Assessing the impact of bicycle infrastructure on cyclists' route choice in the Dutch province of Noord-Brabant

Student
Kai Meijning
k.meijning@gmail.com
Solis-id: 6023703

Supervisor
Dr.C. (Kees) Maat

Responsible professor
Prof.dr.ir. P.J.M. van Oosterom

04-06-2019

Word count: 13343

Table of Content

Foreword

Since I started the bachelor Human Geography in 2013, I have been particularly interested in transport issues. When I am cycling through Amsterdam, I think about how infrastructure could be optimized or why certain cycle routes are more crowded than others. I always wanted to do more research into transportation issues in relation to geography, but these wishes were unfortunately not fulfilled during my bachelor.

When I had to select a research topic for my master thesis, I looked through the possible topics and found the topic on route choice of cyclists. This topic immediately peeked my interest and after discussing with my supervisor, I decided to do research on the impact of infrastructure on cyclist route choice. These decisions have now finally led to the finished product that I hereby present.

This thesis would have not been possible without the required data. Thus, I first want to thank the Fietsersbond for supplying the road network that was used. This network also contained a lot of information about infrastructural factors. Secondly, the cycling GPS data that was supplied by the B-riders program was also crucial for the analyses. Thirdly, I want to thank the researchers at the Breda University of Applied Sciences and in particular Paul van Coeverden that linked the B-riders data to the Fietsersbond network. Without this initial preparation of the data this research would have been much more challenging to complete.

I also want to thank my supervisor, Kees Maat from the University of Delft, for his continued support and feedback. When I was stuck on certain matters of my research, Kees offered new insights which allowed me to continue my research on the right path.

Finally, I want to thank my friends and family for supporting me and providing me with feedback throughout the process.

Abstract

The assumption in traditional route choice models focused on motorized vehicles is that people take the shortest available route to their destination. This assumption however, is not applicable to cyclists. For them, other factors might be more important than the shortest distance or shortest travel time. One of these factors is infrastructure. For some cyclists, the infrastructure may be the reason why a certain route is chosen over the shortest route. Understanding cyclist route choice and generating fitting cyclist route choice models is key to introduce new policies that improve the experience for current cyclists and incentivize more people to cycle. By getting more people to use their bicycle as their main mode of transport instead of their car, large reductions to the emission of greenhouses gasses can be achieved while simultaneously improving public health by a large margin.

In this research, GPS data collected by the B-riders program in the Dutch province of Noord-Brabant is used to find out if the route choice behaviour can be explained by infrastructural factors. This route choice behaviour was studied by comparing the difference between the observed routes and the shortest route in distance with the difference in infrastructural factors along these routes. Studies have shown that there are several infrastructural factors that influence cyclist route choice. These are the road type, type of road surface, quality of the road and intersections and traffic signal control. The data on these infrastructural factors was sourced from the Fietsersbond road network and was added to the routes by using several GIS methods.

When comparing the shortest routes to the observed routes, the observed routes were indeed significantly longer. All but one of the infrastructural factors were also significantly different between the two types of routes. However, when adding the amount of deviation and the difference in infrastructural factors to a multiple linear regression model, only a small amount of the difference between the two types of routes could be explained by the infrastructural factors. It turns out that the one infrastructural factor that significantly influences a cyclist route is traffic signal control at an intersection. This is in contrast with previous research on cyclist route choice, where cyclists preferred to cycle on separated bicycle paths and roads with a certain surface and road quality.

The fact that the multiple linear regression models and the infrastructural factors analysed in this research explained only a small amount of cyclist route choice behaviour can perhaps be used an indication that cyclist route behaviour should be researched using different (non) linear models. To better predict cyclist route choice behaviour in the future, other factors such as the type of landscapes and personal characteristics should also be included.

1. Introduction

1.1 Context and problem statement

One of the main challenges that the world is facing today is climate change induced by growing energy consumption. In the Paris Agreement (2015) corresponding parties have acknowledged that the emission of greenhouse gasses needs to decline globally (Rogelj et al., 2016). One of the main areas the targets and actions are related to is the transportation sector. This sector accounts for 26 percent of global carbon dioxide emissions and it is one of the few sectors where emissions are still growing (Chapman, 2007).

In the transportation sector, the motor car is the second biggest contributor after road freight to greenhouse gas emissions. Passenger cars are increasingly the preferred mode of transport, while they are by far the most damaging to the environment. The potential for alternative modes of transport are, however, abundant. It is estimated that for 40 percent of current passenger car trips a viable alternative is available (Chapman, 2007). Herein, cycling is the most environmental friendly and viable alternative. Furthermore, many other direct and indirect positive effects have been associated with cycling (de Nazelle et al., 2011) such as improvement of public health, decrease in obesity rates and less road injuries. Therefore, scholars and policy makers call for the need to encourage the usage of bicycles as a mode of transport.

Necessary to the understanding of how to encourage cycling as a daily mode of transport are insights into preferred infrastructure for cyclists. When good infrastructure for cyclists is available, the threshold for cycling will lower, causing more people to use their bicycle instead of their car (Sener, Eluru, & Bhat, 2009). However, the knowledge on how bicycle infrastructure influences cyclist route choice behaviour is still fairly limited and little attention has been paid to cycling in comparison to other modes of transport (Dill, 2009; Heinen, van Wee, & Maat, 2010). There are many existing traditional route choice models for motorized traffic available but these cannot be used for cyclists as these models almost always assume that traffic runs along the fastest roads and that drivers take the shortest path. This is not the case for cyclists (Winters, Teschke, Grant, Setton, & Brauer, 2010). Furthermore, accurate GPS data that can be used in cyclist route choice research has only recently become widely available (Broach, Dill, & Gliebe, 2012; Dill & Gliebe, 2008; Hood, Sall, & Charlton, 2011). In comparison to cyclist route choice research employing theoretical data, results from research employing GPS data is much more reliable and devoid of strategic bias (Broach et al., 2012). This ultimately means that the results from the research perform much better in real world situations compared to research based on theoretical data. One of the earliest studies employing actual GPS data was conducted by Menghini et al. (2010) in the Swiss city of Zürich. They revealed that besides an aversion for signal-controlled junctions, cyclists also avoided steep maximum gradients in their route. Since then, a few other studies employing GPS data have been conducted. A study by Broach et al (2012) in the state of Oregon in The United States revealed that, besides travel time, distance, the frequency of turns, slope, the presence of traffic lights along the route, traffic intensity and the presence of cycling infrastructure such as cycle paths or cycle boulevards influenced a cyclists' route choice.

Taking the above facts into consideration, there is a relatively small amount of cyclist route choice research available in comparison to route choice research focused on motorized traffic. In addition, because GPS datasets have only recently become available for use in large scale cyclist route choice research, this type of research is still quite underrepresented in academic literature. Therefore, this research will try to fill these gaps by looking at the influence of infrastructural factors on cyclist route choice in The Netherlands. The Netherlands is an interesting case as it is the country that leads that world when it comes to cycling with 26 percent of all trips made by bicycle (Schepers, Twisk, Fishman, Fyhri, & Jensen, 2017). By using up to date and accurate GPS data to determine the impact of infrastructural factors on cyclist route choice, this study will attempt to fill these research gaps and simultaneously try to provide solutions that will increase the use of bicycles as a mode of transport.

This research starts with an introduction of the research objectives, research questions and research scope. Following this will be the theoretical framework. In this theoretical framework, general theory on cyclist route choice will be introduced. In addition to general theory on cyclist route choice, different theories on which infrastructural factors influence cyclist route choice will be laid out, creating a solid base for the final analysis of the results. After the theoretical framework, the methods and data chapter follows. The methods and data chapter includes the study area of this research, an overview of the data, an operationalisation of the data and the types of analyses that are used. Then, the results from the analysis are presented and the hypotheses are tested. Based on these results, the main and sub questions of this research will be answered in the conclusion, after which the research will be finalized with a discussion of the results and possible recommendations for further research on cyclist route choice in The Netherlands and in general.

1.2 Research objective and research questions

The general research objective for this thesis is to find out which infrastructural factors influence a cyclists' decision to take a certain route over other possible alternative routes. By analysing these infrastructural factors, parties like De Fietsersbond, infrastructure planners and urban planners will be able to act on this information by possibly adjusting the current infrastructure or by adding infrastructure that will improve the situation for cyclists. This is in line with current plans by the Dutch government to invest at least 250 million euros in bicycle infrastructure from 2018 to 2040 (Rijksoverheid, 2018). This research will also be able to increase the knowledge on cyclist route choice in The Netherlands, as research on cyclist route choice currently available is mostly conducted in other countries where the context when it comes to cycling is vastly different.

This research will only focus on cycling trips made for commuting purposes. This choice was made because these trips are usually repetitive and made out of necessity, often leading to a cyclists' choice to undertake these kinds of trips in the most efficient way possible (Van Duppen & Spierings, 2013). This makes these trips very suitable for cyclist route choice research. Additionally, increasing the amount of bicycle commuters can create many advantages for a society such as a decrease in pollution and congestion and the improvement of public health (Heinen et al., 2010).

The main research question that accompanies the general research objective and that I will aim to answer in this thesis is:

To what extent do infrastructural factors and network characteristics influence cyclist route choice in The Netherlands?'.

Accompanying the general research objective are two sub questions that need to be answered before the general research objective can be reached and the main research question can be answered. Below, the two sub questions accompanying the main research question are listed.

1. *Which infrastructural and network factors have the biggest effect on cyclist route choice?*
2. *To what extent is there a relationship between the amount of deviation from the shortest routes and the differences in infrastructural factors between the observed and shortest routes?*

The first sub question looks at the nature of the influence that the infrastructural analysed in this research have on cyclist route choice. The second part of this sub question looks at which of these infrastructural factors has the greatest effect on cyclist route choice. The second sub question statistically compares the observed GPS routes collected by participants in the B-Riders program to the shortest routes with the use of the indicators of infrastructural factors. By doing so, it will be investigated if the deviation from the shortest route can be explained by the difference in infrastructural factors on these routes.

1.3 Research scope

With the use of GPS data, this research will investigate how infrastructural factors and network characteristics influence a cyclists route choice when comparing observed GPS routes to shortest routes. This research focuses solely on trips made for commuting purposes. While the outcomes of this research give implications on how infrastructural factors influence cyclist route choice, it does not give direct advice on how the cycling infrastructure in The Netherlands could be improved. With the information from the results of this study, urban planners and infrastructure planners could draw conclusions on how the current cycling infrastructure could be adjusted, but this would have to be derived from current or new policies on cycling infrastructure. Also, this research will not include the underlying motives for cyclists to choose a particular route. This limitation is made due to time constraints as well as a lack of data that is able to represent these motives.

2. Theoretical framework

This theoretical framework provides an overview of current theory on cyclist route choice. As an introduction, more general theory on why cyclists choose a certain route over other routes is provided. Additionally, an overview is given of different infrastructural factors that influence the route choice of cyclists. These include general bicycle infrastructure and bicycle infrastructure that is specific to The Netherlands, intersections and traffic signals, road surface and road quality.

2.1 Route choice behaviour

In route choice modelling, it is widely acknowledged that commuter cyclists choose their route very differently compared to commuters that make use of motorized vehicles (Ehrgott, Wang, Raith, & Van Houtte, 2012). Route choice modelling for motorized vehicles is often modelled with one objective: to minimize general travel cost. This is the combined travel time and vehicle operating costs. Considering this, the shortest route between destinations profoundly is the same or very close to the observed routes taken by commuter drivers. When commuter cyclists choose their route, they consider multiple objectives that their route has to meet. Combined, these objectives can be seen as the suitability of a route for cycling. This suitability can be seen as a combination of the non-subjective factors that influence the suitability of a route for cycling. Factors along a route that can influence this suitability include but are not limited to factors such as safety, traffic volumes, traffic speeds and presence of bicycle infrastructure along a route such as bicycle lanes. Taking these factors that comprise this suitability of a route into account, cyclists generally do not travel on the shortest possible route (Dill, 2009).

The fact that cyclists generally do not take the shortest route between two destinations does not mean that cyclists do not value the objective to minimize their travel time. In some cases, cyclists even place the highest importance on minimizing the travel time between destinations (Dill, 2009). It is possible that sometimes a longer route has a lower travel time than a shorter route. Route decisions are a trade-off between the time a cyclist has, the motives of a cyclist and the need a cyclist feels to choose a route that complies to a particular factor. This behaviour that cyclists exhibit when choosing their route can be found in research by Zimmermann, Mai, & Frejinger (2017). They found that cyclists are highly put off by long distances and that the length of certain parts of a route has a negative influence on cyclist route choice. However, the perceived length of a route is not only determined by absolute length, but also by the characteristics of a route such as slope or bicycle infrastructure. Taking this into account, when choosing their cycling route people always make a trade-off between distance and road characteristics. A road with a high amount of slope will exhibit a negative value, making the road less attractive to a cyclist. However, a road with an available bicycle facility such as a separated cycle path will exhibit a positive value, meaning that this road is perceived to be shorter in distance compared to roads without a separated bicycle path.

This behaviour is one of the main reasons why cyclists are willing to deviate from the shortest route. Cyclists are willing to cycle along routes with certain characteristics. Of these characteristics, infrastructural factors significantly influence a cyclists' decision to cycle along a specific route (Winters et al., 2010). These infrastructural factors include specific bicycle facilities such as bicycle lanes but also more general infrastructural factors such as traffic lights and intersections. The amount of deviation a cyclist is willing to add to its route is relative to the total travel distance or travel time. The deviations that cyclists are willing to take can also be dependent on location. In countries where cycling is less prevalent than in The Netherlands, cyclists may be more willing to deviate from the shortest route to cycle along specific bicycle infrastructure. There also is a difference between the amount of deviation a cyclist is willing to add to its route between different cyclists (Hood et al., 2011). Findings amongst different studies indicate that infrequent cyclists seem to have a higher preference for dedicated bicycle lanes in comparison to frequent cyclists (Broach et al., 2012; Hood et al., 2011; Su, Winters, Nunes, & Brauer, 2010). Some evidence is also available that states that the same applies to women in comparison to men (Winters et al., 2010). Women, in general, avoid risks more often than men and often choose routes where travel conditions are safer (Emond, Tang, & Handy, 2009). However, when looking at the overall literature on cyclist route choice, it is apparent that women will choose to take the shortest route more often than men, who are more willing to take detours (Heinen et al., 2010). In the following sections, the infrastructural factors that influence cyclist route choice decisions will be further elaborated.

2.2 Bicycle infrastructure

In planning practice, it is often assumed that it is safer for cyclists to separate them from other traffic (Heinen et al., 2010). Also, cyclists themselves tend to prefer separate bicycle paths to bicycle lanes on a road shared with cars or roads without bicycle facilities. The reasons for cyclists to cycle along designated bicycle facilities often relate to safety and traffic intensity (Heinen et al., 2010; Winters et al., 2010). Safety can either be considered ‘objective’ or ‘subjective’ (Heinen et al., 2010). Objective safety refers to the actual safety of cyclists, measured in the amount of accidents involving bicycles per one million inhabitants. Subjective safety refers to how individuals perceive safety and is often measured as the feeling of safety that individuals experience on different roadways. While the effect of objective safety on the use of bicycle facilities is not quite clear, cyclists experience a higher level of subjective safety while cycling along designated bicycle facilities. This subjective feeling of safety can be seen as one of the most important reasons why cyclists are willing to deviate from the shortest route to travel along dedicated bicycle infrastructure.

Research on cyclist route choice confirms the assumption that cyclists prefer cycling along routes with bicycle facilities. In a study by Broach et al (2012) in Oregon, it was found that cyclists have a preference first for separated bicycle paths, followed by bicycle boulevards. Non-separated bicycle paths, described in the study as striped lanes, were only preferred when cycling along a low-traffic neighbourhood street was not an option. Research by Winters et al (2010) further confirms the notion that cyclists prefer cycling along designated bicycle facilities. In their research, 49 percent of the trips made by cyclists were along designated bicycle facilities, while the shortest routes predicted that only 21 percent of the trips by bicycle would be made along such facilities. This means that the mean increase in distance travelled along bicycle facilities was 1583 meters compared to the corresponding shortest route. In addition to a general preference for bicycle facilities, cyclists also prefer these facilities to be connected (Caulfield, Brick, & McCarthy, 2012). Connected bicycle infrastructure allows cyclists to minimize their interaction with other traffic and avoid hazardous junctions further increasing the notion of subjective safety. In many countries however, the current infrastructure does not allow for cyclists to cycle along connected bicycle infrastructure (Sener et al., 2009). In these countries, cyclists would have to make a significant detour to stay on bicycle infrastructure along their route. However, there is a limit to the amount of deviation a cyclist is willing to add to his route to cycle along bicycle infrastructure (Broach et al., 2012). This means that in some cases cyclists simply choose to cycle along main roads that they deem to be safe enough. In other cases, people simply will choose not to cycle at all.

2.2.1 Bicycle infrastructure in The Netherlands

In The Netherlands, a long history of policies dedicated to increasing the well-being of cyclists has created a cycling climate where cycling is relatively safe and problem free. An integrated system of bicycle facilities allows cyclists to cover any trip completely on bicycle facilities or traffic-calmed areas (Pucher & Buehler, 2008). This integrated system of bicycle facilities in The Netherlands is based on the concept of the Dutch street hierarchy (Schepers et al., 2017). This street hierarchy is built upon both the homogeneity and functionality principle and contains a hierarchy of roads. The first level of roads in the Dutch street hierarchy are access roads. These roads provide access to origins and destinations. On these access roads, cyclists are mixed with other traffic. The homogeneity principle implies that differences in speed, mass and direction should not be too large. On roads without bicycle facilities, where cyclists are mixed in with other traffic, the speed limit should always be 30 km/h. The second level of roads are distributor roads. These roads distribute traffic from access roads to trough roads. As distributor roads have a speed-limit of 50 or 70 km/h, the speed differences between cyclists and motorized traffic are too large. On these distributor roads, cyclists should always be separated from other traffic by bicycle lanes or bicycle paths. The third level of roads are through roads. These roads, which are mostly freeways, have a traffic flow function. This means that most of the motorized traffic should be directed to these roads. On these through roads, which have a speed-limit of 100 to 120 km/h, cycling is not allowed. This Dutch system of street hierarchy greatly reduces cyclists' exposure to motorized traffic. This in turn greatly increases the objective and subjective safety for cyclists, creating an environment which encourages people to cycle.

There are a few different types of bicycle infrastructure resulting from this street hierarchy (Schepers et al., 2017). The first type of bicycle infrastructure is the bicycle boulevard or bicycle street (Pucher & Buehler, 2008; Schepers, Hagenzieker, Methorst, Van Wee, & Wegman, 2014). These bicycle streets are implemented on access roads where there is a relatively high number of cyclists compared to motorized traffic. These bicycle streets do not indicate where cyclists or motorized traffic can or can't be on the road, but merely indicate that motorized traffic should be extra aware of cyclists on these roads and that motorized traffic shouldn't rush cyclists or interfere with them (Pucher & Buehler, 2008). The second type of bicycle infrastructure is the non-separated bicycle path. These are marked lanes on distributor roads that indicate that only cyclists are allowed on this lane. These non-separated bicycle paths are either separated from other traffic by a solid line or a broken line. In the case of a solid line, other traffic is not allowed to cross the bicycle path. In the case of a broken line, other traffic is allowed to cross the bicycle path. A broken line is often used when there is a need for cars to cross the bicycle path, which is the case when there are parking spots located next to the bicycle path. The second type of bicycle infrastructure can be classed as the separated bicycle path. These are bicycle paths that are completely separated from other traffic by physical barriers. Separated bicycle paths can either be one-way bicycle paths or two-way bicycle paths. Separated bicycle paths can be found along distributor roads but are sometimes also located along through roads. In the case of trough roads, separated bicycle paths are always completely separated from the actual road by ways of large physical barriers.

2.3 Intersections and traffic signals

Intersections and traffic signals have a combined effect on cyclist route choice (Broach et al., 2012). On their own, traffic signals and stop signs can cause irritation for a cyclist. This is because stopping and accelerating costs cyclists a big amount of effort (Heinen et al., 2010). This causes cyclists to avoid traffic signals on their route. In general, research on cyclists route choice shows that cyclists go out of their way to avoid traffic signals on their route (Broach et al., 2012; Winters et al., 2010). However, traffic signals do not always deter cyclists. When traffic volumes at intersections are high, cyclists might actually prefer a signalized intersection to a non-signalized intersection. Cyclists were willing to add a deviation of 10.4 percent to their original route to avoid an unsignalized intersection with a high traffic count. To avoid a left turn at a busy intersection, cyclists were willing to take a detour of 16.4 percent. In these cases, the positive effects of signals may outweigh their negative effects. Cyclists thus may want to include traffic signals on their route as it may help them to cross particular intersections. Making a left turn at a signalized intersection may actually cause less delay than making a left turn at an unsignalized intersection. Furthermore, not all cyclists attach the same value to traffic signals (Heinen et al., 2010). Overall, experienced cyclists might have a more negative perception of traffic signals than in-experienced cyclists and cyclists in urban areas tend to place a higher value on traffic signals than cyclists in non-urban areas. This can be explained by the fact that experienced cyclists may feel more confident in their abilities and may consider reducing travel time to be more important, negating the need for traffic signals. Commuters may also find street crossings to be less bothersome and bear a greater aversion to traffic signals, although street crossings may still have a negative effect on cyclist route choice. The effect of roundabouts on cyclist route choice is also discussed in research on cyclist route choice (Campbell, Urisich, & Dunn, 2006). In some cases, roundabouts may decrease the safety that a cyclist experiences causing them to be a deterrent to cyclist route choice. However, roundabouts can also prove to be a useful addition to routes for a cyclist as they reduce delays compared to normal intersections.

As mentioned before, in some cases, such as when turning left at a busy intersection or going straight on at a busy intersection, traffic signals may have a positive effect on cyclist route choice. Whether this effect is caused by an actual reduction in delay or an increased perception of safety is unclear (Broach et al., 2012). For commuters and experienced cyclists, traffic signals along a route may also bear a more negative effect than other types of cyclists.

2.4 Road surface and quality

Compared to other infrastructural factors, relatively little research has been conducted on the effect of the type of road surface or the quality of the road surface on cyclist route choice (Heinen et al., 2010). However, in recent research on cyclist route choice in The Netherlands cyclists placed the second most importance on the quality of the road surface when choosing their route (Genugten & Overdijk, 2016). Research on the amount of comfort cyclists experience when riding on different road surfaces shows that the type of road surface and the quality of the road surface could have an influence on cyclist route choice (Hölzel, Höchtl, & Senner, 2012). Different surfaces provide different rolling resistances and transmit different levels of vibrations, making it easier or harder for people to cycle on. In their research, Hölzel et al (2012) found that asphalt and concrete slabs provide the least amount of perceptible vibrations, making it the most suitable surface for cyclists to ride on. In comparison, cobblestones and self-binding gravel provide levels of comfort that fall far behind the levels of asphalt and concrete slabs, making these two types of surfaces far less suitable for cyclists to ride on. Even compared to worn down asphalt and worn down concrete slabs, cobblestones and self-binding gravel fall behind in levels of comfort.

The type of road surface and the quality of the road surface is especially important for women and older cyclists, as these two groups attach greater value to riding on a smooth surface (Heinen et al., 2010).

2.5 Network characteristics

Factors based on the network that cyclists ride on can also influence their route. In their research on route choice, Papinksi and Scott (2011) created several network characteristics based on observed routes and shortest routes to find out how these characteristics differed between these routes. Although this route choice research was based on cars, these characteristics can also be applied to the cyclist route choice analysis and can be considered factors that in turn influence the infrastructural factors that cyclists encounter along their route (Broach et al., 2012; Zimmermann et al., 2017). An example of this is the fact that cyclists sometimes avoid a left turn at an unsignalized intersection and would deviate from their route to include more signalized intersections on their route (Broach et al., 2012). Based on studies investigating cyclist route choice, the network characteristics that Papinski and Scott included in their research and can be considered potentially relevant to cyclist route choice are the number of turns and the direction of these turns, the time and distance of a route, the speed of a cyclist along a route and the longest leg based on distance and time (Broach et al., 2012; Zimmermann et al., 2017).

2.6 Other influences on cyclist route choice

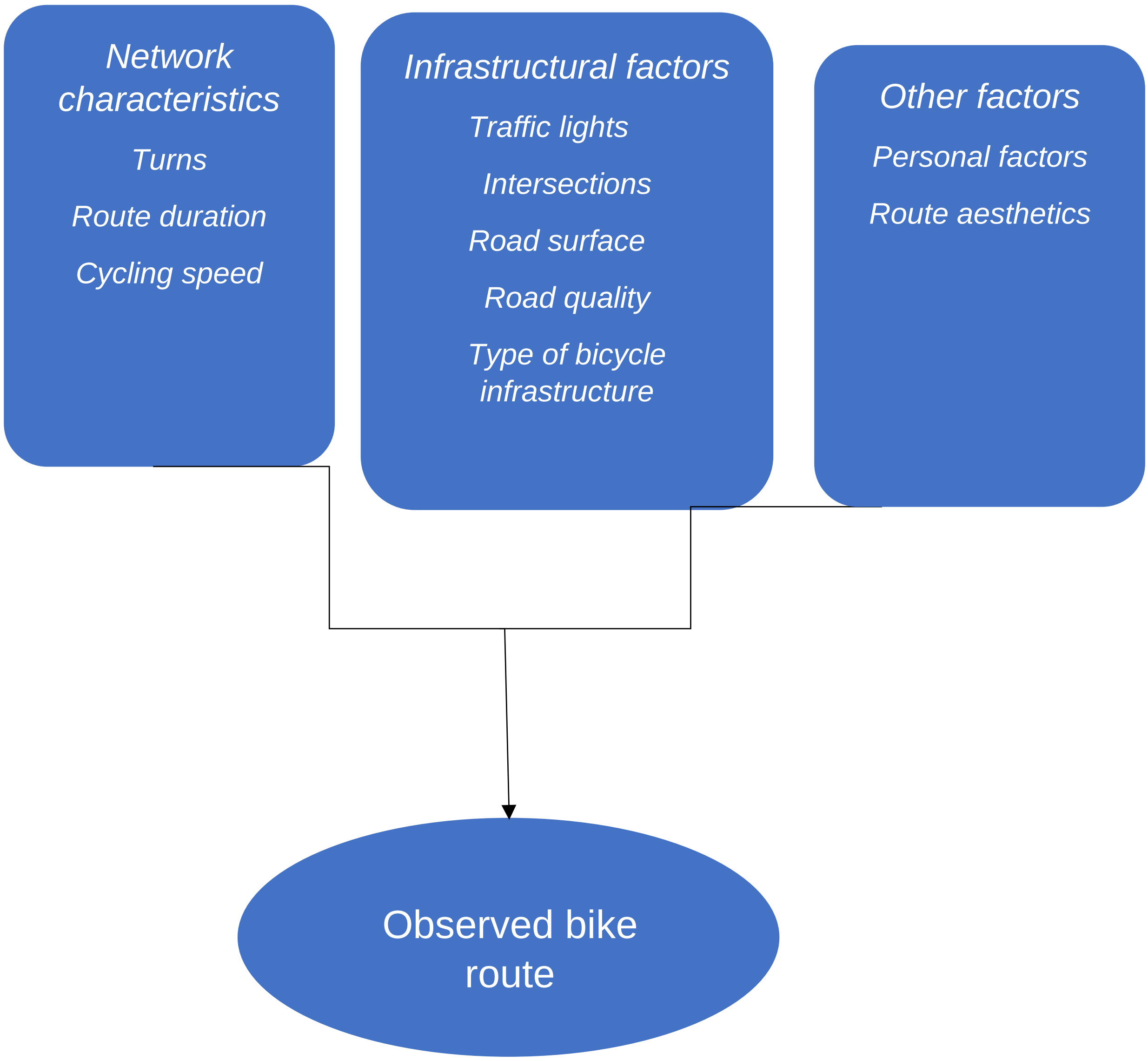
Aside from infrastructural and network factors, there are many other factors that influence a (commuting) cyclist' route choice (Heinen et al., 2010). Cycling for commuting purposes is often part of a daily routine (Van Duppen & Spierings, 2013). This makes cycling a necessary and repetitive activity, often causing people to be less aware of their surroundings, Because of this fact, it can be assumed that cyclists do not choose their route based on subjective characteristics such as aesthetics or personal preference. However, there is also more complexity to be found in the daily commute. Commuting by bicycle can also be a way of getting to know someone's surroundings or exploring a new city. People may also take detours through quiet and green environments to avoid the chaos of their typical commuting route. Regarding these quiet and green environments, cyclists regard cycling in an attractive environment as an important aspect of cycling (Heinen et al., 2010). In study conducted in three cities in Iceland and Norway, 46% of the respondents selected trees and vegetation along their route to be the most favoured feature (Stefansdottir, 2014). Cycling through specific environments may also have special implications for some cyclists, as the smell or sight of these particular environments may hold positive memories for a particular cyclist (Van Duppen & Spierings, 2013). Thus, when people choose their cycling route to commute, they often choose routes that they personally perceive to be as aesthetic or as an overall positive experience. This also means that cyclists avoid environments that they consider to be unpleasant or distressing, such as roads with large amounts of car traffic or environments that are generally described as boring and grey (Stefansdottir, 2014). With further regards to environment, slopes also have a negative effect on cycling (Heinen et al., 2010). Cyclists may often avoid roads with a high amount of slope, as these roads require a lot of additional effort compared to flat roads (Menghini et al., 2010).

In addition to the fact that environmental factors greatly affect cyclist route choice, choosing a route when commuting to work is also still a very personal affair. As mentioned before, women may often choose a different route to cycle than men, as women may prefer routes that they perceive to be safer (Emond et al., 2009). Women may also be less willing to take detours than men. When comparing experienced cyclists to in-experienced cyclists, there is also a clear distinction between the types of routes these two groups will choose. Experienced cyclists may often choose a route which is faster but which is less safe and is considered more dangerous, while in-experienced cyclists will want to ride on roads that are safer and easier to ride on (Winters et al., 2010). Regarding safety, cyclists may also perceive certain roads to be more unsafe than other roads, even if this is not objectively proven. (Heinen et al., 2010). When cyclists witness an accident on a certain road, they will associate that road with a feeling of unsafety, causing them to avoid this road when commuting to work. This aspect of personal perception of certain roads and environment, causes cyclist route choice to be a phenomenon that can never be fully explained. As choosing a cyclist route is part of human behaviour, this uncapturable factor of personal preference for certain roads remains difficult to capture in research.

2.7 Conclusion and conceptual model

In comparison to motorized traffic, the suitability of a route for cyclists is made up of multiple factors. Figure 2.1 presents these factors in a conceptual model. Cyclists are willing to deviate from their route to include certain factors on their route and do not always take the shortest available route to their destination. According to research, infrastructural factors, network characteristics, the aesthetics of a route and personal factors greatly influence cyclist route choice behaviour. Because this research focuses solely on infrastructural factors, only hypotheses on the influence that infrastructural factors have on cyclist route choice are drawn up. These hypotheses are the following: Cyclists are willing to deviate from their route to include more separated bicycle paths. Cyclists will also want to deviate from their route to avoid roads where they have to share the road with motorized vehicles. Regarding intersection and traffic signal control, cyclists will want to deviate to encounter less intersections with traffic signals. However, when crossing busy roads with a lot of traffic, they may often want to take detours to include these on their route as a traffic signal will sometimes cause less delay in this situation. As asphalt, concrete and good quality roads provide the most amount of comfort for cyclists, they will also want to deviate to come across more of these roads.

Figure 2.1: Conceptual model

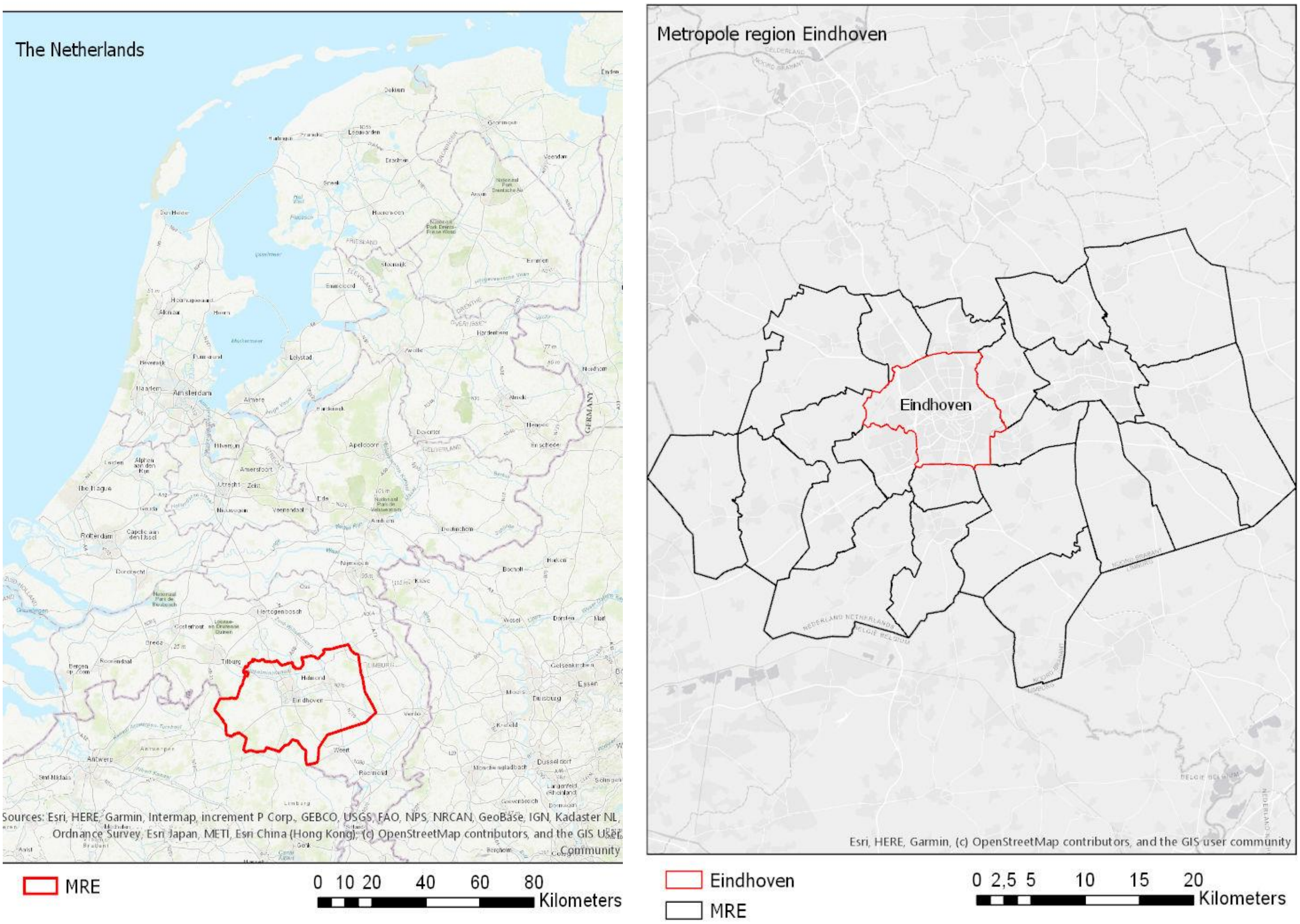


3. Methodology

3.1 Study area

To determine the influence of infrastructural- and network factors on cyclist route choice, this research focuses on the Metropole Region Eindhoven (MRE after this). The MRE is located in the province of Noord-Brabant in the Netherlands and plays a central role in the economy. In 2017, the MRE was the fastest growing region of The Netherlands with an economic growth of 4,9 percent. With 19,7 billion euros it is also the third largest export region, after Amsterdam and Rotterdam. It consists of 21 municipalities around the city of Eindhoven and has 753.426 inhabitants over an area of 1.457,81 km^2 This area was selected because of the availability of GPS data collected by the B-riders program. The B-riders program is a program that encourages people to cycle to work instead of using their car by offering incentives for every cycle trip. Furthermore, it is a suitable case study to analyse commuters choices because of the area's central role in the economy as mentioned before, which induces high levels of commuter interactions between the main city 'Eindhoven' and the surrounding municipalities. In figure 3.1, the location of the MRE in the Netherland is presented, as well as the location of the municipality of Eindhoven within the MRE.

Figure 3.1: Location of the MRE region in The Netherlands and the location of Eindhoven within the MRE



3.2 Data

3.2.1 B-Riders data

In this research GPS cycling data from the B-Riders program is used. This is data recorded by commuting cyclists on electric bicycles in the Dutch province of Noord-Brabant. The B-riders program was created as an incentive to encourage people to use their bicycles instead of their cars for their commute and has now attracted over 7.500 users (B-Riders, 2018.). Participants use an application on their smartphone to track their commute to work. They are encouraged to use the application by means of points or monetary rewards. With earned points participants can either choose to turn in their points for lottery tickets or donate an amount of money to a charity of their choice. The participants start to register their trip by manually starting their trip when they leave for work and stop the trip when they arrive at their workplace or vice versa.

Included in the B-riders dataset is a shapefile containing GPS points, a shapefile containing links (sections of roads that participants in the program have ridden on), a CSV file that includes the sex and age per participant, a CSV file that contains the GPS matches to the links in the dataset and a CSV file containing the user-ID, the route-ID and the purpose for each trip. In table 3.1, an overview of the files included in the B-riders dataset is presented.

Table 3.1: Files included in the B-riders dataset and their description

Name of file	Description
Links	Shapefile containing the road segments that participants have ridden on
GPS-punten	Shapefile containing the GPS points for each route
GPS-match	CSV file containing the map-matching for the links shapefile
Routes_user	CSV file containing the trip purpose per route and the respective participant
Gebruiker-leeftijd-geslacht	CSV file containing the age and sex of each participant

In total, the B-riders dataset contains more than 45 million GPS points, more than 290000 links and more than 2000 individual users.

3.2.2 Fietsersbond cycling network

The cycling network dataset for The Netherlands that is used in this research is the Fietsersbond (cycling union) cycling network (Fietsersbond, 2019). This organisation has created a route planner that is updated by a group of 500 volunteers who manually enter very detailed information about roads and cycle tracks. In comparison to regular route planners created with motorized vehicles in mind, the Fietsersbond route planner contains all cycle tracks but also small alleyways which are only fit for use by non-motorized traffic. In addition to a complete set of roads for cyclists, the Fietsersbond network also contains additional information about each road such as the pavement type, the type of the surrounding environment and temporary detours. Because of the completeness of the Fietsersbond route planner, the underlying network was chosen to be used in the network analysis part of this research. The additional information about infrastructural factors per road segment readily available in the Fietsersbond route network will be also be used in this research. These infrastructural factors will be added to the road segments in both the observed routes and the shortest routes.

3.3 Data processing

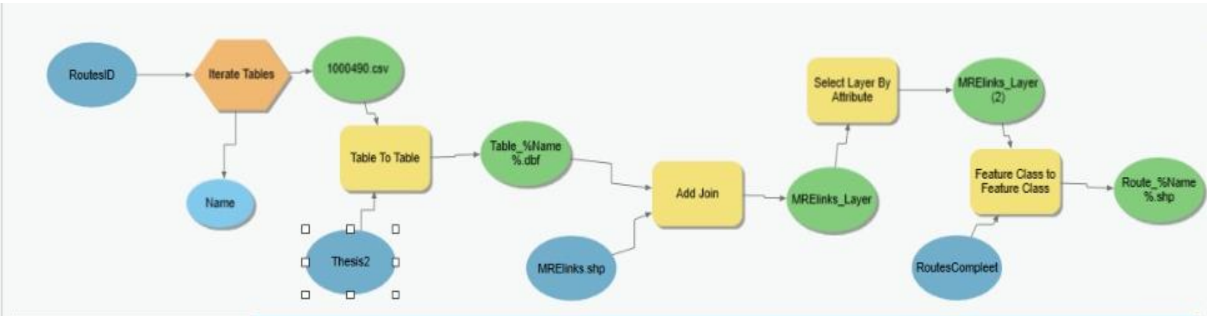
3.3.1 Preparing the cycling network

To prepare the Fietsersbond cycling network for network analysis, a network dataset based on the cycling network was created in a geographical information system (GIS), using ArcGIS Pro. By creating this network dataset, it is possible to generate the shortest routes from the origins and destinations of the observed routes by using the network analysis tools in GIS. All of the roads where cyclists are not allowed were removed from the Fietsersbond network prior to generating the network dataset. This was done to ensure that the shortest routes generated on the network were based on the roads that could actually be ridden on by cyclists. When analysing the network, it was found that some road segments did not connect, which would cause complications when using the network for routing. These inconsistencies in the network were fixed by using the *integrate* and *planarize* tools in ArcGIS. The integrate tool connects vertices that fall within a specified tolerance of one another. Where certain road segments of the Fietsersbond network did not connect, the integrate tool connected these road segments. Afterwards, the planarize tool was used on the integrated network to remove overlapping road segments.

3.3.2 Generating the observed routes

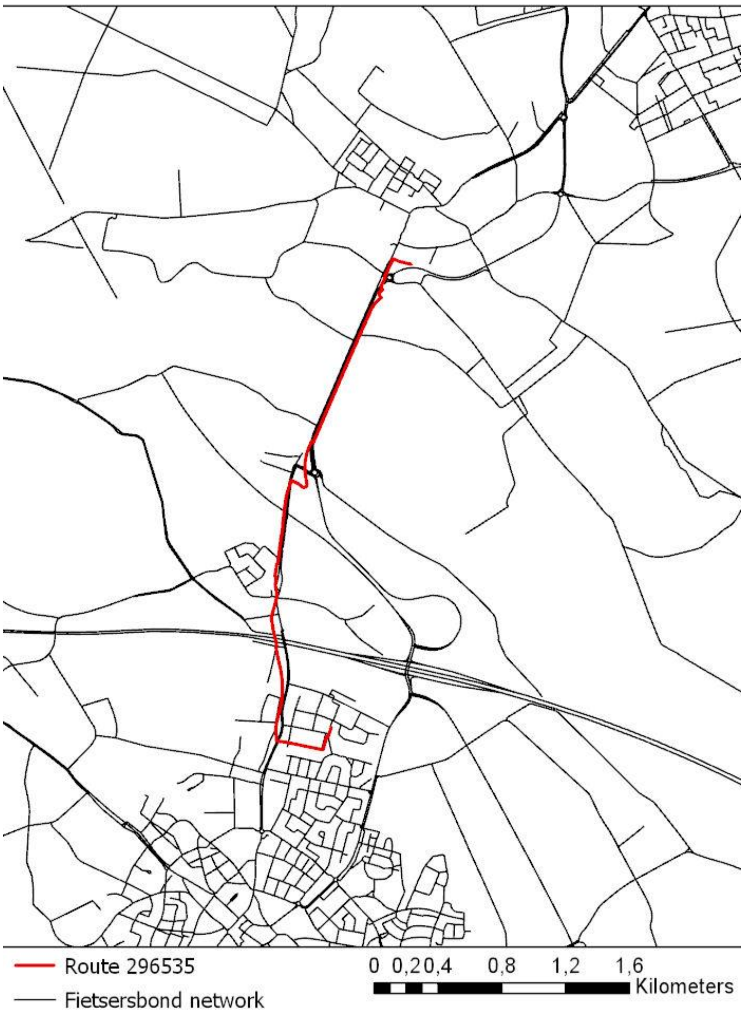
The B-riders dataset contains both data on the GPS points that each route consists of, as well as the road segments that each route consists of. This means that there are two possible ways to generate an observed route from the dataset. For all the steps in the preparation of the B-riders data, ArcGIS Pro was used as well. As the study area of this research is limited to the MRE, the first step in the process of preparing the data is to clip the GPS points or the links to the study area. When the clip function was used on the GPS points, it was found that the large sample size of over 45 million points caused complications. Based on these findings, the road segments were chosen to be used as input for the generation of the observed routes.

Figure 3.2: Model builder for the generation of the observed routes.



First, the links or road segments were clipped to the MRE region. Each of these route segments has a unique ID. In the GPS-match table, it is possible to see which road segment IDs each route consists of. From this GPS-match table, one route for commuting purposes per individual user was sampled using a python script. The result of this script was a table showing the road segments IDs that each sampled route consisted of. This table was then split into 736 individual tables that each represented a route and the respective road segments that each of these routes consisted of. Each of these tables then had to be individually joined to the road segments. To make this process possible, a model was used. This model, which can be seen in figure 3.2, individually joined each route-ID table to the links shapefile, after which all of the links that a route-ID consisted of were selected inside of this links shapefile. To complete the model, these links were then copied to a new shapefile, resulting in 736 individual shapefiles that each represented a unique route. In figure 3.3 an example of a generated observed route on the Fietsersbond network is given.

Figure 3.3: An example of an observed route on the Fietsersbond network



3.3.3 Generating the shortest routes

In the network analysis, the shortest routes based on the origins and destinations of the observed routes were generated. The underlying network that was used for the analysis is the Fietsersbond network. To generate the shortest routes, the origins and destinations of the generated observed routes were added to the network as stops. Each of these stops represented an origin and destination for a particular route-ID. With the solve tool, the shortest route between the origin and destination of each route-ID was generated. 730 shortest routes out of the 730 possible routes were able to be generated. The eight remaining routes were not generated due to the remaining errors in the network that were not solved in the preparation of the network dataset.

Table 3.2: Results of the shortest route analysis

Amount of shortest routes shorter than observed routes	Amount of shortest routes longer than observed route	Total shortest routes
645	85	730

After examining the length of the shortest routes compared to the observed routes, it was found that 85 out of the 730 shortest routes were actually longer than the observed routes. A possible cause for this error can be found in the way a route is generated on a network compared to an observed route generated from GPS data. When cycling, we can presume it is possible to make a U-turn at any given moment. When a shortest route is generated on a network, U-turns are only possible at intersections, increasing the total length of a route. Another possible reason for the fact that 85 out of 730 shortest routes are longer than the observed routes are network errors. As mentioned before, only 730 shortest routes out of the 736 observed routes that they were based on were able to be generated. The network errors that caused these 4 routes to be unable to be generated are also a possible cause for fact that some shortest routes are longer than the observed routes. To prevent the analysis from being skewed, the 83 shortest routes that were actually longer than the observed routes were removed from the analysis. The route-IDs of these 85 routes were also removed from the observed routes.

3.4 Analysis

3.4.1. Joining the network to the observed routes

Once both the observed routes and the shortest routes were generated, the Fietsersbond network could be joined to the respective routes. By joining the routes to the network, the information on infrastructural factors for each route will become available. The attributes in the Fietsersbond network were joined to the observed by using the a-node and b-node variables. The a-node and b-node variables consist of a unique ID for each intersection in the network. The a-node is the intersection where the road segment begins and the b-node is the intersection where the road segment ends. Both the B-riders data and the network contain the a-node and b-node variables. By using query tables, the observed routes could be joined to the network based on the a-node and b-node.

3.4.2 Joining the network to the shortest routes

To join the network to the shortest routes, the shortest routes had to be split into multiple segments. Once the shortest routes were split into multiple segments, the Fietsersbond network could be joined to the shortest routes by joining them by location. As the shortest routes were generated based on the network, each of the routes ran along the network. For the observed routes this was not the case, as the GPS tracks were not matched to the network.

By using the intersect function in the Join by location tool, the information about each road segment was added to the segments that each individual shortest route consisted of.

3.4.3 Statistical analysis

The influence of infrastructural factors and network characteristics on cyclist route choice will be investigated using statistical analysis. The statistical analysis method that is used in this research is a multiple linear regression model. The variables will first be studied using a paired samples t-test before building the final statistical model. With a paired samples t-test, it can be tested if the differences in the variables between the observed routes and the shortest routes are significant (Field, 2013).

With linear regression, the influence of one or multiple independent variables on one or more dependent variables can be tested (Field, 2013). A requirement for linear regression analysis is that the dependent variable must either be an interval or ratio variable. As the dependent variable in this research is a ratio variable, it is possible to perform a linear regression analysis. The software used to perform the paired samples t-test and the linear regression analysis is SPSS.

The dependent and independent variables used in this research are:

- Dependent variable: The amount of deviation from the shortest route (absolute and relative).
- Independent variable: The difference in infrastructural factors and network characteristics between the observed routes and the shortest routes (absolute and relative).

The difference in infrastructural factors and network characteristics between the observed routes and the shortest routes will be tested against the absolute and relative deviation from the shortest route. The absolute and relative deviation from the shortest route will be calculated using the following formulas:

Absolute deviation (meters) : Δ*length = length(observed) – length(shortest)*

Relative deviation (ratio) : ratio= length(observed)/length(shortest)

The difference in infrastructural factors and network factors is expressed as the following formula:

Difference in infrastructural factors/network characteristics: Δ*factors= factors(observed) – factors(shortest)*

The multiple linear regression model will determine if there is a significant relationship between the variables (Field, 2013). In addition, the multiple linear regression model will also show how big the effect of the independent variables on the dependent variable is. This effect can either be a positive or negative value, indicating a positive or negative relationship of the independent variable on the dependent variable.

3.5 Data operationalisation

In this section, the operationalisation of the data available in the Fietsersbond network will be presented. The infrastructural variables are based on studies that have found which infrastructural factors influence cyclist route choice. Found indicators in the literature are the road type, road quality, road surface, the type of intersections, network characteristics, the aesthetics of a route and personal factors. Measuring the aesthetics of a route and the personal factors are outside the scope of this study and measuring network characteristics is not possible due to data limitations.

Road types

For the bicycle infrastructure factor, several indicators are derived from the Fietsersbond cycling network. In table 3.4, all of the classes in the 'road type' variable found in the Fietsersbond cycling network are presented. In the Fietsersbond cycling network, information about the type of road for each road segment is given. In the statistical analysis part of this research, each attribute in the Fietsersbond dataset that represents a type of road is classified to represent an indicator of bicycle infrastructure.

Table 3.4: Classification of the road type variables

<table>
<tr><th>Attribute in dataset</th><th>Classification</th></tr>
<tr><td>Unknown</td><td>Unknown road type</td></tr>
<tr><td>Normal road</td><td rowspan="3">Mixed roads</td></tr>
<tr><td>Pedestrian area</td></tr>
<tr><td>Pedestrian pass trough</td></tr>
<tr><td>Bicycle boulevard</td><td rowspan="2">Non-separated bicycle path</td></tr>
<tr><td>Weg met fiets(suggestie) strook</td></tr>
<tr><td>Solitary bicycle path</td><td rowspan="4">Seperated bicycle path</td></tr>
<tr><td>Solitary bicycle and scooter path</td></tr>
<tr><td>Bicycle path along road</td></tr>
<tr><td>Bicycle- and scooter path along road</td></tr>
</table>

Roads where information about the road type is not available have been reclassified to unknown type, meaning that these roads will not be taken into account in the analysis. In addition to the indicators that have been derived from literature, an extra class has been added. This class, called mixed traffic, encompasses all the roads where bicycles have to share the road with other road users, as bicycle infrastructure is not available on these roads. The non-separated bicycle path indicator is made up of two attributes, while the separated bicycle path is made up of four attributes. This is because these four attributes can all be considered to be separated bicycle paths and thus have been grouped together into one class. Each of the three bicycle infrastructure values will be calculated as a percentage of the total trip made by the respondent. For example, when a trip distance from origin to destination is 10 kilometres, and 6 kilometres of the trip has been covered on separated bicycle paths, the separated bicycle path value will be 60%.

Road surface and quality

For the infrastructural factor 'road surface and quality', the variables called "road surface'' and ''road quality'' were used. In table 1.3 the classes for the road surface variable are presented. Literature on how road surface and road quality affects cycling comfort proved that asphalt and concrete slabs were the most comfortable surfaces to ride on, while cobblestones, self-binding asphalt and other road surfaces proved to be far less comfortable. Also, it was proven that the quality of the road surface also affects the comfort a cyclist experiences while riding along a road.

Table 3.5: Classification of the road surface variables

Attribute in dataset	Classification
Unknown	Unknown road surface
Asphalt/concrete	Asphalt or concrete
Tiles	Other road surface
Clinker tile	
Unpaved	
Half-paved	
Other	
Shell path	

After adding the road surface and road quality to the observed routes and the shortest routes, the road surface attributes in the Fietsersbond cycling network were statistically classified to two classes. This classification is demonstrated in table 3.5. These classes are asphalt and other road surfaces. Based on literature, asphalt and concrete surfaces can be classed as smooth road surfaces, while tiles, cobblestones, gravel and other similar road surfaces can be classes as less smooth road surfaces. It is expected that cyclists will want to ride along smoother road surfaces like asphalt and concrete. The road surface variable will be expressed as the percentage of the total trip length. Thus, if a respondent has made a 10 kilometre trip and 7 kilometres of this trip consisted of smooth road surfaces, the smooth road surface value will be 70%.

Table 3.6: Classification of the road quality variables

Attribute in dataset	Classification
Unknown	Unknown road quality
Bad	Bad road quality
Reasonable	Reasonable road quality
Good	Good road quality

In table 3.6, the classes of the road quality variable in the Fietsersbond cycling network are presented. These road quality attributes will be classified into three classes in the statistical analysis part of this research. The bad road quality class will remain, as well as the reasonable and excellent road quality variables. This is because it is expected that cyclists will want to avoid bad quality roads in favour of reasonable or good quality roads. The road quality indicator will be expressed as a percentage of the total trip length. So, if a respondent has made a 10 kilometre trip and 4 kilometres of this trip has been completed on a road surface with a bad quality, the variable value will be 40%.

Intersections and traffic signals.

Intersections and traffic signals together have a combined effect on cyclist route choice. In some cases, cyclists may want to avoid traffic signals, while in other cases traffic signals may actually have a positive effect on cyclist route choice (Broach et al., 2012). In case of a left turn at a busy intersection, cyclists may actually prefer to have traffic signals present at such an intersectional it helps them to cross such an intersection faster and more conveniently. The same may be true for cyclists that go straight on at an intersection with or without traffic signals. For right turns, no distinction is made between unsignalized and signalized intersections, as these turns avoid most traffic conflicts and delays (Broach et al., 2012). To reflect these findings about the combined effect of intersections and traffic signals in this study different types of intersections will be identified in the observed routes and shortest routes. The network characteristics are presented in table 3.7

Table 3.7: Network characteristics

Indicators of network characteristics
Roundabouts
Intersections
Intersections with a traffic signal

The Fietsersbond network contains information about the different types of intersections for each road segment. These types of intersections include roundabouts, intersections and intersections with a traffic signal. This information was added to the observed and shortest routes to include these variables in the analysis. These factors will be expressed as the amounts of intersection types per kilometre. So if a route of 2 kilometres has a total of 4 roundabouts, the value of the roundabout variable will be 2,00.

3.6 Missing data and outliers

3.6.1 Missing data

Table 3.8: Missing data for the road type, road quality and road surface variables.

	Missing road type observed	**Missing road quality observed**	**Missing road surface observed**	**Missing road type shortest**	**Missing road quality shortest**	**Missing road surface shortest**
Minimum	-0,01	0,00	0,00	0,00	0,00	0,00
Maximum	1,00	1,00	1,00	1,00	1,00	1,00
Average	0,26	0,07	0,07	0,14	0,38	0,32
St.deviation	0,19	0,099	0,15	0,14	0,22	0,17

When examining the total percentage of missing road type, road quality and road surface for all of the three variables in table, the missing road type for the observed routes and the missing road quality and missing road surface for the shortest routes have the highest means. When comparing the missing data for the two route types, the shortest routes have a higher mean value for the missing road quality and missing road surface. Looking at the roads segments in the road network that the shortest routes consist of, errors were found in the network. Some of these routes were said to be located in different provinces of The Netherlands than the study area, like Noord-Holland or Zuid-Holland. An explanation for these errors can probably be attributed to the fact that the information about the road segments in the network was collected by volunteers. These volunteers have probably made several errors when collecting this information, resulting in missing or wrong data for some of the road segments. To control for these errors, all of the routes where the missing data for the road type, road quality or road quality was over 50 percent were removed. In total, 222 routes were removed from the data.

3.6.2 Outliers and data errors

Most of the observed routes do not deviate by a very large amount, but there are outliers and variance in the data. Figure 3.2 shows the largest relative deviation, which is 4.2 times longer than the shortest route. Figure 3.3 shows the largest absolute deviation. Both of the routes seem to have illogical deviations. For the route with the largest relative deviation, the origin and destination point for the observed route seems to be placed on the wrong part of the road, causing the shortest route to be much shorter than the shortest route would actually be when taking the correct origin and destination points into account. Another route which had the same type of error was also removed from the data. For the route with the largest absolute deviation, the route also seems to illogical for a route that was cycled for commuting purposes. Looking at the route and the deviation of 12,01 kilometres, it looks more like the route a motorized vehicle would take due to the fact that most of the road segments that are part of the observed route are through roads with high speed limits. This large deviation could have been caused by a participant of the b-riders program who has accidently tracked their route while using a motorized vehicle. As these routes seem to be illogical, they will be removed from the data.

Fig 3.2: Largest relative deviation from the shortest route

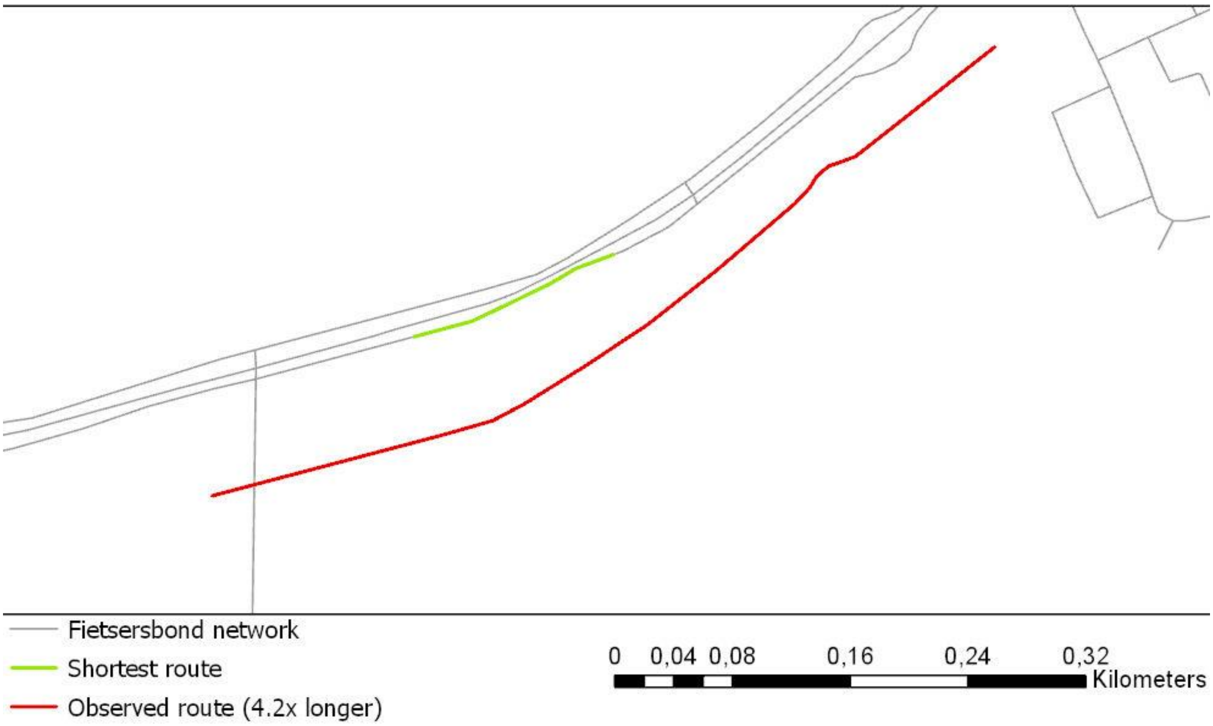
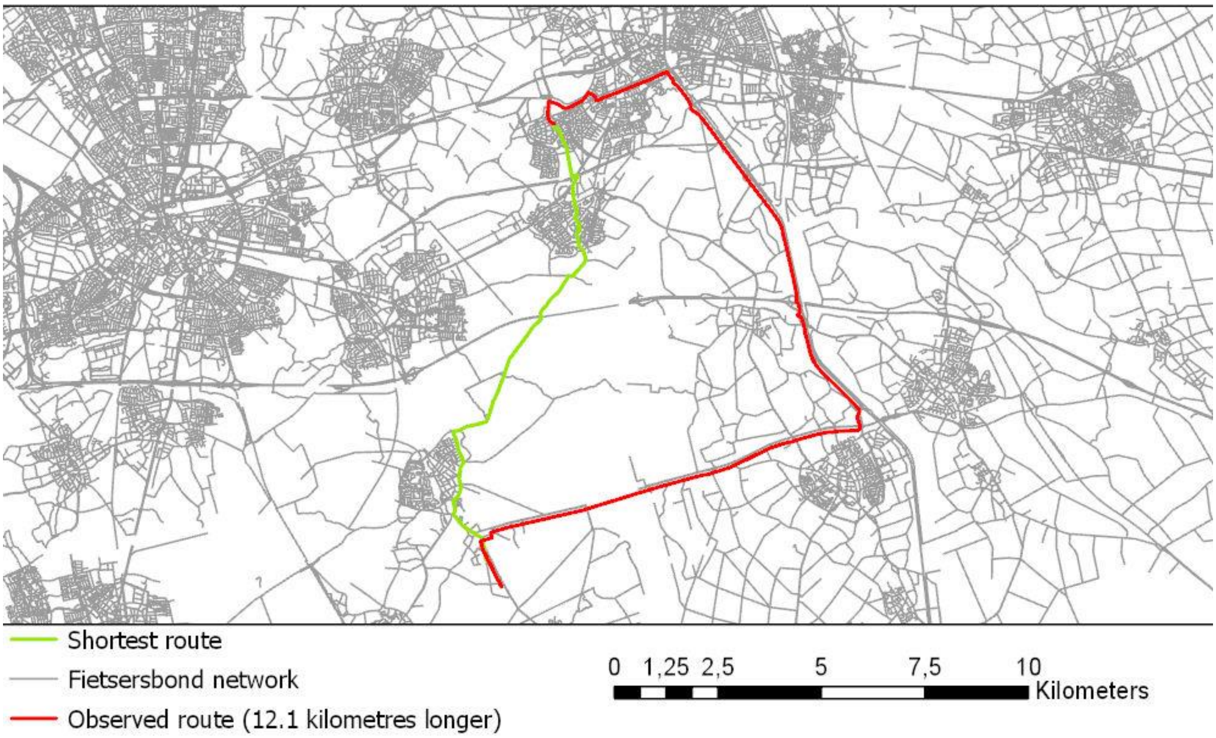


Fig 3.3: Largest absolute deviation from the shortest route



4. Results

In this chapter, the amount of deviation from the shortest route and the difference between the infrastructural factors in the observed routes and the shortest routes are investigated individually by using a paired samples T-test. These results will be accompanied by the descriptive statistics of the variables used in the analysis which can be found in appendix 1. After the results of the paired samples T-test have been discussed, the results of the multiple linear regression analysis will be presented.

4.1 Paired samples T-test

4.1.1 General route characteristics

Initially, the observed routes are compared to the shortest routes to gain insight into the exact differences between the routes. Looking at the initial descriptive statistics for the routes (see table 4.1), it shows that the observed routes are longer than the shortest routes but not by a large margin. The first 50 percent of the observed routes deviate less than 0,5 kilometres from the shortest route. However, the total average deviation from the shortest routes is 0,66 kilometres. Looking at the relative deviation, the average observed route is 1.07 times as long as the shortest route.

Table 4.1: Descriptive statistics for the observed and shortest routes

	Length observed	Length shortest	Abs. deviation	Rel. deviation
Minimum	0,63	0,15	0.00	1.00
Maximum	26,75	26,45	5.06	1.36
Average	10,90	10,23	0,66	1,07
St.deviation	5,30	5,00	0,63	0,06
Q1	7,27	6,53	0,24	1,03
Q2	10,4	9,66	0,46	1,06
Q3	14,23	13,67	0,89	1,10

Looking at the descriptive statistics for the absolute and relative deviation, it seems that the shortest routes are actually shorter than the observed routes. However, this difference could be based on chance, which is why the average deviation from the shortest route should be statistically tested. To test if the difference between two variables is significant, a paired samples T-test should be used (Field, 2013). With this test, the following hypothesis will be tested:

H_0: The average length of the shortest and the observed routes are not different from each other.

The corresponding alternative to this hypothesis is:

H_a: The average length of the shortest and the observed routes are different from each other

Looking at results for the paired samples T-test for the observed routes (M= 10,90) and shortest routes (M= 10,23) in table 4.2, it shows that the observed routes are significantly longer than the shortest routes, T(422) = 16,61, p < 0,01. This means that the h_0 can be rejected and the H_a can be accepted.

Table 4.2: Results of paired samples T-test for the length of the routes

	Mean	St.Dev	T	Sig
Route length	0,68	0,83	16,61 (419)	0,00

4.1.2 Infrastructural factors

Before it will be investigated if infrastructural factors can explain the deviation from the shortest routes, the factors themselves will be investigated. As the hypothesis is that the deviation from the shortest routes can be explained by the difference in infrastructural factors on these routes, we should see if there is actually a significant difference in infrastructural factors. The factors will be investigated by comparing the means of the factors. To test if these means are then significantly different, a paired samples T-test will be used (Field, 2013). The results of this paired-samples T-test for the infrastructural factors will be presented in this chapter.

Road type

Table 4.3: Paired samples T-test for the road type variables

	Mean	St.Dev	T	Sig
Separated bicycle path	0,28	0,17	33,71(419)	0,00
Non-separated bicycle path	0,00	0,086	0.040 (419)	0,968
Mixed roads	-0,40	0,17	-48,27 (419)	0,00

Looking at the means for the road type variables, it seems that the observed routes contain 28 percent more separated bicycle paths than the shortest routes. The mean for the non-separated bicycle path shows that this variable is almost evenly distributed between the routes. For the mixed road variable, a mean of -0,40 shows that on average the observed routes contain 40 percent less mixed roads. These results also show in the paired samples T-test in table 4.3.

The difference in separated bicycle path and mixed roads is significant, with a p value of <0,01. The difference in non-separated bicycle path is not significant with a p value of 0,968. This result also corresponds with the literature, where it was found that cyclists will want to deviate from their route to ride on more separated bicycle infrastructure and avoid mixed roads, but did not deviate from their route to ride on non-separated bicycle infrastructure (Broach et al., 2012; Winters et al., 2010) .

Road quality

Table 4.4: Paired samples T-test for the road type variables

	Mean	St.Dev	T	Sig
Good road quality	0,24	0,19	25,85 (419)	0,00
Reasonable road quality	0,018	0,14	2,74 (419)	0,06
Bad road quality	-0,015	0,07	-4,26 (419)	0,00

In table 4.4 the mean value for good quality roads is 0,24, meaning that on average an observed route has 29 percent more good quality roads than the shortest routes. Looking at the mean values for the reasonable and bad quality roads, these have a very small effect. This small effect can probably be explained by the high number of road segments where information about road quality is missing. Looking at the paired samples T-test for the road quality variables, it shows that the difference for the good road quality and bad road quality variables is significant with p- values of <0.01. The difference in reasonable quality roads is not significant. This corresponds with literature, where it was found that people place high importance on the quality of the road when choosing their route (Genugten & Overdijk, 2016). The fact that the difference in reasonable quality roads is not significant can probably be explained by the fact that people will want to deviate from their route to include good quality roads, but are indifferent towards reasonable quality roads.

Roundabouts and intersections

Table 4.5: Paired samples T-test for the roundabouts and intersections variables

	Mean	St.Dev	T	Sig
Roundabout ratio	0,34	0,60	11,48 (419)	0,00
Intersections ratio	0,32	0,76	8,47 (419)	0,00
Intersections signal ratio	0,40	0,78	10,40 (419)	0,00

The paired samples T-test results for the variables related to roundabouts and the different types of intersections are presented in table 4.5. The difference between the variables are all significant with a p-value of <0,01. The positive means for all three variables show that people will want to deviate from the shortest route to include more roundabouts, regular intersections and intersections with a traffic signal. In literature, it was presented that people will sometimes want to include more intersections with a traffic signal on their route as this may sometimes mean less delay when crossing a busy intersection (Broach et al., 2012). However some cyclists will want to avoid intersections with traffic signals as it may cause delays. The results of the paired samples T-test show that the observed routes contain more intersections in general. This means that cyclists do not have a particular preference towards one type of intersection but will want to deviate from their route to include more intersections. The fact that the roundabout variable has a positive mean indicates that cyclists will deviate from their route to include roundabouts. In literature, it was found that cyclists will sometimes want to include roundabouts on their route as they decrease delay compared to a regular intersection (Campbell et al., 2006).

Road surface

Table 4.6: Paired samples T-test for the road surface variables

	Mean	St.Dev	T	Sig
Asphalt/concrete	0,25	0,16	32,54 (419)	0,00
Other road surfaces	-0,05	0,14	-7,23 (419)	0,00

The results of the paired samples T-test for the road surface variables are presented in table 4.6. Looking at the means for the road surface variables, it shows that asphalt and concrete surfaces have a mean value of 0,25, meaning that on average an observed route contains 25 percent more asphalt or concrete surfaces. For the other road variable the mean value is -0,05, meaning that an observed route only contains 5 percent less other road surfaces. These findings are in accordance with literature on road surface relating to cyclist route choice, where asphalt and concrete slabs were found to be the most suitable for cyclists to ride (Hölzel et al., 2012). Other road surfaces fell far behind these surfaces on comfort levels, but the findings show that cyclists do not deviate a lot from their route to avoid these other road surfaces.

In this section, all variables have been compared individually. It was found that the observed routes are indeed significantly longer than the shortest routes. When looking at the results of the paired samples T-test for the infrastructural factors, all were significantly different between the two routes except for the non-separated bicycle path variable and the reasonable road quality. In the next section, we will assess if the deviation from the shortest route can be explained by the difference in infrastructural factors between these routes by way of multi-linear regression analysis.

4.2 Multiple regression analysis

In this section, the relationship between the deviation from the shortest route and the difference in infrastructural factors will be explained by multi-linear regression models. Multiple different multi-linear regression models will be run to find the best model fit. The multiple linear regression model was chosen as the analysis method because the causal linear relationship between a dependent variable and more than one independent variable will need to be investigated (Vocht, 2015).

This section will investigate to what extent the amount of deviation from the shortest route can be explained by the differences in infrastructural and network factors on these routes. This is examined on the scale of 420 routes. While the original dataset had 730 routes, 85 routes were removed where the shortest route was longer than the observed route. 222 routes were also removed where a variable contained more than 50 percent missing data. Finally, three routes were removed that had illogical deviations from the shortest route.

The null and alternative hypotheses of the linear regression are:

H_0: There is no relationship between the amount of deviation from the shortest route and the differences between the network and infrastructural factors on the observed and shortest route.

H_a: There is a significant relationship between the amount of deviation from the shortest route and the differences between the network and infrastructural factors on the observed and shortest route.

The null hypothesis will be rejected with 95% certainty when p =< 0.05

Several different regression models will be tested. After one model has been run, the explanatory variable with the highest p-value will be removed to eventually achieve the best model fit

4.2.1: Testing the assumptions for linear regression analysis

Dependent variables

Before performing a multi-linear regression analysis, the dependent and independent variables must be tested statistically. First, it should be investigated if the residuals of the regression follow a normal distribution. The residuals of the regression need to follow the normality line in order to draw conclusions from the results of the regression. The test for residuals of the regression for the absolute difference variable and the relative difference variable can be found in appendix 2. Looking at this test, it appears that the residuals of the regression for both variables do not follow the normal distribution, although both variables follow the normality line close enough to assume normality. The assumption for homoscedasticity should also be tested. Both of these assumptions can be tested by running the multiple-linear regression with the dependent variable and the independent variables. The results of this analysis can be found in appendix 2. Looking at this test, the absolute difference variable seems to suffer from heteroscedasticity and fails the assumption for homoscedasticity. In comparison, the relative difference variable passes the assumption for homoscedasticity with a few outliers.

Independent variables

To use the independent variables in a linear regression model, these variables should be tested for multicollinearity. The multicollinearity can be tested by looking at the bivariate correlations between the independent variables. A Pearson correlation value of 0.8 or higher indicates multicollinearity between variables. This is an issue, as your regression model will not be able to accurately associate variance in your outcome variable with the correct predictor variable, leading to muddled results and incorrect inferences. Looking at the correlation coefficients for the independent variables, these do not show a correlation coefficient higher than 0.8. This means that these variables can all be used in the multiple regression models.

4.2.2 Regression for the relative difference variable

In table 4.6, the results of the multiple regression analysis using the relative difference as dependent variable can be seen. The r-square value 0,30 for the model means that at most 3 percent of the relative amount of deviation from the shortest routes is explained by the explanatory variables included in the regression. The anova test for the model shows a p-value of <0.05, which means that alternative hypothesis can be accepted and there is a significant relationship between the dependent and independent variables. In the model, the variable for the relative difference in intersections with a traffic signal shows a significant relationship with a beta coefficient of -0.108. This negative value indicates that as the relative difference between the observed and shortest routes increases, the amount of intersections with a traffic signal per kilometre decreases. In this case, people will deviate from their route to avoid intersections with a traffic signal. All of the other explanatory variables do not show a significant beta-coefficient, which means that the relationship that these variables have with the dependent variable is highly based on chance and cannot be applied to the wider population. In addition, the effects that these variables have on the dependent variable are also very weak, which further proves that the effects that these variables have on the relative difference between the observed and shortest route is negligible.

Table 4.7:Multiple regression analysis with dependent variable *Relative deviation.*

Independent variables	Beta
Separated bicycle path	0.019
Non-separated bicycle path	
Mixed roads	0.022
Good quality	-0.083
Reasonable quality	
Bad quality	0.065
Relative difference roundabouts	0.074
Relative difference intersections	0.089
Relative difference intersections with signal	-0.108*
Asphalt and concrete	0.029
Other road surfaces	
R-square	0.037

*P<0.05, **P<0.01

4.2.3 Regression for the absolute difference variable

In table 4.8 the result of the regression models using the absolute difference as the dependent variable can be seen. The anova test for the models shows a p-value of >0.05, which means that alternative hypothesis can be rejected and that there is not a significant relationship between the dependent and independent variable. The model has an r-square value of 0,030, which means that 3% of the variance in the absolute difference between the routes can be explained by the independent variables included in the model. Like in the regression model employing the relative difference variable as the dependent variable, the intersection with a traffic signal variable is the only variable that remains significant. The negative beta value of -0.108 for this variable remains relatively unchanged compared to the effect that this variable has on the relative deviation from the shortest route. This negative beta value indicates that as the absolute difference between the routes increases, the amount of intersections with a traffic signal per kilometre decreases. This model does thus also indicate that people deviate from the shortest route to avoid intersections with a traffic signal. Similar to the regression model using the relative difference variable, the other remaining variables show a very small effect and are not significant.

Table 4.8: Multiple regression analysis with dependent variable *absolute deviation.* Model 1 includes the absolute differences between the intersection types and model 2 includes the relative differences between the intersection types.

Independent variables	Beta
Separated bicycle path	-0.074
Non-separated bicycle path	
Mixed roads	0.052
Good quality	0.11
Reasonable quality	
Bad quality	0.074
Relative difference roundabouts	-0.002
Relative difference intersections	0.033
Relative difference intersections with signal	0.132**
Asphalt and concrete	0.089
Other road surfaces	
R-square	0.030

*P<0.05, **P<0.01

5. Conclusion and discussion

This research has analysed GPS data and utilized various different GIS methods to find (1) which infrastructural factors have an effect on cycle routes and (2) to see to what extent the shortest possible route can be explained by these factors. By doing so, this research has attempted to add to the current knowledge on cyclist route choice. A deeper understanding on cyclist route choice is essential to promote cycling as the most viable and environmentally friendly alternative for commuting (Chapman, 2007). Not only is cycling environmentally friendly, it also decreases obesity rates, increases public health and decreases road injuries. Therefore, this study explored the following question:

To what extent do infrastructural factors and network characteristics influence cyclist route choice in The Netherlands?'.

To answer this main research question, first the sub questions will be answered.

Which infrastructural and network factors have the biggest influence on cyclist route choice?

Throughout all of the infrastructural factors analysed in this research only one factor proved to have a significant relationship with the amount of deviation from the shortest route. This negative relationship of intersections with a traffic signal indicates that as the difference between the observed and shortest routes increases, the amount of intersections with a traffic signal decreases. Cyclists are thus willing to significantly deviate from their route to avoid intersections with a traffic signal. This is in line with earlier research, where cyclists took significant detours to avoid traffic signals on their route (Broach et al., 2012; Winters et al., 2010). Based on the literature, the other factors that were expected to have a significant effect on the amount of deviation from the shortest route, such as separated bicycle paths or the type of road surface, all showed a very small effect and were not significant. This is remarkable, as the differences in infrastructural factors between the observed routes and the shortest routes were all significant except for non-separated bicycle paths and the reasonable road quality variables. These findings also contradict previous research (Broach et al., 2012; Caulfield et al., 2012), where cyclists significantly deviated from the shortest route to include more separated bicycle paths. As only trips for the purpose of commuting were analysed, cyclists have probably prioritized the travel time for a certain route. This also explains the fact that cyclists deviated from their route to avoid intersections with traffic signals, as encountering traffic signals can increase travel time by a large margin (Broach et al., 2012).

To what extent is there a relationship between the amount of deviation from the shortest routes and the differences in infrastructural factors between the observed and shortest routes?

In general, the shortest routes were significantly shorter than the observed routes. Relatively speaking, the routes were 1.08 times longer, while the average absolute difference between the routes was 0.68 kilometres. When trying to explain this difference by looking at the difference between the infrastructural factors on these routes, only 3.7% of the amount of deviation from the shortest route could be explained. The individual factors also had a very small effect, as discussed in the previous section. Taking these facts into consideration, the extent to which the amount of deviation from the shortest route can be explained by infrastructural factors is less than was expected considering previous research. Other factors, such as personal preference, safety, landscape and individual factors such as sex and age, may be more important for cyclists when choosing their route (Heinen et al., 2010; Stefansdottir, 2014; Van Duppen & Spierings, 2013).

Following the answers of the sub questions, it can be concluded that the influence of infrastructural factors on cyclist route choice in the Netherlands is relatively small. Only intersections with a traffic signal were intentionally avoided by cyclists, while the other infrastructural factors such as the road quality or type of bicycle infrastructure showed little to no effect. This is not in accordance with previous research, where cyclists were found to significantly deviate from their route to include separated bicycle infrastructure and comfortable road surfaces (Broach et al., 2012; Heinen et al., 2010; Hölzel et al., 2012). The results of this research showed that infrastructural factors explain only a small amount of the difference between the observed routes and shortest routes. This may indicate that there are many other factors that cyclists consider when choosing their optimal route. Firstly, previous research on cyclist route choice found that people choose to deviate from the shortest route to include visually pleasing environments and to avoid unappealing busy streets with a lot of car traffic (Stefansdottir, 2014). Secondly, some cyclists may also have a personal preference for certain roads because they may have a had a positive experience on this particular road (Van Duppen & Spierings, 2013). Thirdly, women may cycle different routes than men, as they perceive certain roads to be unsafe or want to avoid dangerous situations (Emond et al., 2009). Finally, cyclist may also choose certain routes to avoid turns, because they have a higher average speed on a route or because a route is longer but takes less time to complete (Broach et al., 2012; Papinski & Scott, 2011).

In further research on cyclist route choice, more explanatory factors should be considered, such as the aesthetics of a route, the amount of car traffic on a route or the amount of turns on a route. By doing so, a more accurate portrayal of cyclist route choice would be given. Also, combining the method of analysing GPS data with surveys on these particular routes can provide new interesting insights. Researchers may also consider different methods than linear regression models to gain new insights into cyclist route choice behaviour. These methods could include non-linear methods such as logistic regression. The route choice of different social and economic groups could also be compared. As mentioned before, men often take different routes than women. It would also be interesting to see if parents when cycling with their children take different routes than without them present. Another interesting study would be to determine if people deviate more or less from the shortest route as they age. In general, more research should be done on cyclist route choice behaviour. Creating complete and accurate cyclist route choice models could further encourage people to choose the bicycle as their daily mode of transport, which has many benefits for the environment and for the public health.

6. Data constraints and limitations

In addition to the conclusion, this section will discuss some of the constraints and limitations of the data.

In this thesis, cyclist route choice was investigated using data from the B-riders program and the Fietsersbond. As the information in the Fietsersbond network is based on volunteered geographic information, the dataset lacked accuracy in some parts. An example of this is that, upon closer inspection, the data on the road segments in the Fietsersbond network was not completely correct. Some roads were said to be in different provinces of The Netherlands. These discrepancies in the data caused road segments to also have missing data or the wrong kind of data. To try to resolve this problem, all of the routes were removed where a variable consisted of more than 50% missing data. Of course, this only partly resolves the problem. If a road network was used that had data for all of the road segments available, the outcomes of the research may have been different and more representative of actual cyclist route choice processes.

Another problem with the data arose with the B-riders dataset. As this data was not map-matched to the network, the routes in the dataset were not situated on the roads. When adding the attributes of the network to the routes this was not a problem, as this information could be joined to the routes via corresponding variables between the two datasets. However, when extracting the origin and destination points from the observed routes to create the shortest routes, some of these points were placed on the wrong points in the network due to these points not laying correctly on the network. When analysing the data, one such route was found and removed from the dataset. Nevertheless, it is possible that other shortest routes may also have been created with origin and destinations placed not completely correctly placed on the network. This causes shortest routes to be shorter than they actually should be.

Taking these other factors into consideration, the fact that the infrastructural factors analysed in this research only explain a small amount of the deviation from the shortest route is not unexpected. Nevertheless, the results from this research can still be seen as viable, because the focus of this thesis was not to try to create a regression model which would completely explain why people choose to deviate from the shortest route. The fact that infrastructural factors only have a small influence on the route people choose when commuting to work is also interesting, meaning that people that commute by bicycle in The Netherlands do not consider infrastructural factors to be very important. This could also have been caused by the fact that The Netherlands already has a lot of high quality bicycle infrastructure, and that there are not a lot of roads where bicycle infrastructure is lacking. Results from this research would probably have been different if they were conducted in a country where cycling is less prevalent.

Bibliography

Briders. *B-Riders, fietsen naar het werk | B-Riders | Fietsen naar het werk*. [online] Available at: http://www.briders.nl/b-riders-fietsen-naar-het-werk [Accessed 23 Oct. 2018].

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*, *46*(10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Campbell, D., Urisich, I., & Dunn, R. (2006). *Improved Multi-Lane Roundabout Design for Cyclists: The C-Roundabout*.

Caulfield, B., Brick, E., & McCarthy, O. T. (2012). Determining bicycle infrastructure preferences - A case study of Dublin. *Transportation Research Part D: Transport and Environment*, *17*(5), 413–417. https://doi.org/10.1016/j.trd.2012.04.001

Chapman, L. (2007). Transport and climate change: a review. *Journal of Transport Geography*, *15*(5), 354–367. https://doi.org/10.1016/j.jtrangeo.2006.11.008

de Nazelle, A., Nieuwenhuijsen, M. J., Antó, J. M., Brauer, M., Briggs, D., Braun-Fahrlander, C., … Lebret, E. (2011). Improving health through policies that promote active travel: A review of evidence to support integrated health impact assessment. *Environment International*, *37*(4), 766–777. https://doi.org/10.1016/J.ENVINT.2011.02.003

Dill, J. (2009). Bicycling for Transportation and Health: The Role of Infrastructure. *Journal of Public Health Policy*, *30*(S1), S95–S110. https://doi.org/10.1057/jphp.2008.56

Dill, J., & Gliebe, J. (2008). Understanding and measuring bicycling behavior: a focus on travel time and route choice. *Transportation Research Record: Journal of the Transportation Research Board*, (December). https://doi.org/10.3141/2140-17

Ehrgott, M., Wang, J. Y. T., Raith, A., & Van Houtte, C. (2012). A bi-objective cyclist route choice model. *Transportation Research Part A: Policy and Practice*, *46*(4), 652–663. https://doi.org/10.1016/j.tra.2011.11.015

Emond, C. R., Tang, W., & Handy, S. L. (2009). Explaining Gender Difference in Bicycling Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2125*(1), 16–25. https://doi.org/10.3141/2125-03

Field, A. (2013). *Discovering Statistics Using SPSS* (4th ed.). London: SAGE publications Ltd. https://doi.org/10.1234/12345678

Fietsersbond (2019). Networkdataset Fietsersbond, retrieved 29 Januari 2019 from www.bikeprint.nl

Genugten, W. Van, & Overdijk, R. Van. (2016). Het optimaliseren van fietsgedrag in verkeersmodellen, 1–27.

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by bicycle: An overview of the literature. *Transport Reviews*, *30*(1), 59–96. https://doi.org/10.1080/01441640903187001

Hölzel, C., Höchtl, F., & Senner, V. (2012). Cycling comfort on different road surfaces. *Procedia Engineering*, *34*, 479–484. https://doi.org/10.1016/j.proeng.2012.04.082

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, *3*(1), 63–75. https://doi.org/10.3328/TL.2011.03.01.63-75

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice*, *44*(9), 754–765. https://doi.org/10.1016/j.tra.2010.07.008

Papinski, D., & Scott, D. M. (2011). A GIS-based toolkit for route choice analysis. *Journal of Transport Geography*, *19*(3), 434–442. https://doi.org/10.1016/j.jtrangeo.2010.09.009

Pucher, J., & Buehler, R. (2008). Cycling for Everyone: Lessons from Europe. *Transportation Research Record: Journal of the Transportation Research Board*, *2074*, 58–65. https://doi.org/10.3141/2074-08

Rijksoverheid. (2018). *Staatssecretaris Van Veldhoven: 200.000 mensen uit de file op de fiets*. [online] Available at: https://www.rijksoverheid.nl/actueel/nieuws/2018/06/12/staatssecretaris-van-veldhoven-200.000-mensen-uit-de-file-op-de-fiets [Accessed 2 Nov. 2018].

Rogelj, J., den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., … Meinshausen, M. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature*, *534*(7609), 631–639. https://doi.org/10.1038/nature18307

Schepers, P., Hagenzieker, M., Methorst, R., Van Wee, B., & Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis and Prevention*, *62*, 331–340. https://doi.org/10.1016/j.aap.2013.03.032

Schepers, P., Twisk, D., Fishman, E., Fyhri, A., & Jensen, A. (2017). The Dutch road to a high level of cycling safety. *Safety Science*, *92*, 264–273. https://doi.org/10.1016/j.ssci.2015.06.005

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, *36*(5), 511–539. https://doi.org/10.1007/s11116-009-9201-4

Stefansdottir, H. (2014). Urban routes and commuting bicyclists' aesthetic experience. *FORMakademisk*, *7*(2). https://doi.org/10.7577/formakademisk.777

Su, J. G., Winters, M., Nunes, M., & Brauer, M. (2010). Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada. *Transportation Research Part A: Policy and Practice*, *44*(7), 495–505. https://doi.org/10.1016/J.TRA.2010.03.015

Van Duppen, J., & Spierings, B. (2013). Retracing trajectories: The embodied experience of cycling, urban sensescapes and the commute between "neighbourhood" and "city" in Utrecht, NL. *Journal of Transport Geography*, *30*, 234–243. https://doi.org/10.1016/j.jtrangeo.2013.02.006

Winters, M., Teschke, K., Grant, M., Setton, E., & Brauer, M. (2010). How Far Out of the Way Will We Travel? *Transportation Research Record: Journal of the Transportation Research Board*, *2190*(October 2015), 1–10. https://doi.org/10.3141/2190-01

Zimmermann, M., Mai, T., & Frejinger, E. (2017). Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies*, *75*, 183–196. https://doi.org/10.1016/j.trc.2016.12.009

Assessing the impact of bicycle infrastructure on cyclists' route choice in the Dutch province of Noord-Brabant

Appendix

Appendix 1: Descriptive statistics

General route characteristics

	Length observed	Length shortest	Abs. deviation	Rel. deviation
Minimum	0,63	0,15	0.00	1.00
Maximum	26,75	26,45	5.06	1.36
Average	10,90	10,23	0,66	1,07
St.deviation	5,30	5,00	0,63	0,06
Q1	7,27	6,53	0,24	1,03
Q2	10,4	9,66	0,46	1,06
Q3	14,23	13,67	0,89	1,10

Road type variables

	Separated bicycle path	Non-separated bicycle path	Mixed roads
Minimum	-0,17	-0.62	-1,00
Maximum	0,89	0,46	0,10
Average	0,28	0,01	-0,40
St.deviation	0,17	0,09	0,17

Road quality variables

	Good road quality	Reasonable road quality	Bad road quality
Minimum	-0,88	-0,64	-0,40
Maximum	0,86	0,67	0,13
Average	0,24	0,02	-0,01
St.deviation	0,19	0,14	0,07

Intersection variables

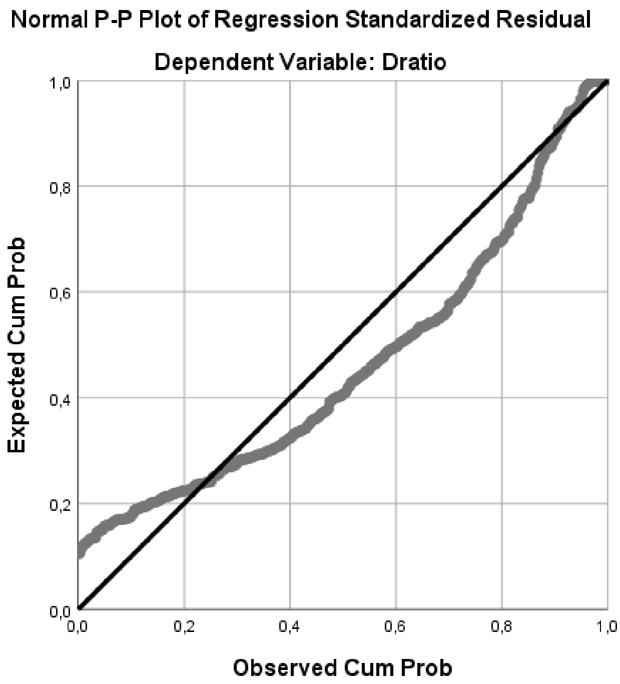
	Roundabout ratio	Intersection ratio	Intersection signal ratio
Minimum	-0,76	-2,53	-2,08
Maximum	5,20	5,26	6,37
Average	0,34	0,32	0,39
St.deviation	0,60	0,77	0,78

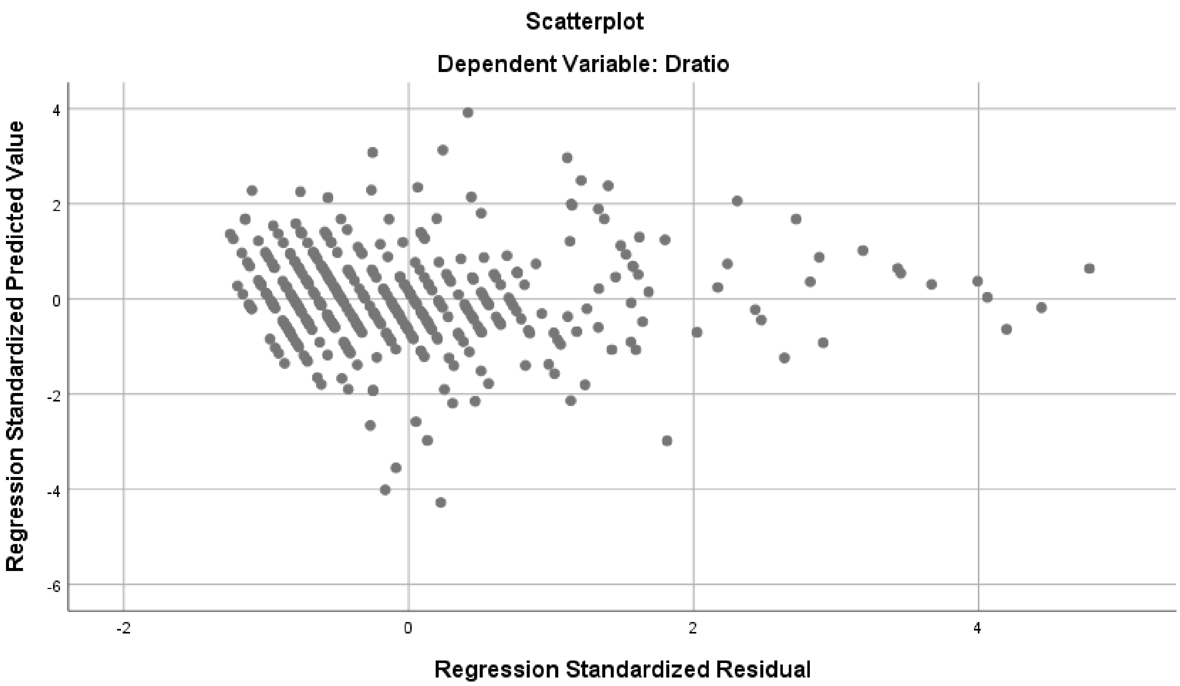
Road surface variables

	Asphalt/concrete surfaces	Other road surfaces
Minimum	-0,49	-0,55
Maximum	0,86	0,55
Average	0,25	-0,05
St.deviation	0,16	0,14

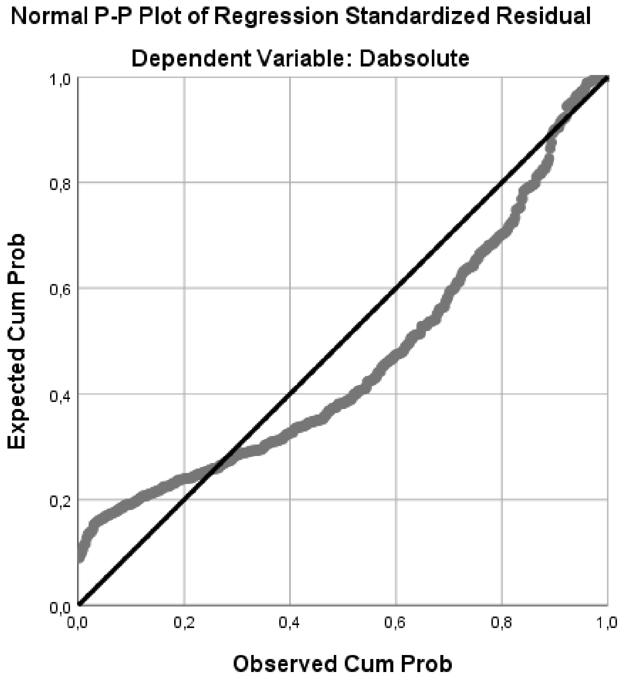
Appendix 2: Testing the assumption for multi-linear regression

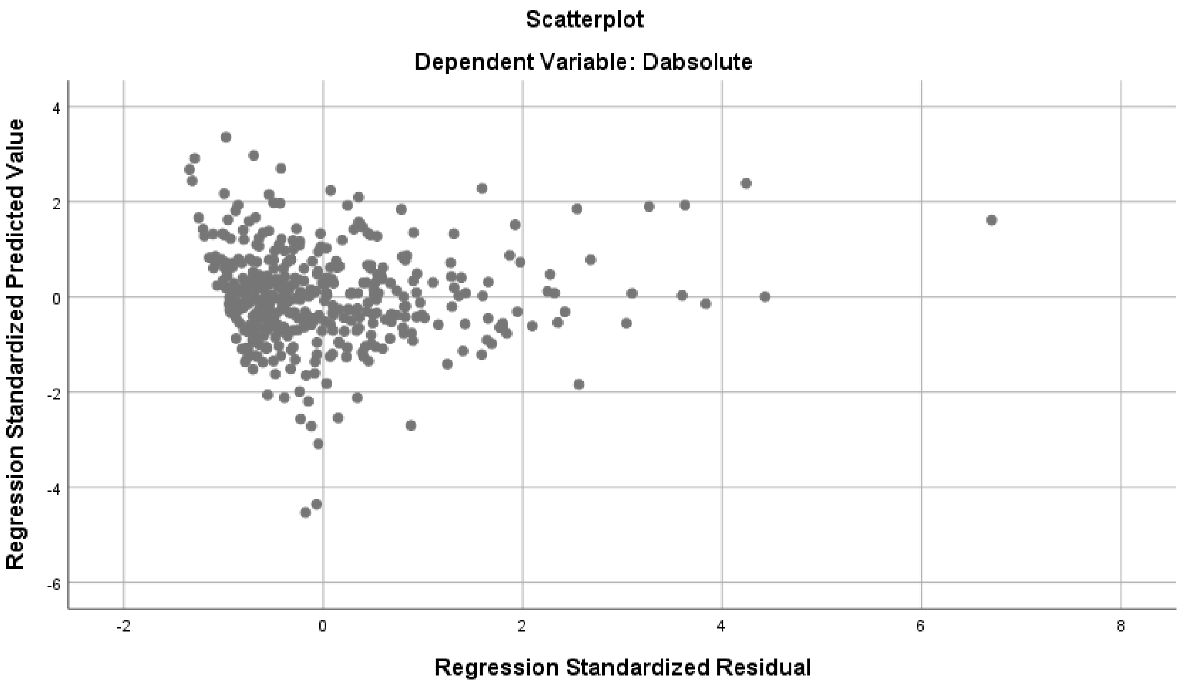
Relative difference variable





Absolute difference variable





Appendix 3: Multicollinearity

Correlations

		Separated bicycle path	Non-separated bicycle path	Mixed roads	Good road surface	Reasonable road surface	Bad road surface	Roundabout R	Crossing R	Crossing sig R	Asphalt	Other road surfaces
Separated bicycle path	Pearson Correlation	1	-,261**	-,574**	,063	-,114*	-,185**	,088	,082	,005	,063	-,136**
	Sig. (2-tailed)		,000	,000	,197	,020	,000	,072	,095	,924	,195	,005
	N	420	420	420	420	420	420	420	420	420	420	420
Non-separated bicycle path	Pearson Correlation	-,261**	1	-,183**	,254**	-,127**	-,044	-,039	-,026	,041	,103*	-,055
	Sig. (2-tailed)	,000		,000	,000	,009	,370	,430	,590	,397	,035	,259
	N	420	420	420	420	420	420	420	420	420	420	420
Mixed roads	Pearson Correlation	-,574**	-,183**	1	-,091	,107*	,245**	-,057	-,081	,158**	,024	,203**
	Sig. (2-tailed)	,000	,000		,061	,028	,000	,248	,098	,001	,627	,000
	N	420	420	420	420	420	420	420	420	420	420	420
Good road surface	Pearson Correlation	,063	,254**	-,091	1	-,573**	-,217**	-,100*	-,012	,090	,638**	-,460**
	Sig. (2-tailed)	,197	,000	,061		,000	,000	,040	,810	,067	,000	,000
	N	420	420	420	420	420	420	420	420	420	420	420
Reasonable road surface	Pearson Correlation	-,114*	-,127**	,107*	-,573**	1	-,081	,103*	-,086	-,101*	-,283**	,383**
	Sig. (2-tailed)	,020	,009	,028	,000		,099	,035	,077	,038	,000	,000
	N	420	420	420	420	420	420	420	420	420	420	420
Bad road surface	Pearson Correlation	-,185**	-,044	,245**	-,217**	-,081	1	,080	,014	,048	-,271**	,369**
	Sig. (2-tailed)	,000	,370	,000	,000	,099		,102	,777	,322	,000	,000

	N	420	420	420	420	420	420	420	420	420	420	420
Roundabout R	Pearson Correlation	,088	-,039	-,057	-,100*	,103*	,080	1	-,042	-,023	-,125*	,054
	Sig. (2-tailed)	,072	,430	,248	,040	,035	,102		,392	,632	,011	,269
	N	420	420	420	420	420	420	420	420	420	420	420
Crossing R	Pearson Correlation	,082	-,026	-,081	-,012	-,086	,014	-,042	1	-,044	-,129**	-,013
	Sig. (2-tailed)	,095	,590	,098	,810	,077	,777	,392		,368	,008	,786
	N	420	420	420	420	420	420	420	420	420	420	420
Crossing sig R	Pearson Correlation	,005	,041	,158**	,090	-,101*	,048	-,023	-,044	1	,107*	-,122*
	Sig. (2-tailed)	,924	,397	,001	,067	,038	,322	,632	,368		,029	,012
	N	420	420	420	420	420	420	420	420	420	420	420
Asphalt	Pearson Correlation	,063	,103*	,024	,638**	-,283**	-,271**	-,125*	-,129**	,107*	1	-,773**
	Sig. (2-tailed)	,195	,035	,627	,000	,000	,000	,011	,008	,029		,000
	N	420	420	420	420	420	420	420	420	420	420	420
Other road surfaces	Pearson Correlation	-,136**	-,055	,203**	-,460**	,383**	,369**	,054	-,013	-,122*	-,773**	1
	Sig. (2-tailed)	,005	,259	,000	,000	,000	,000	,269	,786	,012	,000	
	N	420	420	420	420	420	420	420	420	420	420	420

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).