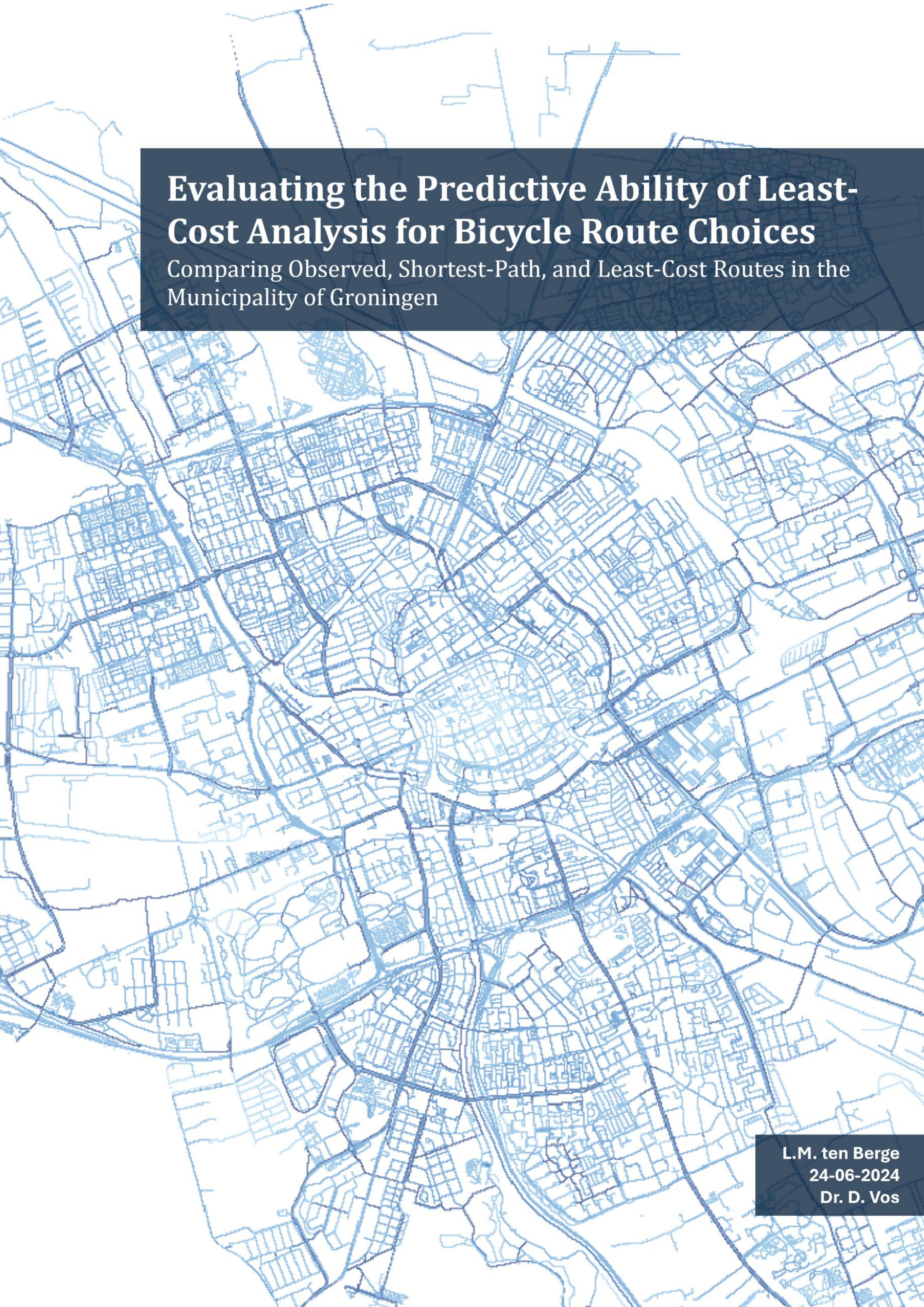
Evaluating the Predictive Ability of Least-Cost Analysis for Bicycle Route Choices

Comparing Observed, Shortest-Path, and Least-Cost Routes in the Municipality of Groningen



L.M. ten Berge
24-06-2024
Dr. D. Vos

Colophon

Bachelor Project Spatial Planning and Design

Title: Evaluating the Predictive Ability of Least-Cost Analysis for Bicycle Route Choices
Subtitle: Comparing Observed, Shortest-Path, and Least-Cost Routes in the Municipality of Groningen
Author: L.M. ten Berge (S4924606)
Contact: l.m.ten.berge@student.rug.nl

University: Rijksuniversiteit Groningenarc
Faculty: Faculty of Spatial Sciences
Supervisor: Dr. D. Vos

Date: 24-06-2024
Version: Final

Abstract

The Netherlands is facing increasing congestion and traffic accidents on cycle paths, necessitating a better understanding of cyclists' behaviour for effective policy and urban design interventions. This study aims to provide insights into cyclists' behaviour through a literature review that identifies factors such as infrastructure and land use that influence route choice. These factors are incorporated into a Weighted Raster Network (WRM) for Groningen, using regression coefficients from a previous study in Enschede. Data from TalkingBikes was used, resulting in 50 GPS trips after a careful selection. Three routes were compared: Observed Route (OBR) based on GPS data, Shortest Path (SHP) representing the minimum possible distance and Least Cost Path (LCP) representing the minimum cost according to the WRM. This study uniquely examines cycling routes using an LCP, unlike previous studies that often compare SHP routes, which do not take into account the factors considered by an LCP. Visual and cost comparisons show that cyclists do not strictly follow LCP or SHP routes, with the OBR showing significant variation. Differences in total costs between routes suggest that context-dependent weights, such as the over-emphasis on segregated cycle path (72.9% of LCP length overlaps with segregated cycle lanes), do not fit well with Groningen context. This study highlights the complexity of accurately predicting cycling behaviour. In order to determine whether an LCP can effectively predict route choice, specific weight coefficients need to be developed for the city of Groningen, and a mixed methods approach should be used to gain deeper insights into cyclists' route choices through interviews.

Key words: *Cycling behaviour, Route choice modelling, Least-cost path analysis, infrastructure factors, land-use factors*

Table of content

List of abbreviations

Abbreviations	Meaning
FSI	Floor space index
LCP	Leas-cost path
MXI	Mixed space index
OBR	Observed route
OD	Origin-Destination
SHP	Shortest path
WRN	Weighted Raster Network

1. Introduction

1.1. Problem context

Cities around the world are trying to promote cycling as a more sustainable mode of transport (Meireles and Ribeiro, 2020). Because cycling has many benefits at a societal and individual level. These include a reduction in carbon emissions, resulting in a cleaner environment and, lower health costs for individuals and society due to increased activity (Handy, Van Wee & Kroesen (2014), Kiviluoto et al. (2022)). Cycling is also known to be the fastest mode of transport for trips of less than five kilometres (Wei & Lovergrove 2013), making it an ideal mode of transport for compact and medium-sized cities (Stevenson et al., 2016).

The Netherlands is a world leader in cycling, with 28% of all journeys made by bicycle in 2022, making it the second most used mode of transport after the car (Centraal bureau voor de Statistiek, 2022). Cycling is deeply integrated into the Dutch road infrastructure and various authorities are working to improve the position of cyclists in traffic by improving cycling facilities (Province of Noord-Brabant (2016), Province of Drenthe (2023), Municipality of Dalfsen (2016)). Despite this, the safety of cyclists is increasingly at risk due to increasing crowding on cycle paths, which is largely attributed to the increase in the use of e-bikes (Wegman & Schepers, 2024). Reports suggest that cyclists now account for 40% of road traffic accidents (NOS, 2024). As a result, safety measures such as helmet recommendations are being implemented. Policy makers are also exploring changes to the urban fabric to improve cyclist safety, recognising the need to make informed decisions about the factors influencing safety and route use to justify the often large investment required. In order for these investments to be useful, it is important that policy makers gain insight into where and how cyclists travel.

1.2. Research gap

Several studies have been carried out in the field of cycling on the behaviour and route choice of cyclists. Research has shown that land use and infrastructure factors can cause cyclists to deviate from the shortest route (Maat, van Wee & Stead (2005), Heinen, Maat, & Van Wee (2011)). Research in this area is conducted in various ways, including surveys, interviews and various quantitative models (Strauss and Miranda-Moreno (2013), Brand et al. (2017), Veenstra, Geurs, Thomas & Van den Hof (2016)). In these quantitative studies, the modelled route is tested against actual bicycle movements (GPS), and it is attempting to predict the likelihood (or attractiveness) of segments. Cyclists are connected to a network (the infrastructure), so vector data (nodes and edges) are often used (Strauss and Miranda-Moreno, 2013). However, there remains no single answer as to how cyclists behave. Often, these studies look at what the shortest route is and based on that, the influence of factors is determined. Lu, Scott, & Dalumpines (2018) looked at cyclists' route choice and concluded that the cycled route is significant different from the shortest route. This study builds on this knowledge, but with a unique approach. By using already known factors from the existing literature and bringing them together to form a Weighted Raster Network (WRN). Raster models are not often used to study bicycle movements. Based on this Raster Network, the extent to which a Least Cost Path (LCP) analysis can predict cycling behaviour will be investigated by comparing the LCP with Shortest Path (SHP) and observed routes (OBR).

1.3. Research aim

This research project aims to gain insight into the efficacy of a raster network based on infrastructure and land use factors that influence cyclists' route choice in predicting actual cycling behaviour. This is achieved by comparing LCP, SHP and observed routes. Consequently, the objective of this study is:

"To assess to what extend it is possible to translate existing infrastructural and land use factors in a weighted raster model and see how similar it is to actual cycling trips".

1.4. Study area

In order to assess whether this approach can be implemented on a larger scale, it is important to first examine it at a local level to ensure quality and reliability. The municipality of Groningen in the Netherlands, shown in Figure 1, was chosen for this study.

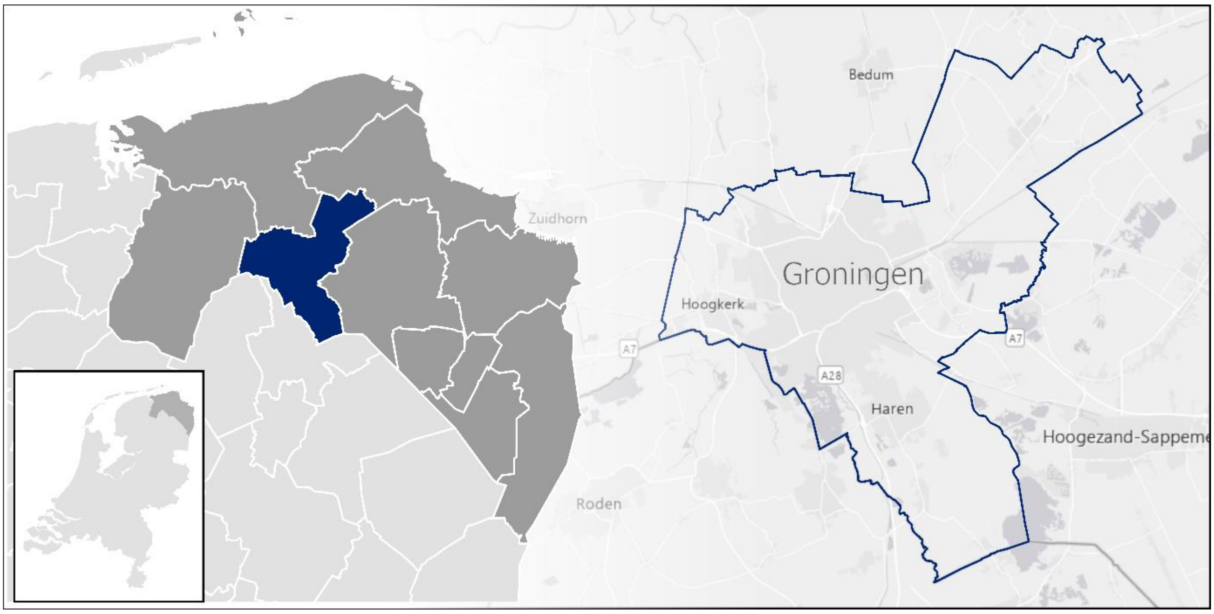


Figure 1: Municipality of Groningen, visualised in blue in the Province of Groningen (by author)

1.5. Research questions

This research examines the infrastructure and land use factors that influence cyclists' route choice. It compares cycling behaviour with LCP and SPH analysis and the observed route on a raster network. To assess this, a weighted grid network is created using factors from existing literature. On this grid network, total costs are calculated for the LCP, SHP and observed routes. The main research question is therefore:

MRQ: *"What is the difference in route choice of cyclists between a least-cost path analysis, the shortest route and the actual route in the municipality of Groningen?"*

The main research question is divided into three sub-questions, which must be answered in sequence in order to answer the main research question. The first sub-question concerns the infrastructural and land use factors that, according to the literature, influence the route choice of cyclists. Based on this, the WRN is created. To obtain the factors that influence route choice, the first sub-question is:

SRQ-1: *"What infrastructure and land use factors can be identified in the literature as having an influence on cyclists' route choice?"*

The infrastructure and land use factors determine the weights for the WRN. This network determines the route and total cost of the LCP. The total costs of the SHP and OBR are calculated to see if there is a difference. Therefore, the second sub-question is:

SRQ-2: *"To what extent do the least-cost path and shortest path analyses correspond to the actual routes chosen by cyclists?"*

Given the exploratory nature of this study, it is necessary to test the usability of LCP analysis to predict route choice. This leads to the third sub-question:

SRQ-3: *"To what extent is it possible to use a least cost path analysis to do predict route choice?"*

1.6. Report outline

This chapter presents the context of the problem, the research gap, the research aim, the study area and the research questions. The subsequent Chapter, Chapter 2 presents the theoretical framework with the existing literature in order to answer the first research question. Chapter 3 then discusses the methodology used to answer the second and third sub-questions. Chapter 4 presents the results of this study and Chapter 5 concludes this study.

2. Theoretical framework

This chapter presents the theories that are relevant to this study. These theories are then brought together in a conceptual model to illustrate how route choice is affected.

2.1 Route choice behaviour

Travelling is an important part of everyday life and is influenced by the mode of transport and the choice of route choice. When travelling, road users evaluate the route alternatives typically choose the one that is most beneficial to them. A number of studies have demonstrated that route choice is influenced by a range of factors, including travel costs, road safety, comfort, travel time, habits, and socio-economic characteristics. (Arslan and Khisty (2005), Prato and Bekhor (2007), Ha, Lee & Ko (2020). From these factors, the most prominent factor is travel time (Bovy & Stern, 1990). There are mainly two dominant groups of factors that are studied, infrastructural factors and land use factors (Koch & Dugundji, 2021) these two groups of factors will be explained in section 2.1.1. and 2.1.2. .

2.1.1. Infrastructural factors

The presence of adequate infrastructure is a significant factor in determining whether cyclists are inclined to take a particular route (Maat, van Wee & Stead, 2005). A number of researchers have found that the attractiveness of a route is significantly influenced by the quality of the cycling infrastructure, as it ensures a good flow and safety (Koch & Dugundji (2021), Mertens, et al. (2016), Winters, Davidson, Kao & Teschke (2011)). This is the case when the bicycle path is separated from other (motorised) traffic (Chen (2016), Li et al. (2012), Ding et al. (2021)). Some researchers even describe this as the most significant factor influencing route choice (Mertens, et al., 2016). Koch & Dugundji 2021 also conducted research in this area and corroborate this assertion. Additionally, they examined the impact of a painted separation between motorised traffic and cyclists on the same road (cycle lane). This also has a positive effect, although it is less pronounced. However, Stinson & Bhat (2003) argue that the extent to which a separated cycle lane affects route choice is less than that of a separate bicycle path. Consequently, the delineation of a bicycle lane on a bicycle path plays a significant role in the route choice of cyclists. Clear delineation and signalling of intersections also contribute to route choice. Schepers et al. (2011) and Wall et al. (2016) concur that marking intersections for cyclists is an effective measure for both navigation and safety of cyclists at an intersection.

Prato, Halldórsdóttir & Nielsen (2018) were able to conclude in their case study in Copenhagen that the presence of traffic lights has a negative effect on cyclists' route choice. Cyclists prefer to try to avoid traffic lights as it takes time (Strauss and Miranda-Moreno, 2013). This effect is underlined by Stinson & Bhat (2003), Koch & Dugundji (2021), Broach, Dill & Gliebe (2012). Khatri, Cherry, Nambisan & Han (2016) also drew this conclusion in their case study in Phoenix, but indicate that traffic lights can be valuable at high traffic intensities and lefthand crossing, as it provides more safety and less time when turning left. The total number of turns also matters on cyclists' route choice (Broach, Dill, & Gliebe (2012). Cyclists are more likely to choose routes with more right turns (due to priority).

The selection of a route is also influenced by the condition of the road surface. A better quality of pavement provides a greater sense of security for cyclists (Gadsby, Tsai & Watkins (2022), Gössling & McRae (2022)). Consequently, cyclists are more likely to select a route with a good surface, which is both safer and requires less energy. Winters, Davidson, Kao & Teschke (2011) reached a similar conclusion in their research using Vancouver as the study area. They found that road maintenance is also of great importance, and that the type of surface and its quality are also significant factors. Other obstacles, such as potholes, which pose significant safety hazards for cyclists, were also identified (Dondi et al., 2011). Such defects in the road surface can lead to accidents by causing cyclists to trip, fall, and lose traction.

In addition, there are two other factors that are less commonly studied, but still have an impact on route choice. First, the presence of street lighting appears to influence route choice (Winters, Davidson, Kao, & Teschke, 2011). This is because street lighting improves the perceived safety of a route when it is dark. During the day, roads with lighting do not affect cyclists' route choice (Uttley, Fotios, & Lovelace, 2020). This effect is often less studied because it is a less significant influence and it is a temporal effect. The final factor that can influence route choice is traffic volume. Li et al., (2012) point out that there is a difference between motorised intensities and cycling intensities. Both have a negative effect on cyclists' route choice. Cyclists are less likely to choose a busy or congested route (Grudgings, Hughes and Hagen-Zanker, 2021). The effect of the presence of motorised traffic is stronger than that of other cyclists.

2.1.2. Land use factors

The environment in which a cyclist cycles has an impact on the route choice of cyclists. There is still much debate in the literature on this topic. For instance, studies by Koch & Dugundji (2021) and Li et al., (2012) conclude that cyclists tend to avoid residential zones. Conversely, Zhao, Ke, Lin & Yu (2020) have concluded that a residential zone has a positive influence on cyclists' route choice. Prato, Halldórsdóttir & Nielsen (2018) specify residential zones even more, distinguishing between high and low density areas. This study indicates that cyclists are more likely to cycle through low density areas than through high density areas. This discrepancy is corroborated by the fact that in high-density areas, one is more likely to encounter conflicts that necessitate waiting. In the same study, Prato, Halldórsdóttir & Nielsen (2018) concluded that cyclists tend to avoid industrial areas, a finding that is supported by Zagorskas and Turskis (2024). The study by Winters, Brauer, Setton & Teschke (2010) presents a contradictory view. This study found no evidence that an industrial zone affects route choice. In addition to residential and industrial zoning, commercial zoning also affects cyclists' route choice. Koch & Dugundji (2021) indicate that commercial areas have a positive influence on cyclists' route choice (Zhao, Ke, Lin & Yu, 2020). However, Winters, Brauer, Setton & Teschke (2010) draw a different conclusion and indicate that commercial areas have no influence on cyclists' route choice. What the literature is more unambiguous about is that mixed land use does have a positive influence on cyclists' route choice (Zhao, Ke, Lin & Yu (2020), Winters, Brauer, Setton & Teschke (2010)). This assertion is supported by the concept of the compact city. Maat, van Wee & Stead (2005) conclude that when multiple facilities are available, the degree of car dependence is reduced.

Furthermore, the floor space index (FSI) also affects the route choice of cyclists. The FSI is defined as the total area a building uses over all floors, divided by the gross area of the building (Paterson, 1949). Chen (2016) describes that cyclists prefer areas with a low FSI. FSI is often related to population density. Once more, there is no definitive answer to this question. Winters, Brauer, Setton & Teschke (2010) describe in their study that areas with higher population density have a more positive influence on cyclists' route choice than areas with lower population density. Saelens, Sallis & Frank (2003) support this.

Next to zoning, blue and green spaces also exert a generally positive influence on route preference (Koch & Dugundji, 2021). Prato, Halldórsdóttir & Nielsen (2018) indicate that scenic areas contribute positively to cyclists' route preferences. Marquart et al. (2020) highlight that the presence of blue space along a cycling route enhances the cycling experience. Furthermore, the study revealed that individuals without time constraints are more likely to select these scenic routes, even if it entails a longer journey time. Nevertheless, the findings of Campos-Sánchez, et al., (2019) indicate that the mere presence of green areas does not significantly influence the propensity of cyclists to utilise such routes. Instead, the proximity to separated cycle paths is a crucial factor in making cycling routes more attractive.

The final land use factor that plays a role in cyclists' route choice is elevation, or slope. Cyclists tend to select routes with less elevation, as they consume more energy (Chen (2016), Stinson & BhatI (2003)). Prato, Halldórsdóttir & Nielsen (2018) found that cyclists' time perception is 4.9

times higher when they have to cycle uphill. Broach, Dill and Gliebe (2012) found that cyclists tend to take significant diversions when cycling uphill at a gradient greater than 2%.

2.2 SHP and LCP

In several studies (Lu, Scott, & Dalumpines (2018), Passmore, Watkins, & Guensler (2024), Meister et al. (2023)), researchers have examined GPS routes and SHP. SHP calculates the shortest possible route between Origin and Destination (OD) based on the existing network, either in terms of distance or time. For example, Prato, Halldórsdóttir, & Nielsen (2018) found that cyclists are willing to take detours to avoid traffic lights, such factor are often not considered in SHP models. Lu, Scott, & Dalumpines (2018) showed that "...routes are statistically different from the shortest path route...".

LCP is partly similar to the SHP approach. The LCP also calculates the shortest possible routes, but instead of distance or time, it minimises route costs. This approach is known as a resistance-based model (Balbi et al., 2020). The LCP route is not calculated on a vector network, but on a raster network, where different factors with assigned weights are associated with each raster cell. Using the origin and destination points, LCP determines the route. Figure 2 illustrates the differences between SHP and LCP routes.

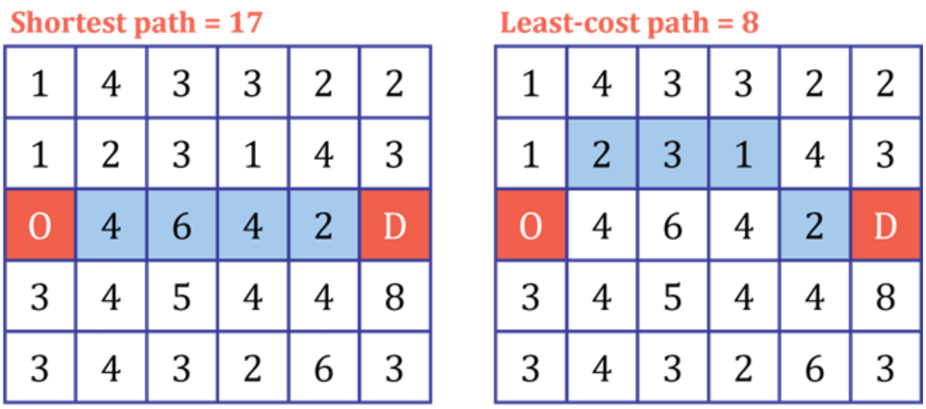


Figure 2: A conceptual representation of how the SHP and LCP determine the route (by author)

2.2 Conceptual model

The conceptual model (Figure 3) shows how the variables are interrelated. First, infrastructure and land use factors determine people's cycling behaviour (route choice). This thesis investigates whether cyclists' behaviour is more similar to a LCP or a SHP. To answer this question, the LCP and SHP are compared with actual cycled routes:

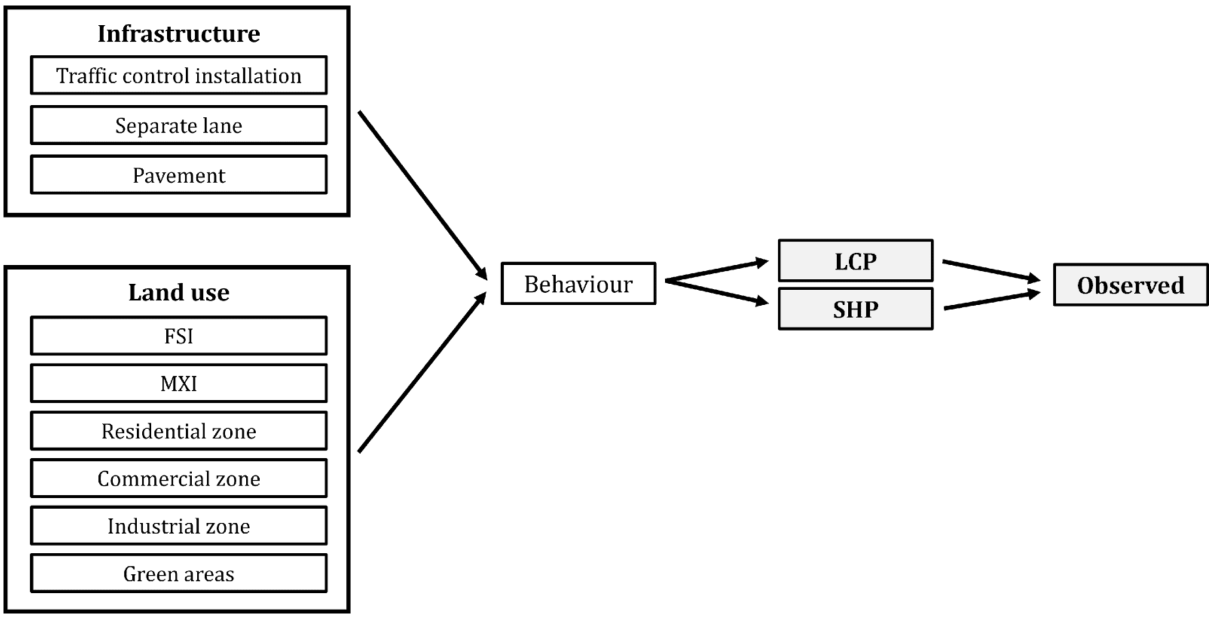


Figure 3: Conceptual model

3. Methodology

The following methodological framework has been developed to provide structure for the methodology:

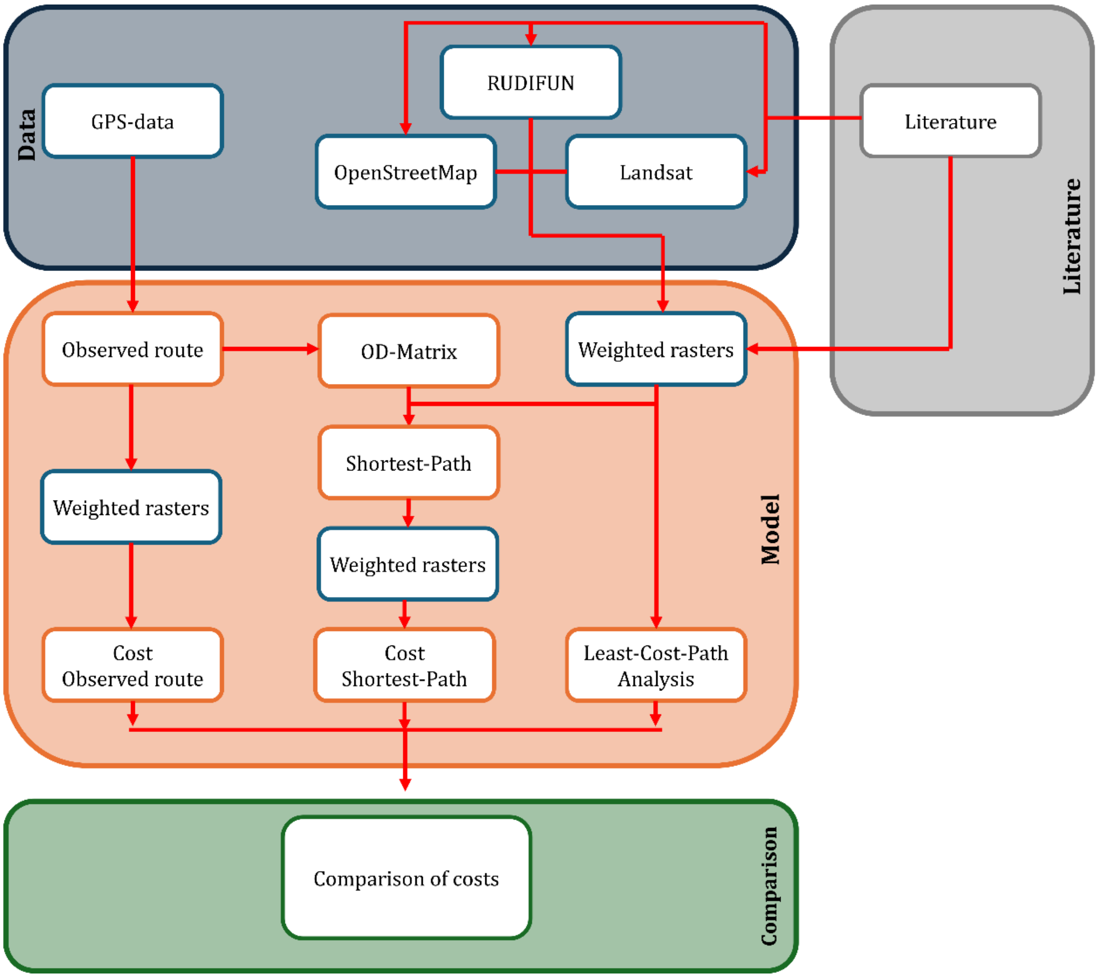


Figure 4: Methodological Framework

The framework consists of four distinct components: literature, data, model and comparison. The literature is discussed in chapter 3. This chapter elaborates on data, model and comparison.

3.1. Data

This research uses a variety of data sources. Open-source datasets will be used primarily. This makes it easier to collect data and continue research on the topic. In addition, one closed-source dataset is used, namely TalkingBikes. The table below provides an overview of the data used, sources and attributes:

Requirements	Spatial Data	Attribute data	Name dataset	Source dataset
Administrative boundaries	Polygon	Gemeentecode	CBS Wijken en buurten 2023	(CBS, 2023.)
Roadnetwork	Line	OSM_ID; Bike_allowed; Car_allowed; Pedestrian_allowed; Surface;	Netwerk_OSM	(OSM, n.d.)
Land use	Polygon	FSI_22, MXI_22, bvo_industrie_22, bvo_winkel_22 and bvo_woon_22	Rudifun_basis_bouwblok_PV 20	(Rudifun, 2022)
NDVI	Raster	Red band	Luchtfoto 2023 Ortho 25cm RGB	(PDOK, n.d.)
NDVI	Raster	NIR band	Luchtfoto 2023 Ortho 25cm Infrarood	(PDOK, n.d.)
Traffic control instalations	Point	OSM_id	highway_traffic_signals_gron ingen	(OSM, n.d.)
Observed routes (gps)	Point	Tripid	TalkingBikes	(Yunex Traffic Nederland, n.d.)

Table 1: Data requirements table

OpenStreetMap' (OSM) data (Openstreetmap, n.d.) is an open-source database where anyone can contribute to create a network dataset, mapping all information regarding infrastructure worldwide and making it accessible to everyone. This study uses an OSM extract of the city of Groningen extracted on [04-02-2024]. Although it can be edited by anyone, it is known as the most reliable network data and is often used in academic studies. The OSM data is also used to identify traffic lights in Groningen.

Apart from the road network, the RUDIFUN dataset (RUDIFUN, 2022) is an important source for this study, as it provides all zoning factors. This dataset was specifically created by the Dutch government to enable urban planners and, spatial researchers to study topic like spatial densities, housing functions and quality of life (PBL, 2022).

In order to calculate the NDVI, the year-average aerial image from PDOK was used. The year-average provides the most representative satellite images possible with regard to green space (PDOK, n.d.).

3.1.1. TalkingBikes

TalkingBikes is the key dataset for this study. This dataset contains bicycle movements (GPS) in the Netherlands between October 2020 and October 2022. It is also the only data source that is not open source. The data is managed by Yunex Traffic Nederland. In 2019, on behalf of the Ministry of Infrastructure and Water Management, a tender was issued for the collection of bicycle movements. Two different companies took on this project: RingRing and Tracefy. Over a period of

2 years, more than 5.6 million bicycle movements were recorded. While Tracefy mainly records (food) delivery activities using GPS trackers installed on bicycles, RingRing focuses on personal trips, including different types of cycling such as commuting, leisure and school trips. Because of its diverse features and extensive amount of data, it provides a highly representative picture of the average cyclist in the Netherlands.

As explained in section 1.4., this study focuses on the municipality of Groningen. All GPS tracks from RingRing are limited to the boundaries of the municipality. Therefore, only trips that took place (partially) within the municipality of Groningen are selected for this case study (see Appendix I). The total number of records within this municipality is 3821. However, not all of these bicycle trips are of good quality. Some trips contain only one GPS point, while others have long intervals between each GPS point, which could indicate a long break or an additional stop during the cycling trip.

The disadvantage of collecting GPS data via a mobile phone in an urban environment is the lack of accuracy (Lindsey et al., 2013). Urban environments can interfere with the GPS signal, resulting in inaccurate route mapping. There are two ways to address this issue: map matching or data filtering. Map matching involves accurately aligning GPS points with a road network using an algorithm (Millard-Ball, Hampshire, and Weinberger, 2019). However, this is beyond the scope of this study. Therefore, the RingRing data is extensively filtered to ensure sufficient accuracy for analysis.

GPS data is measured in points with a certain time interval. For this study, only trips with 20 seconds or less between each GPS registration are used. This ensures that the trip is slightly more accurately matched to the network and that all trips have only one origin and destination (without intermediate stops). This filtering process is shown in Figure 5. A Python script, included in Appendix II, was developed to filter the data.

In addition, another filtering method (Appendix III) was applied, visualised in Figure 6. The GPS signal from a mobile phone can sometimes be disturbed by an urban environment, resulting in random points being recorded. This can lead to the bike movement being routed differently from the actual route taken. For this reason, an additional selection criterion was set, requiring that the distance between each point should not exceed 222 metres (40 km/h).

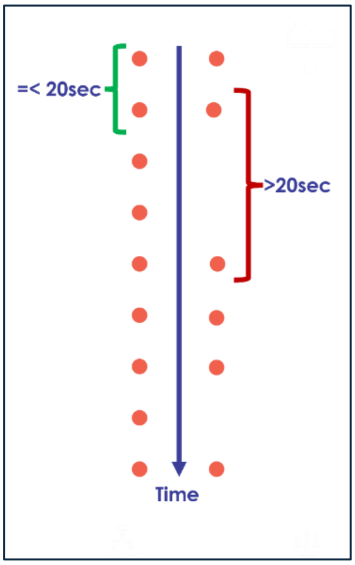


Figure 5: Conceptual illustration of how the GPS filters for accuracy, points must be less than or 20 seconds apart (by author)

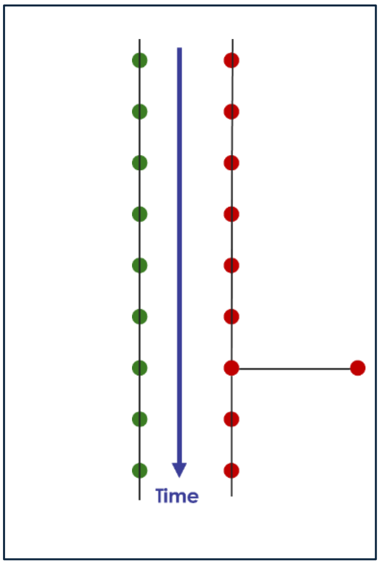


Figure 6: Conceptual illustration of how the GPS outliners are filtered, excluding points with a difference of more than 222m (by author)

3.1.2. Ethical considerations

GPS data is very sensitive information. It shows people's exact movement patterns. that is why it is important to handle it with care. RingRing data are handled according to the guidelines of the Ministry of Infrastructure and Water Management (Ministry of Infrastructure and Watermanagement (2018). RingRing data is collected by users who activate the RingRing app before starting their bike ride. The mobile app then registers the GPS signal. The collection of cycling data raises privacy concerns. To avoid privacy violations, all recorded rides are anonymised. Anonymisation is achieved by assigning a unique GUID (Globally Unique Identifier) to each trip, ensuring that no personal data can be associated. In addition, individuals voluntarily use the RingRing app and can selectively choose when to use it. These measures minimise privacy intrusion.

3.2. Model

The model consists of constructing a WRN and performing a LCP and SHP analysis on it. Subsequently, the total costs of the LCP, SHP, and OBR are calculated on this same WRN. This section describes how the WRN is created and how the costs are calculated.

3.2.1. Factors weighted raster network

In the theoretical framework (Chapter 2) several factors influencing cyclists' route choice have been considered. These factors have to be assigned to each road segment in order to create a WRN. The weights are derived from a previous study by Van Neijen (2022). In this case study in the city of Enschede, a regression analysis was carried out with 13 different factors. As mentioned in Chapter 2, several studies have investigated the influence of different factors on cyclists. While this study in the Netherlands includes many factors, other studies often focus on one specific factor, such as weather influence (Motoaki and Daziano, 2015).

Factor	Standardised β
Infrastructural	
Distance to traffic control installation	-0.124
Cycle lane	0.095
Separate cycle path	1.072
Artificial lighting	-0.246
Paved infrastructure	0.699
Motorised vehicle intensities	-0.675
Bicycle intensities	1.108
Land use allocation	
Residential land use zone	0.689
Commercial land use zone	-0.446
Greenery land use zone	-0.227
Industrial land use zone	-0.232
Land use mix	0.034
Degree of urbanisation	0.365

Table 2: standardized regression coefficients from Van Neijen, (2022).

From this table, 9 of the 13 factors are taken together with their coefficients. Motorised and bicycle intensities are excluded. This is because they are not available in public datasets. There is a dataset available, but it only covers certain segments and not the whole network (NDW, n.d.). If these factors were included in the WRN, it would give a distorted picture when comparing routes with and without cycling intensities. Therefore, intensities are not considered for inclusion in the WRN.

In the OSM dataset is no clear distinction between a cycle lane and an unmarked road. Therefore, this factor is not included in the WRN.

Artificial lighting only influences route choice when it is dark (Winters, Davidson, Kao, & Teschke, 2011). However, it is beyond the scope of this study to separate the route data into day and night periods. Therefore, artificial lighting is not considered in this study.

3.2.1.1 infrastructure

For this study, the following three infrastructure factors were used: Separate cycle path, pavement and distance to traffic lights. A brief description of how these factors are defined is given below.

The OSM dataset indicates the type of use for which the road is intended. This is divided into pedestrian, bicycle and motorised categories. A dummy variable was then created where 0 represents a segregated cycleway segment (bicycle=yes AND pedestrian=no AND motorised=no) and 1 represents a non-separated segment.

The OSM dataset contains information on the surface type of a road segment, categorised as follows:

Class	Type surface
1	Asphalt
2	Concrete, paved & paving_stones
3	Chipseal, sett, unhew_cobblestone, cobblestone, bricks, metal, wood, rubber, fine_gravel, gravel, shells, rocks, dirt & gras.

Table 3: classification of surface types

A buffer analysis was performed for traffic lights. Strauss & Miranda-Moreno (2013) found in their study that the proximity of traffic lights has a negative effect on the route choice of cyclists. In their study they identified 4 buffer categories: 50m, 150m, 400m and 800m. The closer you are to the traffic lights, the less likely you are to cycle there. The cost is then allocated by dividing the distance of the buffer by 50. This gives the following result:

Buffer categories	Distance buffer	Distance normalized
1	50	1
2	150	0,333...
3	400	0,125
4	800	0,0625
5	Other	0

Table 4: classification buffer zones traffic lights

3.2.1.2 Land use

Land use influences cyclists' route choices Koch & Dugundji (2021). Three zoning categories were used in this study: industrial, commercial and residential. In addition, green space and the degree of urbanisation (FSI) and land use mix (MXI) were taken into account.

A pairwise buffer of 250 metres was used to assign zoning, FSI and MXI to each road segment. According to the study by Winters, Brauer, Setton & Teschke (2010), a 250 meter buffer is ideal for an area with urban and more rural areas. The buffer creates a 250m buffer on all sides of the segment. Visually it looks as follows:

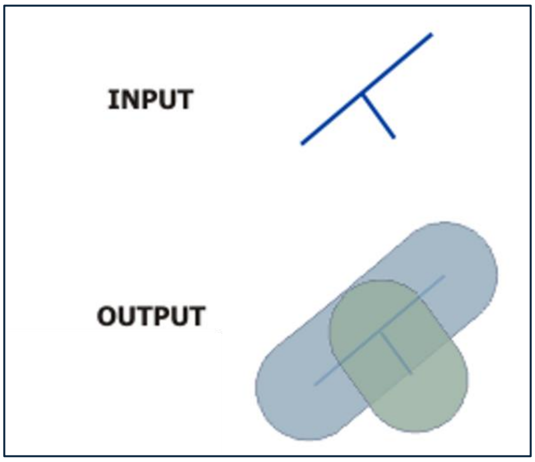


Figure 7: visualisation of the pairwise buffer, adopted from Arcgis.com (2024).

Next, the values of Industrial, Commercial, Residential, FSI and MXI are assigned to each buffer segment using 'summarise within'. Each buffer segment is then linked to the road network so that each road segment is assigned the value of each factor.

Finally, a Normalised Difference Vegetation Index (NDVI) analysis is used to add green land use to the road segments. This is an indicator that can be used with remote sensing to inventory areas with vegetation (Pettorelli et al., 2005).

3.2.2. Cell size

After all factors had been calculated, they were normalised between 0 and 1 to ensure that each factor had an equal influence. A value of 0 indicates relatively low costs, while a value of 1 indicates relatively high costs. Weights were then assigned to each factor according to Table 2. For a WRN, a cell size had to be determined. The cell size refers to the size of each pixel in the raster (in meters). The smaller the pixel size, the higher the resolution and quality of the model. Figure 8 illustrates different cell sizes:

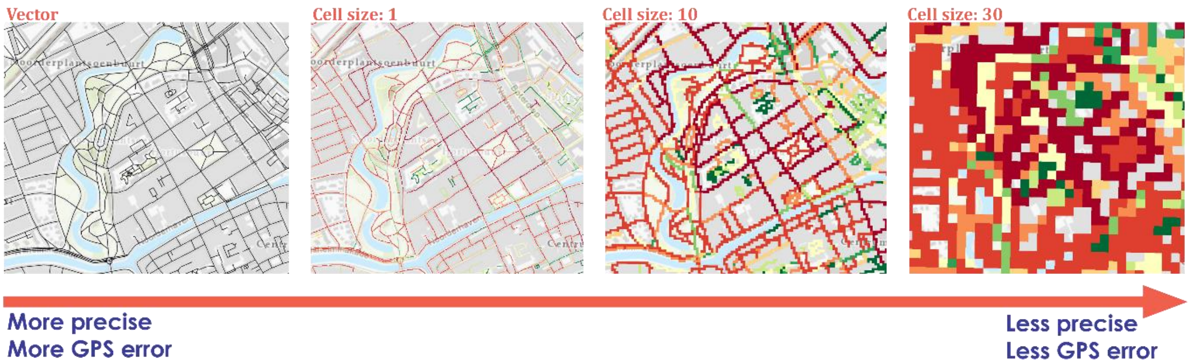


Figure 8: overview of outcome of different cell sizes (by author)

As described in section 3.1.1, no map matching was performed. Therefore, the accuracy of the GPS data is not very precise. For that reason, in this study, a cell size of 10m x 10m was chosen to account for this GPS inaccuracy.

3.3. Constructing LCP, SHP and OBR

After extensive filtering of the data (Section 3.1.1.), 50 routes remain suitable for analysis. The LCP and SHP are constructed for these 50 trips. The SHP is calculated using the route solver in ArcGIS Pro and the LCP is calculated using a custom model with tools from ArcGIS Pro (see Appendix IV). In order to construct the LCP and SHP, an OD-matrix (Origin and Destination) must be created, with the Python script for this calculation available in Appendix V.

The LCP analysis then searches on the WRN for the route with the lowest cost based on the origin and destination. The SHP follows a similar process, but instead of minimising cost, the SHP searches for the shortest distance between origin and destination, based on the OSM road network configured to allow cyclists only on cycle lanes and shared roads. Finally, the costs of the LCP, SHP and OBR are calculated by performing a Zonal Statistics analysis. The Zonal Statistics function sums all the pixel values along the route. This produces a table with the total costs of the LCP, SHP and OBR (Appendix VI).

4. Results

In this section, the results of the described methodology are presented. First, the WRN of the city of Groningen is presented. Then five OBR, LCP and SHP routes are compared.

4.1. Weighed raster network

By assigning weights to each factor, the following weighted raster model was constructed. The map in Figure 9 shows the cost per pixel of a cyclist crossing that cell. The lower the cost, the more likely the model is that a cyclist will pass through that cell. The cost of the weighted grid model ranges from 0.507 to 3.157.

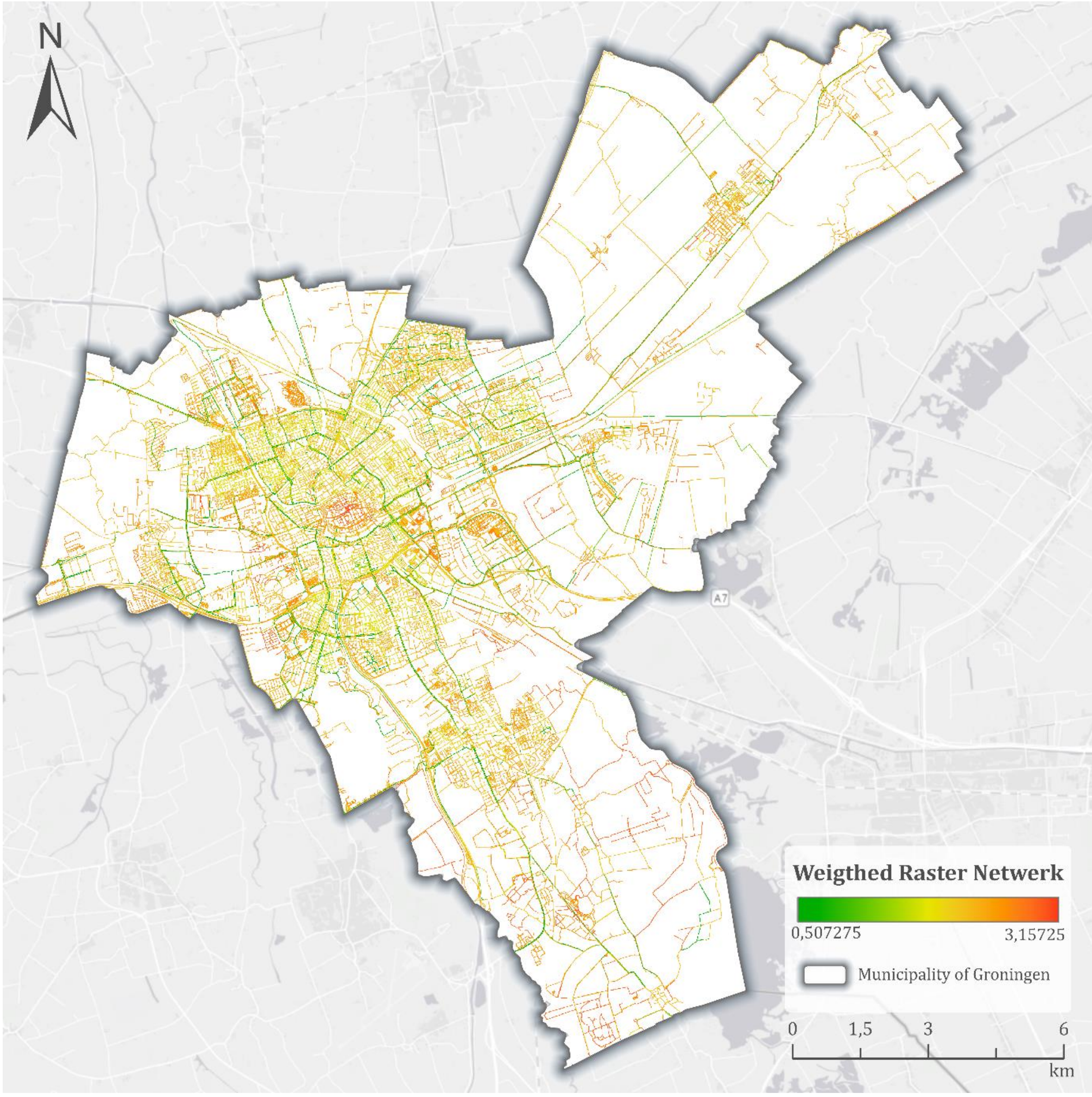


Figure 9: Weighted Raster Network of the municipality of Groningen, by author

The majority of the journeys took place within the city of Groningen, Appendix VII provides a detailed illustration of the city of Groningen.

4.2. Comparison LCP, SHP, OBR

4.2.1. visual

In this section, four randomly chosen trips (T0010, T0040, T0080 and T0270) are presented:

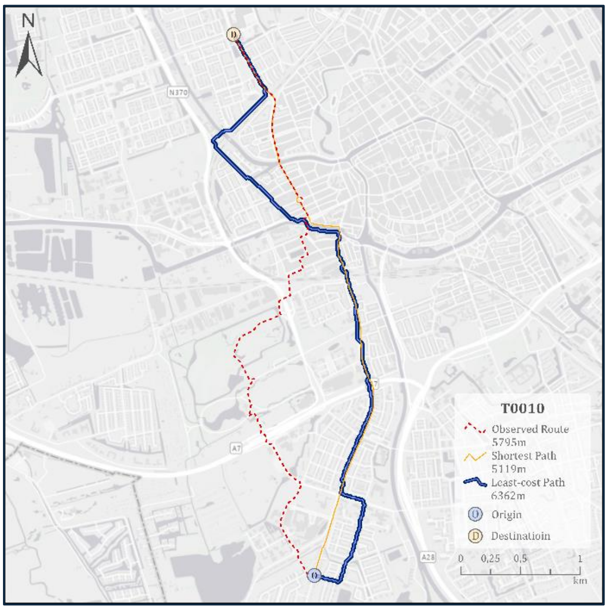


Figure 10: OBR, SHP and LCP of trip T0010 visualised

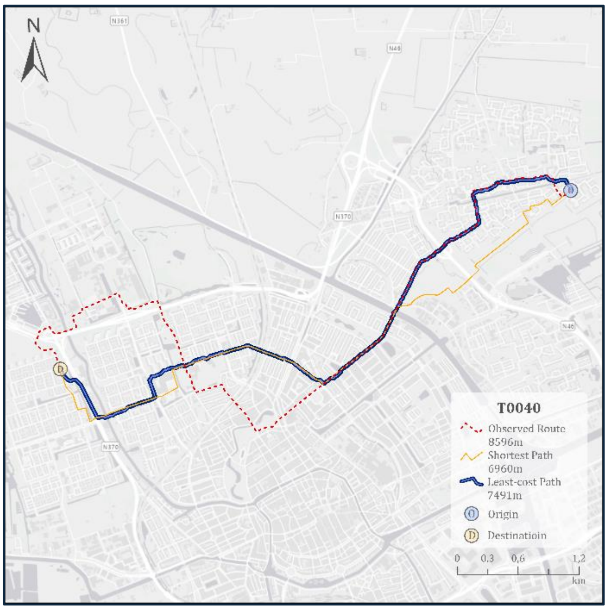


Figure 11: OBR, SHP and LCP of trip T0040 visualised

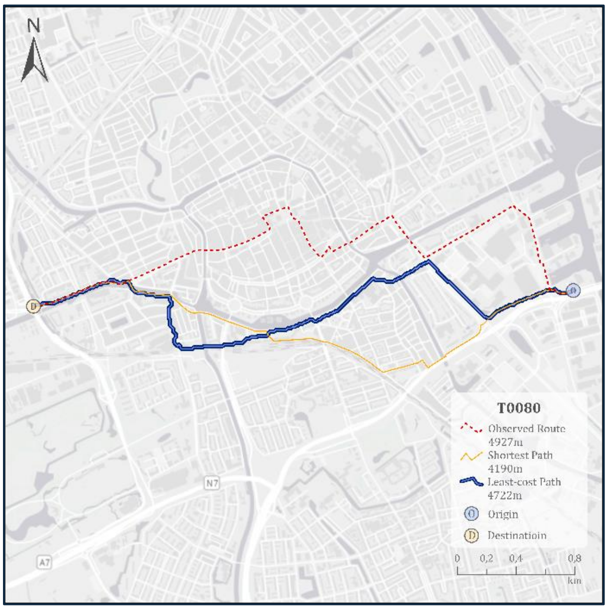


Figure 12: OBR, SHP and LCP of trip T0080 visualised

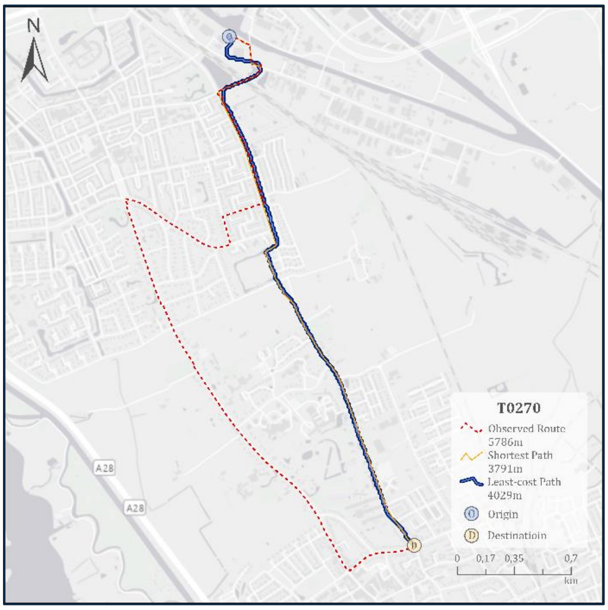


Figure 13: OBR, SHP and LCP of trip T0270 visualised

These four triplicate routes each show different patterns. Neither the SHP nor the LCP closely resemble the OBR. Visually, there are even more similarities between SHP and LCP. Both methods aim to optimise routes: the SHP only considers the shortest distance, while the LCP considers other factors in addition to distance. A possible explanation for the similarities between SHP and LCP could be insufficient differentiation between the weights used in the LCP analysis or incorrect calibration of the weights for the municipality of Groningen.

When comparing the LCP with the OBR and SHP, it is clear that the LCP does not follow a straight line. The LCP tries to optimise costs by taking a zigzag route; with a cell size of 10m, the LCP has more room to zigzag. The model is optimised to find the shortest route and by zigzagging it reduces the cost, distorting the overall cost representation. This is illustrated in Figure 14, where the SHP follows a straight line according to the OSM network, while the LCP zigzags from pixel to pixel.

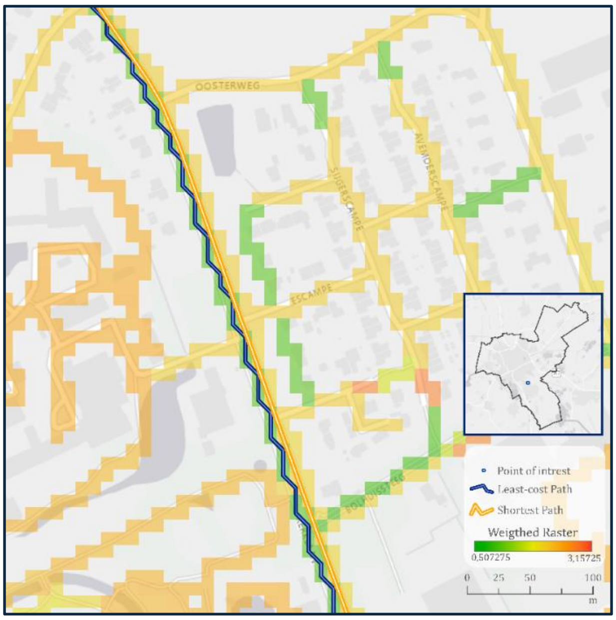


Figure 14: The difference between the straight SHP and the zigzagging LCP within trip T0270

It is also noticeable that the LCP does not follow the traffic rules. The LCP searches for the cheapest route, regardless of traffic regulations. For example, during a single trip, the LCP's route alternates between driving on the right and left side of the road, even though the route is only from start to finish and should consistently follow either the right or left side of the road. This can be seen in Figures 15 and 16.

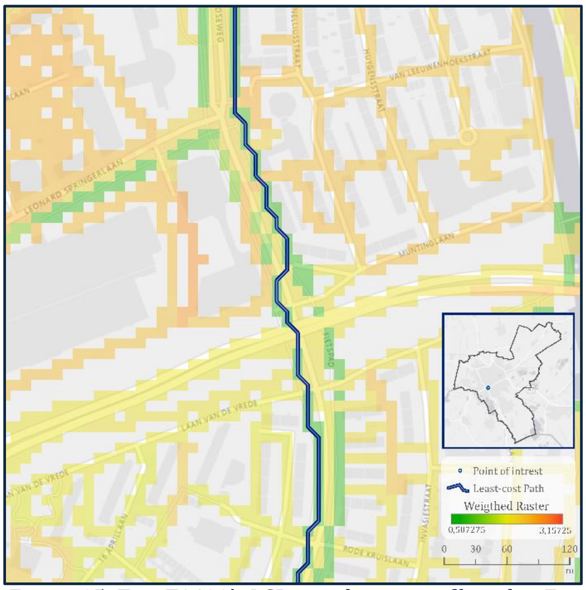


Figure 15: Trip T0400's LCP not obeying traffic rules. First it follows the roads on the left side of the road (south) after which he follows his way on the right side of the road (north)

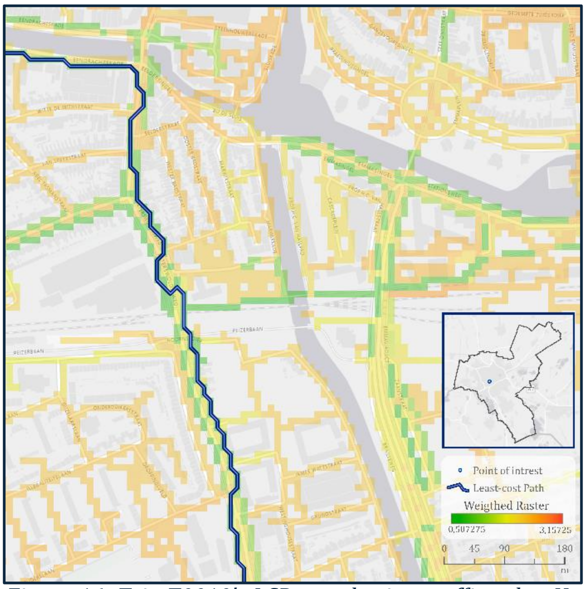


Figure 16: Trip T0010's LCP not obeying traffic rules. He starts on the right side of the road (south) then crosses the road, and goes back and forth in the middle of the map. after which he ends his route on the left side of the road (north)

The LCP's capacity to deviate from the network on occasion results in a shorter route length than that of the SHP. This outcome is unexpected, given that the SHP is designed to represent the shortest possible route. Figure 17 illustrates an example for this phenomenon:

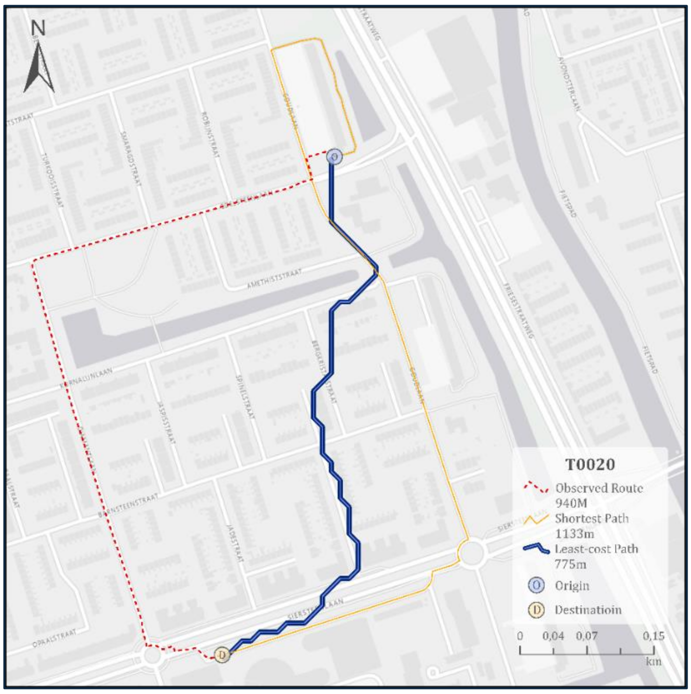


Figure 17: OBR, SHP and LCP of trip T0020 visualised where it can be seen that the LCP passes over buildings. The SHP makes a diversion to the north.

The De SHP adheres to a designated bicycle network, whereas the LCP does so indirectly. Due to its 10m cell size, the LCP has flexibility. When roads are close together, the LCP may deviate from the path, for instance, by passing through an area with buildings nearby, as depicted in Figure 17. Furthermore, in such instances, the OBR may have a shorter distance than the SHP. This is because the OBR may take a path that is technically inaccessible to cyclists (such as a footpath), which the SHP avoids by taking a detour, thereby increasing the route length. Table 5 shows trips where either the LCP or OBR is longer than the SHP:

Trip_ID	Length (m) OBR	Difference (m) OBR-SHP	Length (m) LCP	Difference (m) LCP-SHP	Length(m) SHP
T0020	940,8396828	-193,1123017	775,2691193	-358,6828652	1133,951985
T0050	2445,48217	-21,68229059	2599,655121	132,4906608	2467,16446
T0060	1081,304617	208,0835916	830,1219331	-43,09909234	873,2210254
T0090	1028,669199	93,21041947	823,5533906	-111,9053885	935,4587791
T0160	4469,159001	603,2734065	3767,47258	-98,41301437	3865,885595
T0260	3913,294928	91,63946309	3818,305192	-3,350273301	3821,655465
T0320	936,640666	119,4324631	787,6955262	-29,51267675	817,208203
T0360	1256,980995	219,3502065	1021,543289	-16,08749929	1037,630789
T0370	2280,744035	-34,76208036	2474,802307	159,2961923	2315,506115
T0390	1454,70088	245,4849207	1081,248917	-127,9670427	1209,215959
T0400	2380,485758	-44,3194787	2601,37085	176,5656137	2424,805236
T0420	1348,993428	-152,8222792	1896,812409	394,9967012	1501,815707
T0430	3766,580519	404,856113	3333,624817	-28,09958846	3361,724406
T0480	974,9303211	-131,312431	765,9797975	-340,2629547	1106,242752
T0510	1097,214252	-51,56163276	1128,111832	-20,66405268	1148,775884

Table 5: 15 trips where the SHP does not has the shortest route

A closer examination of the WRN (Figure 18) reveals that the majority of major roads have relatively low costs, indicated with colour green (Figure 19). In most cases, these roads also include a separate bicycle path. This observation can be seen in a comparison between Figures 18 and 19:

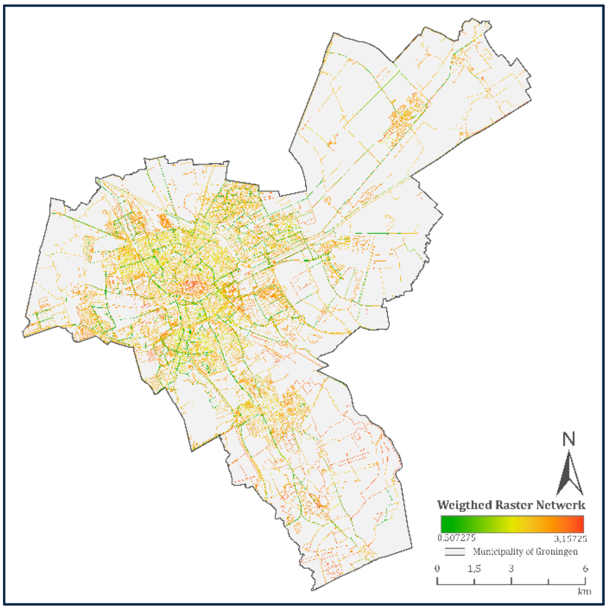


Figure 18: Weigthed Raster Netwerk of the municipality of Groningen

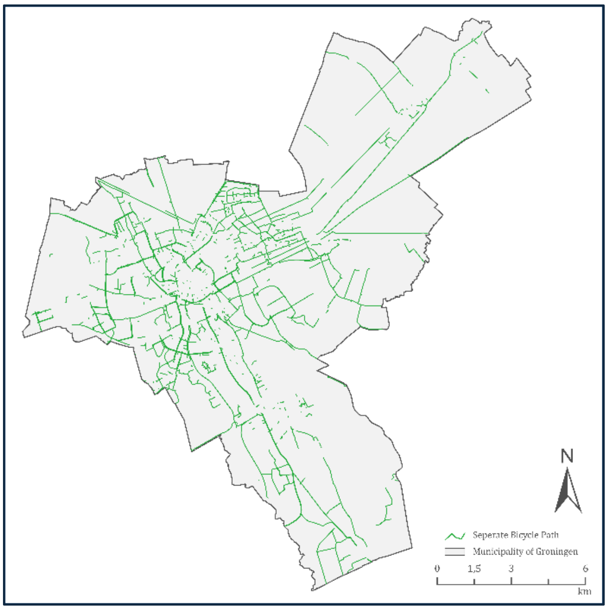


Figure 19: All seperated bicyle lanes in the municpality of Groningen

Figure 18 shows the raster network and Figure 19 shows all the separated cycleways. It can be seen that the low cost areas coincide with the segregated cycle lanes. One possible explanation is that the weight given to the dummy variable of the separate lane (1.072) is too high. To assess this, the percentage of LCP, SHP and OBR routes that were on separate paths was examined. The length of each route that overlaps with a segregated path was divided by the total length of the route to calculate the proportion on a segregated path (Appendix VIII). The following box plot illustrates this ratio for the OBR, SHP and LCP:

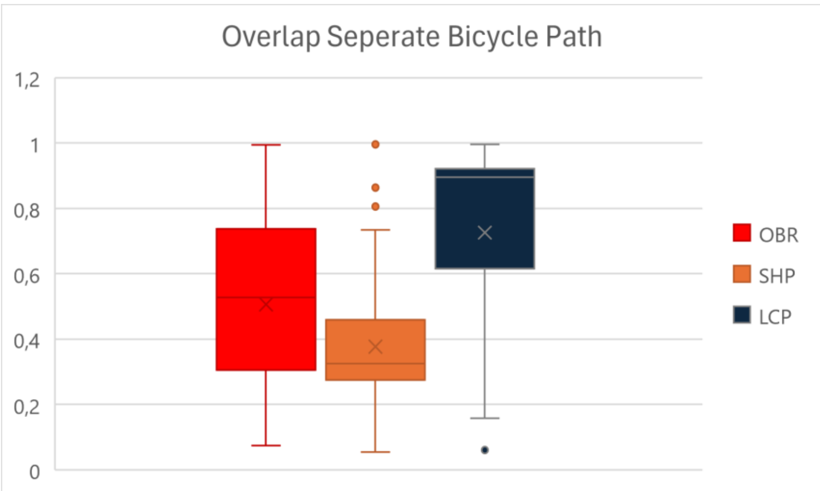


Figure 20: Box plot showing the relationship between the different routes and the length on segregated cycle paths

It is notable that the OBR (average overlap 50,7%) has a greater spread than the SHP or LCP. The SHP has the lowest percentage overlap with separated bicycle paths at 37.7%. In contrast, the LCP has the highest overlap with separated bicycle paths at 72.6%. This suggests that separated bicycle paths are a dominant factor for the LCP.

To examine how the total costs per route compare to each other, the following boxplots were created (see Appendix IX):

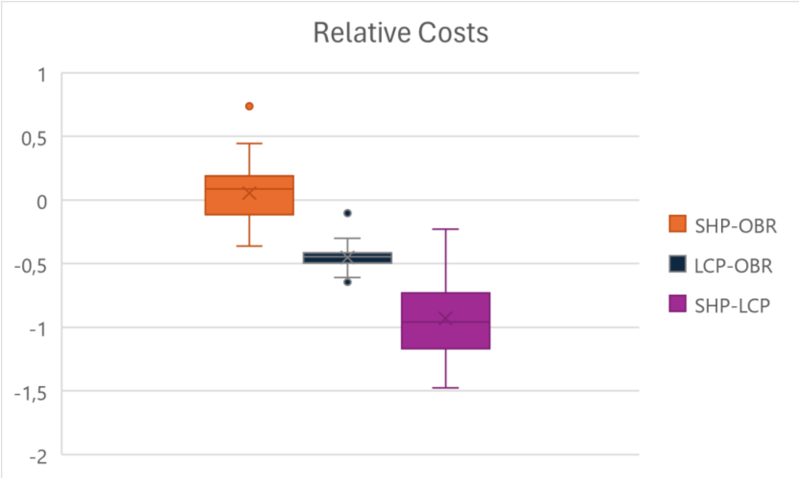


Figure 21: Box plot showing the relationship between costs

The boxplots illustrate the relative ratio between the total costs of SHP and OBR, LCP and OBR, and LCP and SHP. This was calculated by subtracting the total costs of OBR from SHP costs and dividing by the costs of OBR. The same calculation was applied for LCP-OBR and LCP-SHP comparisons. A ratio closer to 0 indicates closer total costs between the compared routes. A positive ratio indicates that the total costs are higher than those of the route being compared. For instance, the average total costs of SHP are 5.32% higher than those of OBR. However, the spread is relatively large, with the first quartile at -11.6% and the third at 18.9%.

A comparison of LCP and OBR indicates that the costs of LCP are, on average, 45.2% lower. It is logical that costs would be lower because the model seeks the lowest possible costs for the route. The variability in costs is relatively small, with the first quartile at -49.3% and the third quartile at 41.3%. A comparison of LCP and SHP reveals a significantly greater difference in total costs. The average cost of LCP is 92.9% lower than that of SHP. However, there is greater variability, with a range of -116.7% to -73.1%. The cost difference is almost twice as much.

Based on a comparison of costs, SHP and OBR are closest to each other. Despite the fact that the costs of the SHP and OBR are relatively similar, a visual comparison (see figures 10-13) reveals that the OBR and SHP are not directly comparable. In fact, the LCP and SHP are more closely related. This is because the cost of the route only indicates the effort a cyclist must make; it does not reflect the direction of the actual route. Therefore, multiple routes may result in the same overall cost.

4.2. Discussion

This study attempted to model route behaviour by examining the LCP, SHP, and OBR. However, it was found that the LCP model in this study is not capable of accurately predicting routes. LCP models are typically used on a larger scale and for different purposes, such as maritime movements, where there is no predefined network (Gustas and Supernant, 2017). A cell size of 10m is actually too small for an LCP model and too large for a detailed bike analysis. This could explain why LCP has not been commonly used to analyse bike behaviour, with other models being preferred (Strauss and Miranda-Moreno (2013), Brand et al. (2017), Veenstra, Geurs, Thomas & Van den Hof (2016)).

Furthermore, predicting the behaviour of cyclists remains challenging due to the strong personal preferences of cyclists when it comes to route choices (Damant-Sirois, Grimsrud, & El-Geneidy, (2014), Félix, Moura, & Clifton (2017)). Each location has its own unique infrastructure that influences how cyclists navigate. Another limitation of this study is the use of coefficients from another case study without validation. Finally, the use of secondary GPS data limits the ability to understand respondents' motivations directly. Conducting interviews to inquire about why cyclists choose specific routes could provide valuable insights and improve the model (Desjardins et al., 2021).

A mixed-method approach would have been a more appropriate methodology for an exploratory study, as it could provide a more nuanced understanding of the model, allowing for a more comprehensive analysis. Additionally, specific weights should be determined for this individual case study. It has also been found that transferring coefficients from one case study to another is not a feasible approach without first validating them. Scaling up this method to reduce traffic casualties is not a viable option based on this study. If unique weighting factors are determined for the municipality of Groningen and a new WRM can be calculated, it is possible that these factors could influence the weights. It appears that the weights are too dependent on a specific case study, which makes it inadvisable to extrapolate these findings to other case studies or to the entire country. For policymakers, an attractiveness map is a more useful tool than a model that can predict route choice. This is because it can lead to the implementation of more concrete measures in the urban fabric.

5. Conclusions

This study demonstrates that predicting cyclist behaviour is very challenging. Based on the model used, cyclists do not strictly follow the LCP or SHP routes. The OBR route shows too much variation to establish consistent patterns. In addition, the use of a raster model with a cell size of 10m proves inefficient for cycle routes. The LCP tends to zigzag in order to optimise costs, disregarding traffic rules by alternating between the left and right side of the road within the same trip. Furthermore, this study has shown that coefficients cannot be transferred from one context to another, leading to a distorted perception of reality.

SQ1: ***"What infrastructure and land use factors can be identified in the literature as having a significant influence on cyclists' route choice?"***

In this study, 22 number of papers were used to identify 16 relevant factors affecting route choice and to assess their suitability for translation into a weighted raster model. Ultimately, 13 factors were identified, of which 9 were selected for inclusion in the raster model.

SQ2: ***"To what extent do the least-cost path and shortest path analyses correspond to the actual routes chosen by cyclists?"***

No visual similarities were identified between the LCP, SHP and OBR. However, the LCP and SHP were found to be more similar in that they both seek a form of the shortest route. The costs of the OBR and SHP were found to be the most closely aligned, with a difference of only 5.32%. In contrast, the LCP was found to differ from the OBR by -45.18% and from the SHP by -93%. No clear relationship was identified between the LCP, SHP and OBR.

SQ3: ***"To what extent is it possible to use a least cost path analysis to do predict cycling behaviour?"***

The study did not identify a clear relationship between cyclists' route behaviour and the LCP. This is attributed to several factors, including the adoption of coefficients from another study without validation, the use of a cell size that is too large, and the LCP not adhering to traffic rules.

MRQ: ***"What is the difference in route choice of cyclists between a least-cost path analysis, the shortest route and the actual route in the municipality of Groningen?"***

Based on the answers to the sub-questions and the results of this study, there is no clear relationship between the LCP, SHP and OBR routes. This study found no evidence to support the use of LCP or SHP to simulate cyclist behaviour. According to the model used in this study, there are clear differences between LCP, SHP and OBR. Furthermore, based on visual interpretation, it can be concluded that the LCP does not conform to legal cycling practice according to this model, as it does not follow the rules of the road.

5.1. Future research

The prediction of cyclist behaviour is an inherently challenging task. While this study did not identify any relationship between the LCP, SHP and OBR routes, this does not imply that such correlation does not exist. This study represents one of the initial attempts to predict cyclist behaviour using an LCP raster model. It should be noted that within the scope of this research, there was no independent analysis conducted to determine the weights of each factor. Instead, an existing study with a different case study was utilised as a foundation for this research, without conducting validation. To accurately predict cyclist behaviour using an LCP, weight validation must be applied before adopting weights from previous studies. Alternatively, future research should conduct its own analysis to assign site-specific weights to the raster, ensuring the model reflects accurate values.

In this study, a cell size of 10m was employed to accommodate the margin of error in GPS signals. As a result, the LCP had more room to move outside the existing network, leading to a zigzag pattern that is impractical for cycling and thus not comparable to an actual bikeable route. Future research should investigate whether reducing the cell size can mitigate zigzagging, preferably it would make use of the same cycling network as the SHP. This would necessitate higher-quality GPS data from accurately measuring or map matching bicycle movements.

Finaly, primary GPS data would be a more appropriate methodology for exploratory research. By directly collecting data, the study can also enquire of respondents, by interviews, why they made certain choices. This approach can provide more profound insights into cycling behaviour by combining both quantitative and qualitative methods.

6. References

Arcgis.com. (2024). *Pairwise Buffer (Analysis)—ArcGIS Pro | Documentation*. [online] Available at: https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/pairwise-buffer.htm

Arslan, T. and Khisty, C.J. (2005) 'A rational reasoning method from fuzzy perceptions in route choice,' *Fuzzy Sets and Systems*, 150(3), pp. 419–435. https://doi.org/10.1016/j.fss.2004.03.021.

Balbi, M. *et al.* (2020) 'Least-cost path analysis for urban greenways planning: A test with moths and birds across two habitats and two cities,' *Journal of Applied Ecology*, 58(3), pp. 632–643. https://doi.org/10.1111/1365-2664.13800.

Bovy, P.H. and Ēliyyahû Stern (1990). *Route choice : wayfinding in transport networks*. Dordrecht U.A.: Kluwer.

Brand, J., Hoogendoorn, S., Niels van Oort and Schalkwijk, B. (2017). Modelling multimodal transit networks integration of bus networks with walking and cycling. [online] doi:https://doi.org/10.1109/mtits.2017.8005612.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transportation Research Part A: Policy and Practice, 1730-1740.

Campos-Sánchez, F. S., Valenzuela-Montes, L. M., & Abarca-Álvarez, F. J. (2019). Evidence of Green Areas, Cycle Infrastructure and Attractive Destinations Working Together in Development on Urban Cycling. Sustainability.

Chen, P. (2016). Built environment effects on bicyclists' route preferences: a GPS data analysis. In P. Chen, Bicycling and the built environment: route choice and road safety (pp. 19-56). Washington: University of Washington.

Damant-Sirois, G., Grimsrud, M. and El-Geneidy, A.M. (2014) 'What's your type: a multidimensional cyclist typology,' *Transportation*, 41(6), pp. 1153–1169. https://doi.org/10.1007/s11116-014-9523-8.

Desjardins, E. *et al.* (2021) '"Going through a little bit of growing pains": A qualitative study of the factors that influence the route choice of regular bicyclists in a developing cycling city,' *Transportation Research. Part F, Traffic Psychology and Behaviour*, 81, pp. 431–444. https://doi.org/10.1016/j.trf.2021.06.005.

Ding, H. *et al.* (2021) 'Role of exposure in bicycle safety analysis: Effect of cycle path choice,' *Accident Analysis and Prevention*, 153, p. 106014. https://doi.org/10.1016/j.aap.2021.106014.

Dondi, G. *et al.* (2011) 'Bike Lane Design: the Context Sensitive Approach,' *Procedia Engineering*, 21, pp. 897–906. https://doi.org/10.1016/j.proeng.2011.11.2092.

Félix, R., Moura, F. and Clifton, K.J. (2017) 'Typologies of Urban Cyclists: Review of Market Segmentation Methods for Planning Practice,' *Transportation Research Record*, 2662(1), pp. 125–133. https://doi.org/10.3141/2662-14.

Gadsby, A., Tsai, J. and Watkins, K. (2022) 'Understanding the influence of pavement conditions on cyclists' perception of safety and comfort using surveys and eye tracking,' *Transportation Research Record*, 2676(12), pp. 112–126. https://doi.org/10.1177/03611981221090936.

Gössling, S.& M.S. (2022) 'Subjectively safe cycling infrastructure: New insights for urban designs,' *ideas.repec.org* [Preprint]. https://ideas.repec.org/a/eee/jotrge/v101y2022ics0966692322000631.html.

Grudgings, N., Hughes, S. and Hagen-Zanker, A. (2021) 'What aspects of traffic intensity most influence cycling mode choice? A study of commuting in Surrey, UK,' *International Journal of Sustainable Transportation*, 17(2), pp. 136–147. https://doi.org/10.1080/15568318.2021.1999539.

Ministry of Infrastructure and Watermanagement (2018). *IMMA-Leidraad-omgaan-met-persoonsgegvens-2018* avaliable : https://www.talking-traffic.com/nl/thema-s/privacy

Gustas, R. and Supernant, K. (2017) 'Least cost path analysis of early maritime movement on the Pacific Northwest Coast,' *Journal of Archaeological Science*, 78, pp. 40–56. https://doi.org/10.1016/j.jas.2016.11.006.

Ha, J., Lee, S. and Ko, J. (2020) 'Unraveling the impact of travel time, cost, and transit burdens on commute mode choice for different income and age groups,' *Transportation Research. Part a, Policy and Practice*, 141, pp. 147–166. https://doi.org/10.1016/j.tra.2020.07.020.

Handy, S., Van Wee, B. and Kroesen, M. (2014) 'Promoting cycling for Transport: Research needs and challenges,' *Transport Reviews*, 34(1), pp. 4–24. https://doi.org/10.1080/01441647.2013.860204.

Heinen, E., Maat, K., & Van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transportation Research Part D: Transport and Environment, 102-109.

Khatri, R., Cherry, C. R., Nambisan, S. S., & Han, L. D. (2016). Modeling Route Choice of Utilitarian Bikeshare Users with GPS Data. Transport Research Record, 141-149.

Kiviluoto, K. *et al.* (2022) 'Towards sustainable mobility – Transformative scenarios for 2034,' *Transportation Research Interdisciplinary Perspectives*, 16, p. 100690. https://doi.org/10.1016/j.trip.2022.100690.

Koch, T. and Dugundji, E.R. (2021). Taste variation in environmental features of bicycle routes. *Data Archiving and Networked Services (DANS)*. doi:https://doi.org/10.1145/3486629.3490697.

Li, Z., Wang, W., Liu, P., & Ragland, D. R. (2012). Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities. Transport Research, 256-261.

Lindsey, G. *et al.* (2013) *Feasibility of using GPS to track bicycle lane positioning*. https://conservancy.umn.edu/items/b567a607-2853-430c-891a-65154b6de10b.

Lu, W., Scott, D.M. and Dalumpines, R. (2018) 'Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths,' *Journal of Transport Geography*, 71, pp. 172–181. https://doi.org/10.1016/j.jtrangeo.2018.07.012.

Maat, K., van Wee, B., & Stead, D. (2005). Land Use and Travel Behaviour: Expected Effects from the Perspective of Utility Theory and Activity-Based Theories. Environment and Planning B: Planning and Design, 33-46.

Marquart, H., Stark, K. and Jarass, J. (2022). How are air pollution and noise perceived en route? Investigating cyclists' and pedestrians' personal exposure, wellbeing and practices during commute. *Journal of transport & health*, [online] 24, pp.101325–101325. doi:https://doi.org/10.1016/j.jth.2021.101325.

Meireles, M. and Paulo (2020). Digital Platform/Mobile App to Boost Cycling for the Promotion of Sustainable Mobility in Mid-Sized Starter Cycling Cities. *Sustainability*, 12(5), pp.2064–2064. doi:https://doi.org/10.3390/su12052064.

Meister, A., Felder, M., Schmid, B. and Axhausen, K.W. (2023). Route choice modeling for cyclists on urban networks. *Transportation research. Part A, Policy and practice*, [online] 173, pp.103723–103723. doi:https://doi.org/10.1016/j.tra.2023.103723.

Mertens, L., Van Dyck, D., Ghekiere, A., De Bourdeaudhuij, I., Deforche, B., Van de Weghe, N., & Van Cauwenberg, J. (2016). Which environmental factors most strongly influence a street's appeal for bicycle transport among adults? A conjoint study using manipulated photographs. International Journal of Health Geographics.

Millard-Ball, A. (2019). Map-matching poor-quality GPS data in urban environments: the pgMapMatch package. *Transportation Planning and Technology*, 42(6), pp.539–553. Available at: https://ideas.repec.org/a/taf/transp/v42y2019i6p539-553.html

Motoaki, Y. and Daziano, R.A. (2015) 'A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand,' *Transportation Research. Part a, Policy and Practice*, 75, pp. 217–230. https://doi.org/10.1016/j.tra.2015.03.017.

Municipality Dalfsen. (2020). Integrale Fietsvisie gemeente Dalfsen

NDW (n.d.). *Verkeersintensiteiten voor verkeersveiligheidsanalyses*. [online] Available at: https://www.ndw.nu/onderwerpen/verkeersveiligheid/intensiteitsgegevens-voor-verkeersveiligheid

Nijen, Van. N. (2022) The influence of infrastructure and land use allocation on the route choice of cyclists avaliable at: https://essay.utwente.nl/91571/1/Nijen_Nick_van.pdf

NOS (2024). *Weer meer fietsers omgekomen, organisaties willen meer helmen zien*. Available at: https://nos.nl/artikel/2516186-weer-meer-fietsers-omgekomen-organisaties-willen-meer-helmen-zien

Openstreetmap. (n.d.). Openstreetmap. Retrieved from Openstreetmap: https://www.openstreetmap.org/#map=19/52.22168/6.86932

Passmore, R., Watkins, K. and Guensler, R. (2024). Using shortest path routing to assess cycling networks. Journal of transport geography, 117, pp.103864–103864. doi:https://doi.org/10.1016/j.jtrangeo.2024.103864.

Paterson, R. W. (1949). The control of urban building development: A review of the floor space index and the code of daylighting control. Planning Outlook Series 1, 1(3), 39–49. https://doi.org/10.1080/09640564908730484

PDOK. (n.d.). Dataset: Luchtfoto. Retrieved from PDOK: https://www.pdok.nl/-/nu-hoge-resolutie-luchtfoto-2023-bij-pdok

Planbureau voor de Leefomgeving. (2022). *RUDIFUN 2022: Ruimtelijke dichtheden en functiemenging in Nederland*. Available at: https://www.pbl.nl/publicaties/rudifun-2022-ruimtelijke-dichtheden-en-functiemenging-in-nederland

Pettorelli, N., Jon Olav Vik, Atle Mysterud, Gaillard, J.-M., Tucker, C.J. and Nils Chr. Stenseth (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, [online] 20(9), pp.503–510. doi:https://doi.org/10.1016/j.tree.2005.05.011.

Prato, C. G., Halldórsdóttir, K., & Nielsen, O. A. (2018). Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space. International Journal of Sustainable Transportation, 770-781.

Province of Drenthe. (n.d.). *Doorfietsroute 'De Groene As' Assen – Groningen*. Assen

Province of Noord-Brabant. (2016). *Uitvoeringsprogramma Fiets in de Versnelling*. 's Hertogenbosch

RUDIFUN. (2022)). Dataportaal Downloads. Retrieved from: Planbureau voor de Leefomgeving: https://dataportaal.pbl.nl/downloads/RUDIFUN1/

Saelens, B., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of Walking and Cycling: Findings From the Transportation, Urban Design, and Planning Literatures. Annals of Behavioral Medicine, 80-91

Schepers, J., Kroeze, P., Sweers, W., & Wüst, J. (2011). Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. Accident Analysis & Prevention, 43(3), 853-861. doi:10.1016/j.aap.2010.11.005

Stevenson, M. *et al.* (2016) 'Land use, transport, and population health: estimating the health benefits of compact cities,' *Lancet*, 388(10062), pp. 2925–2935. https://doi.org/10.1016/s0140-6736(16)30067-8.

Stinson, M. A., & Bhat, C. R. (2003). An Analysis of Commuter Bicyclist Route Choice Using a Stated Preference Survey. Transportation Research, 107-115.

Strauss, J., & Miranda-Moreno, L. F. (2013). Spatial modeling of bicycle activity at signalized intersections. The Journal of Transport and Land Use, 47-58.

Uttley, J., Fotios, S., & Lovelace, R. (2020). Road lighting density and brightness linked with increased cycling rates after-dark. PLoS ONE.

Veenstra, S., Geurs, K., Thomas, T., & Van den Hof, R. (2016). Alle lichten op groen voor fietsmonitoring in Enschede. Verkeerskunde.

Wall, S., Lee, D., Frangos, S. G., Sethi, M., Heyer, J. H., Ayoung-Chee, P., & Dimaggio, C. J. (2016). The Effect of Sharrows, Painted Bicycle Lanes and Physically Protected Paths on the Severity of Bicycle Injuries Caused by Motor Vehicles. Safety, 2(4). doi:10.3390/safety2040026

Wegman, F. and Schepers, P. (2024) 'Safe System approach for cyclists in the Netherlands: Towards zero fatalities and serious injuries?,' *Accident Analysis and Prevention*, 195, p. 107396. https://doi.org/10.1016/j.aap.2023.107396.

Wei, F. and Lovegrove, G. (2013). An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. *Accident analysis and prevention*, [online] 61, pp.129–137. doi:https://doi.org/10.1016/j.aap.2012.05.018.

Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2010). Built Environment Influences on Healthy Transportation Choices: Bicycling versus Driving. Journal of Urban Health, 969-993.

Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling: comparing influences on decisions to ride. Transportation, 153-168.

Yunex Traffic Netherlands. (2024). *Talking Bikes*. [online] Available at: https://nl.yunextraffic.com/projecten/talking-bikes/

Zhao, Y., Ke, S., Lin, Q., & Yu, Y. (2020). Impact of land use on bicycle usage: A big data-based spatial approach to inform transport planning. Journal of Transport and Land Use, 299-316.

7. Appendix

I. Script filter Groningen

```python
# -*- coding: utf-8 -*-
"""
Spyder Editor

This is a temporary script file.
"""

import pandas as pd
import geopandas as gpd
from os import listdir
from os.path import isfile, join
import time
start_time = time.time()

#%% Get list of all files
path = r"C:\Users\luukt\Documents\RUG\SPD3\Bacholer Project SPD\Data\Data_python"
onlyfiles = [f for f in listdir(path) if isfile(join(path, f))]

onlyfiles = onlyfiles[1:]

#%% Process data
GPS = gpd.read_file('Talking_bikes_data/2020-09-30.csv')
GPS = GPS[GPS['field_1']=='Ring-Ring']

GPS=GPS.rename(columns={'field_1':'SuppID','field_2':'TripID','field_3':'RouteID','field_4':'Timestamp','field_5':'Index', 'field_6':'Lat', 'field_7':'Lon', 'field_8':'Heading', 'field_9':'Speed', 'field_10':'Mode', 'field_11':'Accuracy'})

for i in onlyfiles:
    print(i)
    path = 'Data_python/'+str(i)
    df = gpd.read_file(path)
    df=df.rename(columns={'field_1':'SuppID','field_2':'TripID','field_3':'RouteID','field_4':'Timestamp','field_5':'Index', 'field_6':'Lat', 'field_7':'Lon', 'field_8':'Heading', 'field_9':'Speed', 'field_10':'Mode', 'field_11':'Accuracy'})
    df = df[df['SuppID']=='Ring-Ring']
    GPS = pd.concat([GPS, df])

GPS['Lon'] = GPS['Lon'].astype(float)
GPS['Lat']= GPS['Lat'].astype(float)

```

```
GPS = GPS[(GPS['Lon'] >6.40) & (GPS['Lon'] <6.80)]
GPS = GPS[(GPS['Lat'] <53.33) & (GPS['Lat'] >53.10)]

GPS.to_csv('Ring_Ring_Groningen.csv')

end_time = time.time()
verschil = end_time - start_time
print(verschil)
```

II. Script GPS accuracy

```python
# -*- coding: utf-8 -*-
"""
Created on Sun May 12 10:46:51 2024

@author: luukt
"""

from datetime import datetime
import pandas as pd
import numpy as np

# Functie om de tijd in seconden te converteren
def convert_to_seconds(timestamp):
    return datetime.strptime(timestamp, '%Y-%m-%d %H:%M:%S').timestamp()

# Functie om te controleren of de punten binnen 20 seconden van elkaar zijn genomen
def check_time_difference(points):
    timestamps = [convert_to_seconds(point[2]) for point in points]
    return all(timestamps[i] - timestamps[i-1] <= 20 for i in range(1, len(timestamps)))

# Lees de GPS-gegevens van het bestand
data = pd.read_csv(r"C:\Users\luukt\Documents\RUG\SPD3\Bacholer Project SPD\Data\Ring_Ring_Groningen_to_filter.csv")

# Maak een lege lijst om de gefilterde gegevens op te slaan
filtered_data = []

#lijst met alle uniqe tripIDs
unique_id = np.unique(data['TripID'])

for ID in unique_id[1:]:
    data_trip = data[data['TripID'==ID]]


for i in range(len(data)):
    print(i)
    TripID = data.loc[i,'TripID']


# Maak variabelen om de huidige rit-ID en punten op te slaan
current_ride_id = None
current_ride_points = []
# Loop door de lijnen en verwerk de gegevens
for line in lines:
```

```python
    data = line.strip().split(',')
    ride_id = data[1]

    # Als het rit-ID verandert, controleer dan of de punten binnen 20 seconden zijn genomen
    if ride_id != current_ride_id:
        if current_ride_points:
            if check_time_difference(current_ride_points):
                filtered_data.extend(current_ride_points)
        current_ride_id = ride_id
        current_ride_points = []

    current_ride_points.append(data)

# Voeg de laatste set punten toe aan de gefilterde gegevens
if current_ride_points:
    if check_time_difference(current_ride_points):
        filtered_data.extend(current_ride_points)

# Schrijf de gefilterde gegevens naar een nieuw bestand
with open('filtered_gps_data.txt', 'w') as file:
    for data in filtered_data:
        file.write(','.join(data) + '\n')

print("Filtering completed.")
```

III. GPS outliners

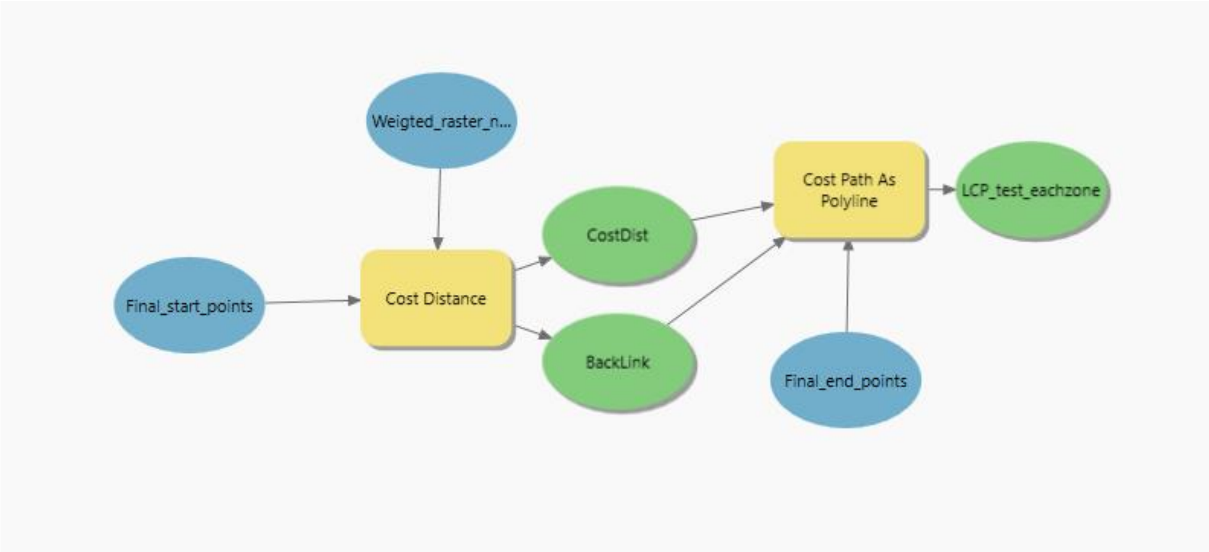
```python
# -*- coding: utf-8 -*-
"""
Created on Sun May 12 14:48:22 2024

@author: luukt
"""

from datetime import datetime
from math import radians, sin, cos, sqrt, atan2

# Functie om de afstand tussen twee punten te berekenen met de haversine-
formule
def calculate_distance(lat1, lon1, lat2, lon2):
    R = 6371.0  # straal van de aarde in km

    lat1_rad = radians(lat1)
    lon1_rad = radians(lon1)
    lat2_rad = radians(lat2)
    lon2_rad = radians(lon2)

    dlon = lon2_rad - lon1_rad
    dlat = lat2_rad - lat1_rad

    a = sin(dlat / 2)**2 + cos(lat1_rad) * cos(lat2_rad) * sin(dlon / 2)**2
    c = 2 * atan2(sqrt(a), sqrt(1 - a))

    distance = R * c * 1000  # Afstand in meters
    return distance

# Functie om de tijd in seconden te converteren
def convert_to_seconds(timestamp):
    return datetime.strptime(timestamp, '%Y-%m-%d %H:%M:%S').timestamp()

# Lees de GPS-gegevens van het bestand
with open('gps_data.txt', 'r') as file:
    lines = file.readlines()

# Maak lege lijsten om de rit-ID's op te slaan die aan de criteria voldoen
within_222m = []
over_222m = []

# Maak variabelen om de vorige rit-ID en punten op te slaan
prev_ride_id = None
prev_lat = None
prev_lon = None
prev_time = None
```

```python
# Loop door de lijnen en verwerk de gegevens
for line in lines:
    data = line.strip().split(',')
    ride_id = data[1]
    lat = float(data[5])
    lon = float(data[6])
    timestamp = data[3]

    # Als dit niet de eerste punt is, bereken dan de afstand en controleer de
tijd
    if prev_ride_id is not None and ride_id == prev_ride_id:
        distance = calculate_distance(prev_lat, prev_lon, lat, lon)
        time_diff = convert_to_seconds(timestamp) -
convert_to_seconds(prev_time)

        # Als de afstand meer dan 222 meter is en de tijd minder dan 20
seconden is, voeg dan toe aan over_222m
        if distance > 222 and time_diff < 20:
            over_222m.append(ride_id)
        else:
            within_222m.append(ride_id)

    prev_ride_id = ride_id
    prev_lat = lat
    prev_lon = lon
    prev_time = timestamp

# Schrijf de resultaten naar twee aparte bestanden
with open('within_222m.txt', 'w') as file_within, open('over_222m.txt', 'w')
as file_over:
    for ride_id in set(within_222m):
        file_within.write(ride_id + '\n')
    for ride_id in set(over_222m):
        file_over.write(ride_id + '\n')

print("Processing completed.")
```

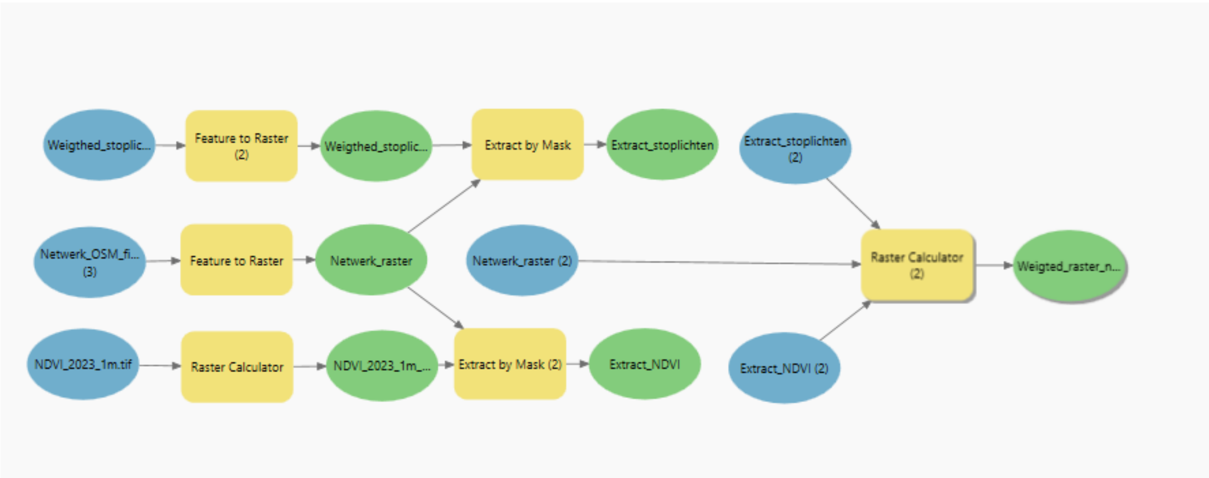
IV. ArcGIS Pro Models

For full details see ArcGIS Pro Project: Bapro_Route_Choice_Luuk_ten_Berge

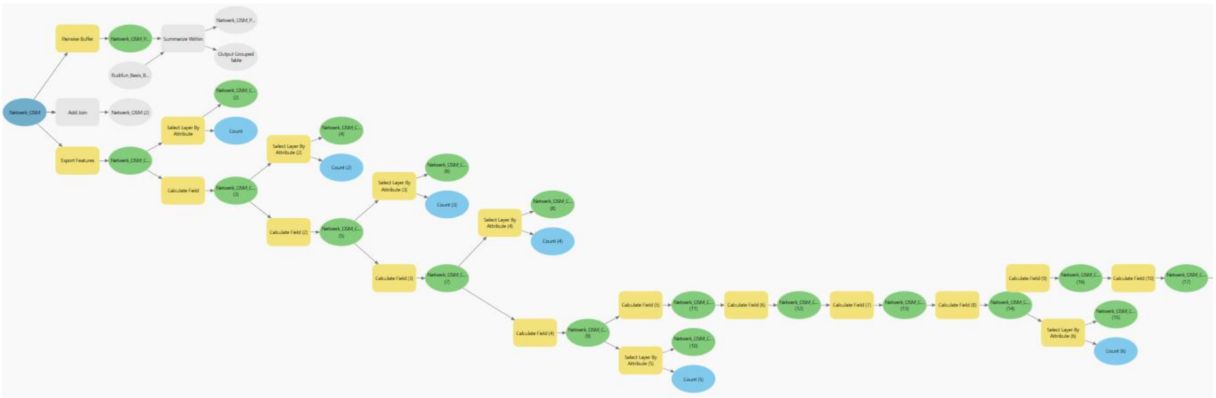
Constructing LCP



WRM (cell size)



Infrastructure factors



Traffic Lights

V. Script OD-matrix

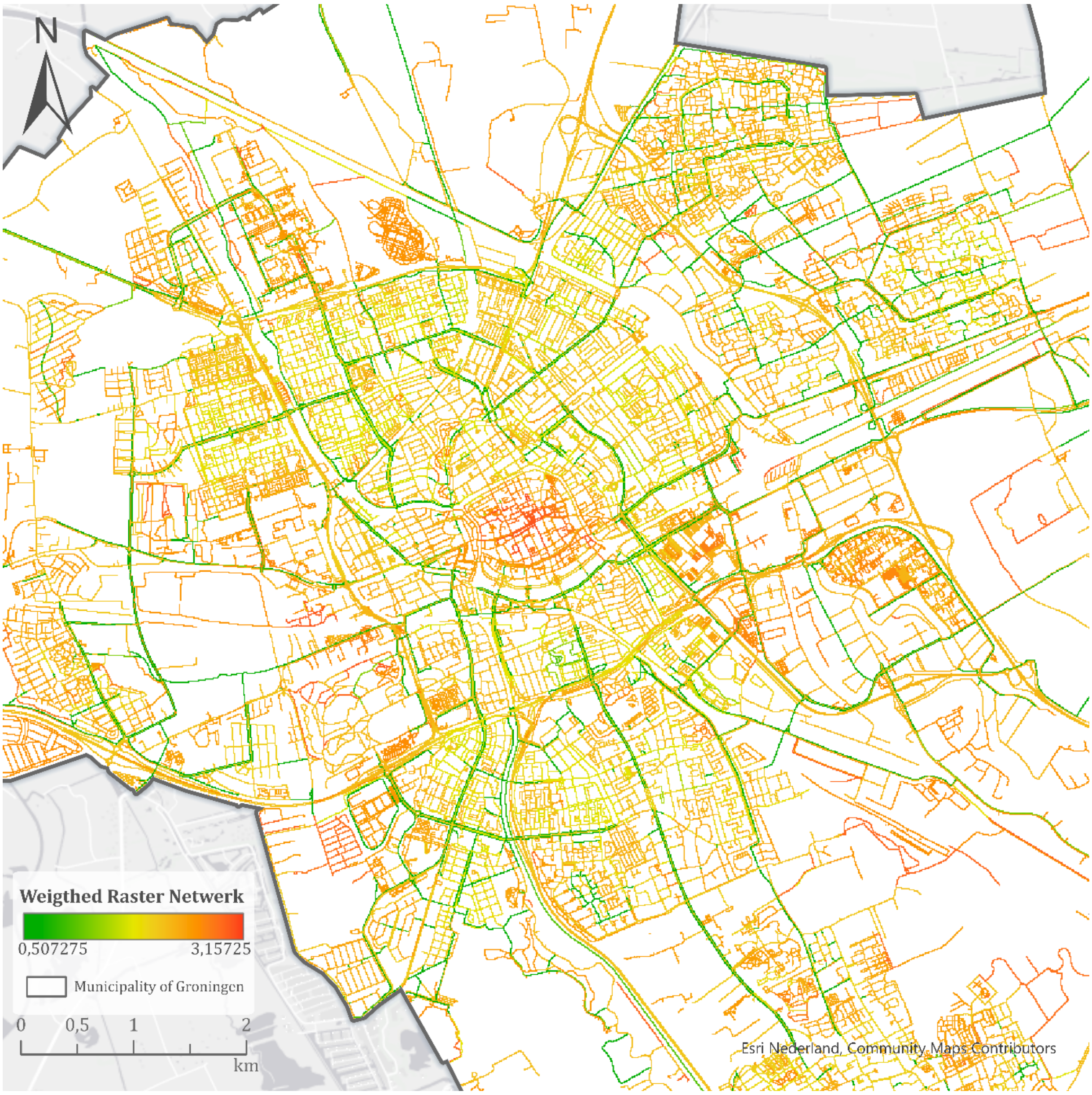
```
import pandas as pd

# Lees de dataset in een DataFrame
df = pd.read_csv(r"C:\Users\luukt\Documents\RUG\SPD3\Bacholer Project
SPD\Data\Filtereddata.csv")

# Controleer de eerste paar rijen van de dataset om te bevestigen dat het
correct is ingelezen
df = df[df.duplicated('TripID', keep=False)]

# Controleer of de vereiste kolommen aanwezig zijn in de dataset
required_columns = {'TripID', 'Timestamp', 'Lat', 'Lon'}
if not required_columns.issubset(df.columns):
    print(f"De dataset mist een of meer vereiste kolommen:
{required_columns}")
    exit()

# Sorteer de dataset op TripID en Timestamp
df = df.sort_values(by=['TripID', 'Timestamp'])

# Haal het eerste en laatste punt per TripID
start_points = df.groupby('TripID').first().reset_index()
end_points = df.groupby('TripID').last().reset_index()

# Optioneel: schrijf het resultaat naar een nieuw CSV-bestand met expliciete
paden
start_points_path = r"C:\Users\luukt\Documents\RUG\SPD3\Bacholer Project
SPD\Data\start_points.csv"
end_points_path = r"C:\Users\luukt\Documents\RUG\SPD3\Bacholer Project
SPD\Data\end_points.csv"

start_points.to_csv(start_points_path, index=False)
end_points.to_csv(end_points_path, index=False)

print(f"Startpunten zijn opgeslagen in: {start_points_path}")
print(f"Eindpunten zijn opgeslagen in: {end_points_path}")
```

VI. Table total costs OBR, SHP, LCP

TripID	Tripid_new	SUM_OBR	SUM_SHP	SUM_LCP
0d6ae944-a525-0305-ac3b-3aed6afb161cbc614d68	T0010	1088,518679	957,8680608	509,7865933
143e74b5-948a-b46d-4a60-d613598035ff2ecd7981	T0020	207,0591826	298,7728276	120,689545
1a95ad3b-ac58-87ee-d918-a6832d1cd33ccdd2bf8d	T0030	895,9381703	934,5866817	475,9660634
1dd63a40-2122-81f6-864b-7b56f79149e051d17edc	T0040	1622,11638	1177,862701	744,0534206
1dd78b1d-6af5-2f61-7e85-b3e26750f33db0070919	T0050	419,3263944	477,9522129	237,8369535
238aac1e-ad6b-3f8d-5095-3658236d2d4840d3a6e9	T0060	243,4918029	215,2690736	138,5467929
3c889952-c9a2-e36d-4e87-6572041aeaf34e54f6a1	T0070	246,2354508	157,3437543	93,75132722
3e4cf57f-0515-fe4b-624f-19d6c1c8edd7c359d65c	T0080	1299,4487	896,536294	461,3234606
42a8478b-7cc5-d144-5aff-ef1eaf37971375e2ba6b	T0090	230,728299	252,6245165	149,1146748
457a5869-1c42-4db8-170c-bde2c98e1ee858df52a9	T0100	441,3623686	487,7013878	258,6921692
4ac6f7af-25ef-f525-8302-26aa8c2986dbe647f7d8	T0110	1096,654463	965,3473746	484,0587862
4c7f566b-a3e5-ba67-50b0-b7b9e8a9d5b2ef372f3d	T0120	365,1569142	327,8795715	208,0600721
4de1b25e-9165-931d-71e5-446f726ac5918a45780a	T0130	969,4022638	1042,444798	513,58025
539759da-73e1-8100-ab24-092d3f5600e673f617be	T0140	930,6831622	1251,834887	541,0753071
54681860-d41e-a2d0-9fbe-4673cb5bca759f1c16d2	T0150	332,0484993	256,9191101	175,9181417
5d160d24-6f2a-66cd-0d17-5d9d57cfb904d2ce1147	T0160	880,8481557	979,0226359	521,4216305
5d6876bf-9fa5-7d3a-0593-4542bff76148e40f4cff	T0170	989,0209373	1173,326217	545,7245039
6134435b-1fc0-31d4-0cb4-db1c80c4c7483f8b70ca	T0180	1092,020526	869,8053042	427,7745906
6a9bc95f-08cc-5a17-e948-6d2d68031fcc737d76d3	T0190	932,3134571	1261,754658	547,1673
6bc7ae6b-0251-eb65-9993-b90e14fd115dc1b29aed	T0200	1021,469941	1181,146429	544,4203334
6c887b64-f4ba-1faf-22a4-681d9e992bc520204b84	T0210	413,052555	487,826586	225,1041449
6dbc961f-92cf-29d8-86cc-681ba035b527b06ec73c	T0220	412,9561905	506,831643	231,9777582
6eab270c-663b-b8bb-fb1e-54878175109bd867f8ea	T0230	998,2649668	1173,326217	545,7245039
70369e79-d883-c7ff-87d6-dca1db6a8e95b5cf9f83	T0240	312,10585	273,1518314	217,6549649
762ee9b2-1afa-dee6-d946-9527cc8cde90607a774a	T0250	323,3794422	247,463941	166,8883695
76566a93-aeee-1e08-d23e-d42b497eca99f1df4a57	T0260	802,0453876	943,512891	405,075812
7b09b664-6b44-bade-a65c-7528d94c0efa4f809b27	T0270	965,788507	733,5884972	400,4099928

7ed17fbf-f2da-9d8c-6b1a-387d0c9d4d18635c23aa	T0280	720,9950914	811,8982025	464,7065684
802feb71-a719-74c9-b443-5bfe772d14ced3cd5697	T0290	509,3020127	542,0509877	321,5754206
868247dd-54bd-c8c3-5218-1e60c5f0bcbd0ca2044e	T0300	1177,448565	1146,341564	527,9853438
8963291e-00b6-5291-0d0a-bc01f42e2b8ea4dfefdb	T0310	915,8504372	1181,236883	544,2513785
8a298a0e-c58d-d9a8-420c-ab74b8fed568e32f3600	T0320	198,701373	164,9098952	133,042483
8ef3a690-e0af-2c71-033f-c4525eb06165a957192e	T0330	826,3937545	801,3345205	418,5224461
915fb74b-3d85-75a5-0480-d26f946f07ac29471edb	T0340	204,9260967	191,7903026	113,6066225
91e6624e-44b5-8c9f-fad0-ec793192cd25cdf6e5c0	T0350	997,0245292	927,5891052	464,7401816
97b005c7-4737-64d1-c90d-8d84cbd939cd961169ed	T0360	280,7693009	203,6662866	165,4713252
996d87d8-2930-5085-cadc-37b67716ee830035c9cd	T0370	397,3826416	441,7635336	222,9439801
b40574d1-cd18-4930-9b79-631cfb3fc4bb5b503727	T0390	295,5754784	298,0872377	167,7371795
c1996df3-14d5-00f0-7c56-0d771ddb354d8548528d	T0400	401,352461	463,9965084	237,712817
c3fda329-96bd-3868-7838-4d43d1d6c7b5fc4843cb	T0410	518,9651961	476,1268066	270,3421174
c8a40cd9-96d6-6501-60b0-566d4ff60d45d1ad13aa	T0420	167,9212124	291,425393	150,4539759
cc490a40-f1e5-027d-3bdd-a509d2453f536cc3059c	T0430	824,7576436	814,8477222	417,7158016
cd173c11-d40b-5c3e-7207-afa4abc12acdf639b217	T0440	983,4050474	1179,255085	544,4203334
e3f30672-8c53-1d44-ec20-1739265d8d7f5c03b10a	T0450	1185,076687	1139,191875	532,9424464
ec00b473-896a-3bde-a21e-27ac66abfc4fa93b79a5	T0460	933,0927509	1022,332743	471,2423012
ec70bdf9-f3d5-8253-723f-b3c95d445ec85c5be265	T0470	938,1896946	1242,537318	535,0670352
ecbc2faf-7e5d-023c-b05f-851d5e5cdd89ccac54f0	T0480	221,2682557	289,4752586	124,6561643
f77176eb-baa4-21fd-7300-9c03ee6cda9e5c938ef4	T0490	966,3006191	1171,019329	543,2856418
f7b85046-f4f1-2fd2-47b6-9c5371a6cf300faa20f8	T0500	1028,942433	1250,08412	542,8221455
fd08f1d6-f4f4-b103-9ba0-bb9709cf1db44eee095e	T0510	214,6386555	258,2323705	148,0971571

VII. Weigthed Raster Netwerk city of Groningen



VIII. Table overlap separate bicycle path

Tripid _new	Shape_Length_OBR	Length_seperate_OBR	Shape_Length_SHP	Length_seperate_SHR	Shape_Length_LCP	Length_seperate_LCP
T0010	5795,010694	5795,010694	5119,348013	4418,15953	6362,396821	6315,057415
T0020	940,8396828	940,8396828	1133,951985	61,56568375	775,2691193	145,8316715
T0030	4020,799249	4020,799249	3754,195327	1571,426905	4155,584412	2934,314128
T0040	8596,394352	8596,394352	6960,464475	5111,571086	7491,930009	6210,954117
T0050	2445,48217	2436,857806	2467,16446	1146,178128	2599,655121	2486,049475
T0060	1081,304617	1081,304617	873,2210254	244,5334017	830,1219331	131,8768461
T0070	1068,767322	1068,767322	688,9009892	86,76403727	756,6904756	385,0035412
T0080	4972,317211	4972,317211	4190,033182	2237,138492	4722,447327	4217,566746
T0090	1028,669199	1028,669199	935,4587791	61,56568375	823,5533906	50,1720283
T0100	2505,420822	2505,420822	2502,279364	1146,178128	2693,797257	2486,049475
T0110	4464,609362	4464,609362	3845,558678	1619,267523	4177,300141	2865,608939
T0120	1604,205591	1604,205591	1395,883187	345,0496317	1472,670273	720,5130568
T0130	5007,826894	5007,826894	4830,137444	2583,991717	5291,56421	4935,302392
T0140	5807,700551	5807,700551	5382,959802	1472,54951	6035,828278	5452,294194
T0150	1567,381516	1567,381516	1455,10357	510,3590598	1567,228714	971,8948562
T0160	4469,159001	4469,159001	3865,885595	897,8969651	3767,47258	2146,641494
T0170	6565,867486	6565,867486	5097,16232	1524,804995	6120,681092	5579,283655
T0180	4418,940929	4418,940929	3535,672088	1619,267523	3904,163056	2952,481179
T0190	5843,171414	5843,171414	5442,07501	1524,804995	6089,970414	5510,578465
T0200	6713,398095	6713,398095	5124,896015	1543,747002	6128,965363	5603,42579
T0210	2277,677092	2277,677092	2320,101551	995,8167137	2464,802307	2372,496484
T0220	2357,167116	2357,167116	2410,297804	1118,727922	2563,086579	2485,222231
T0230	6313,763782	6313,763782	5097,16232	1524,804995	6120,681092	5579,283655
T0240	968,723584	968,723584	945,2413088	71,80820912	962,54834	65,03871315
T0250	1501,615433	1501,615433	1391,807655	510,3590598	1504,802307	971,8948562
T0260	3913,294928	3913,294928	3821,655465	1607,219619	3818,305192	2865,608939
T0270	5786,130269	5786,130269	3791,638237	2481,474911	4029,604615	3710,919564
T0280	4705,391874	4705,391874	4252,191701	3424,044453	5161,025971	4636,821917
T0290	2369,200753	2369,200753	2304,666621	600,6802957	2489,066376	1502,587178
T0300	5282,287442	5282,287442	4973,663738	1472,54951	5989,970414	5520,999384
T0310	5825,68625	5825,68625	5132,386356	1543,747002	6094,11255	5534,720601
T0320	936,640666	936,640666	817,208203	292,5372	787,6955262	271,590485
T0330	3956,894009	3956,894009	3700,521784	1258,061835	4349,188309	4011,842409
T0340	1269,563467	1269,563467	1249,732554	1243,709028	1344,680374	1338,384548
T0350	4594,161769	4594,161769	3735,130757	1619,267523	4111,442277	2934,314128
T0360	1256,980995	1256,980995	1037,630789	258,6598483	1021,543289	221,1755825
T0370	2280,744035	2280,744035	2315,506115	1042,77762	2474,802307	2396,93796
T0390	1454,70088	1454,70088	1209,215959	220,0179249	1081,248917	216,7044651
T0400	2380,485758	2380,485758	2424,805236	1133,457591	2601,37085	2508,934095
T0410	2253,204303	2253,204303	2181,267685	600,6802957	2182,792206	1433,881989
T0420	1348,993428	1348,993428	1501,815707	703,4202284	1896,812409	1783,918526

T0430	3766,580519	3766,580519	3361,724406	897,8969651	3333,624817	2146,641494
T0440	6274,301428	6274,301428	5115,216698	1543,747002	6128,965363	5603,42579
T0450	6522,593946	6522,593946	4978,701736	1543,843173	6029,970414	5600,158162
T0460	4107,959073	4107,959073	4099,258371	1611,152814	4113,158005	2865,608939
T0470	5829,801655	5829,801655	5359,339719	1472,54951	6001,686143	5452,294194
T0480	974,9303211	974,9303211	1106,242752	61,56568375	765,9797975	179,7462525
T0490	5825,724462	5825,724462	5094,751274	1543,747002	6120,681092	5603,42579
T0500	6424,24174	6424,24174	5407,506055	1524,804995	6069,970414	5510,578465
T0510	1097,214252	1097,214252	1148,775884	781,1688578	1128,111832	862,5878239

Tripid_new	%overlap OBR	%overlapSHP	%overlapLCP
T0010	0,752277	0,863032	0,99256
T0020	0,144768	0,054293	0,188105
T0030	0,355356	0,418579	0,706113
T0040	0,732934	0,734372	0,829019
T0050	0,525689	0,464573	0,9563
T0060	0,176568	0,280036	0,158864
T0070	0,074956	0,125946	0,508799
T0080	0,515218	0,533919	0,893089
T0090	0,12346	0,065813	0,060921
T0100	0,537174	0,458054	0,922879
T0110	0,251901	0,421075	0,685995
T0120	0,392126	0,247191	0,489256
T0130	0,600825	0,534973	0,932674
T0140	0,782037	0,273558	0,903322
T0150	0,286431	0,350737	0,620136
T0160	0,604253	0,232262	0,569783
T0170	0,814659	0,299148	0,911546
T0180	0,254423	0,45798	0,756239
T0190	0,78042	0,280188	0,904861
T0200	0,826758	0,301225	0,914253
T0210	0,530338	0,429213	0,96255
T0220	0,563338	0,464145	0,969621
T0230	0,761221	0,299148	0,911546
T0240	0,090324	0,075968	0,067569
T0250	0,300234	0,366688	0,645862
T0260	0,450132	0,420556	0,750492
T0270	0,696352	0,65446	0,920914
T0280	0,646065	0,805242	0,89843
T0290	0,308041	0,260637	0,603675
T0300	0,409175	0,296069	0,921707
T0310	0,786726	0,300785	0,908208

T0320	0,430039	0,357971	0,344791
T0330	0,455566	0,339969	0,922435
T0340	0,994748	0,99518	0,995318
T0350	0,370598	0,433524	0,713695
T0360	0,179588	0,249279	0,216511
T0370	0,542423	0,450345	0,968537
T0390	0,22658	0,181951	0,200421
T0400	0,559707	0,467443	0,964466
T0410	0,338121	0,275381	0,656903
T0420	0,674421	0,46838	0,940482
T0430	0,510102	0,267094	0,643936
T0440	0,767872	0,301795	0,914253
T0450	0,723173	0,31009	0,928721
T0460	0,348181	0,393035	0,696693
T0470	0,774571	0,274763	0,90846
T0480	0,143367	0,055653	0,234662
T0490	0,770887	0,303007	0,915491
T0500	0,749843	0,281979	0,907843
T0510	0,686825	0,680001	0,76463

IX Table Ratio costs

Tripid_new	SHP-OBR	LCP-OBR	SHP-LCP
T0010	-0,12003	-0,53167	-0,87896
T0020	0,442934	-0,41713	-1,47555
T0030	0,043137	-0,46875	-0,96356
T0040	-0,27387	-0,54131	-0,58304
T0050	0,13981	-0,43281	-1,00958
T0060	-0,11591	-0,431	-0,55376
T0070	-0,361	-0,61926	-0,67831
T0080	-0,31006	-0,64499	-0,9434
T0090	0,0949	-0,35372	-0,69416
T0100	0,104991	-0,41388	-0,88526
T0110	-0,11973	-0,5586	-0,99428
T0120	-0,10209	-0,43022	-0,57589
T0130	0,075348	-0,47021	-1,02976
T0140	0,345071	-0,41863	-1,31361
T0150	-0,22626	-0,4702	-0,46045
T0160	0,111454	-0,40805	-0,8776
T0170	0,186351	-0,44822	-1,15003
T0180	-0,20349	-0,60827	-1,03333
T0190	0,353359	-0,41311	-1,30598
T0200	0,15632	-0,46702	-1,16955
T0210	0,181028	-0,45502	-1,16712
T0220	0,227325	-0,43825	-1,18483
T0230	0,175366	-0,45333	-1,15003
T0240	-0,12481	-0,30262	-0,25498
T0250	-0,23476	-0,48392	-0,48281
T0260	0,176383	-0,49495	-1,32923
T0270	-0,24043	-0,58541	-0,83209
T0280	0,12608	-0,35547	-0,74712
T0290	0,064302	-0,3686	-0,68561
T0300	-0,02642	-0,55159	-1,17116
T0310	0,289771	-0,40574	-1,17039
T0320	-0,17006	-0,33044	-0,23953
T0330	-0,03032	-0,49356	-0,91468
T0340	-0,0641	-0,44562	-0,6882
T0350	-0,06964	-0,53387	-0,99593
T0360	-0,27461	-0,41065	-0,23083
T0370	0,111683	-0,43897	-0,9815
T0390	0,008498	-0,43251	-0,77711
T0400	0,156082	-0,40772	-0,95192
T0410	-0,08255	-0,47907	-0,7612
T0420	0,735489	-0,10402	-0,93697
T0430	-0,01202	-0,49353	-0,95072

T0440	0,199155	-0,44639	-1,16607
T0450	-0,03872	-0,55029	-1,13755
T0460	0,095639	-0,49497	-1,16944
T0470	0,324399	-0,42968	-1,32221
T0480	0,308255	-0,43663	-1,32219
T0490	0,211858	-0,43777	-1,15544
T0500	0,214921	-0,47245	-1,30294
T0510	0,203103	-0,31002	-0,74367