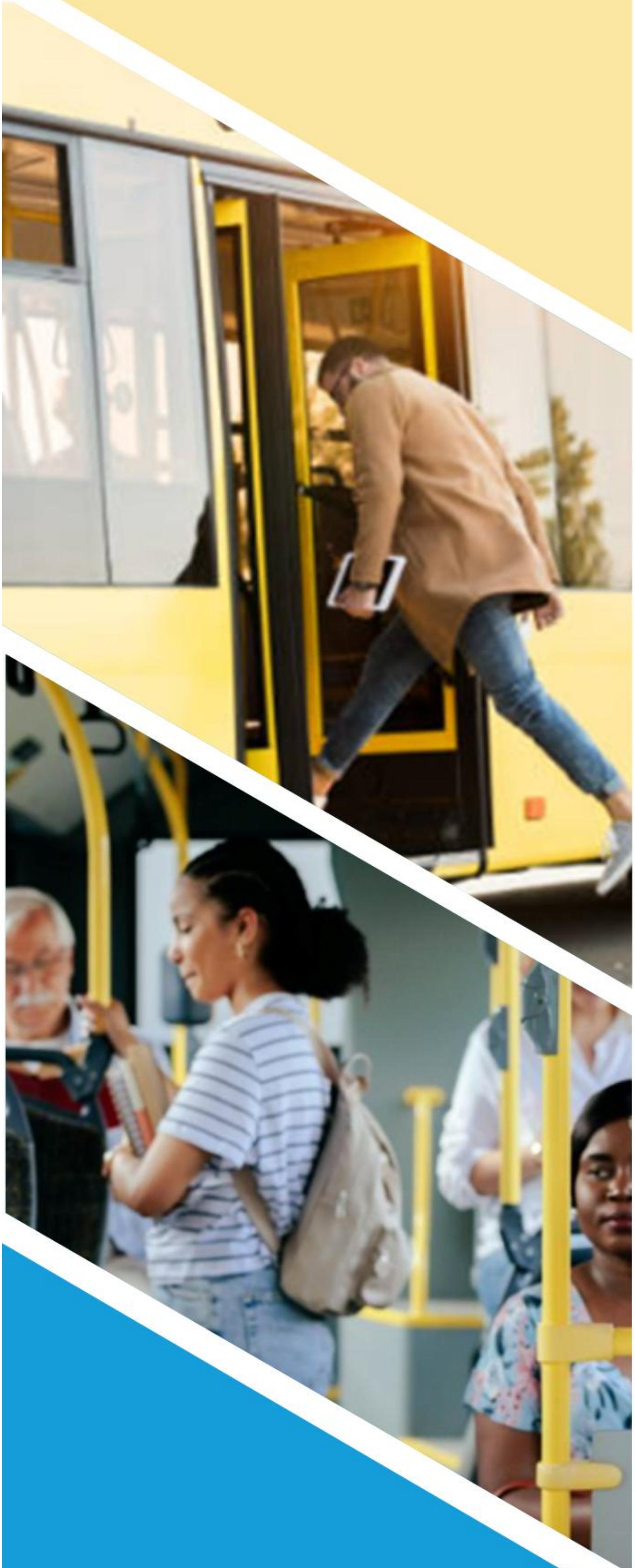

EVALUATING RIDERSHIP PREDICTION METHODS FOR BRT CORRIDORS: BALANCING ACCURACY AND TIME-EFFICIENCY

Bachelor Thesis
13/06/2025
Kirsten Pronk

UNIVERSITY OF TWENTE.

KEYPOINT CONSULTANCY

Evaluating ridership prediction methods for BRT corridors: balancing accuracy and time-efficiency

Bachelor Thesis

Kirsten Pronk

13-06-2025

Institution:	University of Twente
Study:	Bachelor Civil Engineering
Student number:	s2699710
Contact:	c.a.a.pronk@student.utwente.nl
Internal supervisor:	A. (Alejandro) Tirachini PhD
Company:	Keypoint Consultancy
Adress:	Institutenweg 32, 7521 PK Enschede
External supervisors:	C. (Cees) Bakker, R. (Ramon) Kienhuis

Sources front page:

Upper: (Istock, n.d.)

Bottom: (Modeshift, 2023)

PREFACE

You are going to read my Bachelor Thesis “Evaluating ridership prediction methods for BRT corridors: balancing accuracy and time-efficiency”, which I have carried out over the course of ten weeks at Keypoint Consultancy in Enschede. With that I want to thank Keypoint Consultancy for welcoming me and supporting me, in particular my supervisors Cees Bakker and Ramon Kienhuis. During my time at the company, I have learnt about consultancy within the mobility sector, something which seems interesting for a future career. Besides that, I want to thank my supervisor from the University of Twente, Alejandro Tirachini for his support and knowledge about the public transport sector and academic writing.

This thesis has taught me how to work on my own project and be the leader of it, but also how to work together with my supervisors. It also taught me how to work with tools like python, knowledge which will help me in my further studies and maybe even once I get a job.

With that, my Bachelor Civil Engineering at the University of Twente comes to an end and a new chapter will begin with my master which will also be here at the University of Twente.

Kirsten Pronk

13th of June 2025

ABSTRACT

The implementation of Bus Rapid Transit, BRT, within the Netherlands has been a topic of discussion for quite some years, as it is a sustainable and cost-effective way of public transport. BRT systems, characterized by high service levels and their efficiency, have the potential to reduce car dependency and improve regional mobility. While previous studies have identified potential BRT corridors, they have not addressed the critical question of expected ridership. Accurate ridership forecasting is essential to justify investments in new corridors, yet it remains unclear which prediction method offers the best balance between accuracy and practical feasibility.

While various methods exist for predicting ridership, it remains unclear which approach offers the greatest accuracy and efficiency for the prediction of BRT ridership. Therefore, the aim of this research is to find the most accurate and time-efficient prediction method for BRT corridors in the Netherlands. The four methods that are being researched are a comparative case study, the logit model, a regression model and the gravity model. These methods are assessed using a case study of the Assen Kloosterveen-Groningen bus corridor. Once the best method is known Keypoint Consultancy can use the tool for their clients.

The comparative case study relies on identifying a similar corridor using demographic and geographic criteria. The logit model uses modal split and transport-related variables to estimate demand. The regression model incorporates operational variables and coefficients sourced from literature, while the gravity model estimates ridership based on population, employment, travel time, and a calibrated scaling constant. A multi-criteria decision analysis (MCDA) is used to compare the outcomes and determine the most suitable method. The criteria of the MCDA were accuracy, time-efficiency, data availability and computational tools/skills needed.

While the logit model produced the most accurate results, the regression model emerged as the most balanced option according to the MCDA. It offers the best trade-off between accuracy and time efficiency, making it well-suited for practical application. In contrast, the comparative case study method was the least accurate and most time-consuming and is therefore not recommended for future use. The gravity model demonstrated a reasonable accuracy but as it is connected to the comparative case study, it lacks time-efficiency.

TABLE OF CONTENTS

1 INTRODUCTION

The potential implementation of Bus rapid Transit (BRT) in the Netherlands has been a subject of discussion for several years. BRT is a type of Hoogwaardig Openbaar Vervoer (HOV) or high-quality public transport, characterized by high service standards and frequent departures. Integrating BRT into the public transport network aims to increase mobility and decrease car dependency within The Netherlands (Witte & Kansen, 2020).

A crucial step in the development of new BRT corridors is predicting the potential demand. Accurately predicting ridership is an essential step to determine if a BRT corridor is worth the investment. Various methodologies can be used to predict potential ridership, some are easy to compute, other are more complex and difficult to compute. Since transportation projects often have tight deadlines, time efficient prediction methods are essential for effective planning and decision-making. Time efficient prediction allows planners to effectively evaluate the corridor.

The aim of this research is to identify the method which offers the best balance between accuracy and the required time investment and resources. By selecting the most effective forecasting method, policy makers can make informed decisions on where to implement a BRT corridor in the Netherlands.

The structure of this report is the following. First, in Chapter 2, relevant context surrounding the problem is provided as well as the research aim and questions. Then in Chapter 3, theory will be given about the problem. Chapter 4 explains the methodology of each chosen method. After that the results are presented per method that is chosen in Chapter 5. Then the results will be discussed in Chapter 6 and a conclusion formed in Chapter 7. The report will close with some final recommendations.

2 RESEARCH CONTEXT

This section discusses who are involved in this research, what the research motivation is and how this led to the problem statement. After that the study area and research scope will be discussed. At last, the research aim and questions will be formulated.

2.1 INVOLVED PARTIES

The commissioner of this bachelor thesis is Keypoint Consultancy. Their motto, ‘Smartly forward together’, reflects their approach to solving mobility problems through technology and data. Keypoint is a consultancy firm and therefore gives advice to for example municipalities on how to handle mobility problems. The external supervisors at Keypoint are Cees Bakker, team manager mobility and Ramon Kienhuis, mobility advisor. The desire of Keypoint is to gain more insight on the different methods to predict ridership for BRT corridors in the Netherlands. This is then a tool they can use to present predicted ridership to their clients.

2.2 RESEARCH MOTIVATION

Bus Rapid Transit (BRT) combines advantages of several different modes of (public) transport. It reduces travel time, with its rights of way and dedicated bus lane as the two most important travel time improvements compared to car usage. BRT is also less expensive compared to light rail transit (Hess et al., 2005). BRT is also a more environmentally friendly type of transport than car-based mobility. When a BRT operates on petrol or diesel it still gives a reduction of CO2 emissions due to the modal shift from cars to public transport. But more often, buses around the world are electric which makes it even better for the environment (UITP, 2024).

Most BRT systems are found in South America and Asia, but the concept is becoming increasingly popular in Europe. BRT is often a viable solution in lower-income countries as they can not afford expensive rail transport, as seen in South America and Asia (Matsumoto, 2005). But BRT can also work complementary to rail transport, this type of BRT is more common in the Netherlands (Witte & Kansen, 2020).

As the popularity of BRT increases more research needs to be done to this topic. Previous research at Keypoint already explored the possibility of BRT in the Netherlands, mainly the Randstad area. That research identified potential BRT corridors on different criteria, such as population, job availability and the existing transport network (de Wit (2023); Spijker (2024)). These models gave a numerical output for each corridor, after which they were ranked from most promising to least promising.

While these studies provide valuable insights into potential BRT corridors, an important question remained: how many passengers would use these new BRT corridors? Knowing the potential ridership can give valuable insights whether the new BRT corridor will be worth the investment.

To know the potential ridership of a new public transport system, different methods can be used. Some take a lot of time but are expected to provide more accurate results and others are maybe faster to compute but are less accurate. Therefore, this research will evaluate four different methods to determine the potential ridership of new BRT corridors to identify which method is likely to provide the best balance between time-efficiency and accuracy.

2.3 PROBLEM STATEMENT

BRT is increasing its popularity within the public transport sector. It offers reduced travel times compared to cars, is less expensive than rail transport and has environmental benefits. BRT systems are well established in South America and Asia, but they are also found in Europe and the Netherlands. Prior research identified potential BRT corridors based on criteria such as population, job availability and the existing transport network. However, one question is not answered yet: how many passengers would use the new BRT corridor.

There are different methods to predict ridership for new BRT corridors. Examples for prediction methods are regression modelling, the gravity model and the logit model and when all these models are combined, a full four-step model can predict ridership. Some methods provide more precise results but require extensive data collection and analysis and therefore, are time consuming and expensive. Other methods are faster to implement but may lack accuracy in the results. A comparison of different methods would give insights on the prediction of demand and potential revenues, in order to make a good BRT investment.

2.4 STUDY AREA

This research will be conducted along a corridor between a district of Assen called Kloosterveen and the city centre of Groningen. The route of the corridor is illustrated in Figure 1.

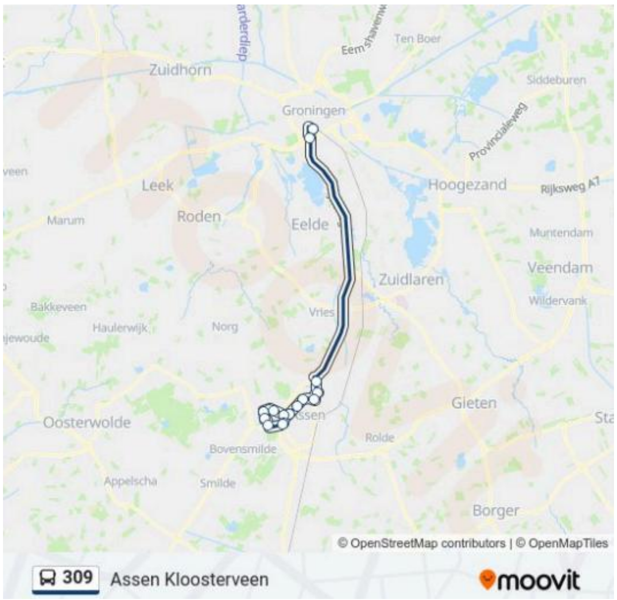

Figure 1: Study area: bus line 309 between Assen Kloosterveen and Groningen (Moovit, 2024)

Bus line 309 has been operational since December 2012. Prior to its existence, people from the district Kloosterveen were required to first travel to Assen central station to reach Groningen. The implementation of this direct bus service eliminated the need for this transfer, enabling passengers to travel directly to Groningen. The route includes 15 stops and operates at a frequency of 4 buses per hour. The bus takes on average 37 minutes to go from Assen Kloosterveen to Groningen centre (Moovit, 2024).

This direct connection plays a crucial role in the regional mobility of citizens. The connection reduces public transport travel times, making it more attractive to travel with public transport instead of a private car.

2.5 RESEARCH SCOPE

As mentioned in the study area the geographical scope of this research will be in the north of the Netherlands, particularly a bus corridor between Assen and Groningen line 309. This connection goes from an outside neighbourhood of Assen, Kloosterveen, to the city centre of Groningen. This corridor will serve as the basis for evaluating the ridership prediction methods for new bus services.

This research investigates various methods for forecasting ridership in the context of newly implemented bus corridors. To simulate a realistic planning scenario, the analysis is done as for the year 2012 as far as data allows it.

The goal of this research is to find the most time-effective but still accurate method for predicting ridership for this corridor. This will be done by comparing four methods.

2.6 RESEARCH AIM

The aim of this research is to find the most cost-effective method, in terms of time investment and amount of data needed, that predicts the potential ridership of new BRT corridors accurately. This will be done by comparing four different methods that can be used to predict potential BRT ridership. The four methods vary in the time it takes to predict and the accuracy of the prediction.

2.7 RESEARCH QUESTIONS

From the information provided in this chapter, a main research question and sub questions are formulated. The main research question is stated as follows:

Which method provides the best balance between accuracy and time efficiency for the prediction of potential ridership for new BRT corridors?

To reach the main research question, multiple sub-questions are formulated. These questions will help with answering the main question. The sub-questions are:

1. *Which methods are there to predict ridership and how do they work?*
2. *How accurate is each method in predicting the potential ridership when comparing it to real world data?*
    a. *How much detail is required for each method to produce reliable prediction?*
3. *How do these methods differ in terms of time investment, required data and complexity?*
    a. *How much time is needed to complete each method?*
    b. *What data is needed for all these methods and how will this data be collected?*
    c. *What are the computational or analytical tools that are required for each method?*

3 LITERATURE STUDY

3.1 WHAT IS BUS RAPID TRANSIT (BRT)

Bus Rapid Transit (BRT) is a bus system that operates at high frequency and speed, combining reliable travel times with high corridor capacity. It offers comfort and is easily distinguishable from regular bus services for passengers (Witte & Kansen, 2020).

BRT contains similar features to a light rail or metro system which makes it more reliable than regular bus services. The three main delays of a regular bus service are boarding and alighting, intersections and traffic congestions. BRT alleviates delays from all three of these elements. In Section 3.2 the criteria of BRT will be discussed, from the perspective of addressing these issues (ITDP, 2024).

3.2 CRITERIA FOR SUCCESSFUL BRT

The BRT standard is an evaluation tool designed by the institute for transportation and development policy (ITDP). It assesses the design and performance of BRT corridors based on performance and sustainability. The tool gives a score out of 100 points based on different criteria explained in a later section.

Based on the total score, a BRT corridor is awarded one of the following certifications:

- Gold 85+
- Silver 70-84
- Bronze 55-69
- BRT basic 20-54

The first category of the evaluation focuses on the core BRT design elements, which are essential for a system to qualify as a BRT. To be classified as a BRT at least 20 points need to be scored in this category of which at least 4 points in the busway alignment.

- **Busway Alignment** (7 Points): The centre of the road of bus-only-lanes keeps the buses away from the busy curb side.
- **Dedicated Right-of-Way** (7 Points): Bus-only-lanes makes traveling faster and makes sure that the buses are not delayed due to congestion.
- **Off-board Fare Collection** (7 Points): Passengers paying on the station instead of on board will make sure the bus reduces delays by picking up passengers.
- **Intersection Treatments** (7 Points): The most important measure for buses to smoothly go through intersections is prohibiting turns for traffic across the bus lanes. This will reduce delays caused by turning traffic.
- **Platform level boarding** (7 Points): The station should be on the same level as the bus. This will ensure quick boarding and will also make the bus more accessible for wheelchair users, other disabled people and strollers with a minimal delay.

The BRT basics are core design elements; four other components are key to a well BRT corridor design:

- **Service planning** (18 Points): A good public transport network starts with service.
- **Stations and buses** (23 Points): The stations and buses of the BRT system need to be of high quality and be placed at the right location for people to use it.

- **Communications** (8 Points): When there are no good communications a passenger will not use the BRT corridor. Communication is a vital part to make the corridor usable.
- **Access and integration** (16 Points): The BRT corridor should be integrated with the other infrastructure within the city. The BRT corridor should connect to this other infrastructure to make sure people reach their destination smoothly after using the BRT.

If a BRT system does not operate well, it can receive point deductions. In practice, the categories that can receive the largest amount of deduction points are poorly maintained infrastructure, overcrowding and low commercial speeds (ITDP, 2024).

3.3 BRT CONNECTIONS IN THE NETHERLANDS

Witte & Kansen (2020) stated that there are three different scale levels of BRT within the Netherlands: Inter-city, short interurban and long interurban.

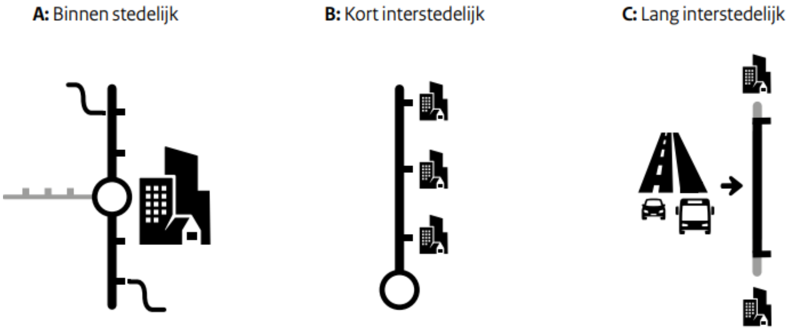

Figure 2: Different types of BRT connections

Inter-city BRT as rail replacement transport is the most common form of BRT. Large cities that did not have the financial resources to build a metro system found this a good alternative solution. A Dutch example of this is the Maxx bus line in Almere.

Short interurban BRT is relatively common within the EU but is also known as BHLS (Buses with High Level of Service). Short interurban BRT connects city centres with suburbs or small towns nearby. An important factor for this kind of BRT is that a direct connection to the city centre needs to be made. Short interurban BRT it is most of the time complementary to rail transport. But also, buses which only operate in peak times can be considered within this scale level. An example for this kind of BRT within the Netherlands is the R-netlijn 300 that goes from the south of Amsterdam to Schiphol to Haarlem.

Long interurban BRT (>50 km) is currently less common in the Netherlands but has a lot of potential. Examples of these routes within the Netherlands are Groningen to Emmen and Breda to Utrecht. Long interurban BRT connects larger cities with a minimal number of stops. This type often uses dedicated bus lanes next to the highway or the emergency lane when there is not that much congestion. Long interurban BRT often serves as either a rail alternative or as a complementary service (Witte & Kansen, 2020).

3.4 DEMAND ANALYSIS

There are several factors that influence the demand for public transport. The three biggest factors that influence the demand are fares, quality of service and car ownership(Paulley et al., 2006a). In the upcoming section all three factors will be explained.

Fares

Fares have a big influence on public transport, particularly on individuals with a lower income. Public transport is often their primary mean of travel, at least for trips beyond walking and cycling distances, making the demand highly sensitive to fare changes because with just a small increase in price, it can push these passengers to find another mode of transport. Fares also influence the overall travel behaviour of people. When fares are high, people are likely to reduce non-essential trips. On the other hand, when fares are lower, people are encouraged to use the mode often, not only for daily commuting but also leisure trips will be made with public transport.

Quality of service

Quality of service is another big factor that can influence the demand greatly. Within the subject of quality of service, many different factors play a role. One of the most important aspects is the frequency of the service. A higher frequency generally leads to a decrease in waiting time, making public transport a more convenient option, increasing demand.

Another crucial factor is transfer-waiting time and the conditions in which people need to wait. People often perceive waiting time as longer than it is, especially when changing modes of transport. If waiting times are long or waiting areas are unclean, unsafe or lack shelter from the weather, this can discourage the use of public transport and therefore reduce demand.

Reliability is closely linked to the two previous factors. An unreliable system leads to longer waiting times and missed transfers. This makes the usage of public transport less attractive and will decrease demand.

Finally, information provision is another factor within the quality of service. While clear and accessible information is important for the experience of users, it impact on demand is relatively limited to the other factors that are mentioned.

Car ownership and income

Income and car ownership have a big influence on bus demand. Higher incomes lead to increased travel frequency, longer trips and car possession, shifting travel from public transport to car use. This will lead to a decrease in demand within the public transport sector in some contexts, especially in buses. Train travel is usually less affected by this factor.

3.5 STUDIES CONDUCTED IN THE PAST

There are several studies conducted within the subject of BRT corridors, in particular regarding the prediction of BRT demand. The first two studies analysed next were also bachelor theses from the University of Twente that have been done at Keypoint. The other studies are on demand analysis for BRT corridors from abroad.

de Wit (2023) developed a framework for finding promising BRT routes between municipalities in the Netherlands using the gravity model as the main criteria. Next to the gravity model also the criteria of current travel times, existing train and metro network, frequency, job availability, number of students, day trips, congestion and airports are considered for the OD matrix. Based on these criteria a ranking was made. This ranking is, however, subject to improvements since existing bus routes were not considered, this was done manually after a promising route was found. At last, two promising BRT routes were analysed in detail.

Spijker (2024) continued the work of de Wit. He automated the search for promising BRT corridors on a smaller scale than municipality scale. The model is based on population, number of jobs, great-circle distance and difference between car and public transportation travel times for assigning a BRT score for each OD-pair. After this, a method was developed to find a complete BRT route with intermediate stops. This was done using Dijkstra's algorithm. Finally, a case study for Schiphol Airport was developed.

There have been several studies on demand analysis for potential BRT ridership. Each study uses its own method of predicting ridership, depending on the availability of data, scope, and the goal of the research. For this research, the main focus will be on discrete models such as discrete choice logit models, regression-based models and spatial interaction models like the gravity model.

Hensher & Golob, (2008) used multivariate linear regression for predicting BRT ridership in an international context. Within linear regression modelling a relationship is estimated between a dependent variable, such as BRT ridership, and various independent variables. The independent variables can be split up in two parts, socio-economic, population density, number of jobs, and transport-related characteristics such as travel time and frequency. To see whether the model fit and has statistical significance the R^2, t-test and F-tests are performed. The study compiles data from 44 big cities around the world, including 2 Dutch cities, Amsterdam and Eindhoven.

Karunakaran et al. (2023), also uses multiple linear regression to estimate demand in Kuala Lumpur. The variables used in this research are in-vehicle travel time, bus stop distance, bus commercial speed, service frequency and rainfall. In this research it is found that bus stop distance is not significant during peak hours and in-vehicle time is not significant during off-peak hours.

Another commonly used approach in predicting ridership is the gravity model. Asmael & Waheed (2018) applied the gravity model to estimate bus ridership. In their research Baghdad is the study area in which 5 routes are looked at. They have gathered their data trough surveys. Origin-destination pairs were made based on population sizes and travel distances between bus stops. With their research they successfully estimated demand for the five routes in Baghdad.

Tirachini et al. (2014) describe a method for forecasting public transport demand by using the logit model. In their research about how traffic congestion and bus crowding influence the design of urban public transport the authors explain how the logit model can be used to estimate bus demand and then design a bus route, with or without dedicated bus lanes, when buses compete for demand with cars and other modes.

A more complete way to forecast demand of a new public transport service would be to use the four-step model. The four-step model, as the name suggests, has four steps in the process of estimating demand: trip generation, trip distribution, mode choice and trip assignment (ITDP, 2018). The trip generation predicts the number of trips originating or destined for the analysis zone. This can be done by regression modelling (Ahmed, 2012). Trip distribution is the next step. In this step the origin and destination are linked to each other making an origin destination matrix. This matrix can be made based on the gravity model or another distribution model. The next step is mode choice. To know how many people will use a certain mode the modal split needs to be calculated. This can be done by using discrete choice models such as the logit model. The last step in the four-step model is the trip assignment. The assignment stage is where the supply of the public transport service is matched with the demand conditions in a simulation, and route choice is predicted. This is done by calculating the travel time and costs required for each possible route. The simulation assigns each passenger to a path and fins the assignment that has the lowest travel time and costs for each traveller (ITDP, 2018).

4 METHODOLOGY

In this chapter the methodology for the different methods that will be analysed will be explained. The four methods that will be explained are: A comparative case study, Logit Model, Regression Model and the Gravity Model.

4.1 COMPARATIVE CASE STUDY

The first method to predict demand is by a comparative case study. To conduct a comparative case study, a similar BRT corridor must be identified. A BRT corridor is defined as an origin and a destination, both are districts. The similar corridor needs to share some key characteristics with the corridor for which ridership is being predicted. The key characteristics that are considered in this research are:

- Population
- Amount of jobs
- Distance between A and B
- Existing direct bus connection

These characteristics also serve as the input data for this method. The goal of this method is to find an existing corridor that closely matches the characteristics of the target corridor. Once a comparable corridor is found, its ridership numbers can be used to estimate the expected number of passengers for the corridor that is being studied. To find a comparable method some steps need to be taken.

4.1.1 Data collection and preparation

The needed data for the method is the following:

- Population per district in the Netherlands (CBS, 2012a)
- Job availability per municipality in the Netherlands (CBS & LISA, 2023)
- Existing bus routes (OpenOV, 2025)

First, the data needs to be prepared. To calculate the distance between the districts, the centre of a district must be found. This is done by finding the centre point of each district based on the population. To find the population centre of each district, the principle of finding the centre of gravity was used. This principle tells that a plane consisting of smaller planes each with their own centre of gravity can be combined to find the total weight and therefore centre of gravity. In the case of this research each district was divided into neighbourhoods. The population of these neighbourhoods are the ‘weights’ and the centroid of each neighbourhood gives the place of the neighbourhood in the district. The python script with the calculation for the centre of the population per district can be found in Appendix A. The centre of population for each district can be seen in Figure 3.

Figure 3: Population centres of districts in the Netherlands

4.1.2 Calculating distances

When the population centres are known, the distance between all districts in the Netherlands can be calculated by using the principle of great circle distance. The great circle distance is a method to calculate the shortest distance over the earth's surface. The great circle distance between Assen Kloosterveen and Groningen is 23 kilometres. Therefore, there was a selection of OD-pairs that are between 15 and 30 kilometres, to increase the analysis speed.

4.1.3 Adding additional parameters

The number of jobs in each municipality is added for potential destinations for passengers. lastly, data that tells whether there is already a direct bus connection between origins and destinations in districts is added, using General Transit Feed Specification data from OpenOV.

4.1.4 Calculating similarity score

To identify the most comparable existing corridor, a similarity score is calculated for each connection based on five different things:

1. Distance between districts
2. Population at origin
3. Job availability at origin
4. Population at destination
5. Job availability at destination

Each parameter is normalized by dividing the absolute difference from the reference corridor by the reference value. This makes sure that big numbers will not dominate the score, but it will take the relative difference between the five parameters.

The final similarity score is the sum of all normalized differences. The closer the value is to 0 the higher the similarity is between the reference corridor and the found corridor, a score of 0 therefore indicates an exact match.

4.2 LOGIT MODEL

The second method for predicting demand uses a multinomial logit model, a common discrete choice model used to estimate mode share based on the utility associated with each transport option. In this study, the model is applied to a binary choice between car and bus, so it will be referred to as the logit model throughout. The method for the logit model is based on the master thesis of Espinoza (2025), which presents parameters of a logit model estimated in the Netherlands. First, the utility for each mode needs to be calculated. The equations for the utility of the bus and car can be seen in Equation 1 and Equation 2.

Equation 1: Bus utility

$$u_b^{ij} = \alpha_b + \beta_a t_{ab}^{i} + \beta_h t_{wb} + \beta_{vb} t_{vb}^{ij}$$

Equation 2: Car utility

$$u_a^{ij} = \alpha_a + \beta_{va} t_{va}^{ij} \frac{\beta_c c_r^{ij}}{o_r}$$

Where,

α_m = Modal constant per mode

β_k = Parameters associated to different attributes

t_{ab}^{i} = Access time at zone i

t_{wb} = Waiting time, estimated as half of the headway between two consecutive buses ($\frac{1}{2} h_b$)

t_{vm}^{ij} = In-vehicle time for mode m between zones i and j

c_r^{ij} = Car running cost to travel between zones i and j

O_r = Average car occupancy rate

When the utility of both modes is known, modal demand can be estimated using the logit model. The formula used for the logit model can be seen in Equation 3.

Equation 3: Logit model

$$y_m^{ij} = Y^{ij} * \frac{e^{u_m^{ij}}}{\sum_n e^{u_n^{ij}}}$$

Where,

y_m^{ij} = Demand for the calculated mode

Y^{ij} = Total demand in OD pair (i,j)

u_m^{ij} = Utility of the calculated mode

$\sum_n e^{u_n^{ij}}$ = Sum of the utility of modes

To calculate the demand based on the logit model the total demand of the corridor, OD pair between i and j, is multiplied with the utility of the bus divided by the sum of the utility of the car and bus (Espinoza, 2025).

4.2.1 Parameters and model constants

To compute the utility values for the modes, it is essential to identify key parameters. These parameters, along with their respective values, can be seen in Table 1. The values are retrieved from the significance report from Kouwenhoven et al. (2023), and are expressed in euros and hours.

Table 1 Parameters per mode retrieved by Significance technical report 2023, the parameters are based in euros and hours

Mode	Parameter	Value
Bus	β_{access_time} (β_a)	-8.68
	$\beta_{waiting\_time}$ (β_h)	-10.66
	$\beta_{in-vehicle\_time}$ (β_{vb})	-9.78
	β_{costs} (β_c)	-0.76
Car	$\beta_{in-vehicle\_time}$ (β_{va})	-8.24
	β_{costs} (β_c)	-0.76

Another important value to calculate the utilities is the modal constant. The modal constant accounts for unobserved preferences and systematic biases do not capture by other variables. The adjusted modal constant is calculated to adjust the observed modal share on the corridor, obtained by OViN (Onderzoen Verplaatsingen in Nederland) data and the normalized modal share within the general population. The equation for the adjusted model constat is shown in Equation 4.

Equation 4: Adjusted model constant

$$\alpha'_m = \alpha_m - \log\frac{q_m}{Q_m}$$

Where,

α_m'= New modal constant

α_m = Original modal constant; Bus = 3, car = 1.79

q_m = Modal share of mode m

Q_m = Normalized modal share of the population

This adjustment ensures that the modal constant reflects both the specific context of the chosen corridor and the broader context of transport preferences within the population (Ortuzar & Willumsen, 2011).

4.2.2 Access time

To calculate the utility the access time is required. The access time refers to the time a passenger needs to walk from their origin to the nearest public transport stop, in this case a bus stop. The furthest an average Dutch person is willing to walk to a bus stop is approximately 700 meters (van der Waerden et al., 2024). This corresponds with a walking time of around 8.4 minutes when a walking speed of 5 km/h is assumed. Conversely, the shortest distances may be covered in less than a minute.

When assuming a uniform distribution of the population, where all passengers are equally likely to live at any point within the service area surrounding the bus stop, the average walking distance can be the middle point of the shortest and longest walking distance. Consequently, a fixed access time is adopted for this study which is 4.2 minutes.

4.2.3 Waiting time

There is no variation in waiting time across the corridor because the waiting time t_{wb}^{i} is half of the headway between two consecutive buses ($\frac{1}{2}h_b$), and the headway is the inverse of the bus frequency ($\frac{1}{f_b}$) (Espinoza, 2025).

4.2.4 In-vehicle time

To estimate the in-vehicle time for both modes, Google Maps was utilized as the source of real-time and route-based travel time estimation.

For the bus mode, the total in-vehicle time was determined as 37 minutes. This duration corresponds from the Aletta Jacobs bus stop, which is in the middle of the district, to Groningen central station as destination. In contrast, the same trip done by car only takes approximately 24 minutes of in-vehicle time (Google, 2025).

4.2.5 Car running costs

Fuel consumption is the basis for calculating the vehicle's operating costs. The costs for 1 litre of petrol were €1.79 (CBS, 2012b), and they use 8 litres per 100 kilometres. This means that the costs of a vehicle are €0.1432 per kilometre.

4.2.6 Total demand

When the utilities are calculated the one thing that is missing is the total demand of the corridor. The total demand is gathered by TomTom data. The data set used was from 2024 as earlier data sets were incomplete or not accessible. From the analysis the traffic flow from Assen Kloosterveen to the city centre of Groningen is 4835 and the traffic flow from Groningen centre to Assen Kloosterveen is 3975. This gives a total demand of 8810 passengers per day.

4.3 REGRESSION MODEL

Another method to predict ridership is by using a regression model as seen in Equation 5. A regression model is a statistical method used to identify the relationships between variables. In this case it held to determine what the potential ridership is and what variables influence that. This model is based on several variables that influence the potential ridership such as speed, frequency and the capacity of the buses. An intercept value is also added, this constant captures other variables that are not specifically mentioned but still have a small influence. To know these parameters a rough design is already needed for the new BRT corridor to know the potential ridership (Currie & Delbosc, 2013).

Equation 5: Regression model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon_i$$

Where:

Y_i = Dependent variable i (predicted ridership)

β_0 = Intercept value

x_i = Independent variables predicting Y_i (e.g. speed, frequency, capacity)

β_{in}= Regression coefficients to be estimated

ε_i = Error

In the model to predict BRT demand presented by Hensher & Golob, 2008, several system variables are found to have a statistically significant influence on the number of passengers per day. These are: Average fare per trip, Average peak headway, number of stations and vehicle capacity. In Table 2 these variables can be found with their correspondent regression coefficient. The dependent variable is the predicted ridership in passengers per kilometre.

Table 2: passenger trips per day regression model variables (Hensher & Golob, 2008)

Variable	Beta parameter
Number of stops	0.0098
Peak headway	-0.1681
Vehicle capacity	0.0052
Average fare per trip	-0.2577
Constant	7.9209
Adjusted R^2	0.742

To predict the dependent variable, ridership, the independent variables must also be known. The first one is number of stations, which is 15 on the line 309 from Assen to Groningen. The second variable is frequency which is four times per hour in peak hours (Moovit, 2024). The third variable is capacity which has a maximum of 80 people (Keolis, n.d.). Lastly, the f, the structure includes a base tariff of €1.12, plus an additional charge of €0.126 per kilometre (Qbuzz, n.d.).

Once the regression model is filled in, Equation 5: Regression model Equation 5, Equation 6 can be filled in to get the predicted ridership. The last step is the multiplication of the result of Equation 6 to the number of kilometres passengers can access the bus route as it was given in passengers per kilometre.

Equation 6: Conversion regression model

$$Predicted\ passengers = e^{Y_i}$$

4.4 GRAVITY MODEL

The gravity model is based on Newton's theory of gravity. The gravity model assumes that the trips produced at an origin and attracted to a destination are directly proportional to the total amount of trips between the origin and destination. Also, when the distance between the origin and destination is smaller more trips will be made and when the travel time is longer less trips will be made. The gravity model for trips has the following form (Ortuzar & Willumsen, 2011):

Equation 7: Gravity model

$$T_{ij} = \alpha O_i D_j f(c_{ij})$$

T_{ij} = Trips produced at i and attracted at j

α = Scaling constant

O_i = Trip production in i

D_j = Trip attraction in j

$f(C_{ij})$ = Deterrence function based on travel costs

The deterrence function captures the travel costs, either in travel time, money spend or a combination of both (generalized cost). The equation for the deterrence function can be modelled in three different ways:

Equation 8: Exponential deterrence function

$$f\left(c_{ij}\right) = exp^{\left(-\beta c_{ij}\right)}\ (Exponential\ function)$$

Equation 9: Power deterrence function

$$f\left(c_{ij}\right) = c_{ij}^{-n}\ (Power\ function)$$

Equation 10: Combined deterrence function

$$f\left(c_{ij}\right) = c_{ij}^{n} * exp^{\left(-\beta c_{ij}\right)}\ (Combined\ function)$$

An important aspect of the deterrence function are the constant values (β or n). These values fit the selected function to a specific location. For this study the exponential function will be used, where β=0.016. This value is based on a public transport study done in Norway (Paulley et al., 2006b).

To calculate the scaling constant (α) the gravity model needs to be done on a similar corridor. The corridor that was used is the solution from the comparative case study. First the gravity model will be calculated for the similar corridor, in which the scaling factor is 1. Then the result of that will be divided by the actual number of passengers of that connection, this number will then be the scaling factor in the equation of the wanted corridor. By doing this the results of the gravity model will be scaled to a real demand (Princeton university, 2008).

4.5 COMPARISON

At the end all four methods will be compared to each other to determine which method is the most time efficient but is still accurate. The methods will be compared on four criteria, the accuracy of each method, the needed data, time-efficiency and the computational tools/skills needed to perform the method.

4.5.1 Accuracy Comparison

To evaluate the accuracy of each method, the estimated output will be compared to the observed passenger count data from the line 309 between Assen and Groningen. All four methods will be compared to the same corridor to ensure consistency for the comparison.

The accuracy of the models will be quantified using statistical error metrics, the absolute error and the absolute percentage error will be used. This will give an objective comparison of each methods performance.

Absolute error:

$$AE = |y_i - \hat{y}_i|$$

Absolute percentage error:

$$APE = \frac{|y_i - \hat{y}_i|}{y_i} * 100\%$$

Where,

Y_i = measured value

$\hat{Y}_i$ = predicted value

(Pahdma M, 2021)

4.5.2 Data Comparison

Each of the methods require different types of data. To ensure a fair comparison a table is made that outlines the required data for each method including:

- Which data is needed (operational data, Historical data, socioeconomic data, demographic data, travel behaviour)
- How accessible is the data
- How much effort is needed to prepare the data

4.5.3 Time-efficiency comparison

Each method needs different time investments. Each method will be compared based on their time-efficiency. The comparison will be based on:

- Model set-up time
- Data processing time
- Computation time
- Interpretation time

4.5.4 Computational Comparison

Fo this question again an overview needs to be made that tells what tolls are needed to compute the method and how easy this tool is to use. For each method the following aspects will be evaluated:

- Required tools or platforms (e.g., Python, Excel, GIS software)
- Level of difficulty
- Accessibility of the tools (e.g., Open source, licenced)

5 RESULTS

In this chapter the results of each model will be explained. After that the comparison between the methods will be explained per subject, first accuracy, then time efficiency and at last the computational skills needed.

5.1 COMPARATIVE CASE STUDY

The top five most similar bus corridors to Assen Kloosterveen and Groningen are visualized in Figure 4. These connections were identified by calculating a similarity score for each direct connection in the Netherlands, based on five parameters: population at both origin and destination, job availability at origin and destination and the distance between origin and destination. The lower the similarity score, the more similar the corridor is to the reference corridor, Assen – Groningen.

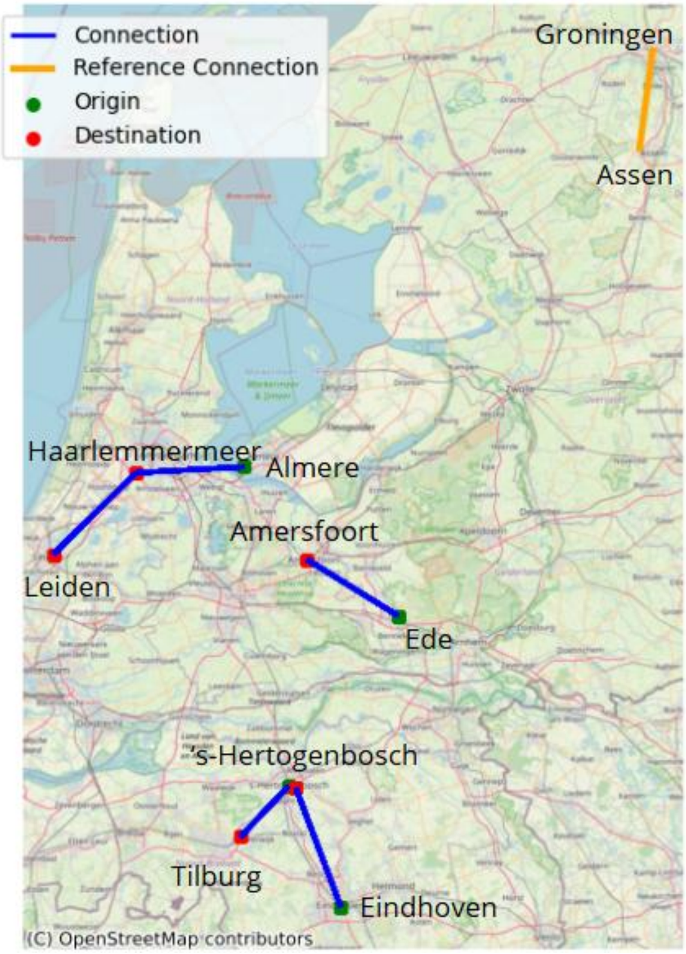

Figure 4: Top 5 similar corridors for Assen and Groningen

Table 3 presents the top five corridors ranked by similarity along with the corresponding bus line numbers. These bus line numbers were identified manually after the ranking was completed.

Table 3: Top similar connections to Assen Kloosterveen and Groningen

Connection	Similarity score	Bus line number
Berkel-Enschot (Tilburg) – 's-Hertogenbosch center	0.884	Line 140
Badhoevedorp – Leiden Noord	0.966	Line 145 (not existing anymore)
Almere Poort – Badhoevedorp	0.995	No connection anymore
Eindhoven center – 's Hertogenbosch zuidoost	1.036	Line156
Ede Oost – Amersfoort Soesterkwartier	1.082	Line 834

Out of 295 possible connections, the most similar existing corridor has a similarity-score of 0.884. Given that most other corridors have a score between 2 and 3 with a mean of 2.471 and a standard deviation of 0.822, as illustrated in Figure 5, the value of 0.884 suggests that there is a high degree of comparability. This corridor's characteristics deviate by less than 20% on average across all five parameters. This makes the corridor between Berkel-Enschot and 's-Hertogenbosch a suitable reference for the ridership prediction between Assen and Groningen.

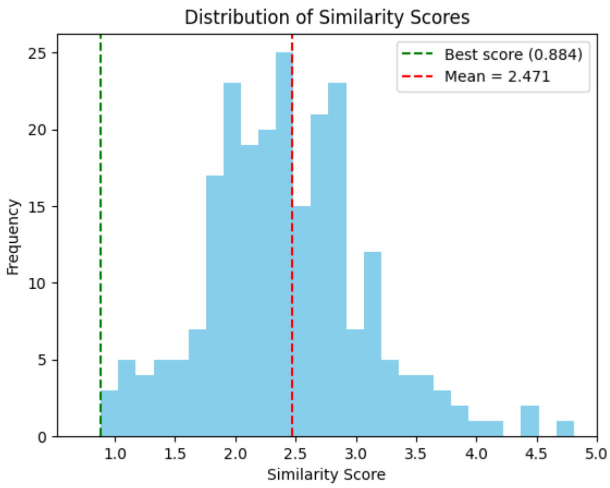

Figure 5: Distribution of similarity scores

As such, bus line 140, which serves this route, is selected as a reference case for the ridership prediction. The ridership data for this line was obtained from Province Noord Brabant. They have provided the ridership per day in November from the year 2019, 2023 and 2024. The results can be seen in Table 4.

Table 4: Ridership numbers of bus line 140

Year	Total ridership in November	Average ridership per workday
2019	25.176	1199
2023	21.363	971
2024	21.220	1010
Average	22.568	1060

5.2 Logit Model

The predicted ridership, using the logit model, are presented below. First, the utility values for each mode were calculated. The utility of the bus has a value of -5.1046, while the utility of the car has a value of -4.2160. These values were than used to calculate the probability of choosing each mode of transport.

The probability of choosing the bus, based on the utility, was calculated to be 29%, while the probability of choosing the car was 71%. These percentages reflect the attractiveness of each mode, so in this case the car is more attractive compared to the bus. The usage of the car has a shorter travel time and is perceived easier to use, as there is no access and waiting time for the car.

Given that the total demand on the corridor was 8810 passengers per day, the multinomial logit model predicts that 2567 passengers will use the bus connection daily. The remaining 6243 people will use the car to go from Assen Kloosterveen to Groningen station.

5.2.1 Sensitivity analysis

To identify which parameters have the greatest influence on the output of the multinomial logit model, a sensitivity analysis was conducted. In Figure 6 the results sensitivity analysis is presented. The model is most sensitive to the in-vehicle time of the bus. When the in-vehicle becomes larger the estimated demand for the bus will decrease.

Similar effects are also observed for the car distance, waiting time and access time, an increase in any of these parameters will result in a decrease in the estimated bus demand.

Conversely, the in-vehicle time for the car and both modal constants have a positive relationship with the estimated bus demand, as their value increases so does the predicted ridership for the bus. Just as the total demand of the corridor, but to this parameter the model is less sensitive than the others with the positive effect.

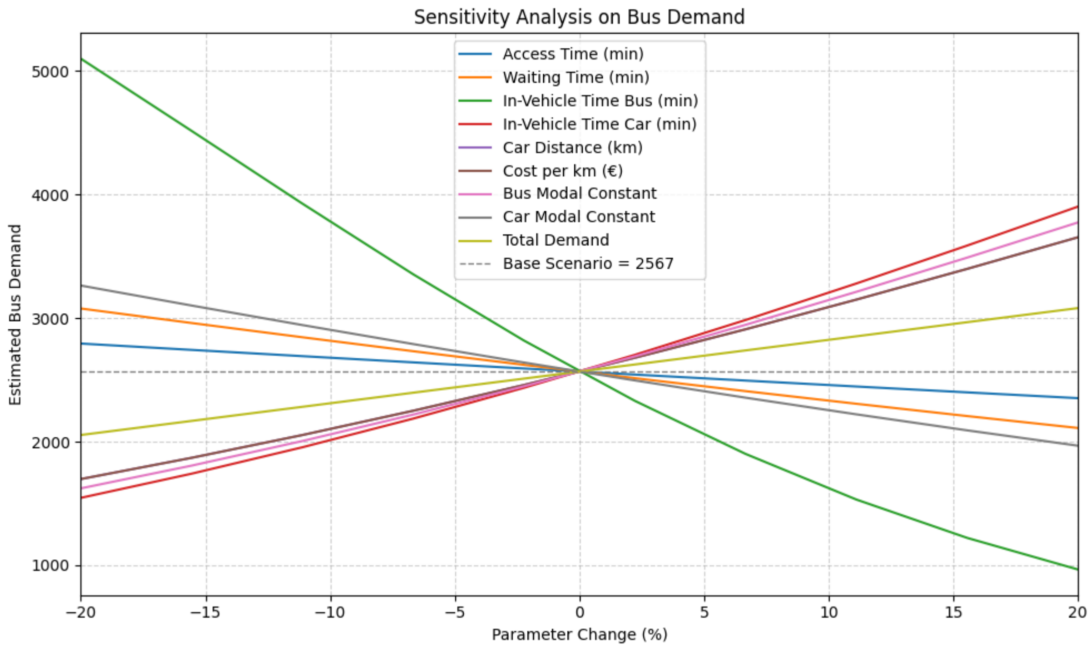

Figure 6: Sensitivity analysis Logit Model

5.3 Regression Model

Using the multivariate regression model as explained in the methodology section, the ridership can be predicted per kilometre for the corridor between Assen Kloosterveen and Groningen. The required inputs for the model are:

- Number of stations: 15
- Peak headway: 15 minutes
- Vehicle capacity: 80 passengers
- Fares: Base fare is €1.12 plus an additional €0.126 per kilometre.

For a full trip covering 30 kilometres, the total fare is 4.90. Since the model is in dollars this should be converted. Assuming €1 is $1.1 which makes the total fare $5.39.

Substituting all the values in the multivariate regression model yields:

$$Y_i = 4.5734$$

Exponentiating this value provides the estimated ridership of:

$e^{4.5734} \approx 97$ Passengers per day per kilometre.

It is important to note that only 11 kilometres of the corridor contains bus stops and therefore allow for on- and off-boarding. The remaining 19 kilometres consist of highway. Consequently, the total estimated ridership across the corridor is 1066 passengers per day in one direction. To get the total estimated ridership numbers, 1066 passengers per day needs to be multiplied by 2 as the bus, and therefore the passengers, go in two directions. This gives a total estimated demand of 2132 passengers per day between Assen Kloosterveen and Groningen station based on regression modelling.

5.3.1 Sensitivity analysis

To identify which parameters, have the most influence on the output of the regression model, a sensitivity analysis was conducted. In Figure 7 the results of the sensitivity analysis are shown. The model is most sensitive to the intercept value. The intercept value has a positive relationship with the ridership, when the value goes up, so does the predicted ridership. The parameters capacity and number of stops have also a positive relation with the ridership but do not have a significant influence.

The parameters headway and fares do have a bigger influence on the ridership with a negative relationship but do not have a big influence compared to the intercept value.

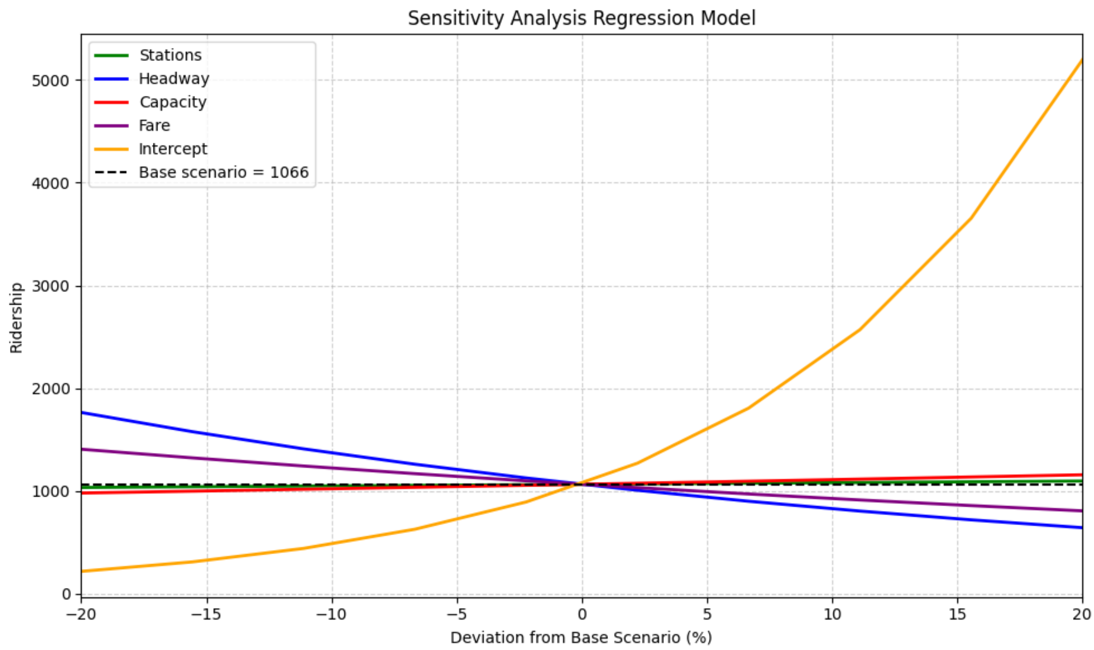

Figure 7: Sensitivity analysis Regression Model

5.4 Gravity Model

Predicting ridership by using the gravity model, as explained in the methodology section, was carried out by calibrating the model with data from a comparable existing corridor. The corridor that has been used is the one found in the comparative case study, the corridor between Berkel-Enschot and ‘s Hertogenbosch. By calibrating the scaling constant (α), the model output represents real-world travel behaviour accurately.

The demand in the corridor of Berkel-Enschot to ‘s Hertogenbosch is 1060 passengers per day. Since the corridor between Assen Kloosterveen and Groningen differs both in travel time and frequency from the corridor between Berkel-Enschot and ‘s Hertogenbosch, scaling factors are applied to estimate changes in the demand. When the scaling factors are known they will be multiplied with the actual demand of the corridor. These factors were derived based on elasticities that are commonly used in transport demand studies. The factors are based on expert knowledge. Demand will increase when the travel times becomes smaller, and the frequency becomes higher. The tables used for the scaling can be found in appendix A.

The travel time for the bus of the Assen – Groningen corridor is 37 minutes, compared to 51 minutes between Berkel-Enschot and ‘s Hertogenbosch. This results in a scaling factor of 1.32. Similarly, the Assen Kloosterveen – Groningen corridor operates at a frequency once every 15 minutes, compared to once every 30 minutes for the corridor between Berkel-Enschot – ‘s Hertogenbosch. This leads to another scaling factor of 1.24.

Multiplying the base demand (1060 passengers per day) by both scaling factors results in an adjusted demand of 1735 passengers per day. This number was used as the actual demand between Tilburg and ‘s Hertogenbosch in the gravity model.

Once the scaling factor (α) is known for the corridor between Assen Kloosterveen and Groningen the predicted demand can be calculated. The predicted ridership of the corridor between Assen Kloosterveen and Groningen is 2207 passengers per day.

5.4.1 Sensitivity analysis

To identify which parameters have the greatest influence on the output of the gravity model, a sensitivity analysis was performed. In Figure 8 the results of the sensitivity analysis are shown.

The model is most sensitive to the inputs of the population of Assen, jobs in Groningen and the demand of the comparative corridor. This is all the purple line in the figure as they all three have the same sensitivity. When one of these parameters increases the predicted demand will also increase. This is also the case when the beta coefficient is increased, but this parameter has less influence on the output of the model than the other three parameters mentioned above.

The travel costs have a negative influence on the predicted bus demand. When the travel costs increase the estimated demand will decrease.

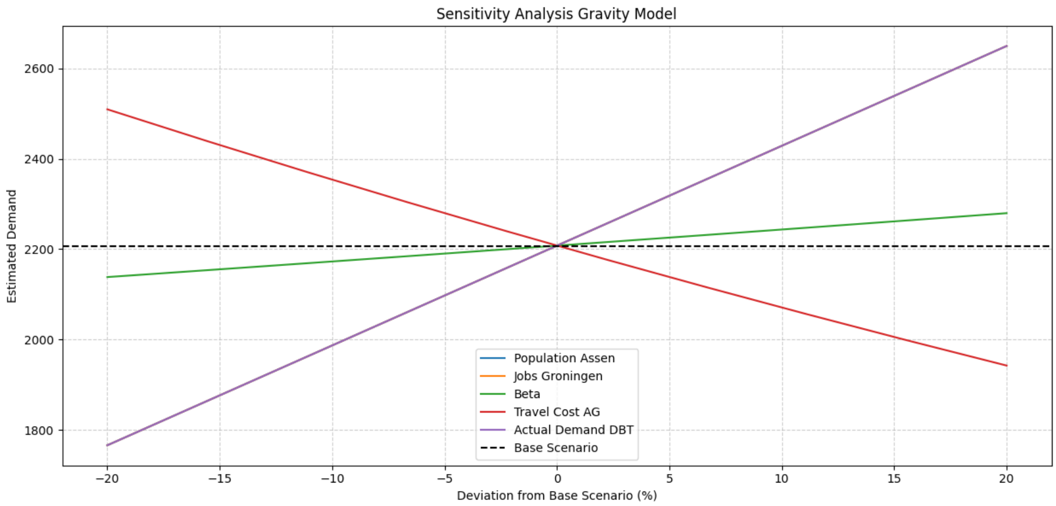

Figure 8: Sensitivity analysis Gravity Model

5.5 ACCURACY COMPARISON

The total ridership between Assen Kloosterveen and Groningen is on average 2420 passengers per day. This number does not only include passengers from line 309, but also those from line 109, which was introduced to reduce the burden on line 309. This number comes from an average of all days in the month of March in the years 2019, 2023 and 2024.

In Table 5 the predicted demand of each method can be seen as well as their absolute and percentage error with respect to the actual ridership between Assen Kloosterveen and Groningen. The actual ridership predicted ridership and absolute error are given in passengers per day.

Table 5: Accuracy comparison

Method	Actual Ridership	Predicted ridership	Absolute Error	Percentage Error
Comparative Case Study	2420	1060	1360	56.2%
Logit Model	2420	2567	147	6.1%
Regression Model	2420	2132	288	11.9%
Gravity Model	2420	2207	213	8.8%

In consultation with my supervisors at Keypoint, a deviation of within 20% is still acceptable for prediction. This means that the Logit model, regression model and the gravity model are of acceptable accuracy, whereas the comparative case study provides a large error of 56%.

5.6 TIME-EFFICIENCY COMPARISON

Each modelling method requires different time investments. To compare each method based on their time investment, this section outlines the time required for the model setup, data processing time, computation time and the interpretation time. At the end a table can be found with a summary of the time investment.

5.6.1 Comparative case study

Implementing the comparative case study required two separate python scripts. The first script calculated the distances between all district centres in the Netherlands and the second ranked the connections based on the input corridor to identify potential similar connections. The full development process took approximately 1.5 weeks but for users with more advanced Python experience, this setup time could be reduced significantly.

The processing of the data is considered high, primarily due to the large number of calculations required to determine the distances between the population centres of the districts, this script alone took about 2 hours to run. The computation time for the ranking script is also high, as it took around three hours to complete on a standard laptop.

The interpretation time is rated as medium. While the output of the script clearly mentioned which connection is the most similar, it did not provide direct ridership numbers. As such, additional input from an external stakeholder was needed.

Overall, the comparative case study method was relatively time-intensive method, particularly in the setup and computation time.

5.6.2 Logit model

The most time-consuming aspect of the logit model was understanding and correctly applying the theoretical formulas. The model requires customization for most of the parameters and especially customizing the model constant with the OViN data took some time. The model setup time is therefore considered high, but if one already understands the logit model, time for this could be significantly reduced, being on the same level as regression and gravity models. Once the theoretical structure is understood, implementing the model in Python is straightforward and can be completed within a day.

The data processing time is classified as medium, mainly due to the effort needed to understand and extract the OViN data to the correct corridor. The computation time is low for the logit model. The Python script runs for approximately two seconds and then there is already an output. Similarly, the interpretation time of the output is low, as the output clearly states the utilities, the probability of choosing a certain mode and the modes estimated demand on the given corridor.

5.6.3 Regression model

The model setup time for the regression model was low, as the regression coefficients were taken directly from an existing academic paper (Hensher & Golob, 2008). If one were to derive these coefficients themselves, setup time would increase significantly and fall into the high category.

No data processing was required, resulting in low data processing time. Similarly, the computation time and interpretation time were also low. The script took about five seconds to complete, and the outputs are directly the estimated ridership.

5.6.4 Gravity model

The setup time for the gravity model is rated as high, as it requires a moderate understanding of the theoretical foundation, including taking the right formula for the gravity model and corresponding cost function, but also the comparative case study needs to be done in order to perform this method which makes it time intensive.

Data processing time is also considered high as again the comparative case study also needs to be done for this method and afterwards that result also needs to be adjusted to the right characteristics of the corridor.

The computation time is considered to be high, just as with the comparative case study. After the script of the comparative case study is done the script for the gravity model does not take a lot of time. This script is done within 20 seconds. The interpretation time of the gravity model is low, as the output is the predicted ridership.

This makes the gravity model an overall high method as the results of the comparative case study are also needed. The gravity model on its own is not a time-intensive method.

5.6.5 Overview

In Table 6 a summary of the time-efficiency of each method can be found. Where the comparative case study and the gravity model have a relatively low time efficiency, and the regression model is the most time efficient. The logit model has approximately the same time-efficiency as the regression model only the logit model has a more complex theoretical background which makes the setup time higher.

Table 6: Summary time-efficiency comparison

Model	Model Setup Time	Data Processing time	Computation Time	Interpretation Time
Comparative Case Study	Medium	High	High	Medium
Logit Model	High	Medium	Low	Low
Regression Model	Low	Low	Low	Low
Gravity Model	Medium	High	High	Low

Legend

- Model setup time: Time to understand the structure of the model
    - *Low:* Simple method with minimal theoretical background, easily implemented
    - *Medium:* Requires moderate theoretical background and customization
    - *High:* Requires a good theoretical understanding of the structure and needs customization
- Data processing time: How long does it take to prepare the data
    - *Low:* Data was ready to use, no time needed
    - *Medium:* Minor things needed to be adjusted, max a workday
    - *High:* Data needed a lot of adjustments, or pre-calculations to use it in the method

- Computation time: How long does it take to run the model and get results
    - *Low:* Up to 30 minutes
    - *Medium:* 30 minutes to 4 hours
    - *High:* 4+ hours
- Interpretation time: Time needed to understand and visualize meaningful results
    - *Low:* Output is easy to interpret
    - *Medium:* Requires moderate efforts to understand the outputs
    - *High:* Complex outputs that need further explanation or visualisation

5.7 DATA COMPARISON

Each modelling method in this study requires different input data. To compare each method this section outlines the type, accessibility, quality and preparation effort of the data used for each method. At the end a table is given where a summary can be found.

5.7.1 Comparative case study

For the comparative case study three types of data sets were needed:

- Population per district in the Netherlands
- Job availability per municipality in the Netherlands
- Existing bus routes

All these data sets are publicly available. Initially, data on job availability on a district level was sought, but this was unavailable, therefore municipal-level data was used. The population dataset was of high quality, while the job data included municipalities that did not exist in 2012. Additionally, the bus route dataset contained outdated services, and services that do not exist, which made it less accurate.

Minimal data preparation was needed. Although the job data was expressed in terms of jobs per 1000 inhabitants and ideally should have been converted to absolute figures, inconsistencies in municipality names between the population and job data sets prevented this transformation.

5.7.2 Logit model

To use the logit model several parameters needed to be known:

- Modal split of the corridor, prior to the implementation of the new public transport service
- Normalized modal split of the population
- Several beta parameters associated with travel time and costs
- Access time
- Headway
- In-vehicle time
- Car running costs
- Average car occupancy
- Total demand of the corridor

To obtain modal split data, the 2011 OViN (Onderzoek Verplaatsingen in Nederland) survey was used. Access to this dataset had to be requested, and its quality was low due to a low number of recorded trips for the Assen-Groningen corridor, 18 trips from Assen to Groningen and 26 trips from Groningen to Assen over the whole year. Although trip weights were applied the data was not considered highly representative. Filtering the OViN dataset for only the relevant trips required some data cleaning and therefore some preparation.

To get the total demand of the corridor TomTom data was used. To get TomTom data a TomTom account is needed which is only accessible through the municipalities of the Netherlands. This makes that TomTom data is not really accessible.

All other parameters, such as access time, headway and in vehicle-time, were obtained from public sources and by the master thesis of Espinoza, which served as a methodological foundation for this method.

5.7.3 Regression model

To use the regression model to predict demand several parameters need to be known:

- Regression coefficients
- Number of stations
- Peak headway
- Vehicle capacity
- Fares

All the data needed for the regression model was easily accessible. The regression coefficients originate from a paper of Hensher & Golob (2008). Operational variables such as headway, vehicle capacity, number of station and the fares were obtained from public sources. The quality of the data is good and little to no preparation was needed to use the data. Only the fares needed to be transformed from euros to dollars.

5.7.4 Gravity model

To get a good result a comparable corridor also needed to be found. This was done using the result of the comparative case study, so everything mentioned in chapter 5.7.1 is also needed data. In addition to what is mentioned at the comparative case study, some additional parameters are needed for the gravity model.

- Population of the districts of Assen and ‘s Hertogenbosch
- Number of jobs in Groningen and Tilburg
- Travel time Assen to Groningen & Tilburg to ‘s Hertogenbosch
- Beta parameter value
- Conversion tables for the constant in the formula

For the population of the districts in Assen and ‘s Hertogenbosch and the number of jobs for the municipalities of Groningen and Tilburg the same data sets as for the comparative case study were used. This makes that the data is accessible, and the quality was also good because the selected municipalities are correctly represented. This made that there was some data preparation needed as the wanted municipalities and districts needed to be subtracted from the data set and the jobs to be transformed to an actual number and not per 1000 inhabitants.

The travel times were extracted from Google Maps which makes it highly accessible, the accuracy might be a little less as congestion can occur, but on the moments of measuring no congestions occurred. But no preparation was needed for this data. At last, a beta coefficient was needed for the cos calculation, this value was obtained from a paper, this makes it highly accessible, of good quality and no preparation needed.

The conversion tables used for the conversion of the demand of the Tilburg – ‘s Hertogenbosch corridor were given by my supervisor, who got them from a transport company. But they are widely spread and therefore accessible for everyone.

5.7.5 Overview

An overview of all that is mentioned above can be found in Table 7. The regression model scores the best as it has high data accessibility and quality and not preparation needed. The logit model has a somewhat lower level of data accessibility, and some of the data used for the logit model is of less quality. The comparative case study has just as the regression a high accessibility for the data, but the quality is a bit less and more preparation is needed. The gravity model is the same as the comparative case study, only a little more preparation of the data is needed.

Table 7: Data comparison summary

Model	Data Accessibility	Quality of the data	Data preparation Effort
Comparative case study	High	Medium	Medium
Logit Model	Medium	Medium to high	Low to medium
Regression model	High	High	Low
Gravity Model	High	Medium	Medium

Legend:

- Data Accessibility: How accessible is the data
  - *Low:* Restricted access to data files
  - *Medium:* Some data sets are public others require access
  - *High:* All data is available
- Quality of the data: What is the quality of the data
  - *Low:* Data has poor quality, is incomplete
  - *Medium:* Data has moderate completeness and consistency
  - *High:* Data has good quality and is complete
- Data preparation efforts: How much does the data need to be prepared
  - *Low:* Data is ready to use, no preparation needed
  - *Medium:* Some structuring or small transformations needed
  - *High:* Requires a lot of adjustments, pre-calculations needed

5.8 COMPUTATIONAL COMPARISON

Each modelling method requires different programs and therefore computational skills. To compare each method based on the computational skills needed this section outlines the required tools/platforms needed, the difficulty level of the tool and how accessible the tool is. At the end a table can be found the summarizes the computational skills needed.

5.8.1 Comparative case study
The comparative case study relies on two Python scripts, one to calculate the distances between all districts in the Netherland and one that ranks potential connections based on the input corridor. Creating and modifying these scripts requires a moderate understanding of Python, although to execute the model basic knowledge is enough.

In addition, QGIS is used to extract the population data at a district level from Pdok. This tool, just as Python, is an open source. To use the GIS tool some knowledge is needed to execute the required task. Microsoft Excel is the last tool that is used for the comparative case study. This tool is used for the jobs per municipality.

5.8.2 Logit model
The logit model uses several tools. The core model is made in Python, where basic skills are needed to perform the script. To determine the model constat, Excel is used. In Excel basic skills are needed such as the filtering, summing and determining the average. QGIS supports the identification of the needed districts to determine the total demand of the corridor. This is then used in the TomTom tool to extract the corridor specific data.

TomTom’s platform requires a registered account, access may also be limited depending on the licence. Overall, this method combined multiple platforms and demands a moderate level of computational skills.

5.8.3 Regression model
The regression model is computationally straightforward. It can be implemented in using a Python script, but due to the simplicity of the formula Excel or even a calculator can be used. Developing the Python script only requires basic knowledge and executing is even simpler.

This method is considered low in computational skills and the tools needed are highly accessible and user-friendly.

5.8.4 Gravity model
The gravity model is implemented in Python and relies on basic mathematical Python skills. Extracting population data is done using QGIS while Excel is used to process the job availability data. Using these tools requires basic knowledge and just as with the comparative case study the tools are highly accessible.

5.8.5 Overview
In Table 8 an overview can be seen of the computational skills/tools needed. The regression model is the most user-friendly option as the formula is not hard to implement. The logit requires a more programming (Python) skills and need some other tools to complete. The comparative case study and, therefore also the gravity mode, were the most difficult to make in Python. This method required a good understanding of Python.

All the tools needed to perform the methods are highly accessible only the TomTom tool needs a license which makes the logit model a bit less accessible than the other three models.

Table 8: Summary computational comparison

Model	Required tools/platforms	Difficulty level	Tool Accessibility
Comparative Case Study	Python, QGIS, Excel	Medium to high	High
Logit Model	Python, Excel, QGIS, TomTom	Medium	Medium
Regression Model	Python	Low	High
Gravity Model	Python, QGIS, Excel	Medium to high	High

- Required tools/platforms: Software or environments needed to run the model
- Difficulty level: How user friendly is the tool/program
    - *Low:* Simple formula based
    - *Medium:* Basic coding skills needed
    - *High:* Lots of coding
- Tool Accessibility: How accessible is the tool/program
    - *Low:* License needed to use the tool/program
    - *Medium:* Free but requires knowledge to use the software
    - *High:* Easily accessible, open source
    -

5.9 MULTI CRITERIA DECISION ANALYSIS

To determine the most appropriate method to predict ridership for a new BRT corridor, a Multi-Criteria Decision Analysis (MCDA) was conducted. The analysis evaluates each method based on the previous explained comparison methods; accuracy, data availability, time-efficiency and computational skills/tools.

Each criterion has an assigned weight, reflecting the importance of the criteria to the output it gives. The weights are based on the perspective of a consultant, as this research is done for a consultancy firm. The total score of each method is then calculated by multiplying the weight with the given score. The score of each method is based on the ranking between the four methods. When a method is given four points, this is the best performing method on that criterion. The weights of each criterion will be explained below:

- Accuracy (Weight: 3)
  Accuracy is considered the most important criterion, as the reliability of the predicted demand directly influences the decision-making process for transport planners. An inaccurate prediction can lead to costly outcomes when a line is made while this will not generate enough passengers.
- Data Availability (Weight: 3)
  The availability and accessibility of the input data is crucial for the implementation of the methods. When no input data is available not method can predict the ridership.

- Time-efficiency (Weight: 1)
  Time-efficiency accounts for all the time needed to perform a method. This is not the most crucial factor but when sources are limited the method that requires the least amount of time is preferred.
- Computational skills/tools (weight: 1)
  This criterion reflects the technical difficulty and software requirements of each method. Methods that require hard to get software or require advanced programming may be a barrier for the implementation of some of the methods.

Each method for predicting ridership is ranked based on each criterion. When a model is given four points this means that this method performs the best based on this criterion. The ranking is based on the comparison explained in chapter 6.5 to 6.8. The evaluation can be found in Table 9.

Table 9: Multi-Criteria Decision Analysis

Criteria	Weight	Comparative case study	Logit model	Regression model	Gravity model
Accuracy	3	1	4	2	3
Data availability	3	3	2	4	3
Time-efficiency	1	2	3	4	1
Computational skills/tools	1	2	3	4	2
	Total	16	24	26	21

Based on the weighted scores, the regression model emerges as the most favourable method with a total score of 29 points. This is mainly due to the strong performance on the data availability, time-efficiency and the computational skills/tools. But it does lack accuracy.

The logit model follows closely with a score of 24. It offers high accuracy, but some more efforts need to be done when computing the method and data is available, but licences are needed to access which makes the method less applicable in some cases.

The gravity model ranks third with 20 points. Since the comparative case study also needs to be done to perform the gravity model it ranks lower. Mainly because this makes it time intensive and more computational skills are needed. But in terms of accuracy, the gravity model sores well.

Finally, the comparative case study scored the lowest with only 16 points. The low number of points is mainly due to the low accuracy of the method. The time investment and data needed for this method is high, but its predictive performance is insufficient. Therefore, it is not recommended to use the comparative case study for predicting ridership numbers.

6 DISCUSSION

Each method gave a result which will be discussed in this section. Just as the assumptions and limitations each method has. At first the overall discussion points for all methods will be discussed. After that the four methods and the comparison will be discussed.

6.1 OVERALL DISCUSSION POINTS

All the methods have some common limitations and overall assumptions.

There are several ways that increases the number of passengers: population growth, people switching from car to public transport and people switching from public transport mode so first they took the train but now they will take the bus. Only the logit model incorporates the switch from car to bus. All other models do not incorporate any of these possible growth factors. This can cause a slight underestimation of the actual number of passengers that will take the bus route.

All the methods are performed on one corridor the Assen – Groningen corridor. Findings from the Assen-Groningen corridor may not generalize to other routes due to contextual differences. Due to low data accessibility and time this has not been done but is a good way to improve the research.

The most influential factors that influences the demand of public transport is car ownership and income. Both are not considered in any of the methods, while according to literature this has a large influence on the ridership of public transport.

6.2 COMPARATIVE CASE STUDY

The comparative case study produced the least accurate prediction of the four methods, underestimating the potential ridership with 56.2%. The biggest limitation of this method is its reliability on available real-world corridors for its comparison. There is rarely a perfect match. The best match for the Assen-Groningen corridor was the one between Berkel Enschot and 's Hertogenbosch. This corridor had a similarity score of 0.884 percent, meaning a deviation of 10-20% on average across the five parameters.

Furthermore, the method is heavily dependent on the assumption that the ridership of a similar corridor directly translates to a predicted ridership for the wanted corridor. With this assumption differences in travel behaviour, land use and other transport alternatives are not considered. Another flaw in this method that it suggests similar corridors that either do not exist anymore or have never existed. Even when a corridor appears similar based on the parameters, it may not actually exist as an operational bus route, limiting its practical reference. This problem could be solved by having a more representable data set, but for this research it could not be achieved. Another data set that could not be found is the number of jobs per district in the Netherlands. So instead, the number of jobs per municipality was used. This can cause errors in the method as municipalities can be big, so the number of workers is probably overestimated.

6.3 Logit Model

The logit model performed the best in terms of accuracy, with only a 6.1% error. But also, in this model some assumptions are made. First, the model only considers two modes: bus and car. Other modes like the train are not considered. The estimation of 29% of people travelling from Assen Kloosterveen to Groningen by bus might be an over estimation because of this.

The second assumption is the linearity of the utility functions. While this is a standard form to use it might not fully capture the behaviour of passengers. The beta parameters also came with some assumptions. They are based on a national data set from 2023, this does not reflect the corridor specific circumstances in 2011 the best. Also, the calculation for the modal constant was based on sparse OViN data making it potentially not representative.

Furthermore, the assumption is made for variables like access and in-vehicle time are static, representing average conditions. Access time can vary on how close a person lives to the bus stop and in vehicle times can be influenced by congestion, the fixed values do not reflect these dynamics. The last assumption made is the total corridor demand. This was based on TomTom data, which only reflects car movements. So public transport or active travel are not considered. TomTom data was chosen because the OViN data was not representative of the corridor. Only one trip was measured for the wanted corridor. Given that only car movement is considered the total demand is probably underestimated. Also, the TomTom data was from 2024 and not 2011 as that dataset did not exist anymore. This may also cause an error in the output.

When taking these assumptions into account and looking into the sensitivity analysis that is performed on the logit model. The assumption that has the biggest impact on the predicted ridership is the in-vehicle time of the bus. This is consistent with what is known about travel behaviour, travel time has often the biggest influence on the number of passengers in public transport. The model was also sensitive to waiting time and access time, reinforcing the importance of service frequency and the coverage of bus stops for BRT planning.

All the parameters influencing the costs of the car have a positive effect on the estimated ridership for the bus. These parameters are the in-vehicle time for the car and the car running costs. When the car usage gets more expensive more people will take the bus. Another parameter that influences the estimated ridership in a positive way is the total demand of the corridor. Although the modal split remains constant, an increase in total demand will still result in higher bus ridership.

When considering the assumptions and looking at the sensitivity analysis the predicted number of passengers is probably an overestimation of the actual number of passengers. This is in line with the validation value. The predicted value was also higher than the actual value. Considering the assumptions, this is correct.

6.4 Regression model

The regression model has an accuracy error of 11.9%. This accuracy is good especially given its low complexity and minimal data needed. However, the model does not take socio-demographic factors into account just as other competing modes. Instead, it estimates ridership based on the system's design: number of bus stops, headway, capacity and fares. This method assumes all other influences as constant.

Another limitation of this method is the intercept value in the model. This constant reflects the average conditions of several BRT systems all over the world. Since the Assen-Groningen corridor is not fully BRT but rather a regional bus route, the constant can result in a higher estimated ridership.

When looking at the sensitivity analysis, the model is highly sensitive to the value of the intercept value. Since the intercept value is derived from international BRT systems rather than Dutch contexts, it may significantly influence the predicted ridership and lead to potential over- or underestimation.

As the model is based on data from full BRT systems, an overestimation of the number of passengers was expected, as BRT should have higher passenger numbers than regional buses. However, in the validation of the model the predicted value was lower than the actual value. This stresses the importance of further research about the application of the regression models in Dutch context.

6.5 GRAVITY MODEL

With a deviation of 8.8% in terms of accuracy the gravity model scores well. But as with all other methods the gravity model also comes with some assumptions. First one is the deterrence function. The form of the function is from literature but the choice among the three possible forms is subjective, and context dependent. Within the function the beta coefficient is a value from transport research from Norway and not from the Netherlands which can cause a slight deviation from the actual numbers.

The second assumption has to do with the reference demand to scale the gravity model. This reference demand is also highly related to the comparative case study. As those predictions are not accurate tables are used to align the frequencies and travel times to make sure a more accurate prediction will be made. The tables that are used that are based on expert knowledge and not based on literature. Although these tables are commonly used in transport modelling it is not known how accurate these tables are. This can lead to a bias in the number of passengers or an oversimplification.

When taking these assumptions into account and look at the sensitivity analysis, the assumption of the demand for the reference corridor has the largest influence on the predicted number of passengers. The assumption for the beta does not have a big influence on the prediction.

6.6 COMPARISON

While the comparison across accuracy, data availability, time efficiency and computational demands offered a structured assessment, some elements are naturally qualitative and not quantitative, and therefore difficult to compare. For example, the quality of data or the complexity of a model is highly context dependent and can vary with the experience and knowledge of a person and the tools available.

The Multi-Criteria Decision Analysis (MCDA) revealed that the regression model scored the highest overall score. The was mainly due to the low complexity, minimal data requirements and ease of implementation of the regression model. However, this outcome should not be interpreted as a one-size-fit-all recommendation. In a context where accuracy is more important the logit model may be more suitable despite the greater resource demands.

The gravity model gives a solid performance in terms of accuracy but as the comparative case study also needs to be used the time-efficiency is lacking. The comparative case study, although it is intuitive, it is not recommended unless an exceptionally close match can be found, and even then, its accuracy is questionable.

In summary, no method is universally superior. The choice will depend on available data, the desired level of accuracy, and the purpose of the ridership prediction. To make an accurate prediction the methods of the logit, regression and gravity model can be combined to give an even more accurate prediction.

7 CONCLUSION

To conclude this research the answers to each sub research question is given below. And at the end the answer to the main research question is given.

7.1 WHICH METHODS ARE THERE TO PREDICT RIDERSHIP AND HOW DO THEY WORK?

From the literature study different methods came forward to predict ridership. These methods are a discrete choice logit model, regression model, the gravity model and the four-step model. Since Keypoint does not have access to a four-step model and developing one is too complex and time consuming this method was excluded in this study, but there were still 4 methods researched. The fourth one being a comparative case study.

In the comparative case study, a similar corridor was found and the number of passengers on the similar corridor served as a prediction for the corridor in question. The logit model uses the modal share to predict demand. The regression model uses several relevant parameters and with linear regression the demand is calculated per kilometre. At last, the gravity model uses population, number of jobs and the costs to go from A to B.

7.2 HOW ACCURATE IS EACH METHOD IN PREDICTING THE POTENTIAL RIDERSHIP WHEN COMPARING IT TO REAL WORLD DATA?

The accuracy of each method can be seen in Table 5. From this table can be concluded that the logit model with a deviation of 6.1% is the most accurate. The comparative case study is the least accurate of the four methods with a deviation of 56.2%. The Gravity model has a deviation of 8.8% and the regression model has a deviation of 11.9%.

Deviations up to 20% are generally considered acceptable for a prediction. Accordingly, all methods except the comparative case study are of sufficient accuracy for a practical application.

This gives that the logit model is the most accurate method to predict ridership. The gravity model is the second best in term of accuracy and the regression model third. All three model have errors lower than 20% deviation range and therefore are considered to be accurate methods to predict ridership. The comparative case study does not fall within the 20% deviation range and therefore is not a good method to predict ridership in terms of accuracy.

7.3 HOW DO THESE METHODS DIFFER IN TERMS OF TIME INVESTMENT, REQUIRED DATA AND COMPLEXITY?

To answer this question three sub questions have been set up, one about the time investment one about the required data and the other about the computational tools required.

7.3.1 How much time is needed to complete each method?

The comparative case study was a relatively time-intensive method. Particularly the set-up of the python script took a long time, about 1.5 weeks. The running time of the script was also relatively high when compared to the other methods, the script took about 2 to 3 hours to run. And at last, the interpretation of the results took also longer since the output of the script only gave a similar corridor, not the predicted ridership, this needed to be asked to an external stakeholder.

The formula for logit model was hard to understand at first, so this took some time. But when the materials were understood the making of the script, the running of the script and the interpretation were relatively fast. Once the material was understood the script could be made within a day, it took about two seconds to run the script, and the outputs of the script were the predicted ridership numbers.

The set-up time for the regression model is considered low, as the regression coefficients were directly taken from an existing paper. If one were to derive these coefficients themselves this would increase the set-up time significantly. The computation time was about 5 seconds, and the script gave the predicted ridership numbers so the interpretation time of the results was also low.

At last, the formula for the gravity model came in all sorts of shapes, so to get the right formula for the gravity model and the cost function took some time. The computation and interpretation time were both low as the gravity model script itself had an output in about 20 seconds and the output was the predicted ridership. But the comparative case study was needed to perform the gravity model, which causes an increase in the time investment needed to perform the gravity model.

Overall, the regression model took the least amount of time, and the gravity model the most time. The comparative case study took, just as the gravity model quite some time, but as the gravity model still needed to adjust the outcome of the comparative case study, the comparative case study ranks third in terms of time investment. The logit model ranks second in terms of time investment.

7.3.2 What data is needed for all these methods and how will this data be collected?

Each method required different kinds of data and data sets. The comparative case study required the population per district in the Netherlands, the job availability per municipality in the Netherlands and the existing bus routes. The data sets for the population and jobs were gathered from CBS and the existing bus routes from OpenOV.

For the logit model the modal split of the corridor pre intervention was needed, this was gathered through OViN data. The access time and average car occupancy was gathered from literature and the waiting time, in-vehicle time and car running costs were gathered through internet sources. The last data needed for the logit model was the total demand of the corridor and this was gathered through TomTom data.

For the regression model, regression coefficients were required. These coefficients were gathered through literature, just like the intercept value. The other data that was needed were the values for the variables, number of stops, peak headway, vehicle capacity and the average fare per trip. This data was gathered through internet sources.

At last, the data needed for the gravity model. For the gravity model, the population of Assen and s' Hertogenbosch was needed and the job availability from Groningen and Tilburg. This data was gathered through CBS. The number of passengers on the bus line 140 between Tilburg and 's Hertogenbosch was needed; this was gathered through an external stakeholder from the province of Noord-Brabant. At last, the travel cost data was needed. The in-vehicle time was gathered through Google Maps, and the beta value was gathered through literature.

When comparing the data based on accessibility, quality and preparation efforts, it can be concluded that the regression scores the best. It has a high accessibility and quality of the data and low efforts needed to prepare the data. For the logit model, the accessibility of the data is medium as the TomTom data is restricted. The quality of the data is medium to high as the OViN data was scare. The preparation of the data for the logit model was low to medium. The comparative case study data's accessibility is high, but the quality of the data is medium as the results also imply routes that do not exist. The preparation needed is also medium for the comparative case study. This is the same for the gravity model, as it uses the comparative case study but also has one additional data source, which is the beta in the deterrence function.

All in all, on the part of data the regression scores the best. After that the comparative case study and the gravity model and at last the logit model.

7.3.3 What are the computational or analytical tools that are required for each method?

For all the methods Python was used to compute the methods. To retrieve data Excel and QGIS were also used in all methods except during the regression model. At last, the TomTom platform was used to retrieve data for the logit model.

The difficulty level of all the methods differed. The needed Python skills were the highest for the comparative case study and the lowest for the regression model. The gravity model and logit model were of a medium level of difficulty in terms of computational skills.

The accessibility of the tools needed to compute each method were high. Python and QGIS are free to use. For Excel was a licence needed and for the TomTom tool an account was needed. Excel can be bought by everyone but for the TomTom account only restricted people can have access.

Overall, the regression model was easiest to use and required the least number of tools and skills. The logit model required the most tools, but all tools needed were easy to use. The gravity model and comparative case study both required the same tools. Only for the comparative case study more skills were needed to make the Python script. Therefore, the comparative case study ranks last in terms of computational skills, the logit model third, the gravity model second and the best one in terms of computational skills needed is the regression model.

7.4 Main question: Which method provides the best balance between accuracy and time efficiency for the prediction of potential ridership for new BRT corridors?

Based on a MCDA involving accuracy, time-efficiency, data availability and required computational skills, the regression model offers the best overall balance. With and accuracy deviation of 11.9%, high data accessibility, and low implementation complexity, it is well suited for early-stage passenger predictions or within a context of limited resources for a consultancy firm.

When accuracy is the top priority and sufficient data is available, the logit model is recommended, as it achieves the highest accuracy with a deviation of only 6.1%. The gravity model offers a middle ground, balancing moderate accuracy, a deviation of 8.8%, with reasonable implementation efforts but it is highly reliable on the comparative case study which makes it a less good option as the comparative case study has a deviation of 56.2% in terms of accuracy and is time intensive. Therefore, the comparative case study should not be considered for predicting ridership only when a truly comparable corridor is found.

8 RECOMMENDATIONS

To improve future outcomes for predicting ridership numbers some improvements can be made. First of all, it would be good to test the models on more corridors within the Netherlands. This would improve the credibility of the methods and validate the consistency of accuracy across different contexts.

Given that the logit, regression and gravity model all fall within an acceptable accuracy range, future applications could combine the predictions of the three methods and therefore provide a range of predictions rather than a single value. This will reduce the impact of individual model assumptions.

The accuracy is highly depended on the quality and availability of the data. In future studies a standardized data source of good quality should be used. The ODiN database of 2023 is already of better quality than the OViN database of 2011. It would also be good to find another way to estimate the total demand of the corridor, not using TomTom data. This will make the use of the logit model more accessible.

Another point about the data used are the intercept values. These values absorb all other traffic circumstances that are not captured in the rest of the formula. The intercept values, together with other parameters of the statistical models, should be calibrated with Dutch data.

When the models are calibrated according to the Assen – Groningen corridor demand that was actually observed, the following intercept values should be used.

For the logit model a new modal constant for the bus was calculated, at first the model constant had a value of 2.8665 and the new value is 2.7842. The modal constant of the car will remain the same with a value of 1.8008. For the regression model the intercept value should be changed to 8.085. At last, in the gravity model, the beta value used in the cost function should be changed to 0.0137, which is similar to the value assumed taken from a Norwegian study (0.016). When using these constants, the real number of passengers will be the outcome of the model.

9 BIBLIOGRAPHY

Ahmed, B. (2012). *The Traditional Four Steps Transportation Modelling Using Simplified Transport Network: A Case Study of Dhaka City, Bangladesh*. www.setscholars.org

Asmael, N., & Waheed, M. (2018). Demand estimation of bus as a public transport based on gravity model. *MATEC Web of Conferences*, *162*. https://doi.org/10.1051/matecconf/201816201038

CBS. (2012a). *PDOK viewer*. https://app.pdok.nl/viewer

CBS. (2012b, March). *Benizne prijzen naar record hoogte*.

CBS, & LISA. (2023). *Banen 2023*.

Currie, G., & Delbosc, A. (2013). Exploring Comparative Ridership Drivers of Bus Rapid Transit and Light Rail Transit Routes. *Journal of Public Transportation*.

de Wit, D. (2023). *A Framework for the proposal of Bus Rapid Transit in the Netherlands* [Bsc Thesis]. University of Twente.

Espinoza, Daquilema. (2025). *OPTIMAL PRICING AND DESIGN OF A BUS CORRIDOR: ANALYZING THE SOCIAL WELFARE* [Msc. Thesis]. University of Twente.

Google. (2025). *Travel time Assen Groningen car*. https://www.google.nl/maps/dir/''/Groningen+Hoofdstation,+Stationsplein,+Groningen/@52.9954447,6.496907,15z/data=!4m14!4m13!1m5!1m1!1s0x47c83ad840b9b765:0x4e147d7aa3e1609d!2m2!1d6.50788!2d52.99875!1m5!1m1!1s0x47c9cd534ba57e77:0xf04c895d20845efe!2m2!1d6.5640642!2d53.2110188!3e0?entry=ttu&g_ep=EgoyMDI1MDUyMS4wIKXMDSoASAFQAw%3D%3D

Hensher, D. A., & Golob, T. F. (2008). Bus rapid transit systems: A comparative assessment. *Transportation*, *35*(4), 501–518. https://doi.org/10.1007/s11116-008-9163-y

Hess, D. B., Yoh, A. C., & Taylor, B. D. (2005). light rail lite or cost effective improvements to bus service evaluating costs of implementing bus. *Journal of the Transportation Research Board, No. 1927*.

Istock. (n.d.). *IStock picture public transport*. Retrieved June 13, 2025, from https://www.istockphoto.com/nl/foto/jongeman-invoeren-van-bus-gm884560268-245904352

ITDP. (2018). *The BRT Planning Guide*. https://github.com/ITDP/the-online-brt-planning-guide

ITDP. (2024). *THE BRT STANDARD 2024 Edition*. https://itdp.org/publication/the-brt-standard/

Karunakaran, R., Yuen, C. W., Chuah, J. H., Zainol, R., & Goh, B. H. (2023). Correlation analysis and regression of transit ridership by means of public buses – A preliminary data analysis. *IOP Conference Series: Materials Science and Engineering*, *1289*(1), 012044. https://doi.org/10.1088/1757-899x/1289/1/012044

Keolis. (n.d.). *Hoeveel reizigers kunnen er in de bussen die Keolis inzet bij treinvervangend vervoer – Klantenservice | Keolis*. Retrieved April 28, 2025, from https://klantenservice.keolis.nl/hc/nl/articles/360000273445-Hoeveel-reizigers-kunnen-er-in-de-bussen-die-Keolis-inzet-bij-treinvervangend-vervoer

Kouwenhoven, M., Muller, J., Thoen, S., Willigers, J., & de Jong, G. (2023). *Values of Time, Reliability and Comfort in the Netherlands 2022 New values for passenger travel and freight transport*.

Matsumoto, N. (2005, March). *Promoting Bus Rapid Transit (BRT)*. https://www.iges.or.jp/en/pub/promoting-bus-rapid-transit-brt/en

Modeshift. (2023). *Why Is There An Increasing Demand For Public Transportation?* . Picture. https://www.modeshift.com/why-is-there-an-increasing-demand-for-public-transportation/

Moovit. (2024). *Richting: Assen Kloosterveen*. https://appassets.mvtdev.com/map/39/l/101/231508.pdf

OpenOV. (2025). *gtfs files bus connections*. https://gtfs.openov.nl/gtfs-rt/

Ortuzar, J., & Willumsen, L. G. (2011). *MODELLING TRANSPORT* (4th edition).

Pahdma M. (2021). *A Comprehensive Introduction to Evaluating Regression Models Introduction Role of Loss Functions in Model Evaluation*.

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., & White, P. (2006a). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy*, *13*(4), 295–306. https://doi.org/10.1016/j.tranpol.2005.12.004

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., & White, P. (2006b). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy*, *13*(4), 295–306. https://doi.org/10.1016/j.tranpol.2005.12.004

Princeton university. (2008). *The Gravity Model*. http://www.webs1.uidaho.edu/niatt_labmanual/Chapters/traveldemandforecasting/theoryandconcepts/GravityModel.htm[10/6/200812:35:45PM]

Qbuzz. (n.d.). *Dienstregeling lijn 309*. Retrieved April 28, 2025, from https://gd.qbuzz.nl/lijn/QBUZZ/g309/1/

Spijker, L. (2024). *Developing a model for finding promising Bus Rapid Transit corridors between city zones* [Bsc Thesis]. University of Twente.

Tirachini, A., Hensher, D. A., & Rose, J. M. (2014). Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding. *Transportation Research Part B: Methodological*, *61*, 33–54. https://doi.org/10.1016/j.trb.2014.01.003

UITP. (2024, November 5). *50 years of BRT: 5 advantages of BRT buses | UITP*. https://www.uitp.org/news/5-advantages-of-brt-buses/

van der Waerden, P., van der Waerden, J., & Burger, M. (2024). Exploring the role of public transport service and walking route related factors to identify maximum walking distances to bus stops in the Netherlands. *Journal of Public Transportation*, *26*. https://doi.org/10.1016/j.jpubtr.2024.100096

Witte, J.-J., & Kansen, M. (2020). *KiM | Kansen voor Bus Rapid Transit in Nederland*. https://www.kimnet.nl/publicaties/rapporten/2020/07/16/bus-rapid-transit-kansrijk-concept-voor-hoogwaardig-busvervoer

10 APPENDIX

10.1 APPENDIX A – PYTHON SCRIPTS

10.1.1 Comparative case study

10.1.1.1 Calculating distances between each district

```python
import matplotlib.pyplot as plt
import geopandas as gpd
import pandas as pd
from shapely.geometry import Point
import contextily as ctx
import pickle

# Reading a GeoPackage file
districts = gpd.read_file("Data/BuurtenNL.gpkg")

subdistricts = districts[["buurtcode","wijkcode","gemeentenaam","aantalInwoners","geometry"]]

# Fix population outlies
subdistricts['aantalInwoners'] = subdistricts['aantalInwoners'].replace(-99997,1)

# calculate population centers for districs
district_centers = []

for wijkcode in districts['wijkcode']:
    district_parts = subdistricts[subdistricts['wijkcode'] == wijkcode]
    total_pop = district_parts['aantalInwoners'].sum()

    if total_pop == 0:
        continue

    weighted_centroids = []
    for geom, pop in zip(district_parts.geometry, district_parts['aantalInwoners']):
        centroid = geom.centroid
        weighted = centroid.x * pop, centroid.y * pop
        weighted_centroids.append(weighted)

    total_x = sum(x for x, _ in weighted_centroids)
    total_y = sum(y for _, y in weighted_centroids)

    center = Point(total_x / total_pop, total_y / total_pop)

    district_centers.append({
        'wijkcode': wijkcode,
        'population_center': center,
        'aantalInwoners': total_pop
    })

```

```
#create DataFrame from results
df_centers = pd.DataFrame(district_centers)

# Merge with districts GeoDataFrame
districts = districts.merge(df_centers, on='wijkcode')
districts['centroid'] = districts.centroid

# plot
fig, ax = plt.subplots(figsize=(12, 12))
districts.plot(ax=ax, edgecolor='black', facecolor='none', linewidth=1, alpha=0.05)
#districts.set_geometry('centroid').plot(ax=ax, color='red', label='Centroid', markersize=5)
districts.set_geometry('population_center').plot(ax=ax, color='green', label='Population center', markersize=5)
ctx.add_basemap(ax, crs=districts.crs.to_string(), source=ctx.providers.OpenStreetMap.Mapnik)
plt.legend()
plt.title('Districts of the Netherlands: Population Centers')
ax.set_axis_off()

#serialize GeoDataFrame
with open('df_centers.pkl','wb') as f:
    pickle.dump(df_centers,f)


```

10.1.1.2 Determining similar corridor

```
import pandas as pd
import geopandas as gpd
import pickle
from gtfs_kit import read_feed

# Load data
districts = gpd.read_file("Data/Inwoners_per_wijk.gpkg")
subdistricts = districts[["wijkcode", "gemeentenaam", "aantalInwoners", "geometry"]]

jobs = pd.read_excel("Banen2012.xlsx")

with open('df_distances', 'rb') as f:
    df_distances = pickle.load(f)

feed = read_feed("Data/gtfs-nl.zip.zip", dist_units="km")

# Reference connection
ref_origin = 'wk010608'
ref_dest = 'wk001400'

# Clean wijkcodes
subdistricts['wijkcode'] = subdistricts['wijkcode'].str.strip().str.upper()
ref_origin = ref_origin.strip().upper()
ref_dest = ref_dest.strip().upper()

# Get reference distance
mask = (
    ((df_distances['Origin'] == ref_origin) & (df_distances['Destination'] == ref_dest)) |
    ((df_distances['Origin'] == ref_dest) & (df_distances['Destination'] == ref_origin))
)
ref_distance = df_distances[mask]['Distance'].values[0] if not df_distances[mask].empty else 1

# Get reference origin population
ref_origin_pop = subdistricts[subdistricts['wijkcode'] == ref_origin]['aantalInwoners'].values[0] if not subdistricts[subdistricts['wijkcode'] == ref_origin].empty else 1

# Get reference destination jobs
ref_dest_gemeente = subdistricts[subdistricts['wijkcode'] == ref_dest]['gemeentenaam'].values[0]
ref_dest_workers = jobs[jobs['Gemeente'] == ref_dest_gemeente]['aantalBanen'].values[0] if not jobs[jobs['Gemeente'] == ref_dest_gemeente].empty else 1
```

```python
# Add population info for Origin and Destination
df_distances = df_distances.merge(subdistricts[['wijkcode', 'aantalInwoners', 'gemeentenaam']], left_on='Origin', right_on='wijkcode', how='left')
df_distances.rename(columns={'aantalInwoners': 'origin_pop', 'gemeentenaam': 'origin_gemeente'}, inplace=True)
df_distances.drop('wijkcode', axis=1, inplace=True)
df_distances = df_distances.merge(subdistricts[['wijkcode', 'aantalInwoners', 'gemeentenaam']], left_on='Destination', right_on='wijkcode', how='left')
df_distances.rename(columns={'aantalInwoners': 'dest_pop', 'gemeentenaam': 'dest_gemeente'}, inplace=True)
df_distances.drop('wijkcode', axis=1, inplace=True)
# Add job data to origin and destination
df_distances = df_distances.merge(jobs, left_on='origin_gemeente', right_on='Gemeente', how='left')
df_distances.rename(columns={'aantalBanen': 'origin_jobs'}, inplace=True)
df_distances.drop('Gemeente', axis=1, inplace=True)
df_distances = df_distances.merge(jobs, left_on='dest_gemeente', right_on='Gemeente', how='left')
df_distances.rename(columns={'aantalBanen': 'dest_jobs'}, inplace=True)
df_distances.drop('Gemeente', axis=1, inplace=True)
# Calculate similarity score
def calc_score(row):
    return (
        abs(row['Distance'] - ref_distance) / ref_distance +
        abs(row['origin_pop'] - ref_origin_pop) / ref_origin_pop +
        abs(row['origin_jobs'] - ref_dest_workers) / ref_dest_workers +
        abs(row['dest_pop'] - ref_origin_pop) / ref_origin_pop +
        abs(row['dest_jobs'] - ref_dest_workers) / ref_dest_workers
    )
df_distances['similarity_score'] = df_distances.apply(calc_score, axis=1)
# Remove connections with same gemeente at both ends
df_distances = df_distances[df_distances['origin_gemeente'] != df_distances['dest_gemeente']]
# Exclude connections involving the reference gemeenten (e.g., Assen or Groningen)
excluded_gemeenten = {ref_dest_gemeente}
ref_origin_gemeente = subdistricts[subdistricts['wijkcode'] == ref_origin]['gemeentenaam'].values[0]
excluded_gemeenten.add(ref_origin_gemeente)
df_distances = df_distances[
    (~df_distances['origin_gemeente'].isin(excluded_gemeenten)) &
    (~df_distances['dest_gemeente'].isin(excluded_gemeenten))
]
#add existing bus connections
# Match GTFS stops to wijk geometries
stops = feed.stops[['stop_id', 'stop_lat', 'stop_lon']]
stops_gdf = gpd.GeoDataFrame(
    stops,
    geometry=gpd.points_from_xy(stops.stop_lon, stops.stop_lat),
    crs='EPSG:4326'
)
# Reproject both to same CRS for spatial join
stops_gdf = stops_gdf.to_crs(subdistricts.crs)
stops_with_wijk = gpd.sjoin(stops_gdf, subdistricts[['wijkcode', 'geometry']], how='left', predicate='within')
# Build mapping from stop_id to wijkcode
stop_to_wijk = stops_with_wijk.set_index('stop_id')['wijkcode'].to_dict()
# Build set of direct wijk-to-wijk connections
```

```
direct_wijk_pairs = set()

trip_stop_times = feed.stop_times.merge(feed.trips[['trip_id']], on='trip_id')
trip_stop_times = trip_stop_times.sort_values(['trip_id', 'stop_sequence'])

for trip_id, group in trip_stop_times.groupby('trip_id'):
    wijk_sequence = group['stop_id'].map(stop_to_wijk).dropna().tolist()
    seen = set()
    for i, origin in enumerate(wijk_sequence):
        for dest in wijk_sequence[i+1:]:
            if origin != dest and (origin, dest) not in seen:
                direct_wijk_pairs.add((origin, dest))
                direct_wijk_pairs.add((dest, origin))
                seen.add((origin, dest))

# Filter to keep only pairs with a direct bus connection
df_direct_bus = df_distances[
    df_distances.apply(lambda row: (row['Origin'], row['Destination']) in direct_wijk_pairs, axis=1)
]

# Sort by similarity and select top 10 with unique origin *municipalities*
sorted_distances = df_direct_bus.sort_values('similarity_score')

unique_origins = set()
top_matches = []

for _, row in sorted_distances.iterrows():
    if row['origin_gemeente'] not in unique_origins:
        top_matches.append(row)
        unique_origins.add(row['origin_gemeente'])
    if len(top_matches) == 10:
        break

df_top_10 = pd.DataFrame(top_matches)
print(df_top_10[['origin_gemeente', 'dest_gemeente', 'Origin', 'Destination', 'Distance', 'origin_pop', 'origin_jobs', 'dest_pop', 'dest_jobs', 'similarity_score']])

#serialize GeoDataFrame
with open('df_top_10','wb') as f:
    pickle.dump(df_top_10,f)

# Export to Excel
top_10_df = df_top_10.head(10)

for _, row in sorted_distances.iterrows():
    if row['origin_gemeente'] not in unique_origins:
        top_matches.append(row)
        unique_origins.add(row['origin_gemeente'])


df_all_matches = pd.DataFrame(top_matches)
print(df_all_matches[['origin_gemeente', 'dest_gemeente', 'Origin', 'Destination', 'Distance', 'origin_pop', 'origin_jobs', 'dest_pop', 'dest_jobs', 'similarity_score']])

print("Number of similar direct bus connections found:", len(df_all_matches))

#serialize GeoDataFrame
with open('df_all_matches','wb') as f:
    pickle.dump(df_all_matches,f)

```

10.1.2 Logit model

```
import pandas as pd
import numpy as np
import math
import matplotlib.pyplot as plt

# --- Constants and Parameters ---

# Modal constants (original)
alpha_bus_orig = 3
alpha_car_orig = 1.79

# Modal shares
q_bus = 0.08    #Modal share for corridor
q_car = 0.92
Q_bus = 0.07    #Normalized modal share
Q_car = 0.93

# Beta values
beta_access = -8.68     # β_a
beta_waiting = -10.66   # β_h
beta_in_vehicle_bus = -9.78  # β_vb
beta_in_vehicle_car = -8.24  # β_va
beta_costs = -0.76      # β_c

# Travel and cost variables
access_time = 4.2 / 60  # hours
waiting_time = 7.5 / 60  # hours
in_vehicle_time_bus = 37 / 60 # hours
in_vehicle_time_car = 24 / 60 # hours
car_km = 30     # distance of trip in km
cost_per_km = 0.1432
car_running_cost = car_km * cost_per_km
car_occupancy = 1.2

# Total demand in OD pair
total_demand = 8810

# --- Calculate adjusted modal constants ---
alpha_bus = alpha_bus_orig - math.log(q_bus / Q_bus)
alpha_car = alpha_car_orig - math.log(q_car / Q_car)

print("Model constat:")
print(f"  Bus constant: {alpha_bus:.4f}")
print(f"  Car constant: {alpha_car:.4f}")

# --- Utility Equations ---
u_bus = (alpha_bus +
         beta_access * (access_time) +
         beta_waiting * waiting_time +
         beta_in_vehicle_bus * (in_vehicle_time_bus))

u_car = (alpha_car +
         beta_in_vehicle_car * (in_vehicle_time_car) +
         (beta_costs * car_running_cost ) / car_occupancy)

# --- Logit Model: Calculate probabilities ---
exp_u_bus = math.exp(u_bus)
exp_u_car = math.exp(u_car)
denominator = exp_u_bus + exp_u_car

p_bus = exp_u_bus / denominator
p_car = exp_u_car / denominator

# --- Demand per mode ---
demand_bus = total_demand * p_bus
demand_car = total_demand * p_car

# --- Output ---
```

```
print("Utilities:")
print(f"  Bus utility: {u_bus:.4f}")
print(f"  Car utility: {u_car:.4f}")
print("\nMode probabilities:")
print(f"  Bus probability: {p_bus:.4f}")
print(f"  Car probability: {p_car:.4f}")
print("\nDemand estimation:")
print(f"  Bus demand: {demand_bus:.2f} passengers")
print(f"  Car demand: {demand_car:.2f} passengers")

```

10.1.3 Regression model

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Define model coefficients
beta_0 = 7.9209
beta_stations = 0.0098
beta_headway = -0.1681
beta_capacity = 0.0052
beta_fare = -0.2577

# Base values
base_stations = 15  # stops
base_headway = 15  # minutes
base_capacity = 80  # persons
base_fare_eur = 4.90  # euro
eur_to_usd = 1.1
base_fare_usd = base_fare_eur * eur_to_usd

corridor_km_with_stops = 11

# Base scenario calculation
log_base = (
    beta_0
    + beta_stations * base_stations
    + beta_headway * base_headway
    + beta_capacity * base_capacity
    + beta_fare * base_fare_usd
)
base_ridership = np.exp(log_base) * corridor_km_with_stops

print(f"Yi: {log_base:.4f}")
print(f"Ridership: {base_ridership:.4f}")

```

10.1.4 Gravity model

```
import pandas as pd
import geopandas as gpd
import numpy as np
import matplotlib.pyplot as plt

# Reading a GeoPackage file
districts = gpd.read_file("Data/Data_per_wijk.gpkg")

subdistricts =
districts[["wijkcode","wijknaam","gemeentenaam","aantalInwoners","gemiddeldeWoningwaarde","gemid
deldInkomenPerInkomensontvanger","percentageNietActieven","geometry"]]

wijkcodes_to_filter = ['WK001400','WK010608','WK079601','WK085507']

filtered_subdistrics = subdistricts[subdistricts['wijkcode'].isin(wijkcodes_to_filter)]

# reading excel S
jobs = pd.read_excel("Banen2012.xlsx")
```

```python

municipalities = ['Assen', 'Groningen', 'Tilburg', 's-Hertogenbosch']

filtered_jobs = jobs[jobs['Gemeente'].isin(municipalities)]

# Input data
# Population
population_assen = filtered_subdistrics.loc[filtered_subdistrics['wijkcode'] == 'WK010608', 'aantalInwoners'].values[0]
population_groningen = filtered_subdistrics.loc[filtered_subdistrics['wijkcode'] == 'WK001400', 'aantalInwoners'].values[0]
population_tilburg = filtered_subdistrics.loc[filtered_subdistrics['wijkcode'] == 'WK085507', 'aantalInwoners'].values[0]
population_denbosch = filtered_subdistrics.loc[filtered_subdistrics['wijkcode'] == 'WK079601', 'aantalInwoners'].values[0]

# Jobs
jobs_assen = filtered_jobs.loc[filtered_jobs['Gemeente'] == 'Assen', 'aantalBanen'].values[0]
jobs_groningen = filtered_jobs.loc[filtered_jobs['Gemeente'] == 'Groningen', 'aantalBanen'].values[0]
jobs_tilburg = filtered_jobs.loc[filtered_jobs['Gemeente'] == 'Tilburg', 'aantalBanen'].values[0]
jobs_denbosch = filtered_jobs.loc[filtered_jobs['Gemeente'] == 's-Hertogenbosch', 'aantalBanen'].values[0]

# Travel costs
beta = 0.016
travel_costDBT = 50 # minuten

# Deterrence function
def deterrenceDBT(c, beta):
    return np.exp(-beta * c)

# Calculate demand Tilburg - Den Bosch
alfa_DBT = 1
cost_factorDBT = deterrenceDBT(travel_costDBT, beta)
attractionDBT = alfa_DBT * population_denbosch * jobs_tilburg * cost_factorDBT

# Acctual demand Tilburg - Den Bosch
Acctual_demand_tilburg_denbosch = 1735 # Change value with the right one

# Calculate new alfa
alfa_2 = Acctual_demand_tilburg_denbosch / attractionDBT

# Travel costs
beta = 0.016
travel_costAG = 40 # minuten

# Deterrence function
def deterrenceAG(c, beta):
    return np.exp(-beta * c)

cost_factorAG = deterrenceAG(travel_costAG, beta)

#Calculate Assen - Groningen
demandAG = alfa_2 * population_assen * jobs_groningen * cost_factorAG

print(f"Demand Assen - Groningen:{demandAG:4f}")

```

10.2 Appendix B – Tables gravity model

In Figure 9 the elasticities for the travel time are shown. The tables come from my supervisor at Keypoint and are often used in the mobility world. The travel time between Berkel-Enschot and 's Hertogenbosch is 50 minutes and between Assen Kloosterveen and Groningen it is 37 minutes by bus. This means that the demand of the corridor between Berkel-Enschot and 's Hertogenbosch will be multiplied with 1.32. So now the demand will be 1399 passengers per day.

		rijtijd bus toekomstig										
		5	10	15	20	25	30	40	50	60	70	80
rijtijd bus huidig	5	1,00	0,81	0,67	0,56	0,47	0,40	0,29	0,22	0,17	0,14	0,11
	10	1,23	1,00	0,82	0,69	0,58	0,49	0,36	0,27	0,21	0,17	0,13
	15	1,49	1,21	1,00	0,83	0,70	0,59	0,44	0,33	0,26	0,20	0,16
	20	1,79	1,46	1,20	1,00	0,84	0,71	0,53	0,40	0,31	0,24	0,20
	25	2,13	1,73	1,43	1,19	1,00	0,85	0,63	0,47	0,37	0,29	0,23
	30	2,51	2,04	1,68	1,40	1,18	1,00	0,74	0,56	0,43	0,34	0,27
	40	3,41	2,77	2,29	1,90	1,60	1,36	1,00	0,76	0,59	0,46	0,37
	50	4,51	3,66	3,02	2,51	2,11	1,79	1,32	1,00	0,78	0,61	0,49
	60	5,81	4,72	3,89	3,24	2,72	2,31	1,70	1,29	1,00	0,79	0,63
	70	7,36	5,98	4,93	4,10	3,45	2,93	2,16	1,63	1,27	1,00	0,80
	80	9,20	7,48	6,16	5,13	4,31	3,66	2,69	2,04	1,58	1,25	1,00
	90	11,31	9,20	7,57	6,30	5,30	4,50	3,31	2,51	1,95	1,54	1,23

Figure 9: Travel time elasticities, factor changing demand by change in travel time

As the frequency of the buses also differs another multiplier needs to be done. The interval of the corridor between Berkel-Enschot and s' Hertogenbosch is 30 minutes while the corridor between Assen Kloosterveen and Groningen has an interval of 15 minutes. This will give another multiplier of 1.24. This will give a total demand of 1735 passengers per day.

		Future interval (in min.) :													
		4	5	6	7	7,5	8	9	10	12	15	20	30	60	90
Current interval (in min.):	4	0	-10	-17	-25										
	5	10	0	-8	-15	-18	-22								
	6	17	8	0	-8	-15	-18	-22							
	7	25	15	8	0	-4	-8	-15	-22						
	7,5		18	15	4	0	-4	-12	-18						
	8		22	18	8	4	0	-8	-15	-19					
	9			22	15	12	8	0	-7	-12	-24				
	10				22	18	15	7	0	-5	-14	-24	-50		
	12						19	12	5	0	-9	-20			
	15							24	14	9	0	-12	-24		
	20								24	20	12	0	-14		
	30										24	14	0	-24	-50
	60													0	-25

Figure 10: Frequency elasticities, factor changing demand by changing frequency