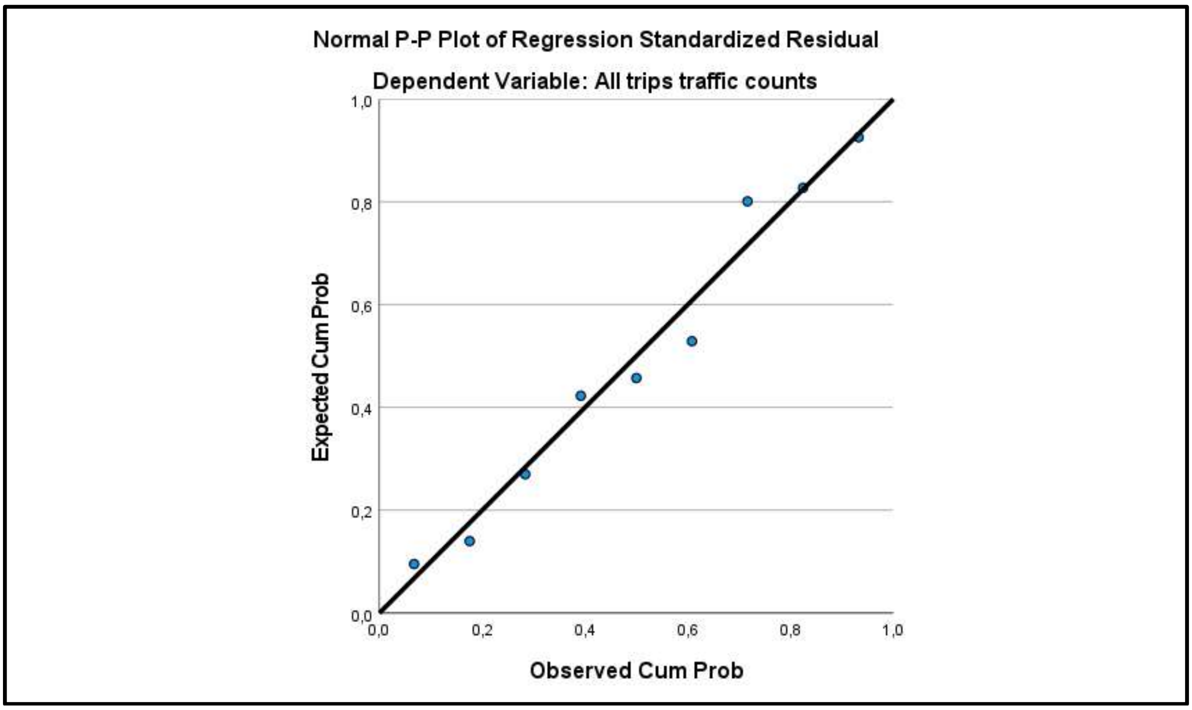



*Figure 41 Normal Predicted Probability plot of all trips*

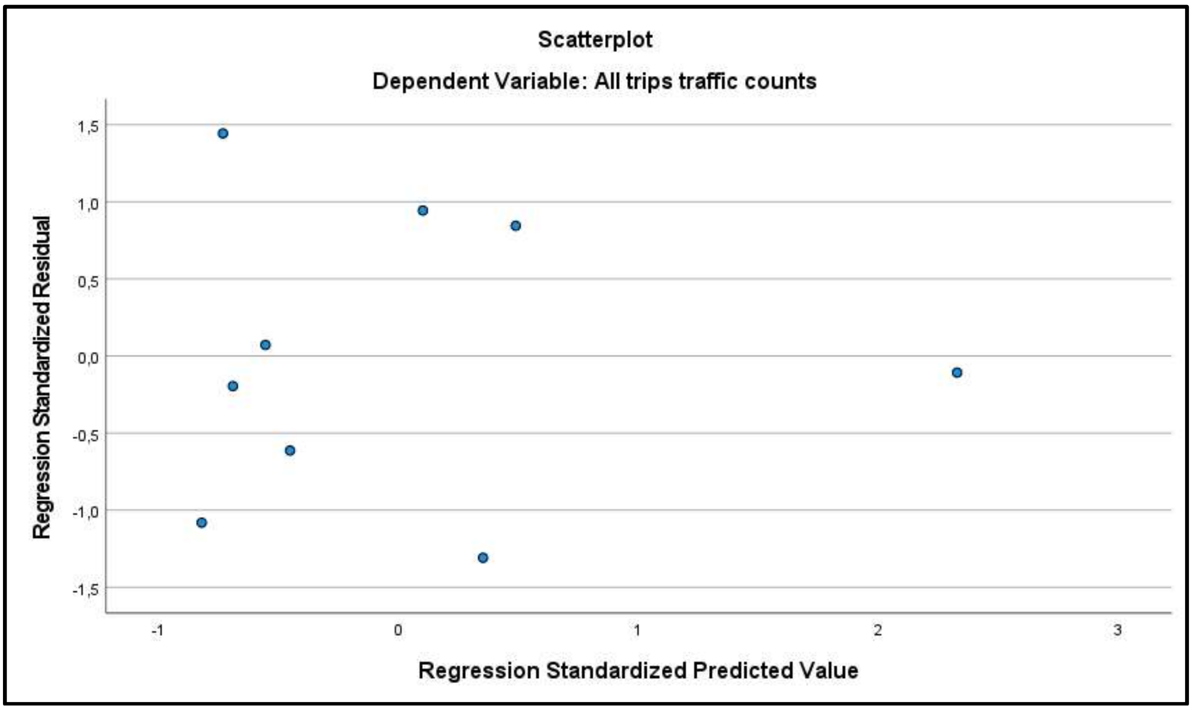



*Figure 42 Scatterplot of predicted values and residuals of the regression of all trips*

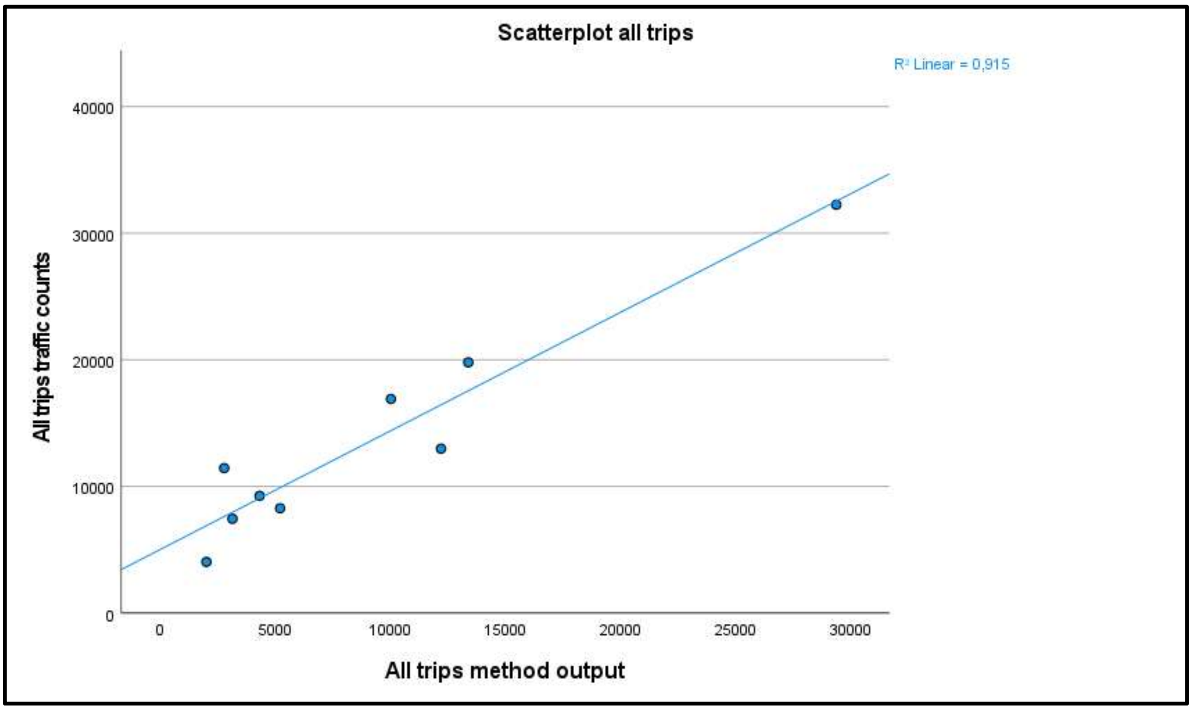



*Figure 43 Scatterplot of the traffic counts and method output with a regression fit line*

### 4.7.3 Interpreting results

The results of the regression analyses show that using the method can predict the traffic counts very well, especially considering the differences between the measurement. The sample size is too small to concrete conclusions about the findings, especially when considering the homoscedasticity issue, but the results are very promising. The regression shows that data set 2 is a little better predicted, this may indicate that the measurement location more closely represents the entry points selected in the case study, but may also be random chance. It has to be considered when interpreting the results that the regression analyses do not compare the absolute values of data sets. Therefore exact values may likely be different in the two data sets.

## 4.8 Conclusions

The conclusion reflects both on the results of the case study and on the method's performance. The conclusions based on the case study's results have as their primary goal to show how the outcome can be used to substantiate policy measures.

### 4.8.1 Results of the case study

The method has three types of results. Namely, the number of trips using an entry point, which are shown in Table 2 and Table 3 (and Appendix 9 for exact values), the maps created in section 4.6 showing the origins, destination, routes and the number of trips (see Figure 37, p. 61, Figure 38, p. 62 and Figure 39, p. 62) and the results of the validation step. A general conclusion is given on these results. Besides the general view, more specific insides are given for the entry point at the Aalsterweg to show how the method can be used to provide more insights regarding specific problems. The problem, in this case, is the small number of users of the Park-and-Ride facility located at the Aalsterweg, as introduced in section 1.3.1 (Burg, 2021).

#### 4.8.1.1 General results

When looking at the total number of trips per entry point, the John F Kennedylaan has the most users, followed by the N270, the Leenderweg and the Antony Fokkerweg; see Table 2. All four roads are high-capacity, multi-lane roads. The four roads are evenly distributed around the city. According to the model, other high-capacity roads like the Noord-Brabantlaan and the Meerveldhovenseweg are used much less. There can be many reasons for the low numbers, e.g., the roads may mainly be used for shorter, local trips; the model may favour other routes because of the made assumptions about route choice; other routes may be favoured because of the aggregation of the destination points; the roads may be used when other routes are congested, which the model does not take into account; or the road may have too high of a capacity. The method does not give answers to what the reason may be. However, it can help identify where more detailed research may be needed to optimize traffic flows in the city.

Of the total number of trips, 43% are made for work- and 57% for non-work trips, see Table 3. The percentages vary significantly between the entry points. The most significant differences are found at the entry points with few trips. However, the number of trips may be too small to make valid conclusions. The ratio between work and non-work trips still varies significantly for the entry points with a larger number of cases. The ratio between peak and off-peak trips and workday and weekend days differs much less between the entry points. Still, some significant differences are visible. Finding these differences could be helpful in policy-making, as seen in section 4.8.1.2.

A limitation of the case study is that trips from neighbouring countries are not included. The Belgian government does perform travel surveys called OVG. However, according to the department, the sample of trips going to Eindhoven is too small to be usable (Departement Mobiliteit en Openbare Werken, n.d.).

The overview map of Figure 37 (p. 61) shows many trips originating close to Eindhoven, reducing when the distance to the city increases. The longer distance routes all aggregate to one of the five major highways to the city, and the shorter distance trips aggregate to collector roads to reach the destination, again following the logic found in the literature study, see section 2.2 (Meyer & ITE, 2016). Some higher populated municipalities further away from Eindhoven have more trips originating than their surrounding municipalities, which is logical as more people live in these municipalities. Thus there are more people that can travel to Eindhoven.

The results of the validation show that the model can, to a significantly good degree, predict the validation data. The regression also shows that a very large part of the variation in traffic counts can be predicted using the method. Because the sample size is low, the measurement of different trips and the differences in measurement location, no concrete conclusions should be made. However, the findings do indicate that the method can be a useful tool for indicating the number of trips using entry points. When interpreting the results, it must be considered that the regression analysis does not compare the absolute values of data sets.

#### 4.8.1.2 Specific insights Aalsterweg

The Park-and-Ride facility located at the Aalsterweg has a small number of users, according to Burg (2021). The proposed method can be used to indicate how many trips pass the Park-and-Ride and find where people come from, why they travel, when they travel and where they are going. The information found can be used to give a possible indication of why the Park-and-Ride location is underutilized. Figure 39 (p. 62) shows the trips made using the Aalsterweg as their entry point into the city. The trips originate primarily to the South of Eindhoven and have their destination point in the South or centre of the city. Looking at the underlying values of the map in Figure 39 (p. 62) shows that 58% of all trips using the Aalsterweg have their destination in the south of Eindhoven, specifically in destination area 37 of Figure 29 (p. 53). 10% goes to the inner ring areas 23 and 24, 22% goes to the city centre and 10% to other areas in the outer ring. The output means that 58% of the trips will unlikely benefit from using the facility as a Park-and-Ride as they are already close to their destination, and 10% are unlikely to take the bus because they are going to outer ring areas where the bus does not go. For these trips, the existence of bike and scooter share may be useful when parking is difficult at the destination. When looking at the ratio between peak and off-peak trips, more trips are made off-peak compared to other entry points, meaning that in peak hour congestion is likely to be lower than in other places, reducing the incentive to use the Park-and-Ride facility for people driving to the city centre or inner ring. More trips are made for non-work purposes, and a higher share of trips are made at the weekend compared to the average of all entry points.

Based on the case study's findings, it makes sense that the Park-and-Ride is underutilized as many passers-by will not go far into the city and will not likely get stuck in congestions because they travel outside peak hours. Solutions based on the findings may include incentivizing people using other entry points, like the Leenderweg, to use the Park-and-Ride facility and focussing on shared micro-mobility to make using the facility more beneficial for people not travelling to the city centre. Advertising to use the Park-and-Ride should be targeted on work and non-work trips as there is no significant difference in the number of trips.

### 4.8.2 Performance of the Method

This section goes into how well the method performed regarding ease of use and accessibility. As shown in the case study, the method can be performed using revealed, publicly available data and with accessible software.

In the validation step, the data used for comparing the method results must have specific characteristics (primarily the location), which may be challenging to find. Using data without the specific characteristics may result in falsely dismissing or accepting the output. It may be useful to use the alternative ways of validating indicated in sections 2.7.4 and 3.7.1.

To conclude, the study's goal has, for the most part, been reached. The method could be simplified to make it easier to use if paid software were used; however, this would make it less accessible.

# 5. Conclusion and Discussion

## 5.1 Conclusion

Problems of congestion caused by bottlenecks at the entry points of cities, the uncertainty of the effectiveness of policies designed to solve the problems, and the lack of studies into traffic flows that specifically focus on entry points warrant studying traffic flows at entry points in more detail than previously done. The lack of knowledge about how to study the traffic flows through entry points and the lack of easy-to-use and openly accessible methods to examine them resulted in the following main research question:

*What would an easy-to-use and openly accessible method look like that makes it possible to examine the traffic flow from outside a city into a city?*

The need and will to participate in activities at different locations is the reason for the demand for transportation. In cities, a high number and density of activity locations exist, and cities thus attract many trips. Several different modes of transport can be used to travel to a city depending on specific circumstances. Cars are especially convenient for travel from rural and scattered areas to activity locations in cities. Many different routes can be taken to reach a city. The fastest route is often taken, generally following arterials or highways with high capacity and speed limits.

Usually, there is ample space to build the infrastructure supply to match the demand for transportation outside of the build-up areas. However, space is generally limited within cities. Because of the limited space, the demand often cannot be met. Mismatches between supply and demand are likely to cause congestion problems. The congestion is frequently seen at places where the traffic enters the city, and thus supply is difficult to increase. If the space does exist to increase supply, it is likely that due to the mechanisms of induced demand, the congestion problems will not be solved.

Methods to create models that predict the demand between origin and destination points exist, but these are often difficult to create or access and do not focus on the entry points of cities. Traffic count data can be used to investigate entry points. However, relevant information for policy-making is generally missing, and traffic counting can be expensive, time-consuming and inflexible, as the data only can give information about specific locations and collecting data on new locations takes time and money. The discussed shortcomings of the existing research options leave a need for an accessible method to examine traffic entering cities which is easy to use and interpret by policymakers.

An important first step to create an accessible method is the use of open data. It was chosen to base the developed method on national travel survey data, which is openly available in several countries. The method consists of five steps, namely filtering, mapping, routing, visualizing and validating. Following these steps, the number of trips per entry point can be found, maps are created to visualize the findings to ease communication, and validation is performed. Scripts are written for large parts of the filtering, routing and visualizing steps, which makes using the method easier. The scripts can be used to perform the method for different cities. Since all cities have a unique spatial road structure, not all the steps could be automated, but they could be performed manually. The mapping and validation steps are mostly not automated because they need to be specifically catered to a city.

The method has been tested on the city of Eindhoven. The process resulted in values for the number of trips using each entry point of Eindhoven and each selected sub-category (work, non-work, peak, off-peak, weekend, workday). By performing the visualization step, maps could be created to communicate the findings easily. The validating step has been performed based on already existing traffic count data. The method could be performed entirely, although several more time consuming and technical manual actions in QGis like defining the city borders and the destination points need to be performed when using the method, impacting the ease of use. This need is primarily due to the heterogeneity of cities but does make the implementation more complex than ideal. The limitations of using the freely available software versions also lead to some difficulties with using the method because of performance issues.

To conclude, a method for examining traffic flow from outside the city to inside the city has been created. The method uses openly available data and widely available software, making the method accessible for everyone to use. The ease of use can be debated and possibly improved in future studies.

## 5.2 Discussion

### 5.2.1 Strengths

The created insight into the usage of the entry points of cities can help substantiate and improve made policy choices regarding the impacts of traffic entering the city. The Park-and-Ride example at the Aalsterweg shows the insights that traffic counters could not give easily, namely that most users of this entry point are heading to locations near the Park-and-Ride facility. With this information, policy options could be created to make the facility more used in a more targeted manner.

Compared to many existing models, the method proposed in this thesis is relatively easy to use. It can relatively quickly provide information about the number of users and the associated insights of these values, for example, the characteristics of potential users of Park-and-Ride locations, as seen in section 4.8.1.2 and Figure 39. The method is accessible to anyone, making it possible for different parties to perform it in many different situations. The findings can therefore be checked and tested by independent parties.

When applied, the developed method can give new insights into an understudied part of transportation, traffic entering cities through entry points. It was found that problems often occur at entry points and that these problems are challenging to solve. No studies were found that discuss entry points in depth. After performing this research and creating the method, information is added on what a possible option is to study the entry points.

### 5.2.2 Limitations and Recommendations

A case study based on Eindhoven, the Netherlands, was conducted to trial the method. The developed method could be entirely performed. However, some software limitations were found during the routing and visualization steps which made the method less easy to use. Furthermore, due to a limited amount of data and the software limitations, origin and destination areas had to be aggregated into larger areas, and travel time needed to be aggregated into two categories. As a result, the insights obtained were less detailed than initially intended. The method’s ease of use may also be affected by the unique characteristics of the transportation network, as well as the geographical and political features of cities. Consequently, choices like defining the city borders and the minimum travel distance could not be standardized. To solve this issue detailed descriptions of how to make the choices are added to the method. Additionally, finding adequate data for validation came with difficulties.

The research question has been answered by developing an accessible method and testing it on a case study. The method mostly fulfils the research objective, but some limitations have been found during the development and testing. In this section, the limitations are discussed, together with recommendations for future research.

The manual tasks in the mapping step and the performing of routing and visualizing steps are time-consuming, slowing down the process significantly. Developing a tool or interface would be ideal when the method is used more often to make the process less time-consuming, addressing the abovementioned issues. This tool would need some variables unique to a city as input which would be used to perform the other steps automatically.

Some software issues were found when performing the case study. The issues mainly manifested when creating the routes and when visualizing the routes. In the routing step, generating the routes takes a long time because of the limit on requests per minute and day to the ORS Tools' server set by the developers. For this reason, the origin and destination areas are aggregated into larger areas. The aggregation meant that some details were lost. In the visualization step, the *'v.clean'* tool from the GRASS plugin uses a lot of computing power, resulting in crashes. When using the given alternative, the *'explode'* tool (see section 3.6.1), many separate features are created, slowing the software down significantly. In future research, it would be recommended to find other solutions to get the same results or to find different software that can perform the needed actions for the method more efficiently to perform the troublesome steps. Using hardware with high computing power might help with a part of the problem.

In section 2.7.1, spatial and temporal aggregation was mentioned as a major shortcoming of the four-step model. Significant aggregation is also part of the development method. Because of the constraints of the input data, origin points, destination points, arrival time, travel purpose, and arrival period are aggregated into a small number of categories. The aggregation is needed because of the limited number of cases in the input data, the way the weights are created, and the software limitation mentioned before. If more input data is available and the software limitations are solved, less aggregation will be needed, and a more detailed analysis can be performed. The aggregation needs to be considered when interpreting the output of the method. For future development of the method, using other revealed data, as described in section 2.7.2, might be beneficial to increase the data input and improve validation. It may also be possible to use other data sets to calibrate the findings of the method, thus improving the reliability of the absolute values.

Validation is an integral part of the method. However, it is likely to be challenging to find suitable data for external validation. Data was found for the case study, but it was not ideal because the measurement locations differed from the defined entry points, as shown in Figure 40 (p. 64). It is recommended look for more specific traffic counts. Another possibility is to perform a sensitivity analysis in future studies to give the output more legitimacy.

The mentioned limitations must be considered when using the developed method and interpreting the results. The given recommendations can be used in possible future research into the topic of entry points of cities. Despite the mentioned limitations, the method does make it possible to generate new insights into traffic entering cities in an accessible manner. The method can help to make or substantiate difficult policy decisions.

(Page intentionally left blank)

# 6. References


Abrahamsson, T. (1998). *Estimation of Origin-Destination Matrices Using Traffic Counts - A Literature Survey* [International Institute for Applied Systems Analysis]. https://pure.iiasa.ac.at/5627

Ahern, A., Weyman, G., Redelbach, M., Schulz, A., Akkermans, L., Vannacci, L., Anoyrkati, E., & van Grinsven, A. (2013). *Analysis of national travel surveys in Europe: OPTIMISM WP2: Harmonisation of national travel statistics in Europe*. Publications Office of the European Union. https://doi.org/10.2788/59474

Apronti, D. T., & Ksaibati, K. (2018). Four-step travel demand model implementation for estimating traffic volumes on rural low-volume roads in Wyoming. *Transportation Planning and Technology*, *41*(5), 557–571. https://doi.org/10.1080/03081060.2018.1469288

Ashcroft, M., Austin, R., Barnes, K., MacDonald, D., Makin, S., Morgan, S., Taylor, R., & Scolley, P. (2016). Expert judgement. *British Actuarial Journal*, *21*(2), 314–363. https://doi.org/10.1017/S1357321715000239

Aumann, C. A. (2007). A methodology for developing simulation models of complex systems. *Ecological Modelling*, *202*(3–4), 385–396. https://doi.org/10.1016/j.ecolmodel.2006.11.005

Azari, K. A., Arintono, S., Hamid, H., & Davoodi, S. R. (2013). Evaluation of demand for different trip purposes under various congestion pricing scenarios. *Journal of Transport Geography*, *29*, 43–51. https://doi.org/10.1016/j.jtrangeo.2013.01.001

Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, *12*(5), 464–474. https://doi.org/10.1016/j.tranpol.2005.06.008

Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*. The University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/C/bo5971175.html

Barcelö, J., Montero, L., Marqués, L., & Carmona, C. (2010). Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring. *Transportation Research Record: Journal of the Transportation Research Board*, *2175*(1), 19–27. https://doi.org/10.3141/2175-03

Bates, J. (2007). History of Demand Modelling. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (Vol. 1, pp. 11–34). Emerald Group Publishing Limited. https://doi.org/10.1108/9780857245670-002

Bell, M. G. H. (1983). The Estimation of an Origin-Destination Matrix from Traffic Counts. *Transportation Science*, *17*(2), 198–217. http://www.jstor.org/stable/25768090

Bera, S., & Rao, K. V. K. (2011). Estimation of origin-destination matrix from traffic counts: the state of the art. *European Transport \ Trasporti Europei*, *49*, 3–23.

Bonnel, P. (2003). Postal, Telephone, and Face-to-Face Surveys: How Comparable are They? In *Transport Survey Quality and Innovation* (pp. 215–237). Emerald Group Publishing Limited. https://doi.org/10.1108/9781786359551-012

Bonnel, P., Fekih, M., & Smoreda, Z. (2018). Origin-Destination estimation using mobile network probe data. *Transportation Research Procedia*, *32*, 69–81. https://doi.org/10.1016/j.trpro.2018.10.013

Borgers, A., & Timmermans, H. J. P. (1986). City centre entry points, store location patterns and pedestrian route choice behaviour: A microlevel simulation model. *Socio-Economic Planning Sciences*, *20*(1), 25–31. https://doi.org/10.1016/0038-0121(86)90023-6

Bos, llona, Molin, E., Timmermans, H., & van der Heijden, R. (2003). Cognition and Relative Importance Underlying Consumer Valuation of Park-and-Ride Facilities. *Transportation Research Record: Journal of the Transportation Research Board*, *1835*(1), 121–127. https://doi.org/10.3141/1835-15

Bostanci, I., Gootzen, Y., & Lugtig, P. (2023). *Data Linkage to Validate and Calibrate Traffic Estimations on a Nationwide Scale: A Framework for Official Statistics*. https://www.cbs.nl/en-gb/background/2023/11/validation-and-calibration-of-traffic-estimations

Bovy, P. H. L., & Stern, E. (1990). *Route Choice: Wayfinding in Transport Networks* (1st ed., Vol. 9). Springer Netherlands. https://doi.org/10.1007/978-94-009-0633-4

Bucsky, P., & Juhász, M. (2022). Long-term evidence on induced traffic: A case study on the relationship between road traffic and capacity of Budapest bridges. *Transportation Research Part A: Policy and Practice*, *157*, 244–257. https://doi.org/10.1016/j.tra.2022.01.018

Burchell, R., Downs, A., McCann, B., & Mukherji, S. (2005). *Sprawl Costs: Economic Impacts of Unchecked Development*. Island Press. https://worldcat.org/nl/title/607787931

Burg, R. (2021, November 21). Verkeer weren uit centrum Eindhoven lukt niet: nieuwe parking kostte 20 miljoen maar blijft leeg. *Eindhovens Dagblad*. https://www.ed.nl/eindhoven/verkeer-weren-uit-centrum-eindhoven-lukt-niet-nieuwe-parking-kostte-20-miljoen-maar-blijft-leeg~a98776ab/

Cascetta, E., & Nguyen, S. (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research Part B: Methodological*, *22*(6), 437–455. https://doi.org/10.1016/0191-2615(88)90024-0

Cascetta, E., & Russo, F. (1997). Calibrating aggregate travel demand models with traffic counts: Estimators and statistical performance. *Transportation*, *24*, 271–293. https://doi.org/10.1023/A:1004968411792

CBS. (2016a). *Avondspits*. https://www.cbs.nl/nl-nl/nieuws/2016/37/studenten-en-scholieren-pieken-in-de-ochtendspits/avondspits

CBS. (2016b). *Ochtendspits*. https://www.cbs.nl/nl-nl/nieuws/2016/37/studenten-en-scholieren-pieken-in-de-ochtendspits/ochtendspits

CBS. (2021). *Wijk- en buurtkaart 2021*. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2021

CBS. (2022a). *Mobility; per person, modes of travel, purposes of travel and regions*. https://www.cbs.nl/en-gb/figures/detail/84710ENG

CBS. (2022b). *Kerncijfers per postcode*. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

CBS. (2022c). *Onderweg in Nederland*. https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/onderweg-in-nederland

Cervero, R., Caldwell, B., & Cuellar, J. (2013). Bike-and-Ride: Build It and They Will Come. *Journal of Public Transportation*, *16*(4), 83–105. https://doi.org/10.5038/2375-0901.16.4.5

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chapin, S. (1974). *Human Activity Patterns in the City: Thing People Do in Time and Space*. Wiley.

Chapleau, R., & Morency, C. (2005). Dynamic Spatial Analysis of Urban Travel Survey Data using GIS. *5th Annual ESRI International User Conference*, 1–14. https://proceedings.esri.com/library/userconf/proc05/papers/pap1232.pdf

Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*, *68*, 285–299. https://doi.org/10.1016/j.trc.2016.04.005

Chu, Z., Cheng, L., & Chen, H. (2012). A Review of Activity-Based Travel Demand Modeling. *CICTP 2012*, 48–59. https://doi.org/10.1061/9780784412442.006

Cools, M., Moons, E., Creemers, L., & Wets, G. (2010). Changes in Travel Behavior in Response to Weather Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, *2157*(1), 22–28. https://doi.org/10.3141/2157-03

Davis, S. C., & Boundy, R. G. (2019). *Transportation Energy Data Book: Edition 37* (37th ed.). https://doi.org/10.2172/1493136

Departement Mobiliteit en Openbare Werken. (n.d.). *Onderzoek Verplaatsingsgedrag Vlaanderen (OVG)*. Retrieved June 26, 2023, from https://www.vlaanderen.be/mobiliteit-en-openbare-werken/onderzoek-verplaatsingsgedrag-vlaanderen-ovg

Department of Transportation UK. (n.d.). *National Travel Survey (NTS)*. Retrieved May 14, 2023, from https://www.gov.uk/government/collections/national-travel-survey-statistics#latest-national-travel-survey-statistics

Dijk, M., de Haes, J., & Montalvo, C. (2013). Park-and-Ride motivations and air quality norms in Europe. *Journal of Transport Geography*, *30*, 149–160. https://doi.org/10.1016/j.jtrangeo.2013.04.008

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271. https://doi.org/10.1007/BF01386390

Domarchi, C., Tudela, A., & González, A. (2008). Effect of attitudes, habit and affective appraisal on mode choice: an application to university workers. *Transportation*, *35*(5), 585–599. https://doi.org/10.1007/s11116-008-9168-6

Dufec. (2021). *Basec: Verkeersintensiteiten en verkeersonderzoek*. https://basec.nl/

Egu, O., & Bonnel, P. (2020). How comparable are origin-destination matrices estimated from automatic fare collection, origin-destination surveys and household travel survey? An empirical investigation in Lyon. *Transportation Research Part A: Policy and Practice*, *138*, 267–282. https://doi.org/10.1016/j.tra.2020.05.021

Ertman, S., Ertman, J., & Zakharov, D. (2016). Adaptation of urban roads to changing of transport demand. *E3S Web of Conferences*, *6*, 01013. https://doi.org/10.1051/e3sconf/20160601013

Ewing, R., Hamidi, S., Tian, G., Proffitt, D., Tonin, S., & Fregolent, L. (2018). Testing Newman and Kenworthy's Theory of Density and Automobile Dependence. *Journal of Planning Education and Research*, *38*(2), 167–182. https://doi.org/10.1177/0739456X16688767

Félix, R., Cambra, P., & Moura, F. (2020). Build it and give 'em bikes, and they will come: The effects of cycling infrastructure and bike-sharing system in Lisbon. *Case Studies on Transport Policy*, *8*(2), 672–682. https://doi.org/10.1016/j.cstp.2020.03.002

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). SAGE Publications.

Geisberger, R., Sanders, P., Schultes, D., & Delling, D. (2008). Contraction Hierarchies: Faster and Simpler Hierarchical Routing in Road Networks. In *Experimental Algorithms* (pp. 319–333). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68552-4_24

Gemeente Eindhoven. (2019). *Buitengebied (P+R Aalsterweg) Toelichting*. https://www.planviewer.nl/imro/files/NL.IMRO.0772.80325-0301/t_NL.IMRO.0772.80325-0301.html

Gemeenteraad Eindhoven. (2013). *Eindhoven op weg*. https://www.eindhoven.nl/sites/default/files/2020-11/Eindhoven%20op%20weg%202014.pdf

Gössling, S. (2020). Why cities need to take road space from cars - and how this could be done. *Journal of Urban Design*, *25*(4), 443–448. https://doi.org/10.1080/13574809.2020.1727318

GRASS Development Team. (2016). *Geographic Resources Analysis Support System (GRASS) Software, Version 7.2.0. Open Source Geospatial Foundation*. http://grass.osgeo.org

Grijzen, J. (2010). *What Do Consultants Do in a Regional Spatial Planning in the Netherlands*. Amsterdam University Press.

Groningen Bereikbaar. (2021, May 27). *P+R's in Groningen steeds populairder*. https://www.groningenbereikbaar.nl/nieuws/4

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5), 646–675. https://doi.org/10.1093/poq/nfl033

Guan, X., Wang, D., & Jason Cao, X. (2020). The role of residential self-selection in land use-travel research: a review of recent findings. *Transport Reviews*, *40*(3), 267–287. https://doi.org/10.1080/01441647.2019.1692965

Hägerstrand, T. (1970). What about people in Regional Science? *Papers of the Regional Science Association*, *24*(1), 6–21. https://doi.org/10.1007/BF01936872

Harris, R. (2004). *Creeping Conformity: How Canada Became Suburban, 1900-1960* (C. Heron, Ed.). University of Toronto Press. https://www.jstor.org/stable/10.3138/j.ctt1287wb6

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, *40*(1), 27–37. https://doi.org/10.1179/000870403235002042

Holtom, B., Baruch, Y., Aguinis, H., & A Ballinger, G. (2022). Survey response rates: Trends and a validity assessment framework. *Human Relations*, *75*(8), 1560–1584. https://doi.org/10.1177/00187267211070769

Horváth, M. T., Mátrai, T., & Tóth, J. (2017). Route Planning Methodology with Four-step Model and Dynamic Assignments. *Transportation Research Procedia*, *27*, 1017–1025. https://doi.org/10.1016/j.trpro.2017.12.127

Hymel, K. M., Small, K. A., & Dender, K. Van. (2010). Induced demand and rebound effects in road transport. *Transportation Research Part B: Methodological*, *44*(10), 1220–1241. https://doi.org/10.1016/j.trb.2010.02.007

Ivanchev, J., Aydt, H., & Knoll, A. (2015). Spatial and Temporal Analysis of Mismatch between Planned Road Infrastructure and Traffic Demand in Large Cities. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, *2015-October*, 1463–1470. https://doi.org/10.1109/ITSC.2015.239

Jackson, L. J., Trebitz, A. S., & Cottingham, K. L. (2000). An Introduction to the Practice of Ecological Modeling. *BioScience*, *50*(8), 694–706. https://doi.org/10.1641/0006-3568(2000)050[0694:AITTPO]2.0.CO;2

Jafari, A., Both, A., Singh, D., Gunn, L., & Giles-Corti, B. (2022). Building the road network for city-scale active transport simulation models. *Simulation Modelling Practice and Theory*, *114*, 102398. https://doi.org/10.1016/j.simpat.2021.102398

Janzen, M., Vanhoof, M., Smoreda, Z., & Axhausen, K. W. (2018). Closer to the total? Long-distance travel of French mobile phone users. *Travel Behaviour and Society*, *11*, 31–42. https://doi.org/10.1016/j.tbs.2017.12.001

Jiao, J., & Dillivan, M. (2013). Transit Deserts: The Gap between Demand and Supply. *Journal of Public Transportation*, *16*(3), 23–39. https://doi.org/10.5038/2375-0901.16.3.2

Jomehpour Chahar Aman, J., & Smith-Colin, J. (2020). Transit Deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, *89*, 102869. https://doi.org/10.1016/j.jtrangeo.2020.102869

Karamychev, V., & van Reeven, P. (2011). Park-and-ride: Good for the city, good for the region? *Regional Science and Urban Economics*, *41*(5), 455–464. https://doi.org/10.1016/j.regsciurbeco.2011.03.002

Katsnelson, A. (2021). Colour me better: fixing figures for colour blindness. *Nature*, *598*(7879), 224–225. https://doi.org/10.1038/d41586-021-02696-z

Ke, J., Zheng, H., Yang, H., & Chen, X. (Michael). (2017). Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. *Transportation Research Part C: Emerging Technologies*, *85*, 591–608. https://doi.org/10.1016/j.trc.2017.10.016

Kelly, J. A., & Clinch, J. P. (2006). Influence of varied parking tariffs on parking occupancy levels by trip purpose. *Transport Policy*, *13*(6), 487–495. https://doi.org/10.1016/j.tranpol.2006.05.006

Kennisplatform CROW. (2015). *Handboek P+R - Kennis voor beleid, beheer en praktijk* (Ed Graumans, Ed.).

Kerr, L. A., & Goethel, D. R. (2014). Simulation Modeling as a Tool for Synthesis of Stock Identification Information. In *Stock Identification Methods* (pp. 501–533). Elsevier. https://doi.org/10.1016/B978-0-12-397003-9.00021-7

Krizek, K. J. (2003). Neighborhood services, trip purpose, and tour-based travel. *Transportation*, *30*, 387–410. https://doi.org/10.1023/A:1024768007730

McNally, M. G. (2007). The Four-Step Model. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (Vol. 1, pp. 35–53). Emerald Group Publishing Limited. https://doi.org/10.1108/9780857245670-003

McNally, M. G., & Rindt, C. R. (2007). The Activity-Based Approach. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (Vol. 1, pp. 55–73). Emerald Group Publishing Limited. https://doi.org/10.1108/9780857245670-004

Meek, S., Ison, S., & Enoch, M. (2008). Role of Bus-Based Park and Ride in the UK: A Temporal and Evaluative Review. *Transport Reviews*, *28*(6), 781–803. https://doi.org/10.1080/01441640802059152

Meyer, M. D., & ITE (Institute of Transportation Engineers). (2016). *Transportation planning handbook* (4th ed.). Wiley. https://worldcat.org/title/953662335

Mingardo, G. (2013). Transport and environmental effects of rail-based Park and Ride: evidence from the Netherlands. *Journal of Transport Geography*, *30*, 7–16. https://doi.org/10.1016/j.jtrangeo.2013.02.004

Mokhtarian, P. L., Samaniego, F. J., Shumway, R. H., & Willits, N. H. (2002). Revisiting the notion of induced traffic through a matched-pairs study. *Transportation*, *29*, 193–220. https://doi.org/10.1023/A:1014221024304

Nieuwenhuijsen, M. J., & Khreis, H. (2016). Car free cities: Pathway to healthy urban living. *Environment International*, *94*, 251–262. https://doi.org/10.1016/j.envint.2016.05.032

OECD. (2010). *The Future for Interurban Passenger Transport*. OECD. https://doi.org/10.1787/9789282102688-en

Oldenziel, R., Emanuel, M., Albert de la Bruheze, A. A., & Veraart, F. (2016). *Cycling Cities: The European Experience*. Foundation of the History of Technology and LMU Rachel Carson Center. http://www.cyclingcities.info/

OpenStreetMap contributors. (2023). *OpenStreetMap database*. https://www.openstreetmap.org/

Pitsiava-Latinopoulou, M., & Iordanopoulos, P. (2012). Intermodal Passengers Terminals: Design Standards for Better Level of Service. *Procedia - Social and Behavioral Sciences*, *48*, 3297–3306. https://doi.org/10.1016/j.sbspro.2012.06.1295

QGIS Development Team. (2020). *QGIS Geographic Information System* (3.16.16). QGIS Association. https://www.qgis.org

Rasouli, S., & Timmermans, H. (2014). Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences*, *18*(1), 31–60. https://doi.org/10.1080/12265934.2013.835118

Rijksoverheid. (2021). *XII Infrastructuur en Waterstaat Rijksbegroting 2022*. https://www.rijksoverheid.nl/onderwerpen/wegen/documenten/begrotingen/2021/09/21/xii-infrastructuur-en-waterstaat-rijksbegroting-2022

Ritchie, H., & Roser, M. (2018). Urbanization. *Our World in Data*. https://ourworldindata.org/urbanization

Rodrigue, J.-P. (2020). *The Geography of Transport Systems* (5th ed.). Routledge. https://doi.org/10.4324/9780429346323

Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological Modelling*, *90*(3), 229–244. https://doi.org/10.1016/0304-3800(95)00152-2

Santos, A., Mcguckin, N., Nakamoto, H. Y., Gray, D., Liss, S., Santos, A., Mcguckin, N., Nakamoto, H. Y., Gray, D., Liss, S., & Systematics, C. (2011). *Summary of Travel Trends 2009 National Household Travel Survey*. http://nhts.ornl.gov/2009/pub/stt.pdf

Shen, X., Zhou, Y., Jin, S., & Wang, D. (2020). Spatiotemporal influence of land use and household properties on automobile travel demand. *Transportation Research Part D: Transport and Environment*, *84*, 102359. https://doi.org/10.1016/j.trd.2020.102359

Statistics Canada. (n.d.). *National Travel Survey (NTS)*. Retrieved May 14, 2023, from https://www.statcan.gc.ca/en/survey/household/5232

Statistics Solutions. (n.d.). *Testing Assumptions of Linear Regression in SPSS*. Retrieved June 25, 2023, from https://www.statisticssolutions.com/testing-assumptions-of-linear-regression-in-spss/

Stinson, M. A., & Bhat, C. R. (2004). *A Comparison of the Route Preferences of Experienced and Inexperienced Bicycle Commuters*.

Stopher, P. R., & Greaves, S. P. (2007). Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*, *41*(5), 367–381. https://doi.org/10.1016/j.tra.2006.09.005

Thomas, T., & Tutert, B. (2015). Route choice behavior in a radial structured urban network: Do people choose the orbital or the route through the city center? *Journal of Transport Geography*, *48*, 85–95. https://doi.org/10.1016/j.jtrangeo.2015.08.013

Tian, L.-J., Lyu, C.-R., & Zhao, Y.-X. (2017). The Park-and-Ride Behavior in a Cumulative Prospect Theory-Based Model. *Journal of the Operations Research Society of China*, *5*(3), 363–376. https://doi.org/10.1007/s40305-017-0153-7

U.S. Department of Transportation. (2021). *Fiscal Year 2024 Budget Highlights*. https://www.transportation.gov/mission/budget/fiscal-year-2024-budget-highlights

van Hal, E., Barten, M., Contreras, V., & Ruinaard, S. (2022). *Actualisatie Eindhoven op Weg 2022*. https://www.eindhoven.nl/bestuur-en-beleid/beleid/visies-plannen-en-verordeningen

Wada, K., Seo, T., & Shiomi, Y. (2021). Bottleneck. In *International Encyclopedia of Transportation* (pp. 134–142). Elsevier. https://doi.org/10.1016/B978-0-08-102671-7.10306-9

Wagner, B., Human, S., & Winkler, T. (2021, January 1). Bias in Geographic Information Systems: The Case of Google Maps. *Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2021.103

Waller, S. T., Chand, S., Zlojutro, A., Nair, D., Niu, C., Wang, J., Zhang, X., & Dixit, V. V. (2021). Rapidex: A Novel Tool to Estimate Origin–Destination Trips Using Pervasive Traffic Data. *Sustainability*, *13*(20), 11171. https://doi.org/10.3390/su132011171

Wallner, G., Kriglstein, S., Chung, E., & Kashfi, S. A. (2018). Visualisation of trip chaining behaviour and mode choice using household travel survey data. *Public Transport*, *10*(3), 427–453. https://doi.org/10.1007/s12469-018-0183-5

Wardrop, J. G. (1952). Road Paper. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers*, *1*(3), 325–362. https://doi.org/10.1680/ipeds.1952.11259

Weiner, E. (2016). *Urban Transportation Planning in the United States* (5th ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-39975-1

Wolf, J., Bricka, S., Ashby, T., & Gorugantua, C. (2004). Advances in the Application of GPS to Household Travel Surveys. *National Household Travel Survey Conference*.

Wolf, J., Guensler, R., & Bachman, W. (2001). Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data. *Transportation Research Record: Journal of the Transportation Research Board*, *1768*(1), 125–134. https://doi.org/10.3141/1768-15

Wolf, J., Oliveira, M., & Thompson, M. (2003). Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey. *Transportation Research Record: Journal of the Transportation Research Board*, *1854*(1), 189–198. https://doi.org/10.3141/1854-21

Yang, F., Jin, P. J., Cheng, Y., Zhang, J., & Ran, B. (2015). Origin-Destination Estimation for Non-Commuting Trips Using Location-Based Social Networking Data. *International Journal of Sustainable Transportation*, *9*(8), 551–564. https://doi.org/10.1080/15568318.2013.826312

Zheng, P., & Mike, M. (2012). An Investigation on the Manual Traffic Count Accuracy. *Procedia - Social and Behavioral Sciences*, *43*, 226–231. https://doi.org/10.1016/j.sbspro.2012.04.095

Zijlstra, T., Vanoutrive, T., & Verhetsel, A. (2015). A meta-analysis of the effectiveness of park-and-ride facilities. *European Journal of Transport and Infrastructure Research*, *15*(4), 597–612. https://doi.org/10.18757/ejtir.2015.15.4.3099

# 7. Appendix

## Contents of Appendix

## Appendix 1: QGis route assignment plugin comparison

Based on six criteria, eight plugins have been tested for the route assignment, as shown in Table 6. The plugins were selected by searching for the relevant plugins on the plugin database on the official website: https://plugins.qgis.org/. ORS Tools plugin conforms to all but one constraint, it is not fully free to use. However, although, ORS Tools has a daily limit of requests on the free version, it was found that it was high enough for performing the method. The question marks in Table 6 indicate that the information on the particular constraint is unknown, as the availability was limited.

*Table 6 QGis route assignment plugin comparison*

| **Plugins** | **Constraints** | | | | | |
|---|---|---|---|---|---|---|
| | **Still supported** | **Open-source and free** | **Loading data set** | **Available in process modeler** | **Many calculations at once** | **Supported modes** |
| *Valhalla* | *Yes* | *No* | *?* | *?* | *?* | |
| *Hqgis* | *Yes* | *Yes* | *No* | *No* | *?* | |
| *ORS Tools* | *Yes* | *Daily limit* | *Yes* | *Yes* | *Yes* | *Car, bike, walking, wheelchair* |
| *OpenTripPlanner* | *Yes* | *Yes* | *Yes* | *No* | *Yes* | *PT, bike, walking* |
| *Online route mapper* | *Yes* | *Yes* | *No* | *No* | *No* | |
| *OSM route* | *Yes* | *Yes* | *No* | *No* | *?* | |
| *TravelTime platform* | *Yes* | *Monthly limit* | *Yes* | *Yes* | *Results in crash* | |
| *OSM Tools* | *No* | *?* | *?* | *?* | *?* | |

# Appendix 2: SPSS syntax for export

## Appendix 2.1 Export cross tab origin municipalities and destination areas general

Figure 44 shows the created general syntax to export to selected data from SPSS to Excel to be further analysed. The red, underlined text (including curly braces) needs to be filled in by the user. For the filter, the chosen criteria need to be added with the specified variables filled in, a general example has been filled in. The other red, underlined text are not influential on the output, and can freely be chosen by the user. The script can be copied and pasted behind each other to export the data subsets for all subcategories at once, the sheet name should be different for each data set to avoid overwriting the data.

```
OUTPUT CLOSE ALL

DATASET ACTIVATE  {dataset name}.
USE ALL.
COMPUTE filter_$=({Area of destination="Chosen postal code areas" AND Area of origin ≠"Chosen
postal code areas" AND Travel distance ≥ x AND One trip = 1 AND Transport mode = x}).
VARIABLE LABELS filter_$ '(select)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

DATASET ACTIVATE DataSet1.
CROSSTABS
  /TABLES=”*origin municipality*” BY PC
  /FORMAT=AVALUE TABLES
  /CELLS=COUNT
  /COUNT ROUND CELL.

* Export Output.
OUTPUT EXPORT
   /CONTENTS  EXPORT=ALL  LAYERS=PRINTSETTING  MODELVIEWS=PRINTSETTING
   /XLSX  DOCUMENTFILE=’{file location/file name}’    OPERATION=CREATESHEET SHEET = "{sheet
name}"
   LOCATION=STARTCELL('A1')  NOTESCAPTIONS=YES.

OUTPUT CLOSE ALL
```

*Figure 44 General SPSS syntax for exporting the selected data from the data set*

## Appendix 2.2 Export cross tab origin municipalities and destination areas case study

Figure 45 shows the created syntax for the case study to export to selected data from SPSS to Excel to be further analysed. The filter has been filled in. The red, underlined text (including curly braces) are not influential on the output, and should be chosen by the user. The script can be copied and pasted behind each other to export the data subsets for all subcategories at once, the sheet name should be different for each data set to avoid overwriting the data.

```
OUTPUT CLOSE ALL

DATASET ACTIVATE {dataset name}.
USE ALL.
COMPUTE filter_$=(DestArea = 1 AND OriginArea = 1 AND AfstV ≥ 50 AND Verpl = 1 AND Car = 1).
VARIABLE LABELS filter_$ '(select)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

DATASET ACTIVATE DataSet1.
CROSSTABS
  /TABLES=VertGem BY PC
  /FORMAT=AVALUE TABLES
  /CELLS=COUNT
  /COUNT ROUND CELL.

* Export Output.
OUTPUT EXPORT
   /CONTENTS  EXPORT=ALL  LAYERS=PRINTSETTING  MODELVIEWS=PRINTSETTING
   /XLSX  DOCUMENTFILE='{file location/file name}'   OPERATION=CREATESHEET SHEET = "{sheet
name}"
   LOCATION=STARTCELL('A1')  NOTESCAPTIONS=YES.

OUTPUT CLOSE ALL
```

*Figure 45 Case study-specific SPSS syntax for exporting the selected data from the data set*

## Appendix 3: Linking the exported data to the municipality map

For every subcategory, a cross-tabulation has been exported from SPSS to Excel. Each subcategory is exported to its own sheet, as shown in Figure 46, with one sheet for all trips towards the city and one sheet for combining the sheet into one. An example of what is in one of the sheets is given in Figure 47. The list of municipality names has been copied from the municipality map from CBS. Using the formula shown in Figure 48, all values from each sheet are automatically connected to the right municipality.

| Combine | All | Peak | NonPeak | Weekend | Workday | Work | NonWork |
|---|---|---|---|---|---|---|---|

*Figure 46 One Excel sheet per subcategory*

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 27 | | | | | | Vertrekgemeente * PC Crosstabulation | | | | | | | | | | |
| 28 | Count | | | | | | | | | | | | | | | |
| 29 | | | | | | | | PC | | | | | | | | |
| 30 | | | 11 | 21 | 22 | 23 | 24 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | Total |
| 31 | Vertrekgemeente | Geen vertrekpunt in Nederland | 0 | 63474 | 233120 | 0 | 130571 | 0 | 126422 | 208243 | 0 | 0 | 114317 | 0 | 285976 | 1162123 |
| 32 | | Groningen | 0 | 133251 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 133251 |
| 33 | | Almere | 0 | 0 | 115485 | 113538 | 0 | 0 | 0 | 108393 | 90171 | 0 | 20216 | 0 | 0 | 447803 |
| 34 | | Stadskanaal | 0 | 0 | 0 | 0 | 19451 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 19451 |
| 35 | | Veendam | 0 | 0 | 0 | 0 | 0 | 0 | 28724 | 0 | 0 | 0 | 0 | 0 | 0 | 28724 |

*Figure 47 Cross tabulation destination areas and origin municipalities for all trips*

B202

```
=IFERROR(INDEX(INDIRECT("'"&B$1&"'!$B$31:$C$2000"); MATCH($A202;(INDIRECT("'"&B$1&"'!$B$31:$B$2000"));0);2);0)
```

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | GM_NAAM | All | All | All | All | All | All | All | All | All | All | All | All |
| 200 | Baarle-Nassau | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 201 | Bergen op Zoom | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 153293 | 0 | 0 |
| 202 | Best | 362612 | 167376 | 655257 | 503479 | 247626 | 328514 | 1674072 | 2978679 | 209966 | 355704 | 38623 | 471481 |
| 203 | Boekel | 0 | 175860 | 0 | 0 | 0 | 164576 | 79110 | 0 | 0 | 0 | 0 | 0 |
| 204 | Boxmeer | 410941 | 0 | 0 | 115165 | 81345 | 0 | 295402 | 0 | 0 | 0 | 60904 | 0 |

*Figure 48 Formula to match all values to the right municipality*

## Appendix 4: Script route generation

The routes are generated using the python script in Figure 49. First, the origin points where more than 0 trips to the chosen destination point originate are selected. Next, the actual routes are created. The red, underlined text must be replaced (including curly braces) and filled in by the user with the correct information.

```
"""
Model exported as python.
Name : Get routes
Group :
With QGIS : 31616
"""

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
from qgis.core import QgsProcessingParameterExpression
from qgis.core import QgsProcessingParameterVectorLayer
from qgis.core import QgsProcessingParameterString
from qgis.core import QgsProcessingParameterFeatureSink
import processing


class GetRoutes(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        self.addParameter(QgsProcessingParameterExpression('destination', 'destination
expression', parentLayerParameterName='', defaultValue='{Selected destination point} = 0))
        self.addParameter(QgsProcessingParameterVectorLayer('destinationpoint', 'destination
point layer', defaultValue=None))
        self.addParameter(QgsProcessingParameterString('Joinedfields', 'Joined fields',
multiLine=False, defaultValue='{All names of fields related to the selected destination point}'))
        self.addParameter(QgsProcessingParameterFeatureSink('JoinedDirections', 'Joined
directions', optional=True, type=QgsProcessing.TypeVectorAnyGeometry, createByDefault=True,
defaultValue=None))

    def processAlgorithm(self, parameters, context, model_feedback):
        # Use a multi-step feedback, so that individual child algorithm progress reports are adjusted
for the
        # overall progress through the model
        feedback = QgsProcessingMultiStepFeedback(3, model_feedback)
        results = {}
        outputs = {}

        # Extract by expression the origin points where more than 0 trips originate
        alg_params = {
```

```
            'EXPRESSION': parameters['destination'],
            'INPUT': '{Origin point layer}',
            'FAIL_OUTPUT': QgsProcessing.TEMPORARY_OUTPUT,
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['ExtractByExpression'] = processing.run('native:extractbyexpression', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(1)
        if feedback.isCanceled():
            return {}

        # Directions from points 2 layers generating the routes
        alg_params = {
            'INPUT_AVOID_BORDERS': None,
            'INPUT_AVOID_COUNTRIES': '',
            'INPUT_AVOID_FEATURES': [],
            'INPUT_AVOID_POLYGONS': None,
            'INPUT_END_FIELD': '',
            'INPUT_END_LAYER': parameters['destinationpoint'],
            'INPUT_MODE': 1,
            'INPUT_PREFERENCE': 0,
            'INPUT_PROFILE': 0,
            'INPUT_PROVIDER': 0,
            'INPUT_SORT_END_BY': '',
            'INPUT_SORT_START_BY': '',
            'INPUT_START_FIELD': '{municipality name field}',
            'INPUT_START_LAYER': outputs['ExtractByExpression']['FAIL_OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['DirectionsFromPoints2Layers'] = processing.run('ORS
Tools:directions_from_points_2_layers', alg_params, context=context, feedback=feedback,
is_child_algorithm=True)

        feedback.setCurrentStep(2)
        if feedback.isCanceled():
            return {}

        # Join attributes by field value
        alg_params = {
            'DISCARD_NONMATCHING': False,
            'FIELD': 'FROM_ID',
            'FIELDS_TO_COPY': parameters['Joinedfields'],
            'FIELD_2': '{municipality name field}',
            'INPUT': outputs['DirectionsFromPoints2Layers']['OUTPUT'],
            'INPUT_2': '{Data file}',
```

```
            'METHOD': 1,
            'PREFIX': '',
            'OUTPUT': parameters['JoinedDirections']
        }
        outputs['JoinAttributesByFieldValue'] = processing.run('native:joinattributestable',
alg_params, context=context, feedback=feedback, is_child_algorithm=True)
        results['JoinedDirections'] = outputs['JoinAttributesByFieldValue']['OUTPUT']
        return results

    def name(self):
        return 'Get routes'

    def displayName(self):
        return 'Get routes'

    def group(self):
        return ''

    def groupId(self):
        return ''

    def createInstance(self):
        return GetRoutes()
```

*Figure 49 Python script for route generation*

## Appendix 5: Merge routes and calculate number of trips using the entry points

The generated routes are merged and the number of trips using the entry points is calculated using the created script shown in Figure 50. The red, underlined text (including curly braces) needs to be replaced with the specific names for field and layers used in file the script is used for.

```
"""
Model exported as python.
Name : entry numbers
Group :
With QGIS : 31616
"""

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
from qgis.core import QgsProcessingParameterFeatureSink
import processing


class Model(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        self.addParameter(QgsProcessingParameterFeatureSink('EntryNumbers', 'Entry numbers',
type=QgsProcessing.TypeVectorAnyGeometry, createByDefault=True, defaultValue=None))

    def processAlgorithm(self, parameters, context, model_feedback):
        # Use a multi-step feedback, so that individual child algorithm progress reports are adjusted
for the
        # overall progress through the model
        feedback = QgsProcessingMultiStepFeedback(2, model_feedback)
        results = {}
        outputs = {}

        # Merge vector layers
        alg_params = {
            'CRS': None,
            'LAYERS': [{ALL route layers}],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['MergeVectorLayers'] = processing.run('native:mergevectorlayers', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(1)
        if feedback.isCanceled():
            return {}
```

```
        # Join attributes by location (summary)
        alg_params = {
            'DISCARD_NONMATCHING': False,
            'INPUT': '{Layer with entry location lines or polygons} ',
            'JOIN': outputs['MergeVectorLayers']['OUTPUT'],
            'JOIN_FIELDS': [''],
            'PREDICATE': [0],
            'SUMMARIES': [5],
            'OUTPUT': parameters['EntryNumbers']
        }
        outputs['JoinAttributesByLocationSummary'] = processing.run('qgis:joinbylocationsummary',
alg_params, context=context, feedback=feedback, is_child_algorithm=True)
        results['EntryNumbers'] = outputs['JoinAttributesByLocationSummary']['OUTPUT']
        return results

    def name(self):
        return 'model'

    def displayName(self):
        return 'model'

    def group(self):
        return ''

    def groupId(self):
        return ''

    def createInstance(self):
        return Model()
```

*Figure 50 Python script for merging routes and calculating the number of trips using the entry points*

## Appendix 6: Preparations visualization of the routes

The preparations for the visualization of the routes are made using the script shown in Figure 51. The red, underlined text (including curly braces) needs to be replaced with the specific names for field and layers used in file the script is used for.

```
"""
Model exported as python.
Name : model
Group :
With QGIS : 31616
"""

from qgis.core import QgsProcessing
from qgis.core import QgsProcessingAlgorithm
from qgis.core import QgsProcessingMultiStepFeedback
import processing


class Model(QgsProcessingAlgorithm):

    def initAlgorithm(self, config=None):
        pass

    def processAlgorithm(self, parameters, context, model_feedback):
        # Use a multi-step feedback, so that individual child algorithm progress reports are adjusted
for the
        # overall progress through the model
        feedback = QgsProcessingMultiStepFeedback(4, model_feedback)
        results = {}
        outputs = {}

        # Merge vector layers
        alg_params = {
            'CRS': None,
            'LAYERS': '{Route layers}',
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['MergeVectorLayers'] = processing.run('native:mergevectorlayers', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(1)
        if feedback.isCanceled():
            return {}

        # Field calculator
        alg_params = {
            'FIELD_LENGTH': 30,
            'FIELD_NAME': '{subcategories}',
```

```
            'FIELD_PRECISION': 5,
            'FIELD_TYPE': 0,
            'FORMULA': 'coalesce({subcategory field to destination 1},0) + coalesce({subcategory field
to destination 2},0) + ... + coalesce({subcategory field to destination n},0)',
            'INPUT': 'MergeVectorLayers',
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['FieldCalculator'] = processing.run('native:fieldcalculator', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(1)
        if feedback.isCanceled():
            return {}


        # Explode lines
        alg_params = {
            'INPUT': outputs['FieldCalculator']['OUTPUT'],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
        }
        outputs['ExplodeLines'] = processing.run('native:explodelines', alg_params, context=context,
feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(2)
        if feedback.isCanceled():
            return {}

        # Create spatial index
        alg_params = {
            'INPUT': outputs['ExplodeLines']['OUTPUT']
        }
        outputs['CreateSpatialIndex'] = processing.run('native:createspatialindex', alg_params,
context=context, feedback=feedback, is_child_algorithm=True)

        feedback.setCurrentStep(3)
        if feedback.isCanceled():
            return {}

        # Join attributes by location (summary)
        alg_params = {
            'DISCARD_NONMATCHING': False,
            'INPUT': outputs['CreateSpatialIndex']['OUTPUT'],
            'JOIN': outputs['CreateSpatialIndex']['OUTPUT'],
            'JOIN_FIELDS': ['{fields of subcategories}'],
            'PREDICATE': [2],
            'SUMMARIES': [5],
            'OUTPUT': QgsProcessing.TEMPORARY_OUTPUT
```

```
        }
        outputs['JoinAttributesByLocationSummary'] = processing.run('qgis:joinbylocationsummary', 
alg_params, context=context, feedback=feedback, is_child_algorithm=True)
        results['RoutesForValidation'] = outputs['JoinAttributesByLocationSummary']['OUTPUT']
        return results

    def name(self):
        return 'model'

    def displayName(self):
        return 'model'

    def group(self):
        return ''

    def groupId(self):
        return ''

    def createInstance(self):
        return RoutesForValidation()
```

*Figure 51 Python script for the preparation of visualizing the results of the routing step*

# Appendix 7: The included and excluded postal code areas of the municipality of Eindhoven in the case study

*Table 7 List of included and excluded postal code areas of the municipality of Eindhoven*

| Included | Excluded |
|---|---|
| 5611 | 5657 |
| 5612 | 5658 |
| 5613 | |
| 5614 | |
| 5615 | |
| 5616 | |
| 5617 | |
| 5621 | |
| 5622 | |
| 5623 | |
| 5624 | |
| 5625 | |
| 5626 | |
| 5627 | |
| 5628 | |
| 5629 | |
| 5631 | |
| 5632 | |
| 5633 | |
| 5641 | |
| 5642 | |
| 5643 | |
| 5644 | |
| 5645 | |
| 5646 | |
| 5647 | |
| 5651 | |
| 5652 | |
| 5653 | |
| 5654 | |
| 5655 | |
| 5656 | |

## Appendix 8: The postal code areas combined to form the destination areas for the case study

*Table 8 The postal code areas combined to form the destination areas for the case study*

| ID | Destination Area | Postal Codes |
|---|---|---|
| *11* | *City centre* | 5611 |
| *21* | *Inner ring – Strijp* | 5616<br>5617 |
| *22* | *Inner ring – Woensel / TU/e* | 5612 |
| *23* | *Inner ring – Tongelre / Stratum* | 5613 |
| *24* | *Inner ring – Gestel / Stratum* | 5614<br>5615 |
| *31* | *Outer ring – Hurk / Het Ven* | 5652 |
| *32* | *Outer ring – Strijp / Woensel* | 5621<br>5622<br>5624<br>5626<br>5627<br>5651 |
| *33* | *Outer ring – Woensel* | 5623<br>5625<br>5628<br>5629<br>5632<br>5633 |
| *34* | *Outer ring – Eckart* | 5631 |
| *35* | *Outer ring – Tongelre* | 5641<br>5642 |
| *36* | *Outer ring – Stratum* | 5643<br>5645<br>5646<br>5647 |
| *37* | *Outer ring – Gennep / Stratum* | 5644 |
| *38* | *Outer ring – Gestel* | 5653<br>5654<br>5655<br>5656 |

## Appendix 9: Findings case study per entry point

*Table 9 Output of the case study, number of trips per entry point*

| id | Entry point | All | Work | Non-work | Peak | Off-Peak | Weekend | Workday |
|---|---|---|---|---|---|---|---|---|
| 1 | JFK laan | 42.909 k | 17.803 k | 25.106 k | 13.601 k | 29.307 k | 11.414 k | 31.495 k |
| 2 | Huizingalaan | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 3 | Boschdijk | 6.335 k | 2.142 k | 4.194 k | 1.966 k | 4.370 k | 1.389 k | 4.946 k |
| 4 | Antony Fokkerweg C | 16.120 k | 8.760 k | 7.360 k | 5.884 k | 10.236 k | 3.386 k | 12.734 k |
| 5 | Tilburgseweg | 12.728 k | 5.851 k | 6.878 k | 3.773 k | 8.956 k | 3.416 k | 9.312 k |
| 6 | Meerenakkerweg | 8.008 k | 4.768 k | 3.240 k | 3.426 k | 4.581 k | 915 k | 7.092 k |
| 7 | Hightechcampus | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 8 | Aalsterweg | 7.637 k | 3.418 k | 4.218 k | 1.984 k | 5.653 k | 2.337 k | 5.300 k |
| 9 | Leenderweg | 17.841 k | 10.356 k | 7.486 k | 6.764 k | 11.077 k | 4.060 k | 13.781 k |
| 10 | N270 | 19.572 k | 7.319 k | 12.254 k | 5.844 k | 13.729 k | 4.997 k | 14.575 k |
| 11 | Soesterbeek | 2.930 k | 292 k | 2.638 k | 675 k | 2.254 k | 670 k | 2.260 k |
| 12 | Geldropseweg | 4.091 k | 1.127 k | 2.964 k | 1.040 k | 3.051 k | 847 k | 3.244 k |
| 13 | Meerveldhovenseweg | 4.615 k | 2.107 k | 2.508 k | 1.615 k | 3.001 k | 1.129 k | 3.487 k |
| 14 | Hurksestraat | 5.318 k | 929 k | 4.389 k | 1.431 k | 3.887 k | 2.122 k | 3.196 k |
| 15 | Velddoornweg | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 16 | Prof Holstlaan | 2.965 k | 1.068 k | 1.896 k | 455 k | 2.509 k | 344 k | 2.620 k |
| 17 | Noord Brabantlaan | 1.935 k | 91 k | 1.844 k | 241 k | 1.694 k | 763 k | 1.172 k |
| 18 | Oirschotsedijk | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 19 | Eindhovenseweg | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 20 | Sterrenlaan | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 21 | Opwettenseweg | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 22 | Molendijk | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 23 | Urkhovenseweg | 525 k | 139 k | 386 k | 127 k | 397 k | 136 k | 389 k |
| 24 | Kanaaldijk-Noord | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| 25 | Rielsedijk | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k | 0 k |
| | Total | 153.529 k | 66.168 k | 87.360 k | 48.825 k | 104.703 k | 37.924 k | 115.605 k |

## Appendix 10: The ratio of entry point usage for each sub-category

*Table 10 The ratio of entry point usage for each sub-category*

| ID | Entry point | All trips | Work | Non-work | Peak | Off-peak | Weekend | Workday |
|---|---|---|---|---|---|---|---|---|
| 1 | JFK laan | 28% | 27% | 29% | 28% | 28% | 30% | 28% |
| 3 | Boschdijk | 4% | 3% | 5% | 4% | 4% | 4% | 4% |
| 4 | Antony Fokkerweg C | 10% | 13% | 8% | 12% | 10% | 9% | 12% |
| 5 | Tilburgseweg | 8% | 9% | 8% | 8% | 9% | 9% | 8% |
| 6 | Meerenakkerweg | 5% | 7% | 4% | 7% | 4% | 2% | 7% |
| 8 | Aalsterweg | 5% | 5% | 5% | 4% | 5% | 6% | 4% |
| 9 | Leenderweg | 12% | 16% | 9% | 14% | 11% | 11% | 14% |
| 10 | N270 | 13% | 11% | 14% | 12% | 13% | 13% | 12% |
| 11 | Soesterbeek | 2% | 0% | 3% | 1% | 2% | 2% | 1% |
| 12 | Geldropseweg | 3% | 2% | 3% | 2% | 3% | 2% | 2% |
| 13 | Meerveldhovenseweg | 3% | 3% | 3% | 3% | 3% | 3% | 3% |
| 14 | Hurksestraat | 3% | 1% | 5% | 3% | 4% | 6% | 3% |
| 16 | Prof Holstlaan | 2% | 2% | 2% | 1% | 2% | 1% | 1% |
| 17 | Noord Brabantlaan | 1% | 0% | 2% | 0% | 2% | 2% | 0% |
| 23 | Urkhovenseweg | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| **Total** | | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

# Appendix 11: Assumptions check regression analysis workday and peak-hour trips

## Appendix 11.1 Assumptions check regression analysis workday 1

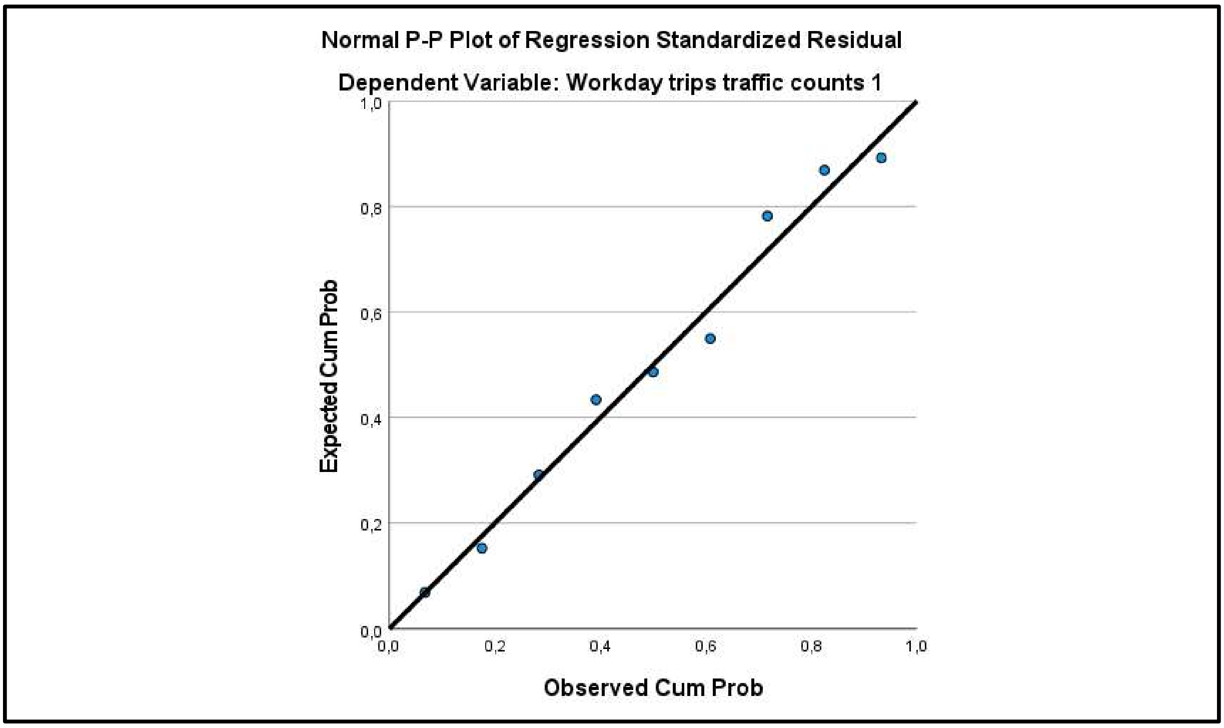



*Figure 52 Normal Predicted Probability plot of workday trips 1*

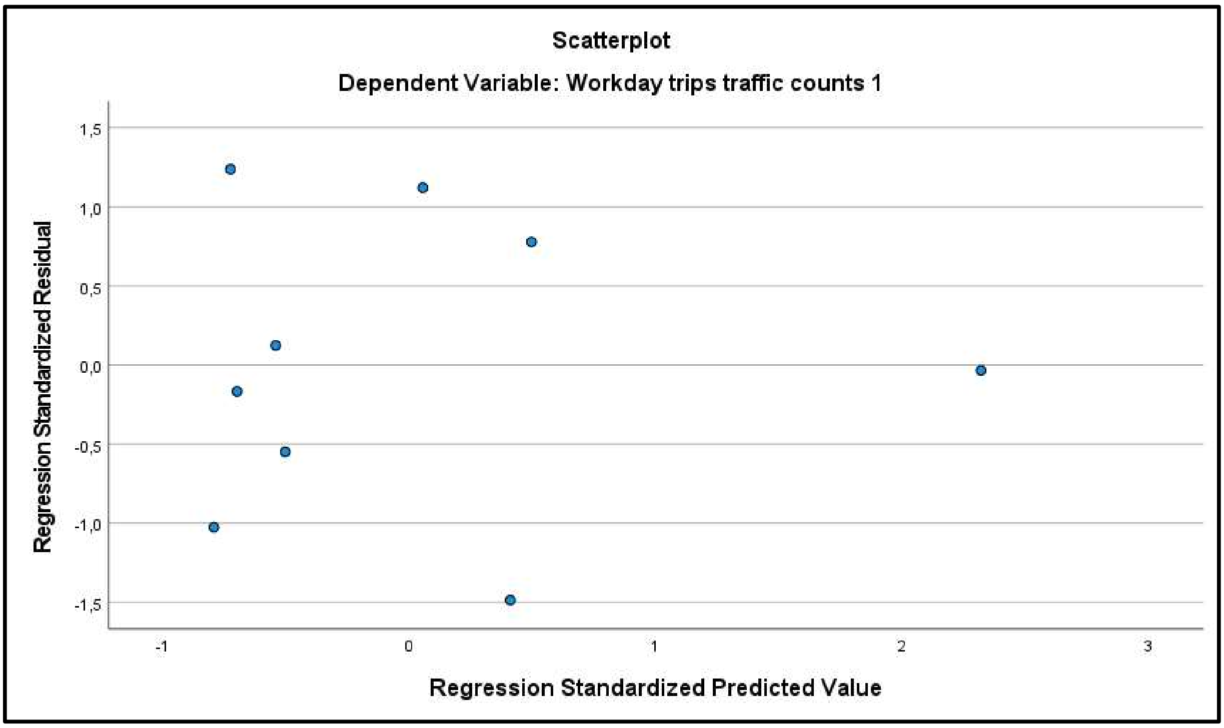



*Figure 53 Scatterplot of predicted values and residuals of the regression of workday trips 1*

## Appendix 11.2 Assumptions check regression analysis workday 2

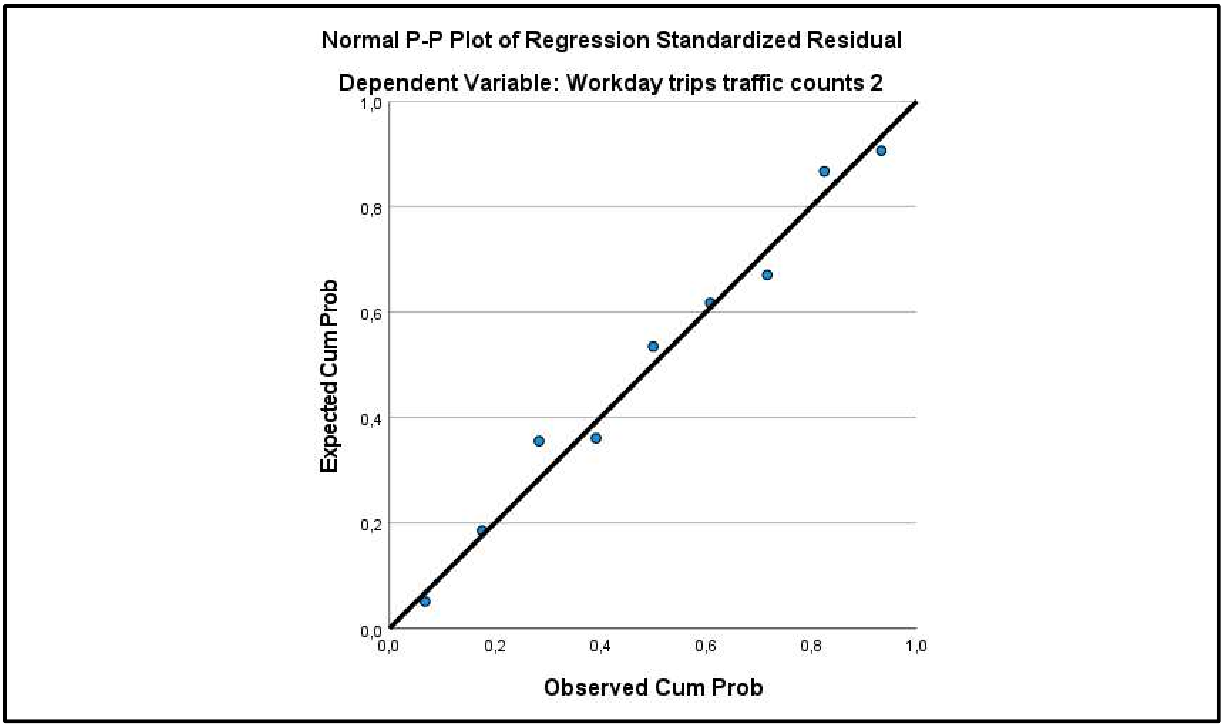



*Figure 54 Normal Predicted Probability plot of workday trips 2*

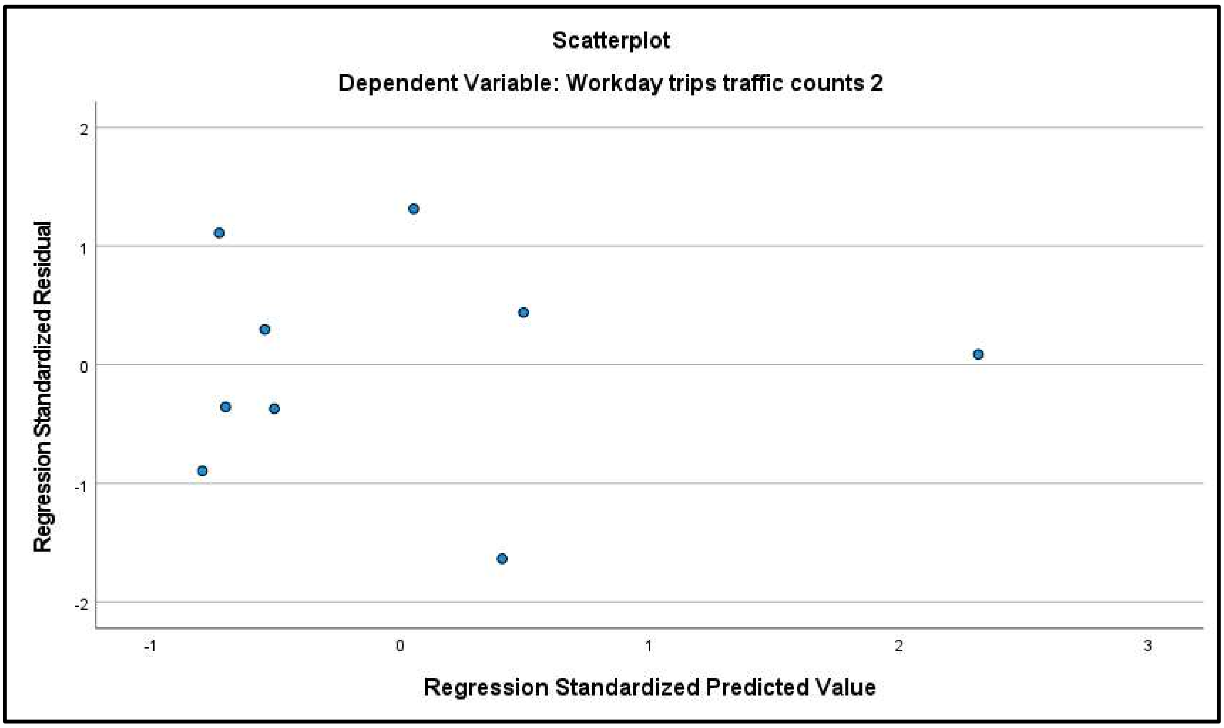



*Figure 55 Scatterplot of predicted values and residuals of the regression of workday trips 2*

## Appendix 11.3 Assumptions check regression analysis workday average

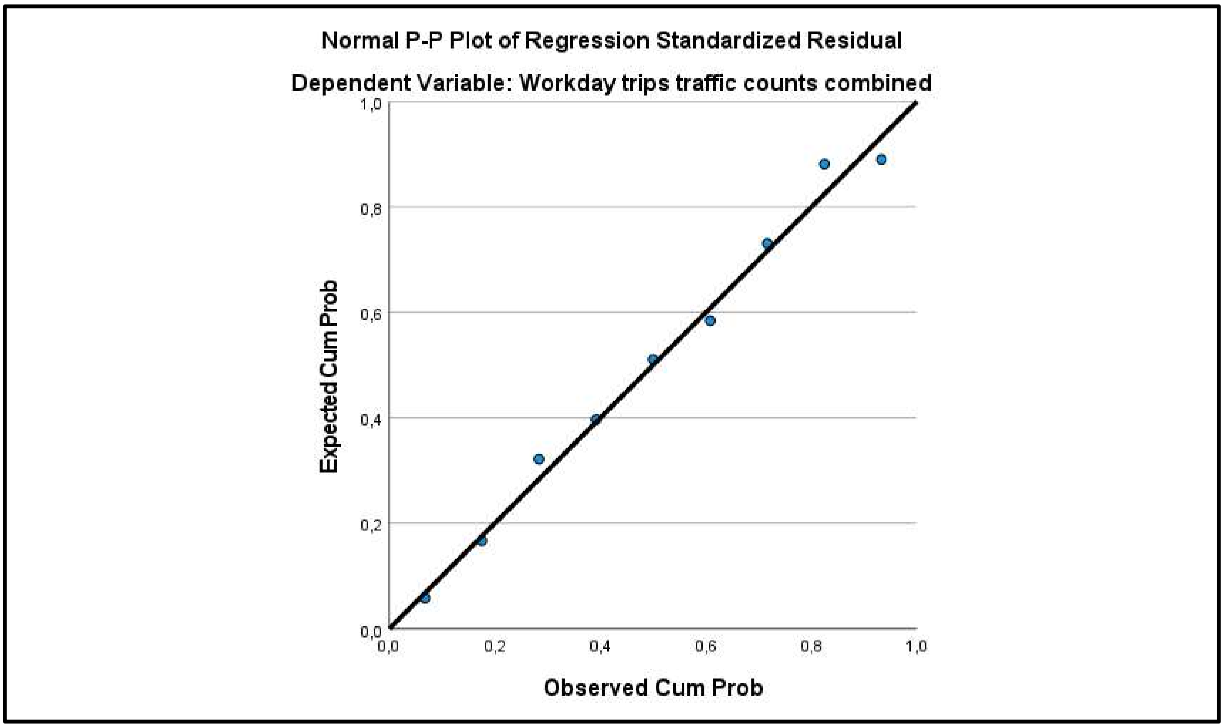



*Figure 56 Normal Predicted Probability plot of workday trips average*

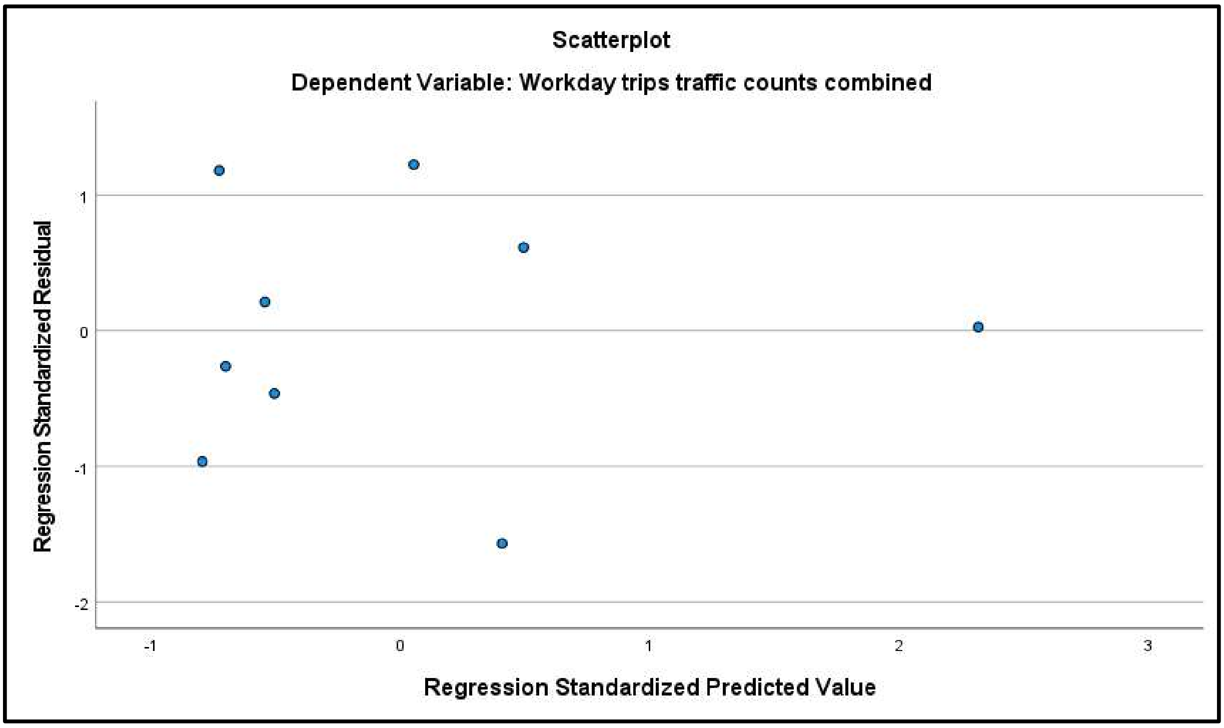



*Figure 57 Scatterplot of predicted values and residuals of the regression of workday trips average*

## Appendix 11.4 Assumptions check regression analysis peak-hour

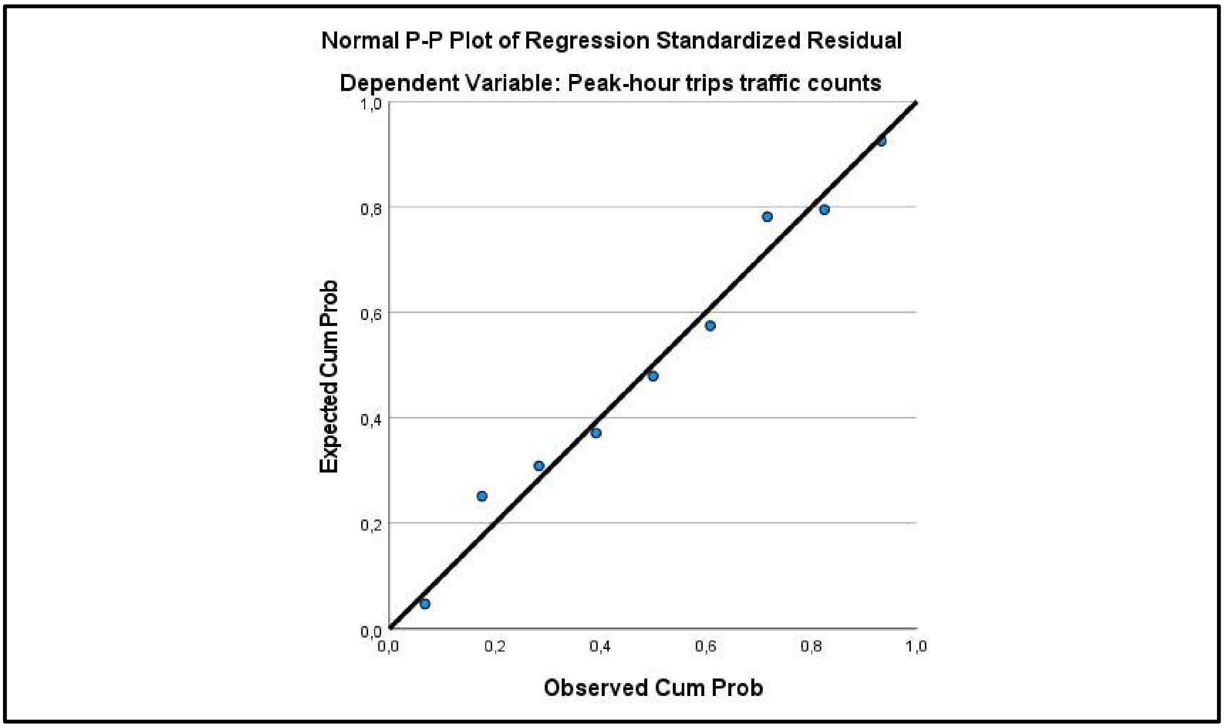



*Figure 58 Normal Predicted Probability plot of peak-hour trips*

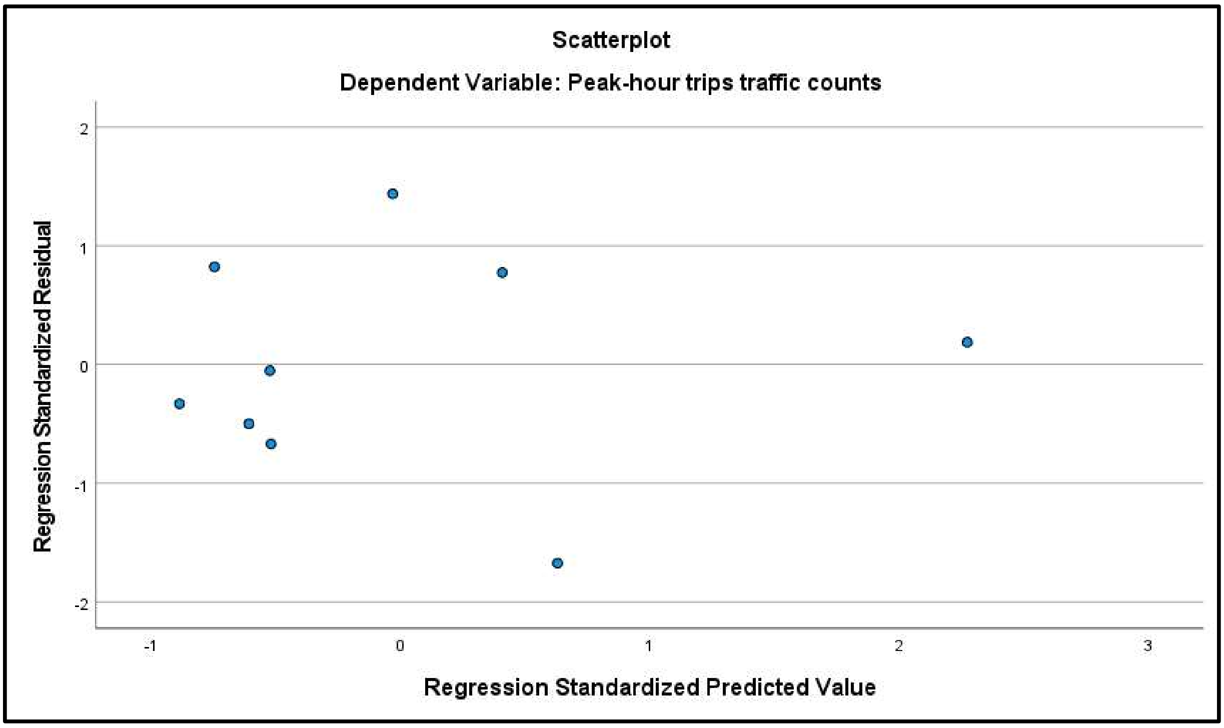



*Figure 59 Scatterplot of predicted values and residuals of the regression of peak-hour trips*

# Appendix 12: Output maps of the case study per sub-categories

## Appendix 12.1: Output maps of the case study, work trips

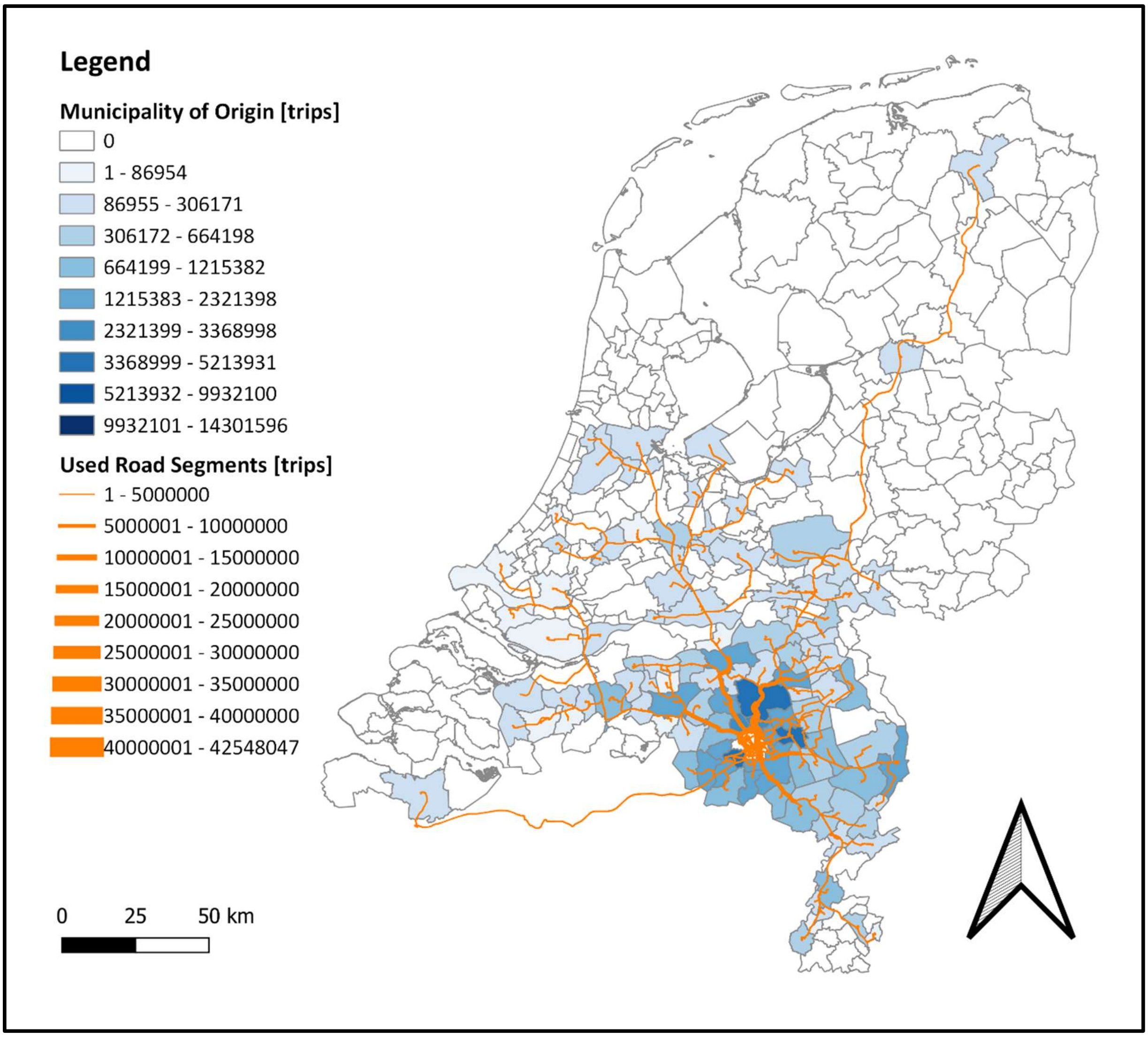



*Figure 60 Case study Eindhoven visualization, work trips towards Eindhoven 2017-2020, overview*

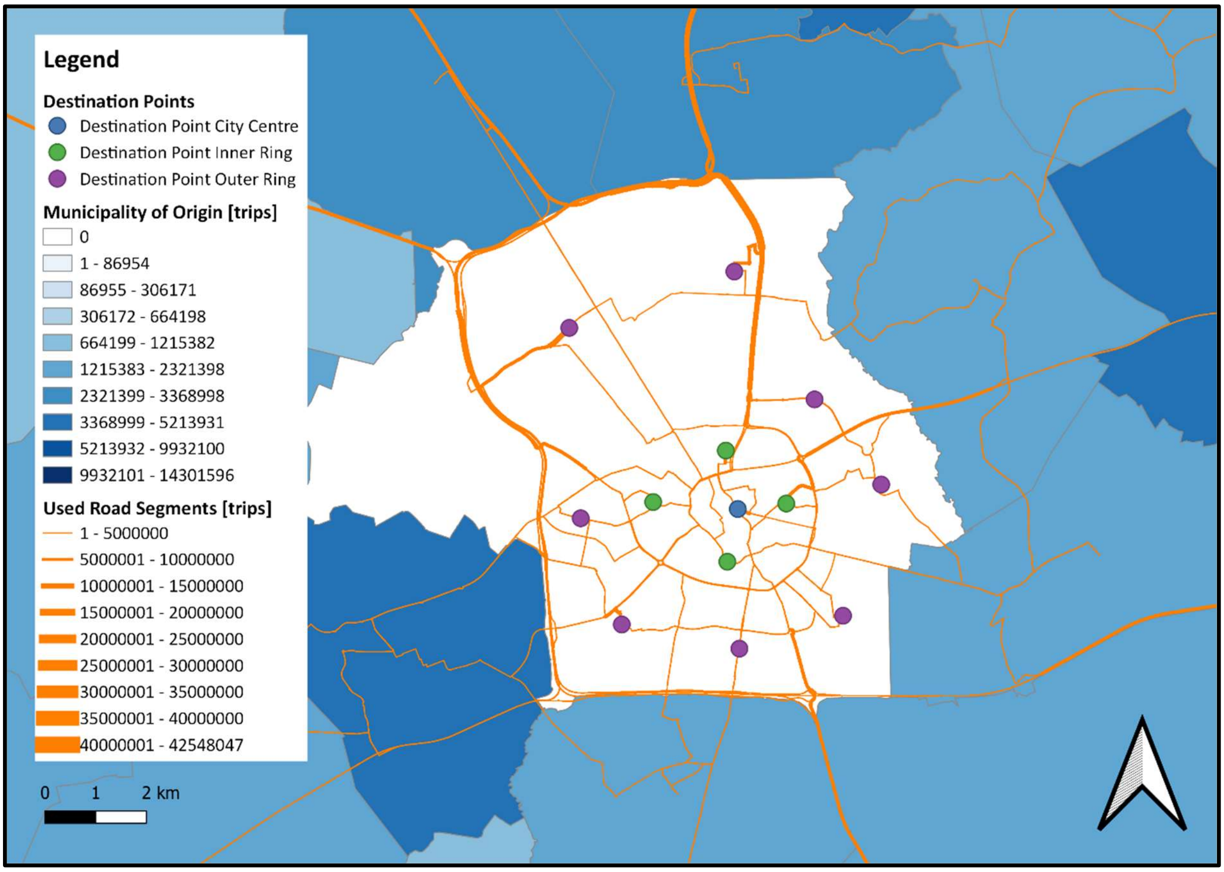



*Figure 61 Case study Eindhoven visualization, work trips towards Eindhoven 2017-2020, zoomed*

## Appendix 12.2: Output maps of the case study, non-work trips

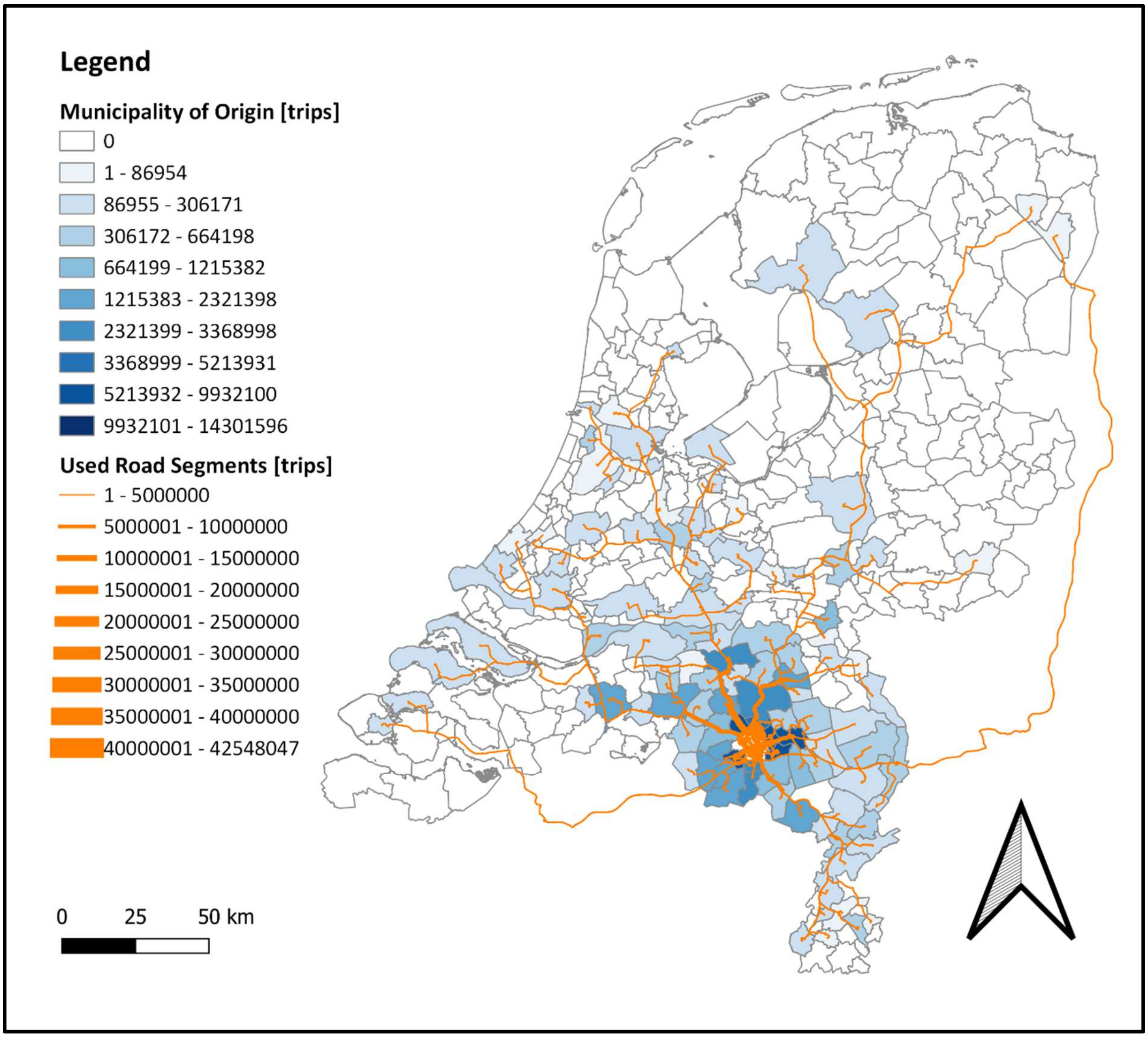



*Figure 62 Case study Eindhoven visualization, non-work trips towards Eindhoven 2017-2020, overview*

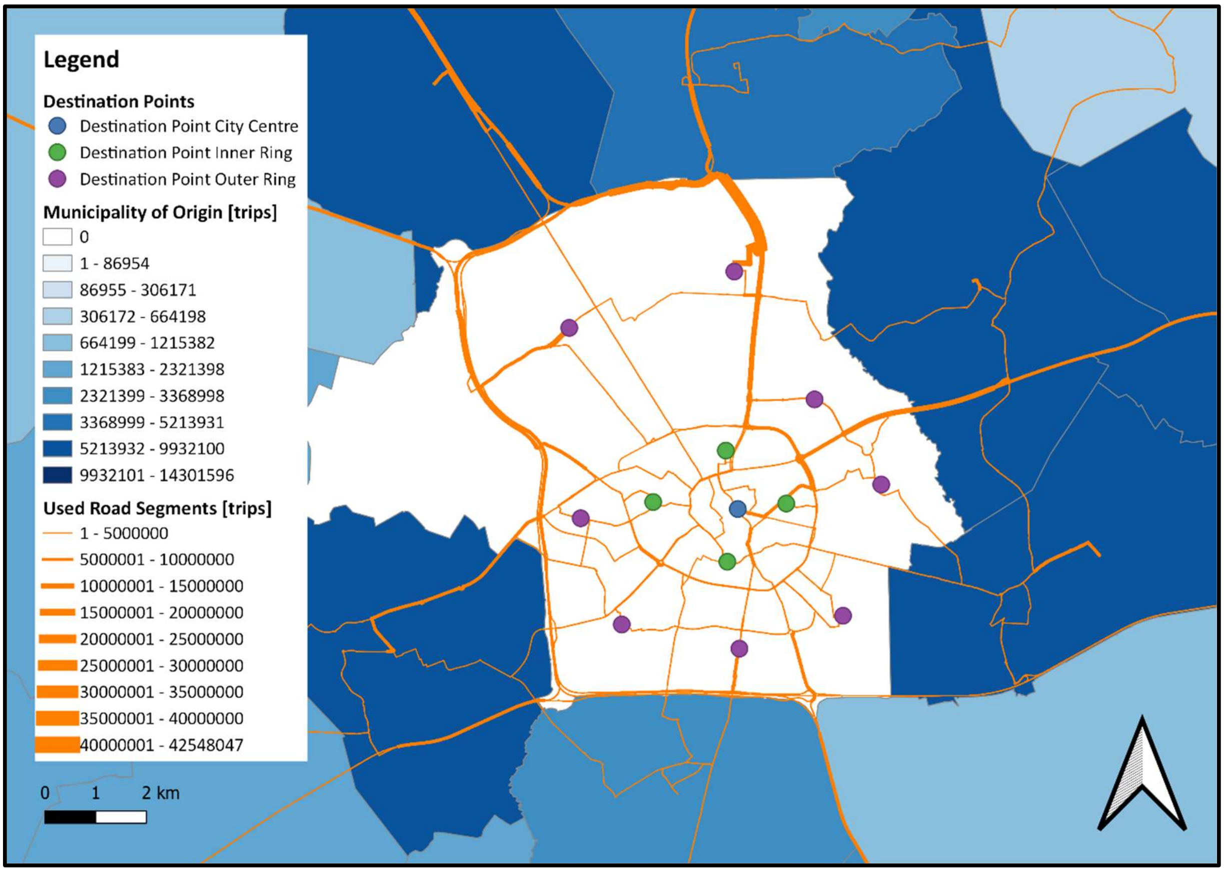



*Figure 63 Case study Eindhoven visualization, non-work trips towards Eindhoven 2017-2020, zoomed*

## Appendix 12.3: Output maps of the case study, peak trips

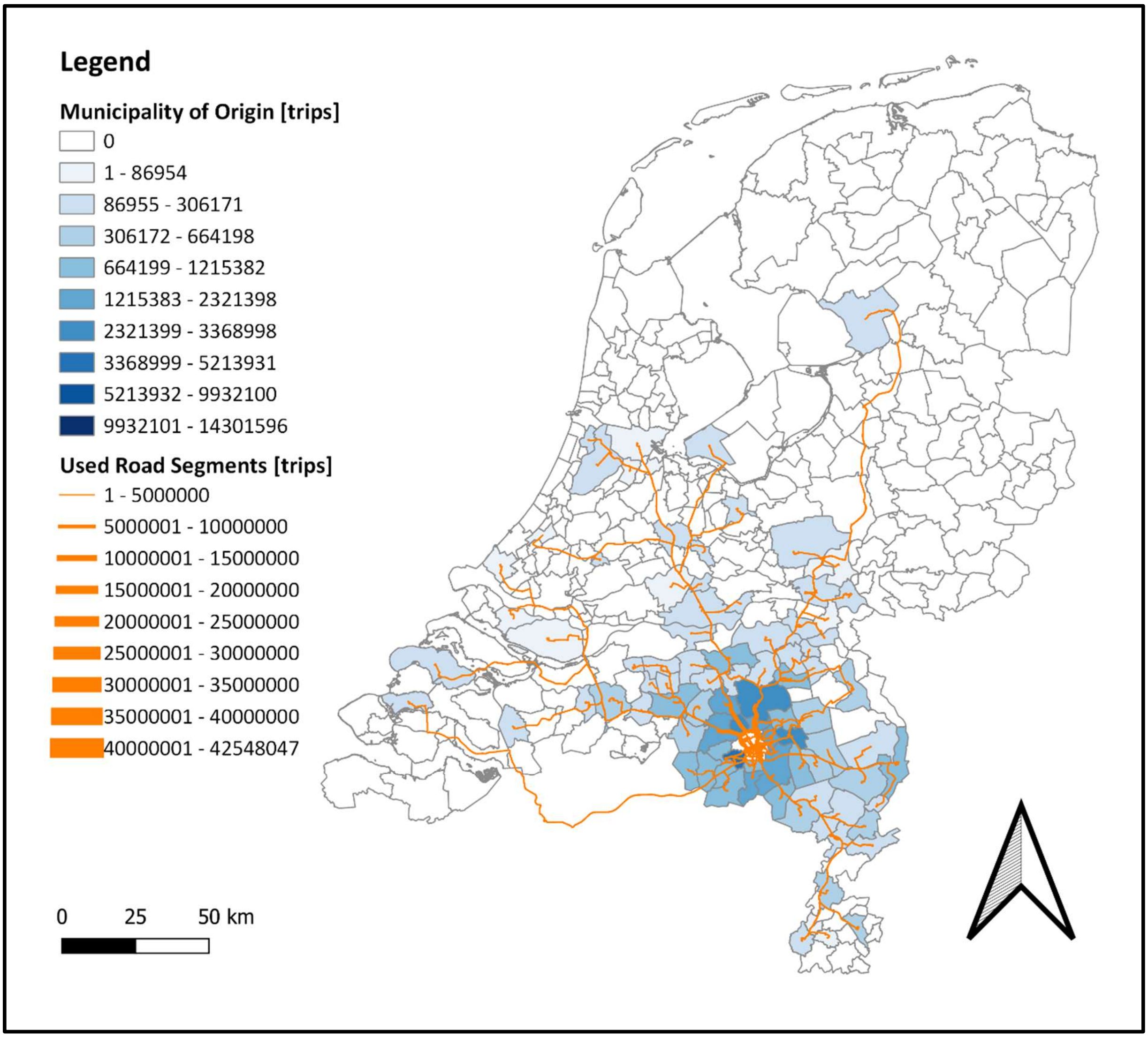



*Figure 64 Case study Eindhoven visualization, peak trips towards Eindhoven 2017-2020, overview*

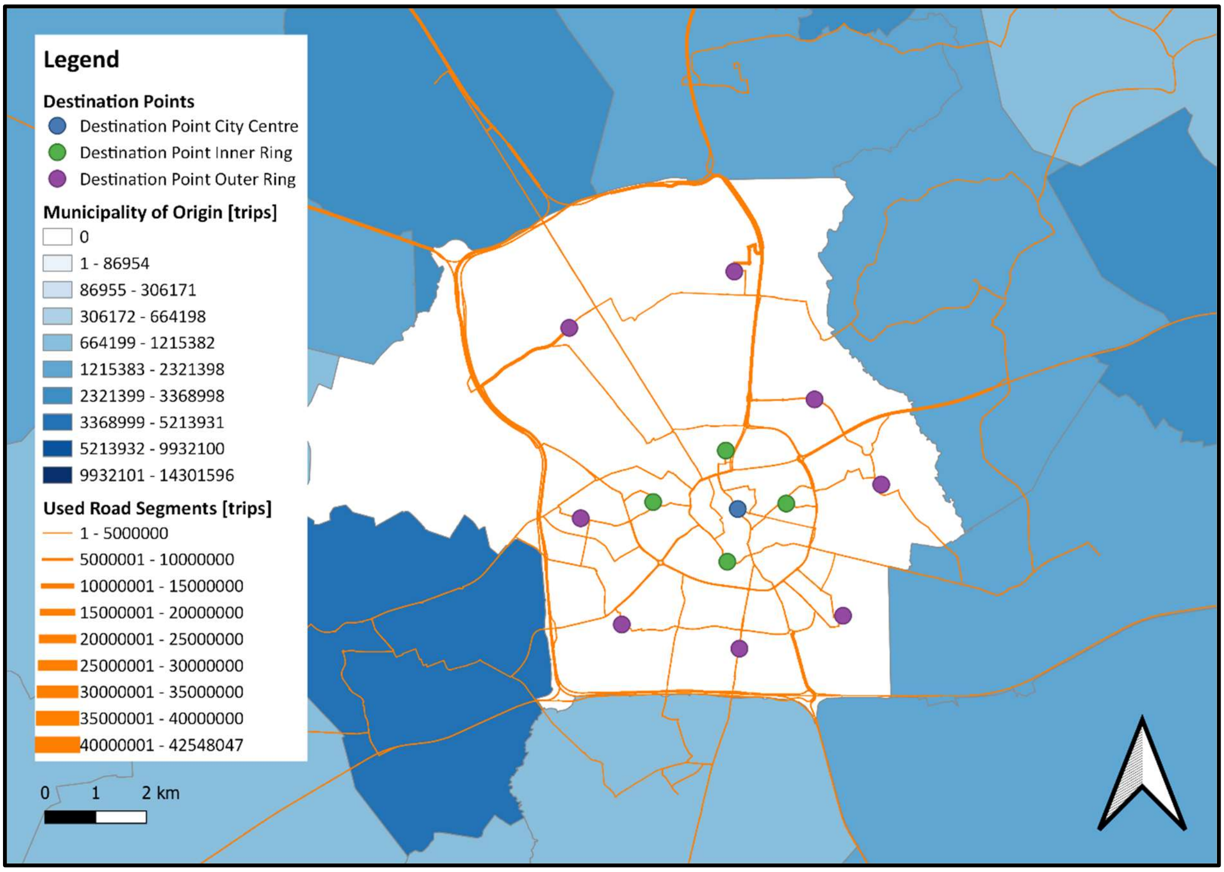



*Figure 65 Case study Eindhoven visualization, peak trips towards Eindhoven 2017-2020, zoomed*

## Appendix 12.4: Output maps of the case study, off-peak trips

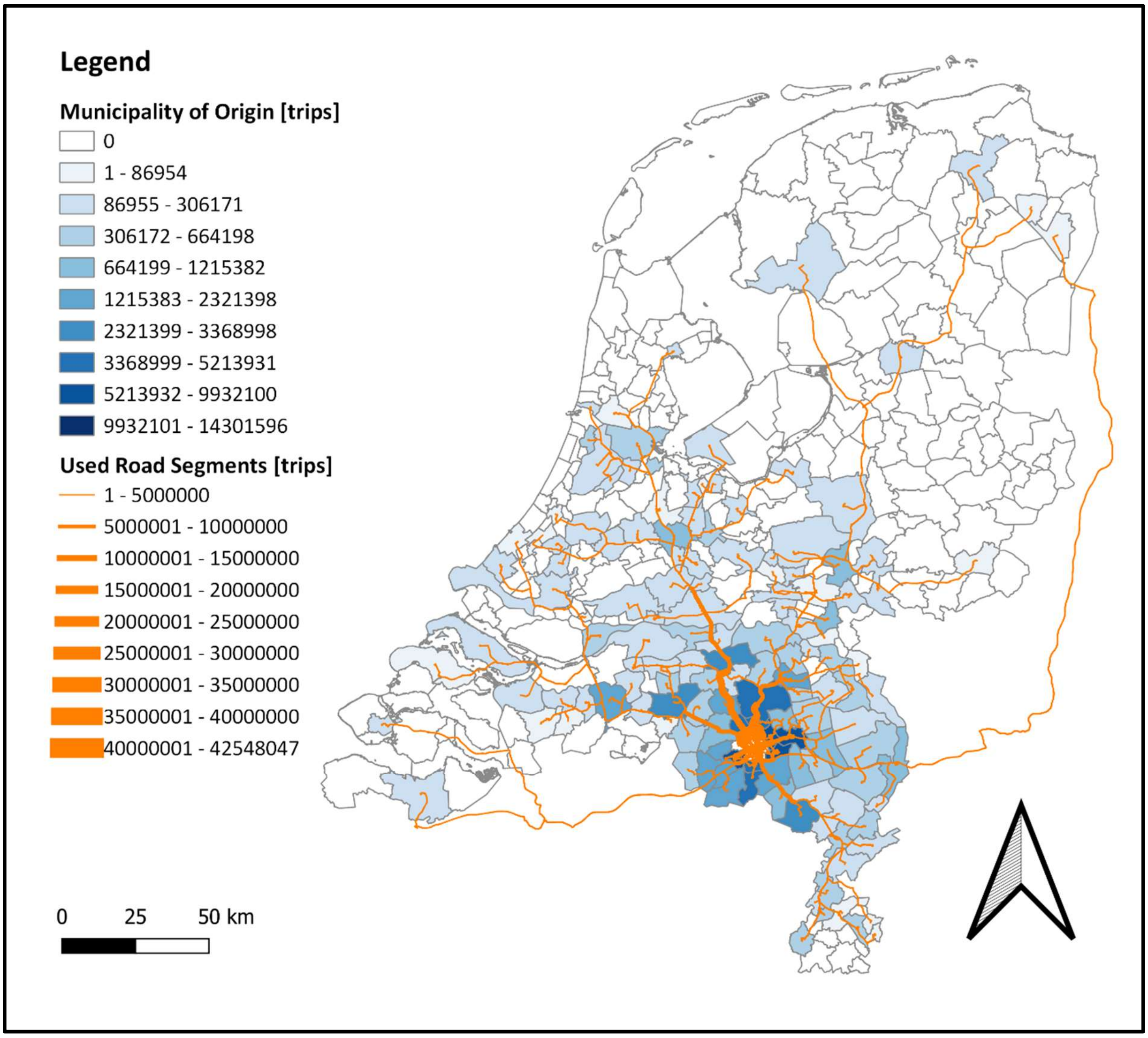



*Figure 66 Case study Eindhoven visualization, off-peak trips towards Eindhoven 2017-2020, overview*

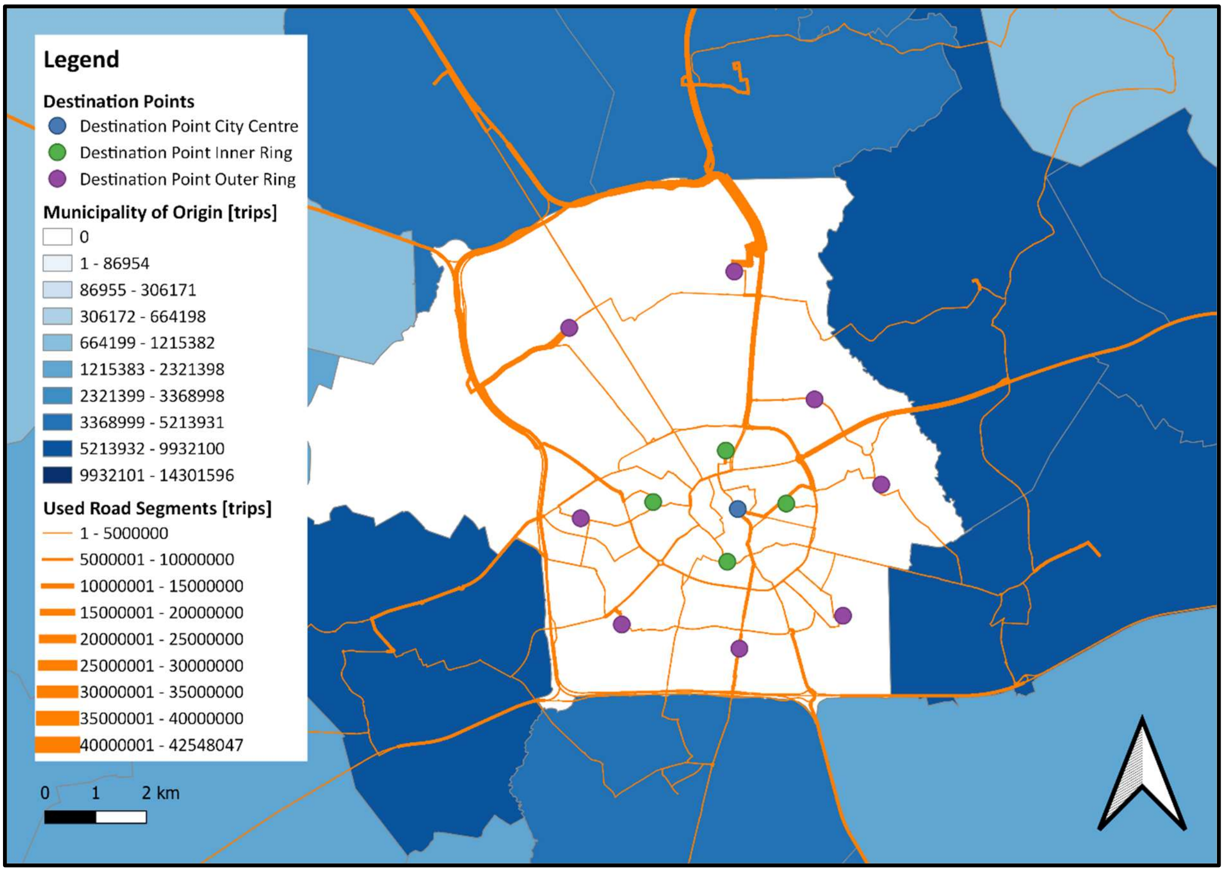



*Figure 67 Case study Eindhoven visualization, off-peak trips towards Eindhoven 2017-2020, zoomed*

## Appendix 12.5: Output maps of the case study, weekend trips

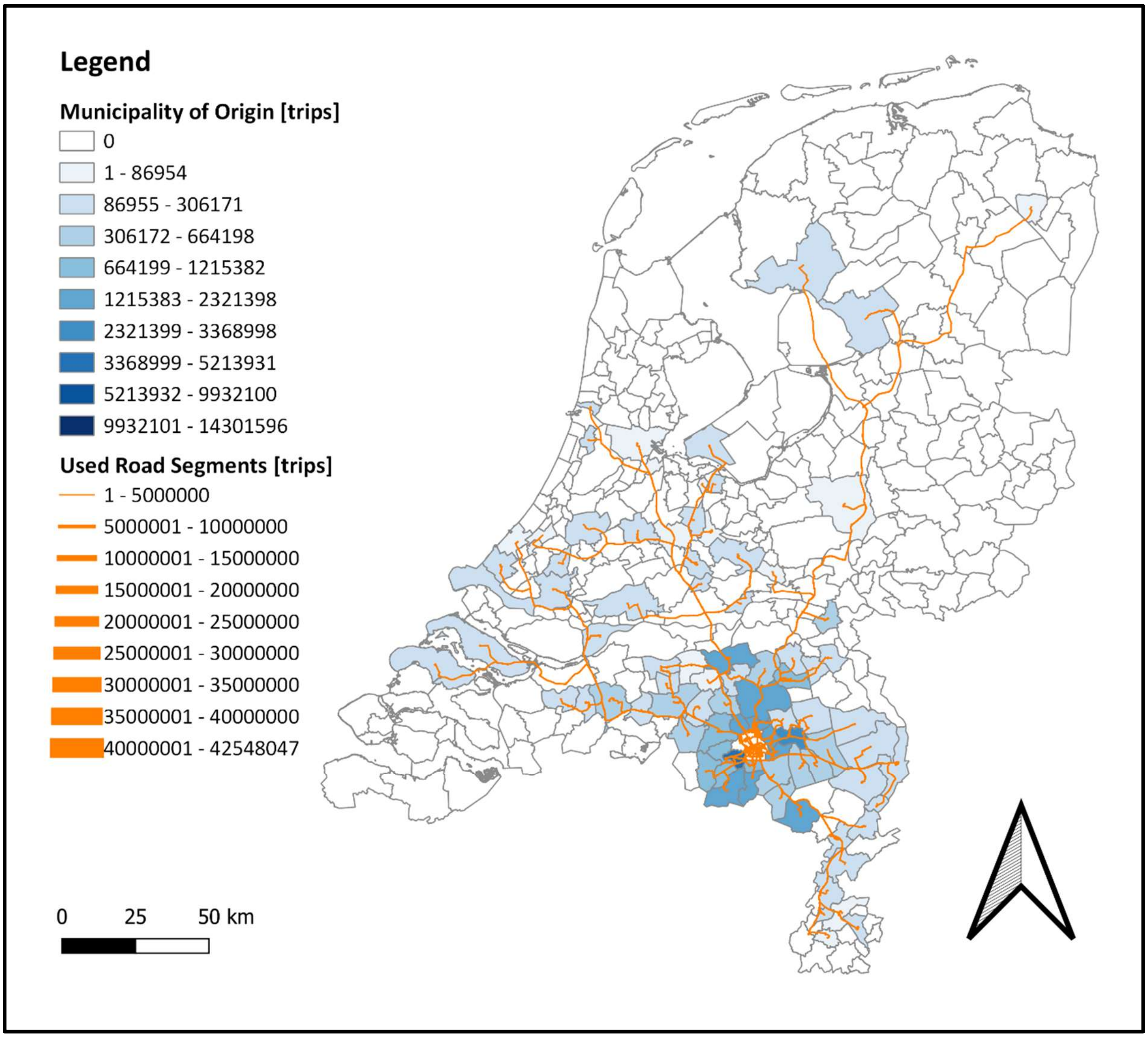



*Figure 68 Case study Eindhoven visualization, weekend trips towards Eindhoven 2017-2020, overview*

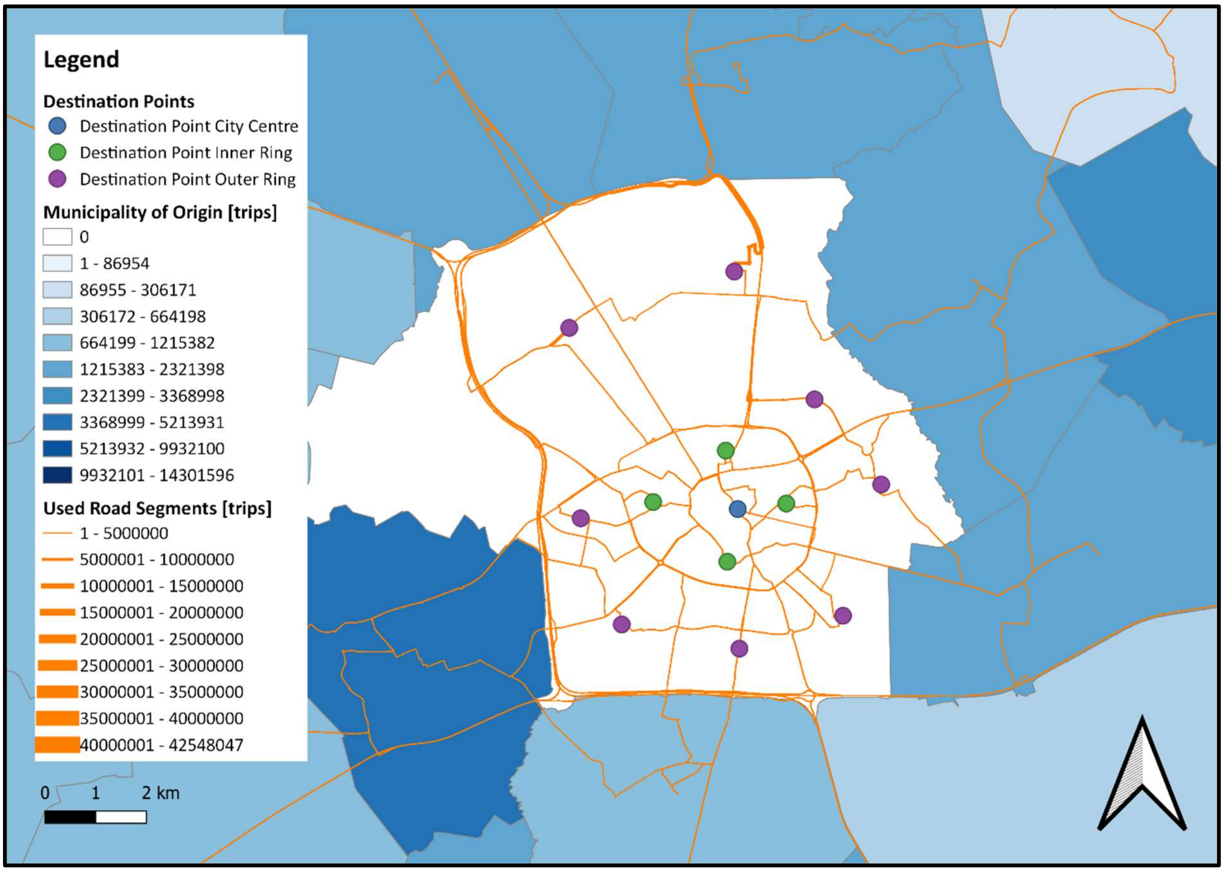



*Figure 69 Case study Eindhoven visualization, weekend trips towards Eindhoven 2017-2020, zoomed*

## Appendix 12.6: Output maps of the case study, workday trips

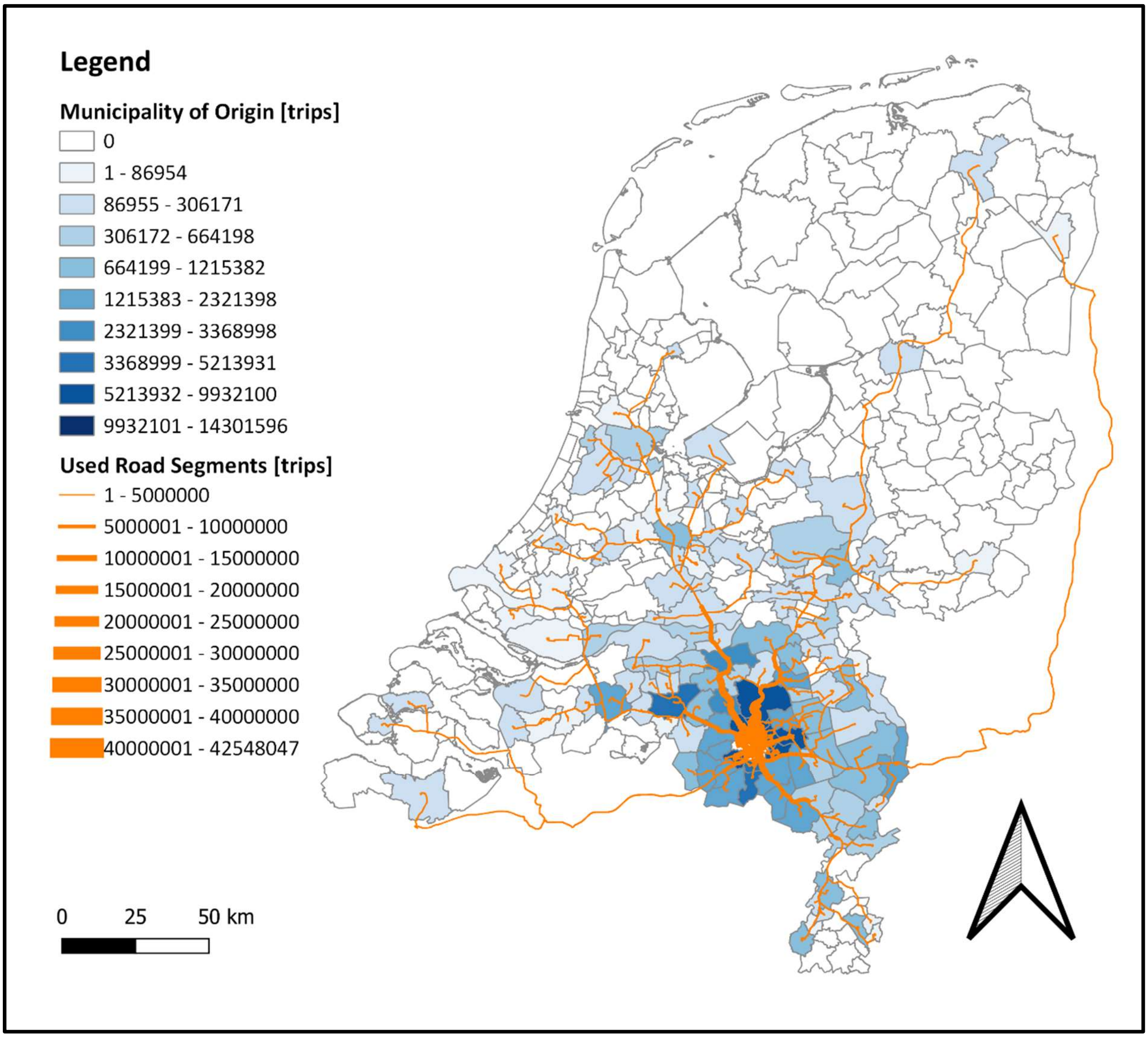



*Figure 70 Case study Eindhoven visualization, workday trips towards Eindhoven 2017-2020, overview*

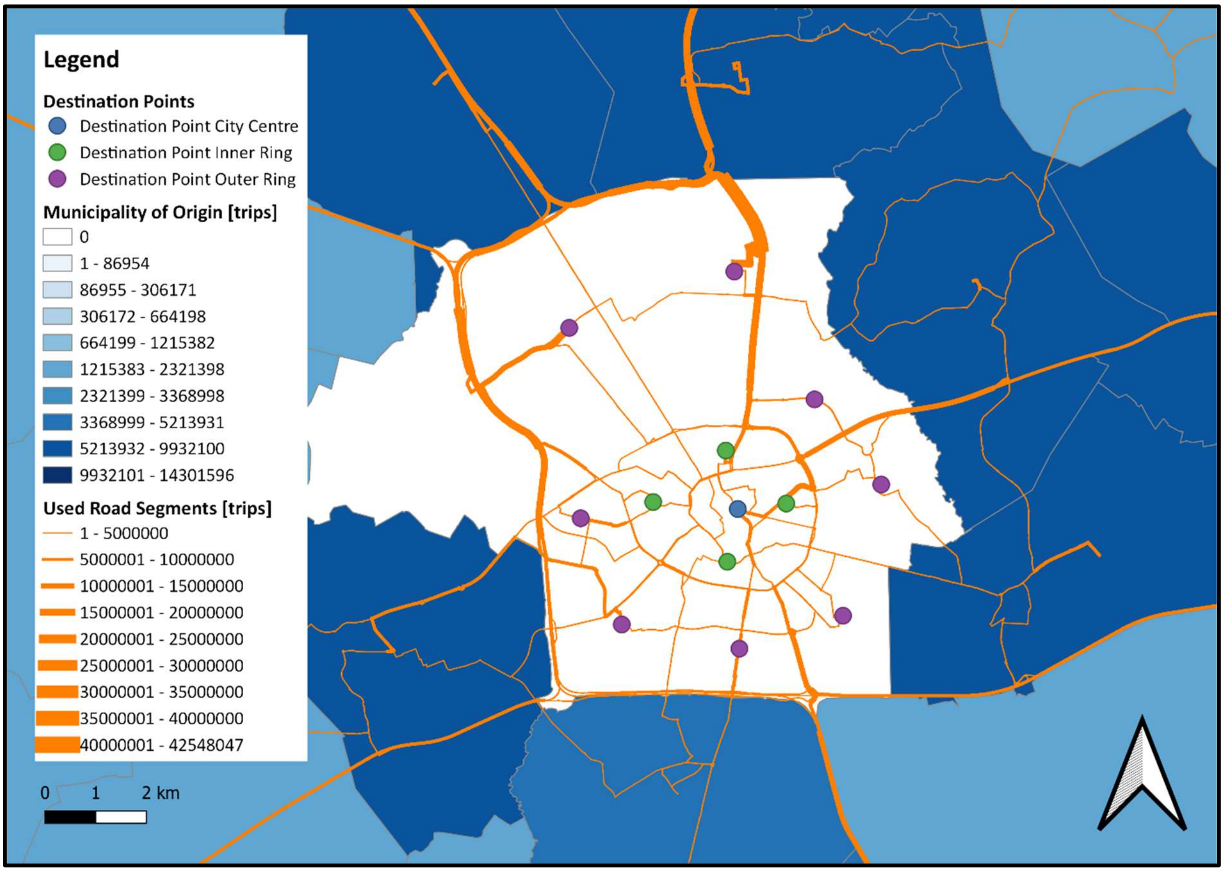



*Figure 71 Case study Eindhoven visualization, workday trips towards Eindhoven 2017-2020, zoomed*