Electric cycling in Flanders

Empirical research into the functional use of the e-bike

Paola Astegiano, Chris M.J. Tampère, Carolien Beckx, Inge Mayeres, Willem Himpe

[xxx]

8/31/2017

Wetenschapspark 5 bus 6 | 3590 Diepenbeek

Consortium UHasselt, KU Leuven en VITO

ISBN: [AANVULLEN] – [AANVULLEN]

Wettelijk Depot: [AANVULLEN]

NUR [AANVULLEN]

Dit rapport kwam tot stand met de steun van de Vlaamse Overheid, programma ‘Steunpunten voor Beleidsrelevant Onderzoek’. In deze tekst komen onderzoeksresultaten van de auteur(s) naar voor en niet die van de Vlaamse Overheid. Het Vlaams Gewest kan niet aansprakelijk gesteld worden voor het gebruik dat kan worden gemaakt van de meegedeelde gegevens.

Het Steunpunt Verkeersveiligheid 2012-2015 voert in opdracht van de Vlaamse overheid beleidsondersteunend Wetenschappelijk onderzoek uit over verkeersveiligheid. Het Steunpunt Verkeersveiligheid is een samenwerkingsverband tussen de Universiteit Hasselt, de KU Leuven en VITO, de Vlaamse Instelling voor Technologisch Onderzoek.

Inhoudstafel

Figuren

Tabellen

Samenvatting

Titel: Elektrisch fietsen in Vlaanderen

Ondertitel: Empirisch onderzoek naar het functioneel gebruik van de e-fiets: methodieken en eerste resultaten

Net als in andere landen stimuleert het beleid in Vlaanderen het gebruik van de fiets als vervoerswijze. In deze context is de komst van de elektrische fiets (e-fiets) een beloftevolle ontwikkeling. De ambitie van het project SPRINT is om te verkennen waar het beleid, bv. middels het aanbieden van gepast infrastructuur, ervoor kan zorgen dat de vele gunstige effecten van de e-fiets in het kader van een duurzaam mobiliteitsbeleid worden waargemaakt, zonder negatieve implicatie op de verkeersveiligheid. Dit project draagt hiertoe bij door inzichten te verschaffen in hoe gebruikers in de praktijk hun e-fiets gebruiken en welke rol die speelt in het mobiliteitsgedrag. De focus ligt hierbij op het functioneel eerder dan het recreatief gebruik van de e-fiets, met inbegrip van woon-werk verkeer. Verder wordt onderzocht hoe bestaande fietsinfrastructuur, waarvan de dimensies vaak relatief beperkt zijn doordat ze uitgebouwd wordt als onderdeel van bestaande wegen, al dan niet past bij de e-fiets: welke kenmerken van routes krijgen de voorkeur van e-fietsers en hoe kunnen we geschikte samenhangende netwerken voor functioneel (e-)fietsen ontwerpen?

Voor deze studie werden enquêtes gehouden en werd GPS data verzameld met behulp van een logger op de e-fiets van de vrijwilligers die aan het project deelnamen. De deelnemers werden gerekruteerd in Gent en Leuven. Zij hadden de beschikking over een gepersonaliseerde webpagina waarop hun persoonlijke gps-tracks gevisualiseerd konden worden, en waarop bijkomende vragen voor de enquêtes konden worden beantwoord.

Uit de enquêtes kunnen we afleiden dat mensen de e-fiets zelden als enige vervoermiddel gebruiken, maar eerder als aanvulling op de klassieke fiets en de auto. Zij gebruiken hun e-fiets vooral voor woon-werkverkeer. Gerapporteerde gevaarlijke situaties verschillen eigenlijk nauwelijks van die met gewone fietsen en zijn veelal te wijten aan interacties met ander verkeer op kruisingen en aan het ontbreken van fietspaden. Deelnemers vinden de beschikbaarheid over veilige fietsenstallingen en publieke laadpunten sterk ontoereikend, maar tonen zich over het algemeen tevreden met de fietsinfrastructuur (fietspaden).

Het routekeuzemodel dat op basis van de verzamelde gps data kon worden geschat, toont dat meer nog dan door afstand de routekeuze wordt bepaald door afslagbewegingen en overgangen in wegenhiërarchie. Wellicht meer nog dan gewone fietsers, zijn elektrische fietsers bereid om om te rijden als ze daarmee minder onderbrekingen ondervinden door kruispunten en (vooral linksaf) bewegingen. Fietsen langs secundaire wegen wordt verkozen boven primaire wegen, maar ook boven lokale wegen en paden waar autoverkeer niet toegestaan is.

De studie legt ook de basis van een set kwantitatieve indicatoren op basis van gps-data waarmee kan bepaald worden hoe e-fietsers de hiërarchie van het fietsnetwerk ervaren. Dit moet het voor het eerst mogelijk maken om te meten hoe goed het fietsnetwerk voldoet aan de eisen die in theorie worden gesteld door het volledige verplaatsingspatroon van een e-fietser.

De vastgestelde routevoorkeuren laten het belang zien van een goed gestructureerd hiërarchisch netwerk. Om de werkelijke noden van e-fietsers op dit vlak nog beter in kaart te brengen, stellen we een nieuwe methodologie voor om uit gps-data een impliciete hiërarchie van het fietsnetwerk vast te stellen, zoals gebruikers deze percipiëren. Hoewel de data die specifiek binnen dit project verzameld werd niet volstond om deze methodiek te valideren, verwachten we dat deze methode een erg waardevol instrument kan worden in op fietsers gerichte vervoersplanning, waarbij structurele veiligheid in een netwerkstructuur kan ingebouwd worden.

English Summary

Title: Electric cycling in Flanders

Subtitle: Empirical research into the functional use of the e-bike

Nowadays, in Belgium as in many other countries, a policy goal is to increase the bicycle modal share. In order to do that, particular attention is given to the electric bike. The ambition of the SPRINT project is to shape an image of the policy and infrastructure requirements needed to safeguard the numerous positive effects of a possible breakthrough of electric bikes as a basic portion of a sustainable mobility system. Indeed, electric bicycles have the potential of being an important driver in the transition towards sustainable mobility. However, it is important to ensure that this paradigm shift takes place under safe conditions. This project contributes to the stream of research in understanding how to support this transition towards more sustainable modes of transport, specifically by analysing mobility habits of e-bike owners. The key point of the SPRINT project is to understand whether the e-bike may represent a valid alternative for commuting (or functional) trips. Moreover (and not of secondary importance) it has the final objective to understand if the cycling infrastructures which are continuously being developed in Flanders (typically retrofitted along existing roads in densely built area with many spatial constraints, hence often of modest dimensions) can fulfil the needs of the e-bikers.

This study is based on the analysis of both survey and GPS data coming from devices installed directly on the e-bike of the people who volunteered to participate in this project. The recruitment phase was launched first in Ghent and subsequently in Leuven. Each participant had a personal webpage where he/she could visualize his/her gps-routes as well as have access to complementary surveys and interesting tools such as a map of public re-charge spots, newsletters, forum, etc.

The survey's results show how, in terms of ownership, the participants tend to accompany an electric bike with an ordinary bike and a car, and how the e-bike is mostly employed for home-work trips. The dangerous situations in which they incurred are very similar to those proper of ordinary bikes and are mainly due to interactions with other vehicles and/or lack of bicycle infrastructure. The participants are also completely dissatisfied with respect to the presence of safe storage and public re-charge spot while they are quite satisfied in terms of bicycle infrastructures.

The route choice model estimated from the collected gps tracks shows the importance of attributes such as turns and changes in road hierarchy (e.g. two consecutive links characterized by two different types of road) in route choice decisions from e-bikers. Probably more than regular cyclists, they prefer uninterrupted routes with less intersections and turning movements and are willing to make detours to achieve this. The most preferred roads are the one labelled as "secondary" roads.

These route choice preferences highlight the importance of a well-structured hierarchical cycling network. In order to further inquire the true needs of the (e-)cyclist population, we propose a new methodology to infer from gps-data an implicit hierarchical cycling network as it emerges from the user's perspective. Although the data available in this project did not allow validating this approach, we anticipate that this can be a highly valuable instrument in cycling transport planning but also in warranting structural safety in the network.

1 Introduction

1.1 Background and problem statement

In recent decades, the promotion of non-motorized modes of transport is increasing as part of more sustainable eco-mobility vision. In particular, cycling is being promoted, due to the fact that it is low-cost, low-polluting and produces great health benefits. Therefore, a broad literature has been recently focusing on better understanding the determinants of bicycle ownership and the way to further promote bicycling [1]. Considering this, different research directions were taken among the experts in this sector.

Groundbreaking studies [2] evaluated the factors that impact bicycle use frequency for an individual's commute to and from work as well as the integration of cycling with public transportation [3]. The first study shows how availability of showers or clothing lockers at the workplace does not appear to inspire bicycle commuters to commute by bicycle more frequently; on the other hand, using a bicycle for non-work trip purposes increase the frequency of commuting by bicycle to work. The second study focuses on highlighting the successful policy implemented by the Netherlands, Denmark and Germany in making bicycling safer: provision of separate cycling facilities along heavily travelled roads, combined with traffic calming of most residential neighbourhoods; full integration with public transport; traffic education and training of both cyclists and motorists; making driving expensive and inconvenient in the city centres; strict land-use policies that generate shorter and thus more bikeable trips. Different from the UK (where only about 1% of trips are by bike) these countries encourage the coordination of different sets of cycling policies. Knowledge about infrastructure preferences for cyclists led to determine that a correlation exists between the level of cycling confidence and preferred types of infrastructure [4], [5]. More recently, Heinen [6], [7] exhibited the influence of bicycle commuters' attitudes on mode choice decisions, under the assumption that cyclists with a more intense commuting journey, either in terms of distance or frequency, in general have a more positive attitude towards cycling. They tested this hypothesis with a factor analysis underlying three main factors: awareness, direct trip-based benefits (measured with Likert scale) and safety. They show that the decision to cycle is influenced by the factor "direct trip-based benefit" at all distances, whereas the "awareness" is influential only over long distances and "safety" over shorter distances.

Although it is generally expected that also in Flanders, the electric bicycle has the potential to increase the modal share of cycling and improve the general attitude towards cycling as a functional mode of travel (recreational use being beyond the scope of this research), hardly any data on this topic is available for Flanders/Belgium. This study intends to contribute to filling this knowledge gap.

1.2 State of the art

1.2.1 Worldwide situation

Unlike the aforementioned studies that address cycling in general, this project focuses particularly on the e-bike, which is nowadays gaining more and more popularity. Because of its higher speeds (compared to the ordinary bike) and longer reach, it extends the capabilities of normal cycling and could be an attractive alternative to the car. In addition, it can be more appealing for people who would otherwise be deterred by the physical effort of unsupported cycling. In the near future, the e-bike could become an important way to incentivize a shift from car to bicycle in order to reduce road congestion, traffic-related air pollution, road accidents and infrastructure costs. Currently, the world's leader e-bike market is China but, in the last few years, a strongly positive trend of the e-bike market share is also observed in north Europe and in the U.S.

In [8] and [9] the first investigation on how and why e-bikes developed so quickly in eastern countries has been performed, providing important insights to policy makers in China and abroad. They showed how timely regulatory policy can influence the purchase choices of millions, incentivizing the use of a new mode of transport introduced in the market. An additional way to increase electric bike use would also be to consider control strategies that limit the number of stops for this mode, through signal coordination or grade separated intersections, thus increasing the travel time advantage of electric bikes. In [10] the environmental and safety impacts of alternative modes, such as public transport or personal cars that are the standard competitor of the e-bike are analyzed. The study confirms that electric bikes are a clean mode of transport with low noise levels and zero tailpipe emissions. In terms of safety, it also showed how the fatality rates are nearly as low as bicycle fatality rates and much lower than cars.

Another important role is played by the U.S. and North America. Their markets are still behind China and Europe, but a strong group of researchers [11], [12], [13] is investigating which factors influence purchase decisions in these countries and, with a comparison between ordinary bikes and e-bikes, trying to understand whether e-bikes can effectively address barriers to bicycling and therefore encourage more sustainable mobility. Their results suggested that e-bike users cycle more often and to more distant locations. Moreover, e-bikes allow people with physical limitations to cycle thanks to electric assist.

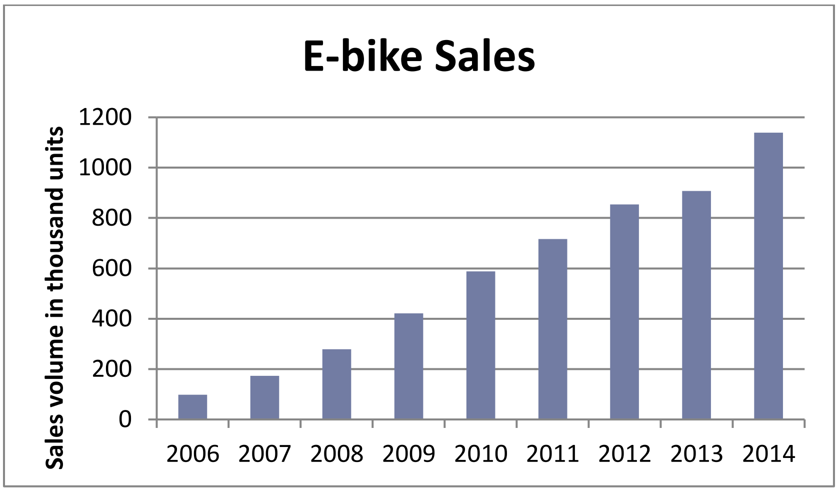
Even though the interest in studying the impact of the e-bike in the daily mobility habits is increasing, the current state of the art suffers of two major drawbacks: the lack of a large-scale analysis on the travel behaviour for e-bikes [14] and the need to quantify the influence of e-bikes on travel behaviour. In the authors' opinion, these gaps lead to underestimate both the role that the e-bike could play in the commuting trip (e.g. higher usage for longer distances) and its different needs (e.g. type of road, private storages, re-charge spots, etc) in comparison to the traditional bikes.

1.2.2 Belgian position in the worldwide context

In recent years and decades, in Belgium as in many other countries, the local authorities have been undertaking a range of activities to stimulate cycling as a daily transport mode. Consequently, decision makers are facing a lot of open questions, ranging from how to develop an effective cycling policy to understanding how to provide high-quality infrastructure.
In order to incentivize cycling as a daily urban transport mode, and rendering it competitive with the other modes, riding a bicycle has to be physically possible, safe and attractive. From here stems the importance of creating an integrated cycling infrastructure policy that takes in consideration cyclists' needs (e.g. design of the street, continuity of the route, etc.).
To do so, investments in making bike use more appealing (improvement of bicycle paths, introducing bike or e-bike sharing systems, etc.) seem necessary, but such interventions should be preceded by an exploratory phase aimed at gaining extra insights with respect to the way people perceive the different alternatives, and under which conditions they are willing to use a bike to substitute a car trip.
Above all, it is important to identify the locations where investment in cycling infrastructure would be most valued, hence improving the overall utilization of the cycling network in the region. This involves understanding travel behavior of cyclists or of potential cyclists, and factors influencing cyclists' decisions on destinations, whether to bike or not, and on route choices.

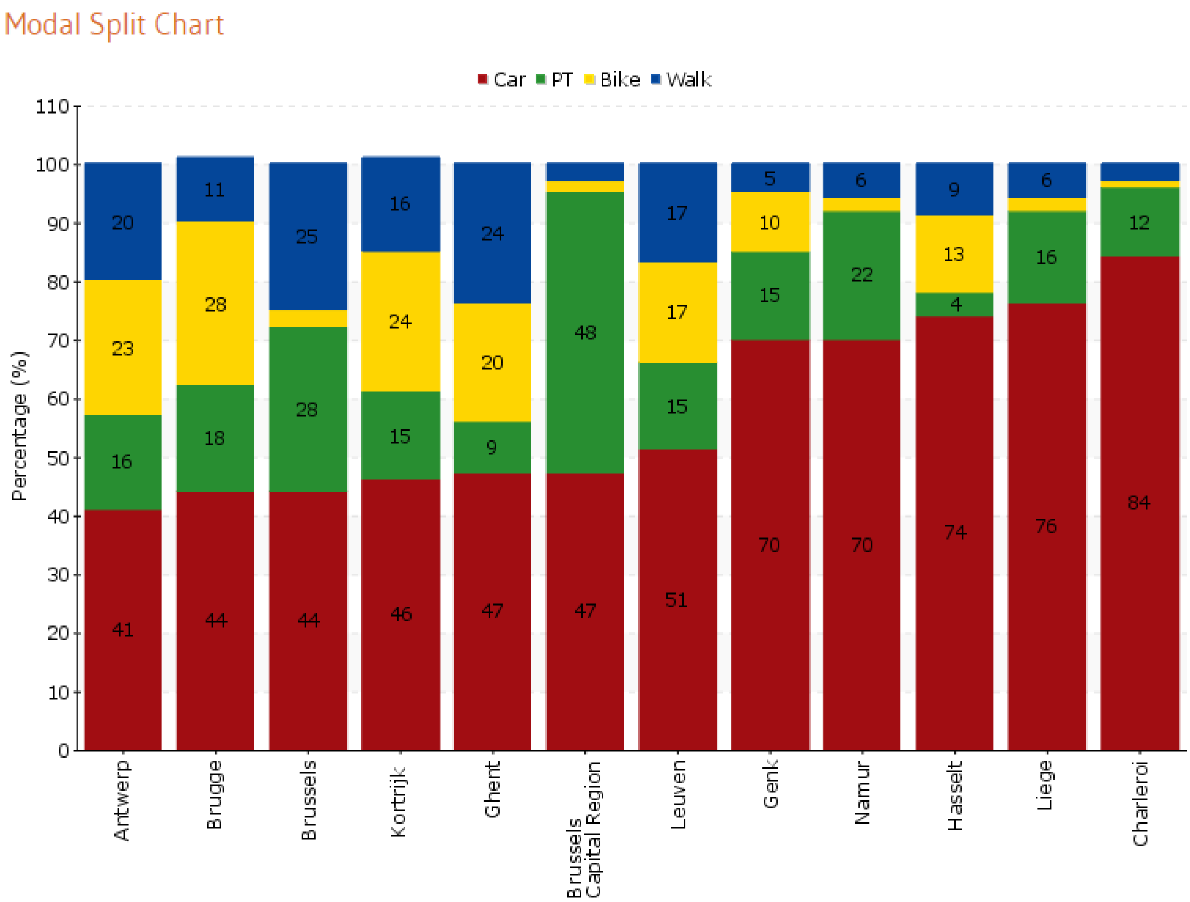
Through the SPRINT project of the Policy Support Centre on Traffic Safety, KU Leuven, VITO and the University of Hasselt aim to understand the role of the electric bike in this context. In fact, the global market for e-bikes continues to expand with the advent of new technological developments and the increasing affordability and availability of product offer. For example in 2015 the European Electric bicycle market grew by 4% [15]. In Belgium, the rise of the electric bike is mainly supported by women: the e-bike accounts for 58.2% of the increase in the ladies bike segment.

In Figure 1 we can also see the impressive trend of the e-bike sales in Europe from 2006 to 2014 (@Statista2016). This shows the number of electric bicycles sold in the European Union (EU) between 2006 and 2014, in thousand units. Sales of e-bikes grew steadily each year growing from 98 thousand units in 2006 to 1.12 million units in 2014.



Figuur 1: Trend of E-bike sales from 2006 to 2014

Moreover, the modal split [16] in the whole country of Belgium (Figure 2) shows how the car is the most used mode in all the cities. However, in the Flemish region the other modes of transport are more competitive than in Wallonia.



Figuur 2: Modal Split in Belgium

Where does the e-bike fit in this scenario? In fact, estimated (and freely accessible) statistics focused on e-bike use are quite scarce. Only from 2012 has the Eurostat (the European data bureau) started to register the import rates of electric bicycles while in 2014 were provided the first statistics [17] on internal production in Germany (of 410,000 e-bike sold in 2013 in Germany, 130,000 were imported and 278,000 were produced internally). In 2012, the EU's member states imported a total of 551,782 e-bikes (including speed-pedelecs). Unfortunately, official detailed sales data are available only for Germany, the Netherlands, Switzerland, France, Italy and Austria (Table 1). However, we can assume that for Belgium trends similar to those of its neighboring countries such as the Netherlands and Germany will hold. Non official statistics for Belgium [18] highlight how, in the last two years, the sales in Belgium increased by 150 percent, from 20,000 to 50,000 units sold.

Tabel 1: E-bike market in Europe

Country	2012	2011	2010
Germany	380.000	310.000	200.000
The Netherlands	171.000	178.000	166.000
Switzerland	52.900	50.000	35.000
France	46.100	40.000	38.000
Italy	48.200	45.000	40.000
Austria	40.000	32.000	20.000
Total EU states	**738.200**	**655.000**	**499.000**

1.3 Research goals and research questions

Based on the argument presented above, the key point of the SPRINT project is to understand whether the e-bike may represent a valid alternative for functional trips (like commuting). Several studies in Belgium focused on the behavior of traditional cyclists: the OVG, for example, reported how the functional bicycle trips performed using a normal bike (non-electric) are, for the 80%, shorter than 5 km, making cycling mainly a local transport mode. Cycling can play, anyhow, a significant role in longer trips as a feeder mode for public transport. In this combination the traditional bike may be useful to support the transition towards more sustainable modes of transport.
Moreover (and not of secondary importance) the SPRINT project has the final objective to understand if the cycling infrastructures which are currently under construction in Flanders (typically retrofitted along existing roads in densely built area with many spatial constraints, hence often of modest dimensions) can fulfil the needs of the e-bikers. Are the different authorities involved in this construction aiming at guaranteeing the continuity or directness (avoiding detours) of the route? Do any differences exist in terms of safety between ordinary bikes and e-bikes? Are e-bike riders for example more inclined to share the road with the general traffic because of their higher speed compared to the ordinary bike? Does the presence of e-bike facilities, such as re-charge spots and storage, have an influence on the usage of the e-bike? Are these facilities currently adequate or is there any room for improvements?

The ambition of the study is to shape an image of the policy and infrastructure requirements needed to safeguard the numerous positive effects of a possible breakthrough of electric bikes as a basic portion of a sustainable mobility system. Indeed, electric bicycles have the potential of being an important driver in the transition towards sustainable mobility. However, it is important to ensure that this paradigm shift takes place under safe conditions. In fact, some research ([19],[20]) indicates that safety risks of electric bikes, depending on infrastructure characteristics, could yield higher accident risks, in proximity of an intersection, compared to ordinary bikes. The first study focused on the behavior at the signalized intersections finding that e-bikers are more inclined to run with red-light while the second one investigated the underestimation of the speed by the drivers when an e-bike is approaching. Their results showed that with increasing cyclist speed, accepted time gaps became significantly shorter.

The ambition of this project is not a direct statistical analysis of e-bike safety aspects. Rather, it looks at preconditions that may structurally affect safety as the transition towards e-cycling continues. It aims at the sequence of questions:

- which users within the travelers' population may decide to use an electric bicycle for which kind of trip pattern?
- what affects their decision to use certain routes for e-cycling?
- and how can stakeholders plan a functional bicycle network suitable for e-bikers, that inherently offers convenience and safety as the number of e-cyclists is expected to rise in the future?

The ambition of the project was to develop methodologies to answer these three main research questions, and to collect data on which these methodologies could be applied. The outcome of the project was eventually more successful in the methodological developments than in the data acquisition and validation. Still, some partial and preliminary insights could be retrieved from the data and are presented in this report. However, as the methodology promises much richer insights if more and better data is available, continuation of this research and of the data collection, as well as data transfer and big data techniques are highly recommended.

Summarizing the SPRINT project aims to answer the following research questions:

1. Which is the profile of the e-bike user?
2. Which type of mode of transport the e-bike owner tends to own together with the e-bike?
3. For which type of activity is the e-bike mainly used?
4. How people reconstruct their mobility habits after the acquisition of the e-bike? And how did it reflect on their ownership decisions?
5. What is e-bike owners' opinion on the e-bike facilities such as re-charge spot and storage?
6. How is the travel behavior of e-bike owners in terms of travelled distance, time and speed?
7. How does the travel speed of e-bikers depend on user's characteristics and weather conditions?
8. Which type of route do the e-bikers prefer to use?
9. In terms of safety, which are the typical accidents in which the e-bikers are involved? Are they different compared to ordinary bikes?
10. Does the bicycle network used by the e-bikers coincide with that one suggested by policy makers?

1.4 Defining the methodology of the research

The area of study focuses on the Flemish region of Belgium, where the highest population density is found. This region counts a population of around 6.5 million inhabitants, mostly concentrated in the area circumscribed by the Brussels-Antwerp-Ghent-Leuven agglomerations.

The project's recruitment campaign started in 2014, initially focusing on the city of Ghent. The idea was to convince 300 people in participating as volunteers in a GPS tracking campaign. In order to be eligible they had to own at least one e-bike and they had to allow us to install a GPS device on it. We immediately encountered several issues: the installation of the device implied voided warranties, the installation was not possible for all brands and types of e-bikes, people had to go personally to a shop in order to get the installation and wait two hours before to have successfully installed the device. These issues convinced us to broaden the area of study also to Leuven and to reconsider the size of the initial sample. This final decision was also supported by the fact that numerous studies abroad met similar difficulties in recruiting e-bike riders, using at the end very small sample (at most 30 participants) [21], [22], [23]. Our ultimate sample was finally composed of 60 users.

The main core of the project is based on a GPS tracking campaign, but the participants could also access a personal web-page [24] in which they could visualize their daily routes and fill in surveys and

access other interesting tools (recharge points map, newsletter, forum, etc.). With the support of the surveys we aimed to answer at research questions number 1-2-3-4-5-9 included in paragraph 1.3. Ideally, in order to answer these research questions one should have submitted the same survey to the same participants before and after the acquisition of the e-bike. In our case, it was not possible since we recruit only people that already had an e-bike, thus we rely on their memories for the information related with their mobility habits before the acquisition of the e-bike. Moreover, we are forced to ask them their currently mobility habits (which mode for which activity) because the GPS device is installed directly on the e-bike and therefore we did not register trips with the other modes of transport. For question number 4, in order to apply our methodology (full details in section 2.3.2) we would have need panel data over different years for the same people with both travel diary and ownership information included. Since we don't have it available for Belgium, we will show our methodology applying it at partial data of a case study in Ghent (when the e-bike was not yet included) where we have the observations only related to a particular year (2008). Thanks to the GPS data we will instead answer to question number 6-7-8-9-10.
For the route choice modeling we needed a routable cycling network to generate the route alternatives for each trip. Somewhat to our surprise this map was hardly available. The transport planning models for policy support in Flanders ("multimodaal model Vlaanderen" and the provincial spin-offs thereof) do exist for the entire Flanders but do not contain any cycling infrastructure data. Other data sources we consulted were either incomplete (e.g. covering only certain municipalities) or available only in encrypted formats for particular commercial purposes (e.g. routing in cycle-gps devices). The only resource free available (except for OpenStreetMap), was the "wegenregister" network dataset owned by AGIV (Agentschap voor Geografische Informatie Vlaanderen). It contains all the roads of Belgium with their related attributes and it is split in different subset. For this study two of these subsets are employed: "Wegsegment" (links) and "Wegknoop" (nodes). A description of the dataset is included in section 2.3.1.1.

For question number 7 we combined the GPS data with the results of the survey while for question number 8 the GPS data are combined with the open source data of weatherbase.com website.

Finally, in order to answer question 10, we propose a new approach to build a hierarchical network as experienced by cyclists starting from gps data and retrieving some metrics directly from them.

The research reported in this report was exploring new terrain. It appeared that the ambition of the project was too high to be fully successful. Whereas we set out with some innovative data processing and modeling methods that would be tailored to the specific case of the e-bike, the data availability forced us to reduce the original ambitions in several ways. For one, less data than expected could be collected during the project itself (in terms of number of participants volunteering to be gps-tracked, and in terms of surveys completed by these participants). Secondly, application of the methodologies and models that we intended to use in this project required availability of other data sources and processing tools that we assumed to be readily available (like: digital maps containing relevant bicycle attributes, routable digital cycle maps, map matching software, route generation algorithms), but proved to be hard or impossible to obtain.

As a result, the results delivered in this project are less decisive and empirically validated, whereas more methodological innovations have been developed and described (even though sometimes tested on other data than the e-bike datasets collected during this project). In simpler phrasing: so far we learned less about e-biking in Flanders than we had hoped initially, but we learned a lot about how to better collect, process and model data to improve this empirical knowledge in the future.

1.5 Outline and contributions of the project

The remainder of this report is split into three parts: (i) Data analysis, containing the full analysis related with both surveys and GPS data; (ii) Safety aspects, in which we will point out the main safety issues related with the use of the e-bike; (iii) Policy implications, in which we will highlight the most interesting findings in terms of analysis of data and safety from a policy perspective.

1.5.1 Data Analysis

1.5.1.1 Descriptive Analysis: Survey Data

Through the surveys included on the website we aimed at investigating the e-biker's profile as well as his/her mobility habits before and after the acquisition of the e-bike; we moreover collect their opinions about electric bike facilities currently deployed in Flanders. Thanks to these surveys, we obtained a complete overview of the socio-demographic information of our sample while also collecting information about which mode of transport they used for a particular trip and with which frequency. We are therefore able to understand how they changed this frequency and the mode used for their daily activities after the acquisition of the electric bike.

1.5.1.2 Descriptive Analysis: GPS Data

For this study a GPS tracking device (GPS logger) was installed directly on the e-bikes of people who volunteered to participate in the SPRINT research program. The tracking devices allowed the collection of GPS data on the cycling trajectories without inferring the normal activities of the participant. Through the analysis of more than 14000 trips we aimed at investigating: the existence of a possible relationship between user's characteristics (e.g. age, experience) and trip features (e.g. speed); the influence of the weather conditions on e-bike rides; how e-bike riders behave during different days of the week and different hours of the day.

1.5.1.3 Methodology

This part of the report will be divided in three sub-sections:

1.5.1.3.1 Investigating route choice decisions

We develop a route choice model that combines trip features and road network characteristics.

1.5.1.3.2 Investigating the influence of daily modal choice on ownership decisions

We quantify the explanatory power of mobility-related attributes in explaining vehicle ownership decision using the data obtained from the BMW (Behaviour and Mobility within the Week) project [25].

1.5.1.3.3 Building a hierarchical network as experienced by (e-)cyclists

Using gps data we suggest a methodology to characterize each link of the network in a cyclist's trip and aggregating this over all observed trips, hence revealing the collectively perceived hierarchy of the network. This can be used to analyze whether the infrastructure provided matches the cyclist's needs.

1.5.2 Safety Aspects

In this section we further investigate the route choice decisions of our users highlighting a possible relation between the choice of a route and the exposure to traffic conflicts. Moreover, with the support of a survey we will show which areas/locations are perceived as dangerous and from which type of riders (e.g. elderly, less expert, etc.). The speed analysis, already showed in the previous section, will be also framed in this context.

1.5.3 Policy implications

To the extent possible, we extrapolate our analyses in the previous sections, providing policy makers with an overview of the infrastructure requirements needed to safeguard the numerous positive effects of a possible breakthrough of electric bikes as a basic portion of a sustainable mobility system.
For instance, investments in increasing the number of safe spots where to publicly store the e-bike are indispensable in order to incentivize the use of the e-bike also for non-commuting trips. Other possible improvements should aim at guaranteeing the continuity of the trip decreasing the number of conflicts with pedestrians more than with the general traffic.
Furthermore, we point out the importance to build the hierarchy of a cycling network starting from the true usage of the network itself and derive quality measures from it.

2 Data Analysis

2.1 Descriptive statistics: Survey Data

The main core of the project is based on a GPS tracking campaign, but the participants could also access a personal web-page [24] in which they could visualize their daily routes and fill in surveys and access other interesting tools (recharge points map, newsletter, forum, etc.).

Through the surveys included on the website we aimed at investigating the e-biker's profile as well as his/her mobility habits before and after the acquisition of the e-bike; we moreover collect their opinions about electric bike facilities currently deployed in Flanders. These surveys provide socio-demographic information of our sample while also collecting information about which mode of transport they used for a particular trip and with which frequency. We are therefore able to understand how they changed this frequency and the mode used for their daily activities after the acquisition of the electric bike.

On a total sample of 60 participants, 43 filled in the survey (of which 3 in anonymous way). Table 2 contains a detailed description of our sample in which all the percentages are computed against the full sample (60 pp) and therefore since we don't know the information for the remaining 17 participants the percentage showed in the table don't sum to 100%. Our sample is homogenously gender distributed; it has an average income proper of the Belgian middle-class; even though a consistent part of the sample has an age included between 40 and 60 years, the younger age (<40) is more represented than the old one (>60).

Tabel 2: Socio-Demographic information

Socio-Demographic Information					
Male	42%	**Gender**	Blue-collar employer	3%	**Employment Status**
Female	50%		White-collar employer	94%	
Age (<=40)	15%		Self-employed	3%	
Age (40-60)	23%	**Age**	Not employed	13%	
Age (>=60)	8%		Live alone	7%	**Living Situation**
Income (0-750 € per month)	7%		Together with partner without kids	15%	
Income (751-1500 € per month)	15%		Together with partner and kids	38%	
Income (1501-2000 € per month)	7%		Lower than college education	10%	**Level of education**
Income (2001-2500 € per month)	22%	**Income**	College	27%	
Income (2501-3000 € per month)	10%		University	27%	
Income (>3000 € per month)	5%		Higher than university education	3%	
City Center	15%	**Postal code**	Driving License	65%	**Driving License**
Surroundings	52%				

In Table 3 and Table 4 a summary of the ownership information is included at individual and household levels respectively. Since the GPS device is installed directly on the e-bike (more details in section 2.2) all of our participants have to own at least one e-bike. Among them only 12% had the e-bike from less of six months at the beginning of the tracking campaign. Almost everybody has also at least an ordinary bike (not surprising for Belgian standards). Car ownership is also highly consistent, while this is not the case for public transport subscriptions (either for bus or train): only 13% of the people declare to own one.

Tabel 3: Vehicle ownership at individual level

Vehicle ownership (individual level)

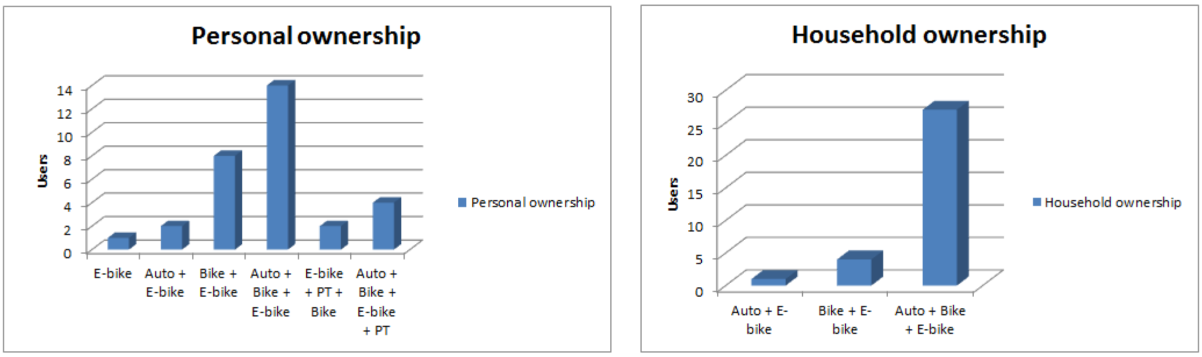
	0	**1**	**2**	**>2**
Bike	12.5%	52.5%	20%	15%
Electric Bike	0	87.5%	12.5%	0
Car	27.5%	70%	0	2.5%
PT	53% (no)	13% (yes)		

Tabel 4: Vehicle ownership at household level

Vehicle ownership (Household level)

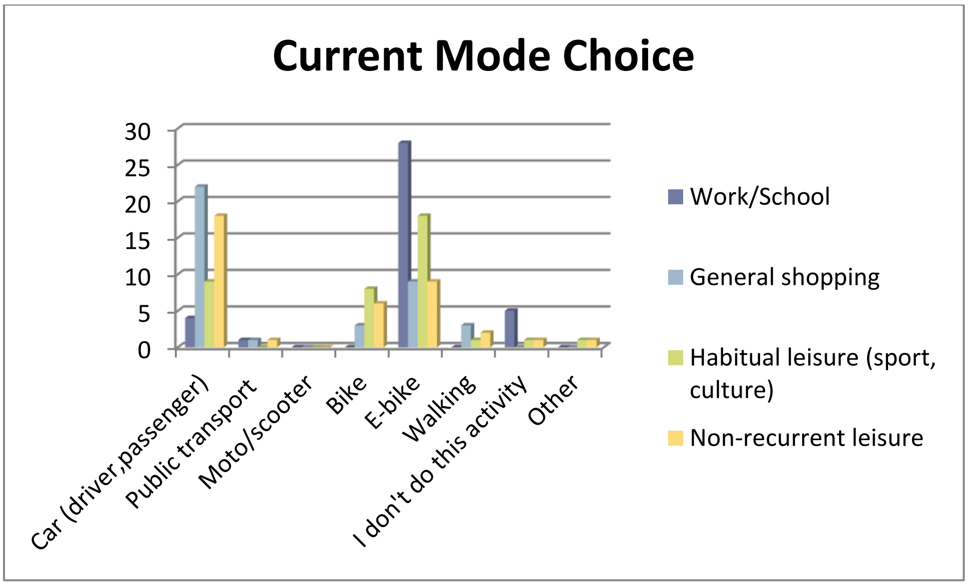
	0	**1**	**2**	**>2**
Bike	3%	10.5%	18%	68%
Electric Bike	0	74%	21%	5%
Car	10%	61%	26%	3%

More precisely they tend to own mainly bike and car, together with the e-bike, both at personal and household levels (Figure 3).



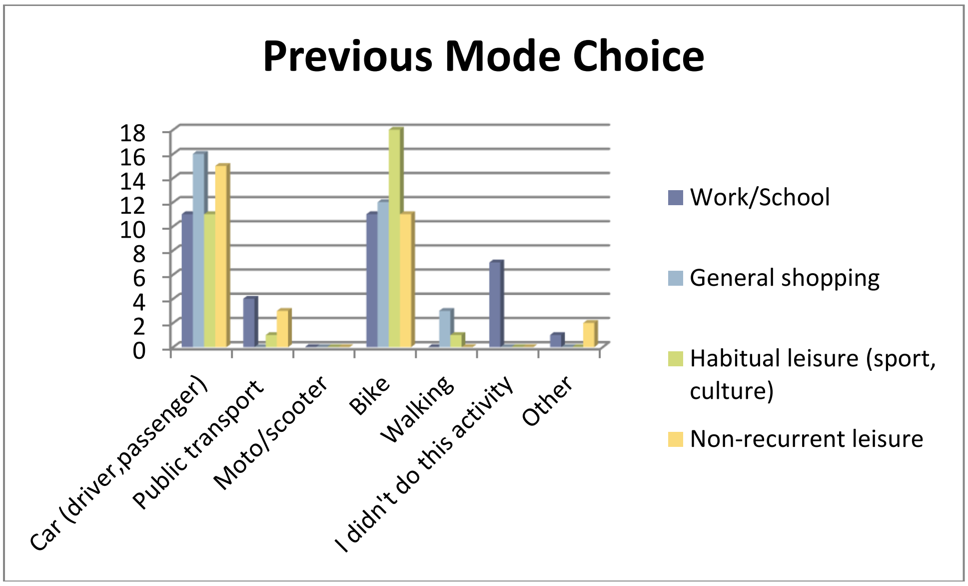
Figuur 3: Combination of modes owned both at individual and household level

Given these insights we decided to further investigate for which activities each mode of transport is used and with which frequency. We could also explore whether they changed their habits with the advent of the e-bike.



Figuur 4: Current mode choice

Figure 4 shows how the e-bike is mostly used for commuting trips (to work/school) and during free time ("habitual leisure"). Instead for "general shopping" and "non-recurrent activities" the car is the preferred mode of transport. Looking at the previous mobility habits (before the acquisition of the e-bike) in Figure 5, we can see how the e-bike mainly substituted the commuting trips previously done by ordinary bike and/or by car. Trips performed via Public Transport (PT), originally few, also diminished, next to disappearance. As already demonstrated in other studies [26], a traditional bike is likely to replace trips originally done by PT, consequently the ownership of the e-bike makes the value of this replacement still higher. For a detailed modal shift report, please refer to Table 5.



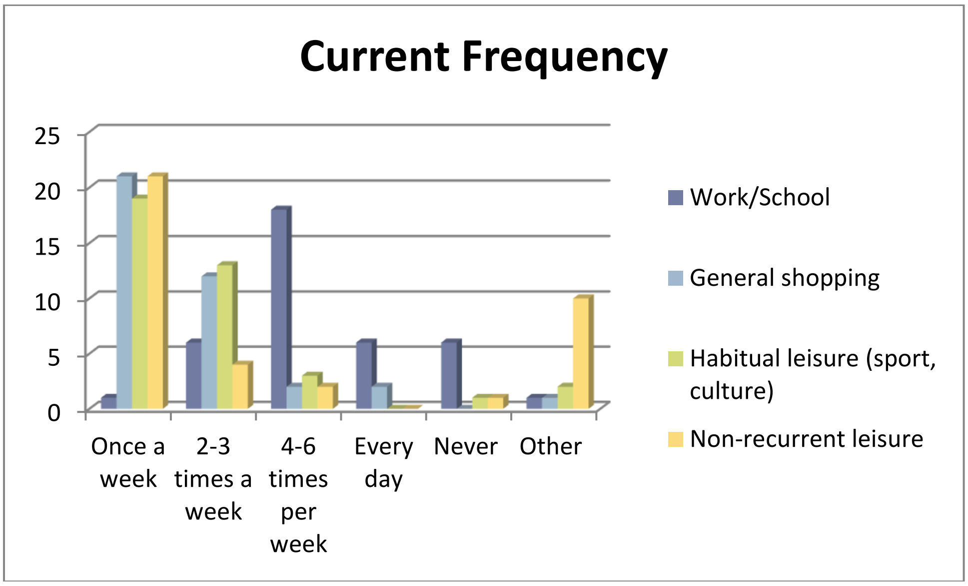
Figuur 5: Mode choice before the acquisition of the e-bike

Tabel 5: Modal shift

Changes in user percentage %

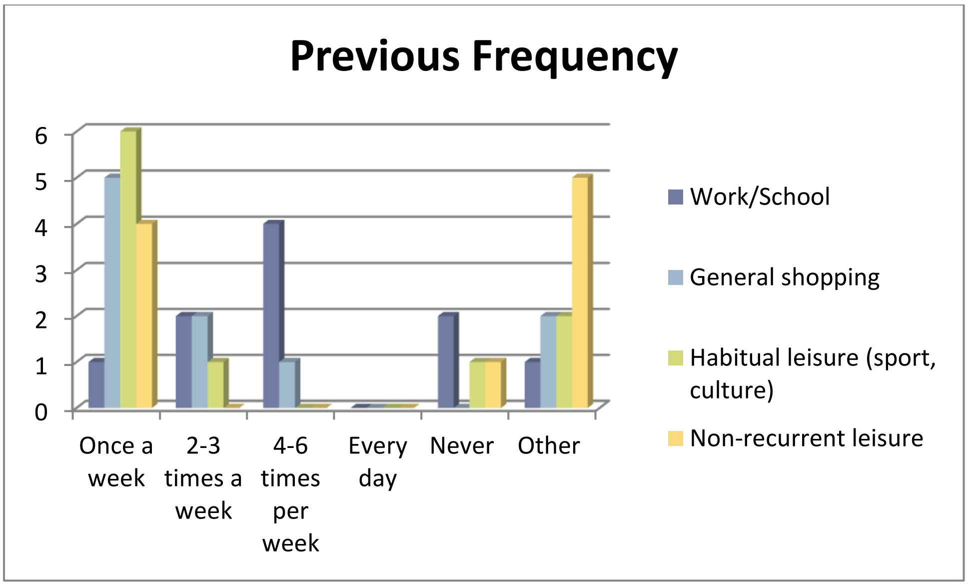
	Work/School	General shopping	Habitual leisure (sport, culture)	Non-recurrent leisure
Car to E-bike	64	No change	45	13
Pt to E-bike	50	50	100	67
Bike to E-bike	100	67	61	45
Walk to E-bike	Nobody used it before	No change	No change	Nobody used it before

In order to investigate any further relationship between the purpose of the trip and the chosen mode, we asked information also related with the frequency of these trips.



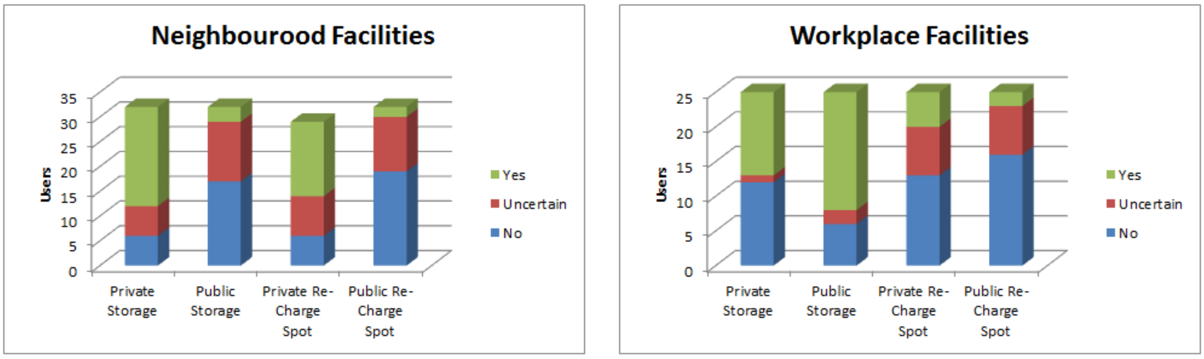
Figuur 6: Current frequency for each activity

Figure 6 shows how the use of modes of transport other than the e-bike (Figure 4) is associated with the activities performed with a lower frequency such as “General shopping” and “Non-recurrent activities”. Analyzing the frequency before the acquisition of the e-bike (Figure 7), only 17% of people declare to have changed it. The most relevant change is in the “habitual leisure”, an activity that, as we have previously (Figure 4) shown, is now mainly performed with an e-bike.



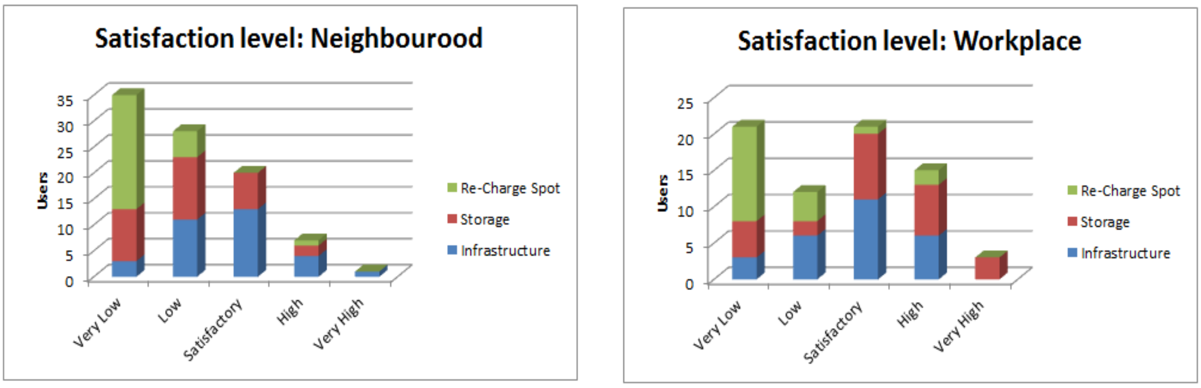
Figuur 7: Frequency before the acquisition of the e-bike

Through this survey we have also explored the presence of some facilities (Figure 8) that could encourage the rise of the e-bike market share.



Figuur 8: Presence of facilities in the neighbourood and at the workplace

In particular, we asked our sample to focus on the presence of storage and charging spots in their neighbourhood and at the workplace. In our opinion these two elements are fundamental for increasing the ownership/usage of e-bikes. The most relevant aspect appears to be the total lack of public re-charge spots. While it is possible to address this weakness at home location (by recharging the bike through private electricity plugs), this is not always the case at the workplace, especially if this lack is accompanied by the absence of a company-mandated storage, where sometimes public plugs are accessible. The share of people that declare not to have a private storage at the workplace is considerably high. In order to incentivize the use of the e-bike and given its initial high investment costs, providing a safe spot in which to store them becomes indispensable.



Figuur 9: Degree of satisfaction of e-cycling infrastructures

According to this analysis, users are entirely dissatisfied by the presence of re-charge spots while borderline satisfied regarding storages and network infrastructure (Figure 9).

2.2 Descriptive statistics: GPS Data

2.2.1 Pre-processing Framework

The automatic collection of travel behaviour data by GPS (Global Position System) systems is often employed to measure people's journeys and consequently to investigate their mobility habits [27]. The advantages of using GPS data are numerous, from the automatic depiction of trip origin, destination and route to the determination of travel times, distances and speeds. Another main advantage is that GPS data eliminate individual's misperception of the travel time and other trip parameters [28]. On the other hand, GPS devices can generate large amounts of raw data. Hence, intensive pre-processing routines are required to extract data suitable for travel behaviour analysis.

Concerning the analysis of GPS data, [29], [30] and [31] have already demonstrated the possibility to use GPS technologies in travel studies and evaluated the effectiveness of GPS in capturing the characteristics of different types of trips. Later studies, such as [32] and [33] evaluated the use of processed GPS data for both trip tracking and transportation-mode detection without the support of questionnaires. Their results showed that trip identification deviates slightly from the census data whereas for mode detection it was not possible to distinguish between transportation modes with similar speed such as bus and bike trips.

Strongly linked to the processing of GPS data, are the map matching methods that are necessary to align a sequence of observed user locations with the road network on a digital map. Map matching of the observed GPS trajectories will allow to extract the features of the network (e.g. road type) and to match them with the GPS data of interest. In [34] the authors proposed a probabilistic map-matching approach to overcome uncertainty problems cause by poor quality of the GPS data coming from smartphones. Previous research work based on GPS data focused on ordinary bikes and cars.

For this study a GPS tracking device (GPS logger) was installed directly on the e-bike of people who volunteered to participate in the SPRINT research program. The tracking device allowed the collection of GPS data on the cycling trajectories without inferring the normal activities of the participant.

Before performing a complete analysis of the GPS tracks, some data cleaning is needed. In particular the raw data frames (Figure 10) are processed and grouped into a single dataset through the following four steps: (i) verify the data integrity (e.g. checksum), since the data may be compromised due to hardware errors (data transmission or storage); (ii) parse the data frame attributes (e.g. latitude, longitude, speed, accuracy, altitude, date and time); (iii) filter invalid data (i.e. data frames without coordinates and other attributes); (iv) merge the valid data into a single dataset. After these steps, the final dataset is ready to be segmented into trips (i.e. the locations point that belong a specific trajectory are labelled as a trip).

Valid data frame:

$GPLOC,A,1,134304.00,5104.15351,N,00341.16019,E,08,1.09,4.0,0.00,1.465,10,271014*3B

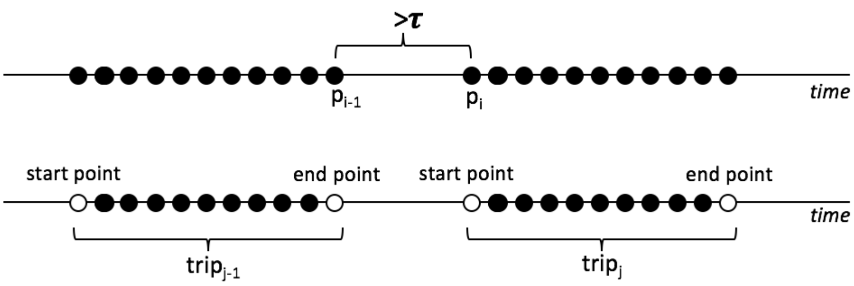
Invalid data frame:

$GPLOC,V,0,000000.00,,,,,,,,,84,000000*13

Figuur 10: Data frames obtained from the device

Given the fact that the GPS data collection was performed using a device installed directly on the e-bikes there is no need to detect the mode of transport. However, the start and end time of each trip are necessary in order to compute trip statistics regarding travelled distance, time spent, and average speed. Therefore, the sequence of GPS points must be segmented into trips with a start and end point.

Let be P a sequence of GPS points $p_i \in P$, $P=\{p_1, p_2, \ldots, p_n\}$ that come from the same device (i.e. GPS logger), we can split the P into trips as depicted on Figure 11 if the time interval between consecutive points exceeds a certain threshold [35], $t(p_i) - t(p_i\text{-}1) > \tau$; $i>1$ where $t(p_i)$ stands for the time at point p_i and τ is a parameter for segmenting, the so called dwell time.

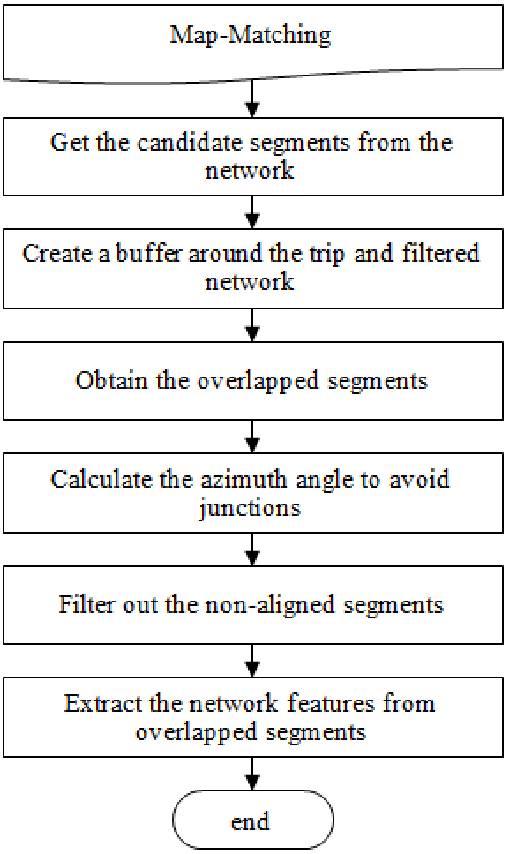


Figuur 11: Trip segmentation using dwell time criterion

Considering the dwell time criterion, i.e. the minimum time difference between two GPS points after which it is assumed that an activity took place [36], we split the GPS points into individual trips using a τ equal to 600 seconds. In previous analysis, we considered a dwell time of 300 seconds [37] but it turned out that several short trips actually were one long trip, thus producing over-segmentation [38].

After the trip segmentation stage we obtain trip segments (lines). This allows us to identify easily the start/end point of each trip; these points are used for computing the shortest path as well as alternative paths.

We rely on the python implementation of the igraph[1] library for computing the shortest path between each origin and destination pair. Up five additional paths were calculated for the same O/D pair based on a set of (generalized) link costs. Further details on this procedure are included in section 2.3.1.2 and in Appendix A.



Figuur 12: Flowchart of map-matching for extracting OSM features

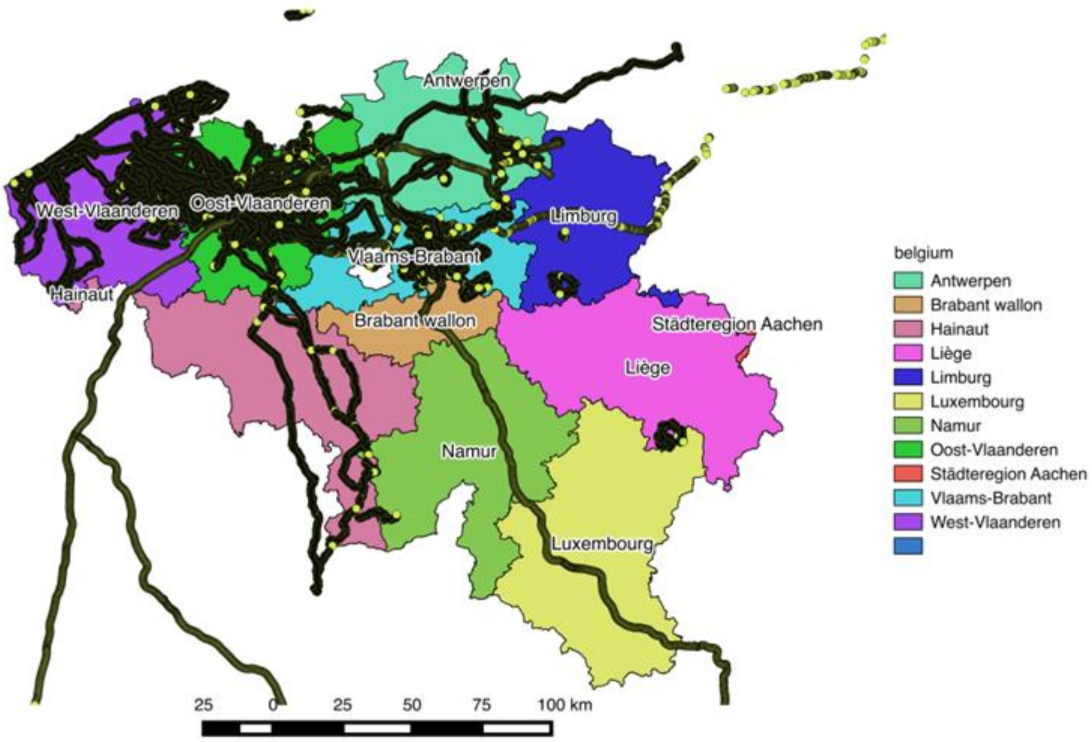
Figure 12 shows the stages involved in the map-matching process. We begin the map-matching algorithm by getting the candidates segments (individual roads) from the road network. This step is necessary to reduce the number of roads to be processed. Candidate segments can be selected either by defining a bounding box containing all the trips or by getting all the roads from the network that overlap any part of a trip. Since our recorded trips have a wide coverage we opted for the second, but we included a buffer of 10 meters around each road, which represents the average GPS accuracy of our data loggers. As a result, a set of overlapped segments was obtained. A problem with this approach is that parallel roads may be matched especially at junctions. To cope with this issue the azimuth angle (i.e. the angle between a reference plane and a point), was used to compare whether or not different segments were in fact aligned. Segments for which the azimuth angle between the trip and road was higher than 10° degrees were filtered out. The resulting aligned segments were used to extract the road type and type of surface.

2.2.2 Practical analysis

The data set employed for this study contains information about 14410 trips performed by e-bikes mainly in Flanders during 2014 and the first part of 2016 for a total of 98 weeks. In this section we first show some descriptive statistics with a focus on speed analysis. We then explore the results of our map-

[1] http://igraph.org

matching algorithm and, finally, we derive insights in route choice modelling for e-bike owners through a Path Size Logit model[39].



Figuur 13: Overview of the trips

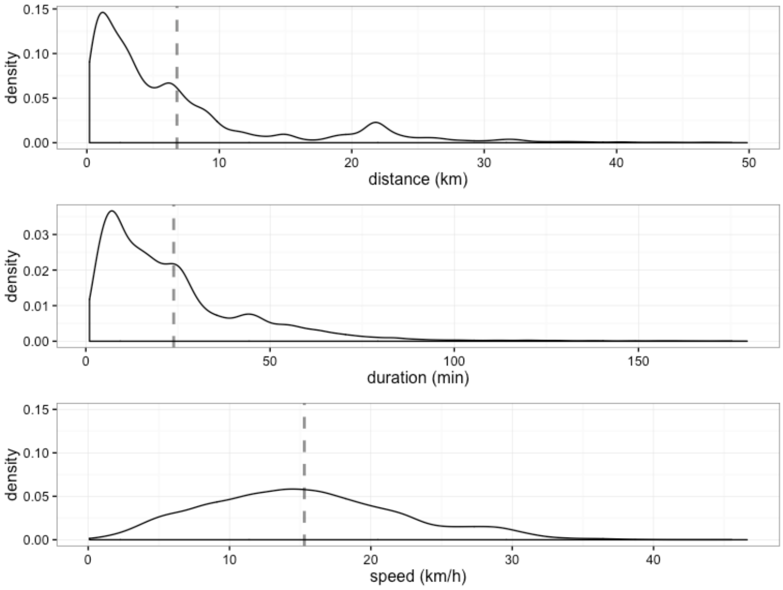
The statistics we provide in this section come from a grand total of 14410 trips. The days in which the e-bikes are mainly used correspond to working days (from Monday to Friday) with a total of 11930 trips (83% of all trips in 72% of time of week), whereas during the weekend this number decreases to 2480 trips (17%).

Trip features such as distance, duration and speed respectively expressed in kilometres (km), minutes (min) and kilometres per hour (km/h) are summarized in Table 6.

Tabel 6: Statistics of the travelled features

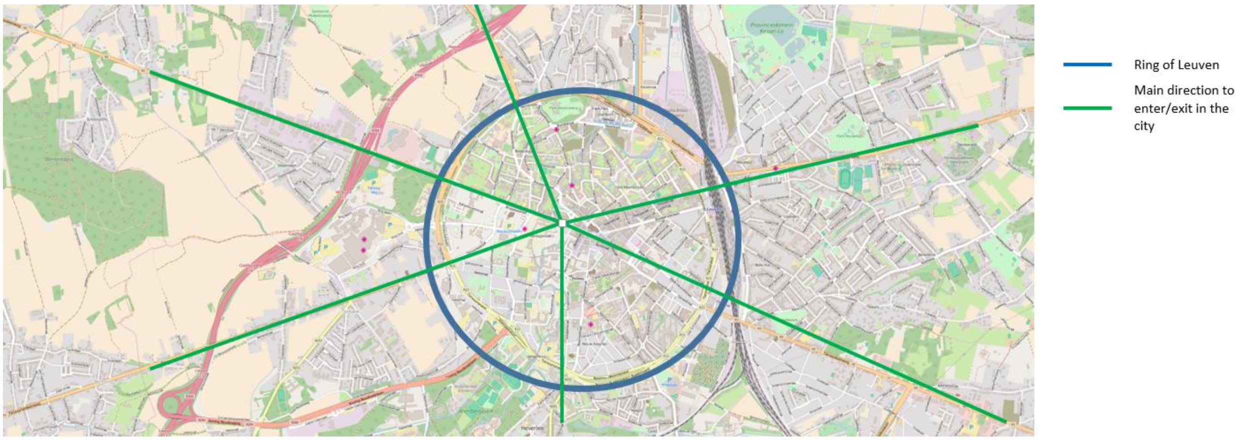
	trips	mean	max	total
Distance (km)	14410	6.8	49.1	98017.5
Duration (min)	14410	23.8	179.3	343651.6
Speed (km/h)	14410	17.1	45.0	

From Table 6 it is straightforward to see that the values of these statistics, apparently low in terms of average speed (17.1 km/h) and travelled distance (6.8 km), are instead in line with previous studies [9] which reported an average speed of 14 km and an average distance of 6 km. Hereafter, we compute the probability density estimation of each feature to further investigate the characteristics of our sample, as shown in Figure 14.

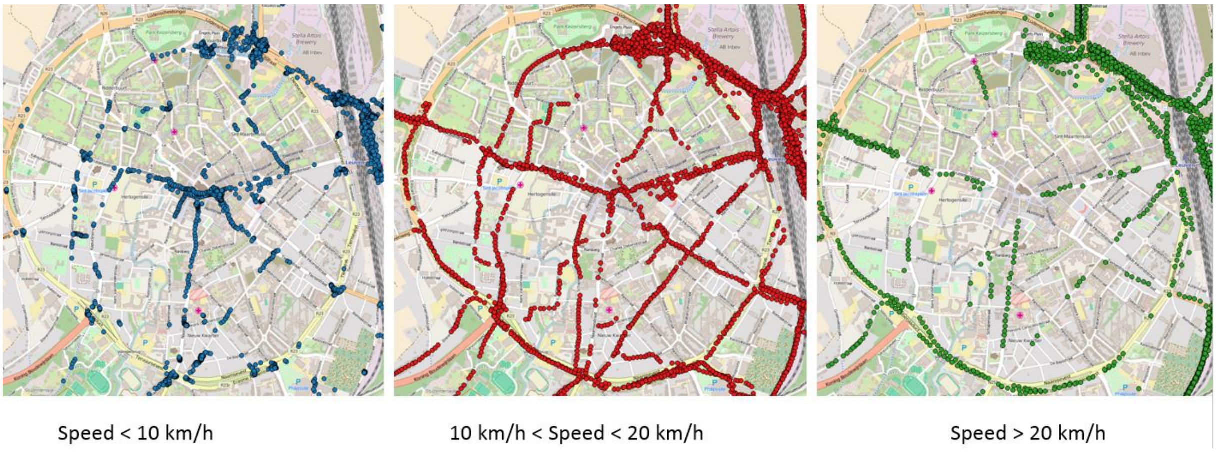


Figuur 14: Densities of the travelled features

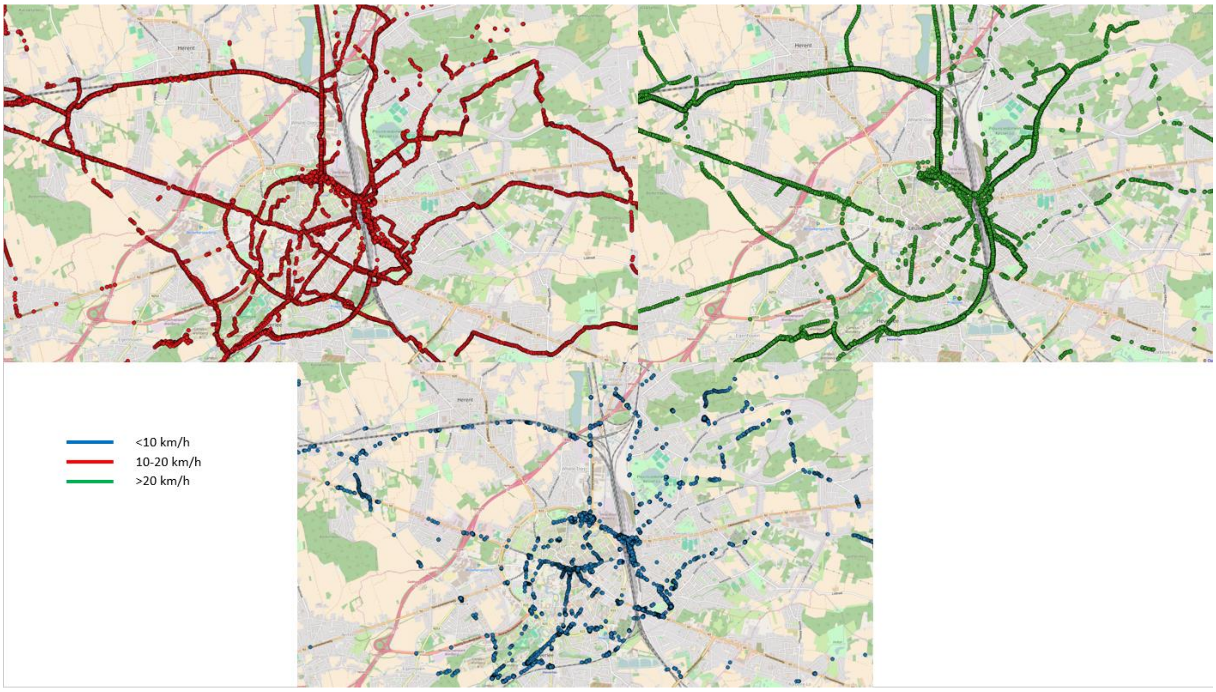
From the probability density estimation we identify at least two “speed-groups”: one distributed around an average of 15 km/h and the other between 25 and 30 km/h. The existence of such groups is partially explained by the area in which these trips are performed, for instance the city centre vs its outskirts. Given the general picture of the Leuven network in Figure 15 where the blue circle defines the ring of Leuven and the green lines the main direction to enter/exit from the city center, Figure 16, zoomed into the ring, shows a higher speed on the ring itself, a range between 10 and 20 km/h on the main direction, while a lower speed more sparse in the city center. Moreover, zooming out (Figure 17), outside the city center higher speeds are revealed.



Figuur 15: Leuven network

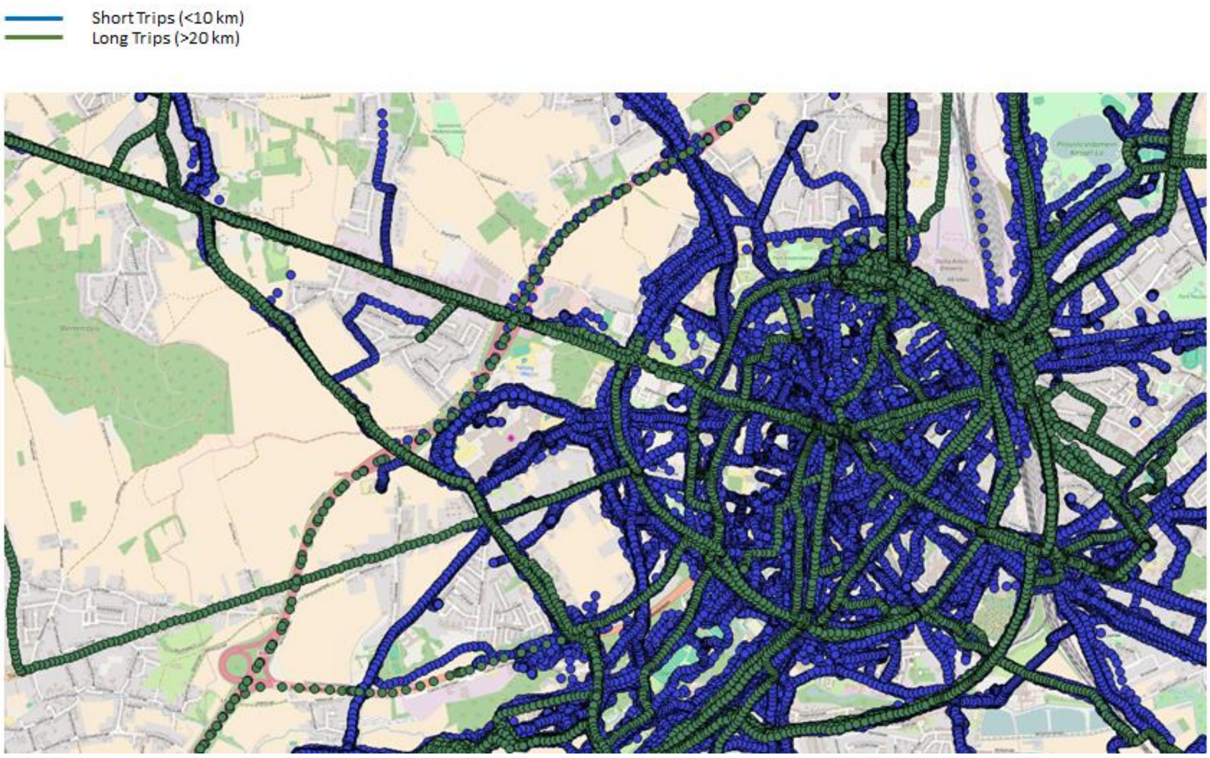


Figuur 16: Speed distribution within the ring



Figuur 17: Speed distribution within and outside the ring

The Ward Hierarchical Clustering approach [40] was applied to group trips based on the average travel distance. This clustering approach minimizes the total within-cluster variance. General statistics per cluster are shown in Table 7. It is interesting to see how the average speed increases with trip length. This could partially be explained as a selection effect: the fitter cyclists may be the ones riding both longer and faster. Another explanation, as showed in Figure 18, is that shorter trips are performed more in residential areas where the trip is interrupted more frequently at intersections and because of more frequent interactions with other road users, whereas longer distances use more infrastructures allowing uninterrupted ride.



Figuur 18: Locations of Short and Long trips

Tabel 7: Speed-groups

group	percent	trips	distance	duration	speed
1	26.1	3762	0.9	7.7	7.0
2	31.6	4551	3.2	15.3	12.8
3	29.7	4280	8.4	30.7	16.4
4	12.6	1817	24.3	62.7	23.2

In the following tables, the trips of each group are separated in terms of the type of day (Table 8) and the on/off peak hour (Table 9). Our participants are more willing to ride longer distance during the working days than during the weekend. This suggests a higher usage of the e-bikes for functional trips more than the recreational ones. However without doing an analysis which would capture the trip purpose this hypothesis cannot be confirmed. This latter analysis require further processing stage since the only way to get the information on trip purpose is inferring it directly from the gps locations.

Considering the type of hours, as expectable they ride faster in off-peak. However, the morning peak is more affected than the evening peak by a lower average travel speed.

Tabel 8: Speed-groups per type of the day

Week day	group	percent	trips	distance	duration	speed
working days	1	20.1	2903	0.9	7.6	7.1

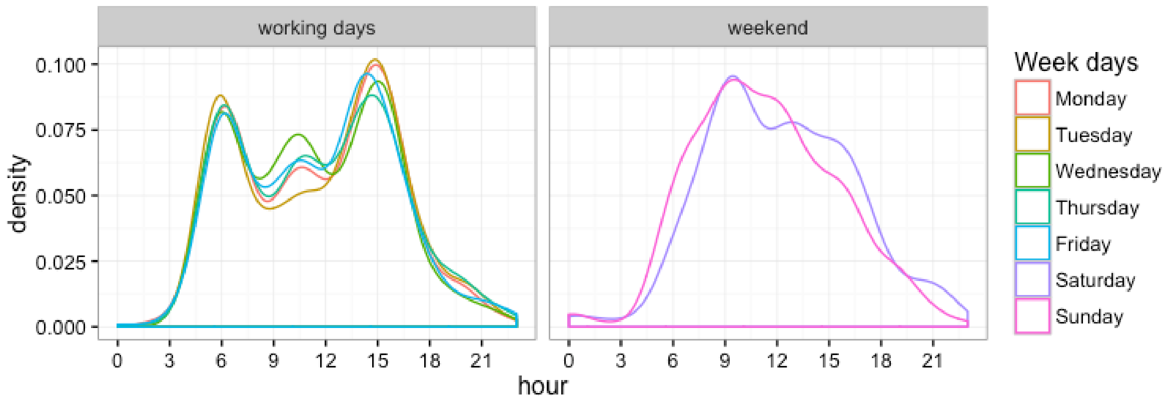
working days	2	24.9	3585	3.2	15.1	12.7
working days	3	26.5	3823	8.4	30.4	16.6
working days	4	11.2	1619	24.1	60.4	23.9
weekend	1	6.0	859	0.9	8.0	6.8
weekend	2	6.7	966	3.2	15.9	12.1
weekend	3	3.2	457	8.4	33.2	15.2
weekend	4	1.4	198	26.1	81.7	19.2

Tabel 9: Speed-groups per type of hour

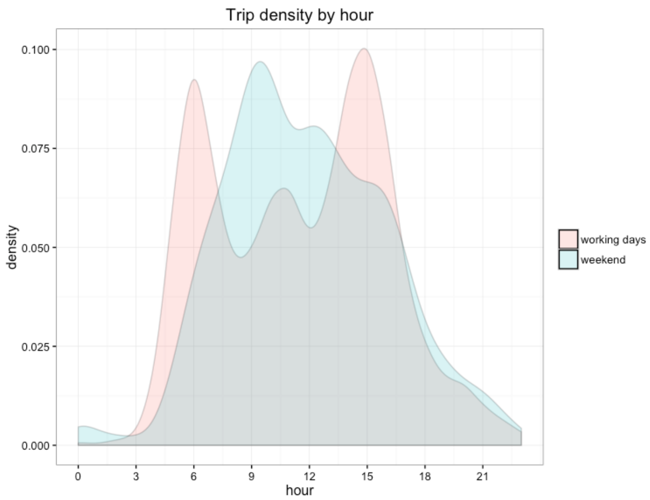
Peak hour	group	percent	trips	distance	duration	speed
off peak	1	16.1	2324	0.9	7.6	7.1
off peak	2	19.5	2815	3.2	15.6	12.3
off peak	3	22.0	3172	8.4	30.7	16.4
off peak	4	10.1	1457	24.4	62.3	23.5
morning peak	1	6.0	866	0.9	8.2	6.6
morning peak	2	6.5	937	3.1	16.3	11.4
morning peak	3	3.9	562	8.3	31.8	15.7
morning peak	4	1.1	152	24.1	69.3	20.9
evening peak	1	4.0	572	0.9	7.4	7.3
evening peak	2	5.5	799	3.2	12.9	14.9
evening peak	3	3.8	546	8.2	29.3	16.8
evening peak	4	1.4	208	23.6	60.7	23.3

In Flanders Region (Belgium), where the most of our trips are performed, the peak hours during the working days range from 07:00-10:00 (morning peak) and 16:00-19:00 (evening peak).

By computing the probability density estimation of trips along the day, as depicted in Figures 19 and 20, we can see that most of the trips are spread wider (in time) in the morning (from 05:00 to 11:00) than in the evening (from 14:00 to 17:00), although the second range is composed of a higher number of trips. Some trips extend during the night especially during the weekend.

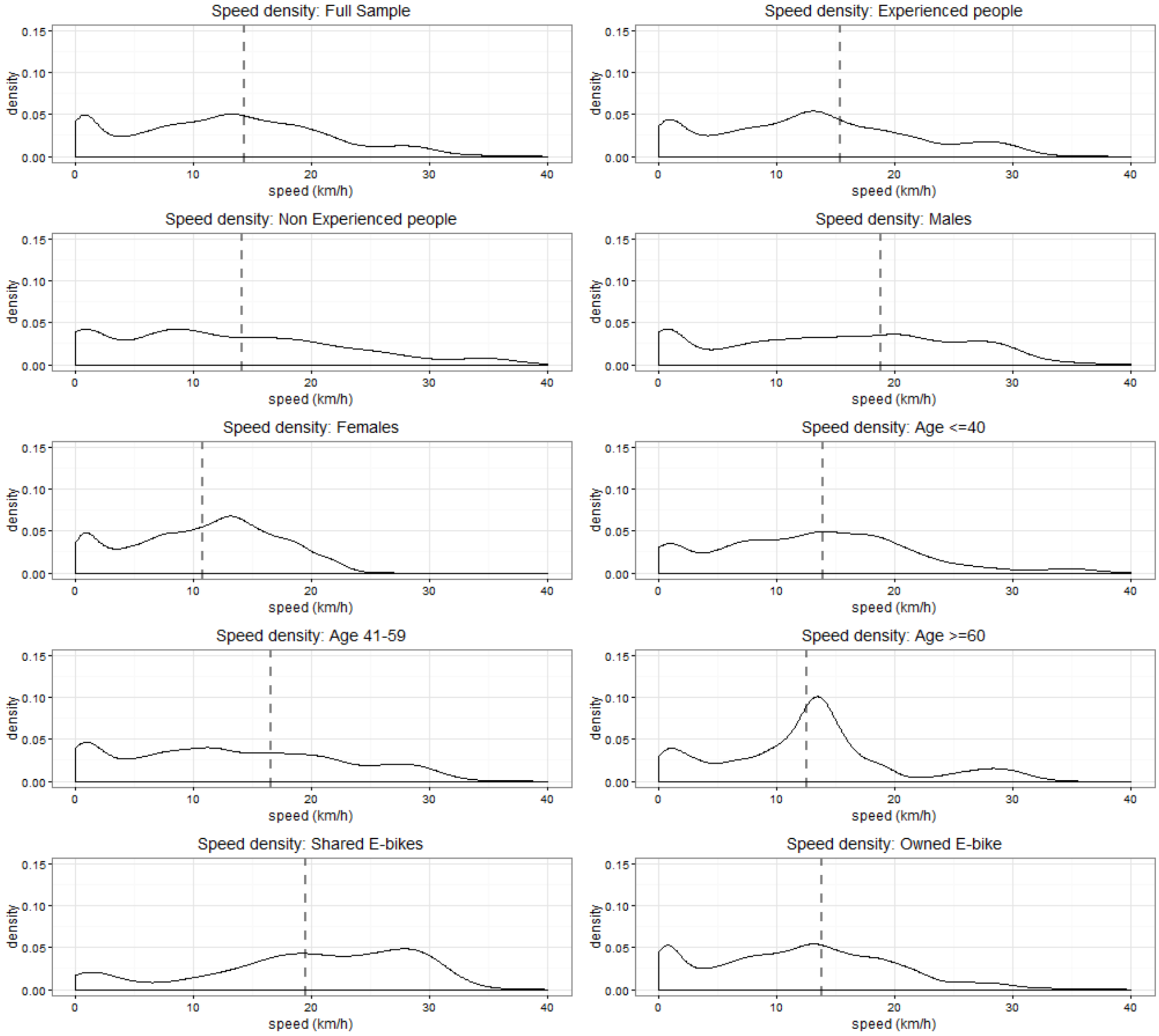


Figuur 19: Probability density of the trips along the days



Figuur 20: Trip density by hour

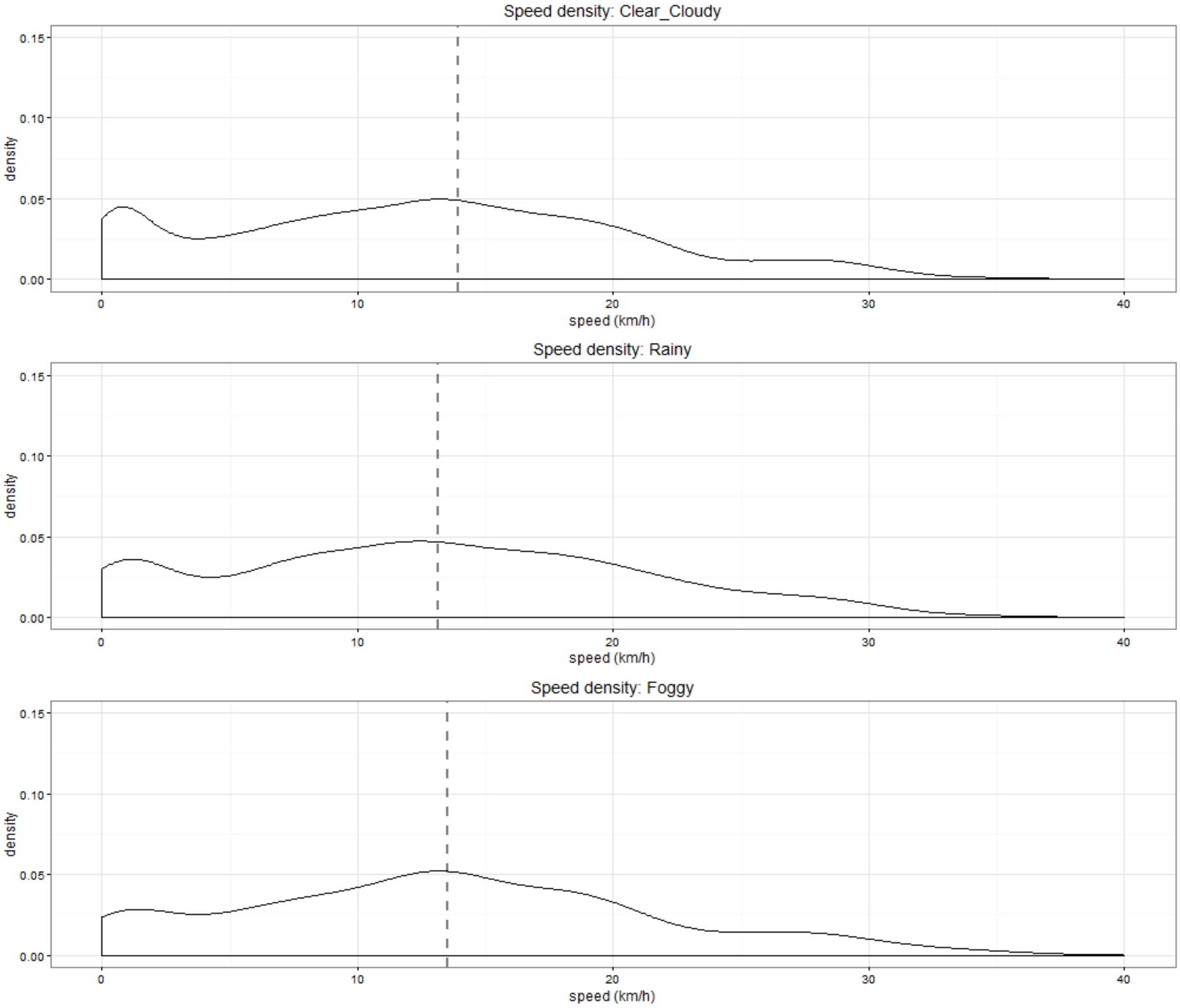
In our opinion the factor “speed” is one of the most crucial elements influencing the usage of e-bikes. For this reason we plotted (Figure 21) the probability density distribution of the speed considering user’s profile. We consider a user “expert” if at the beginning of our campaign he already had at least one year of experience with the e-bike. Moreover, in the final sample also 4 shared e-bikes are included and therefore we also provide a speed comparison between “shared e-bike” and “owned e-bike”.



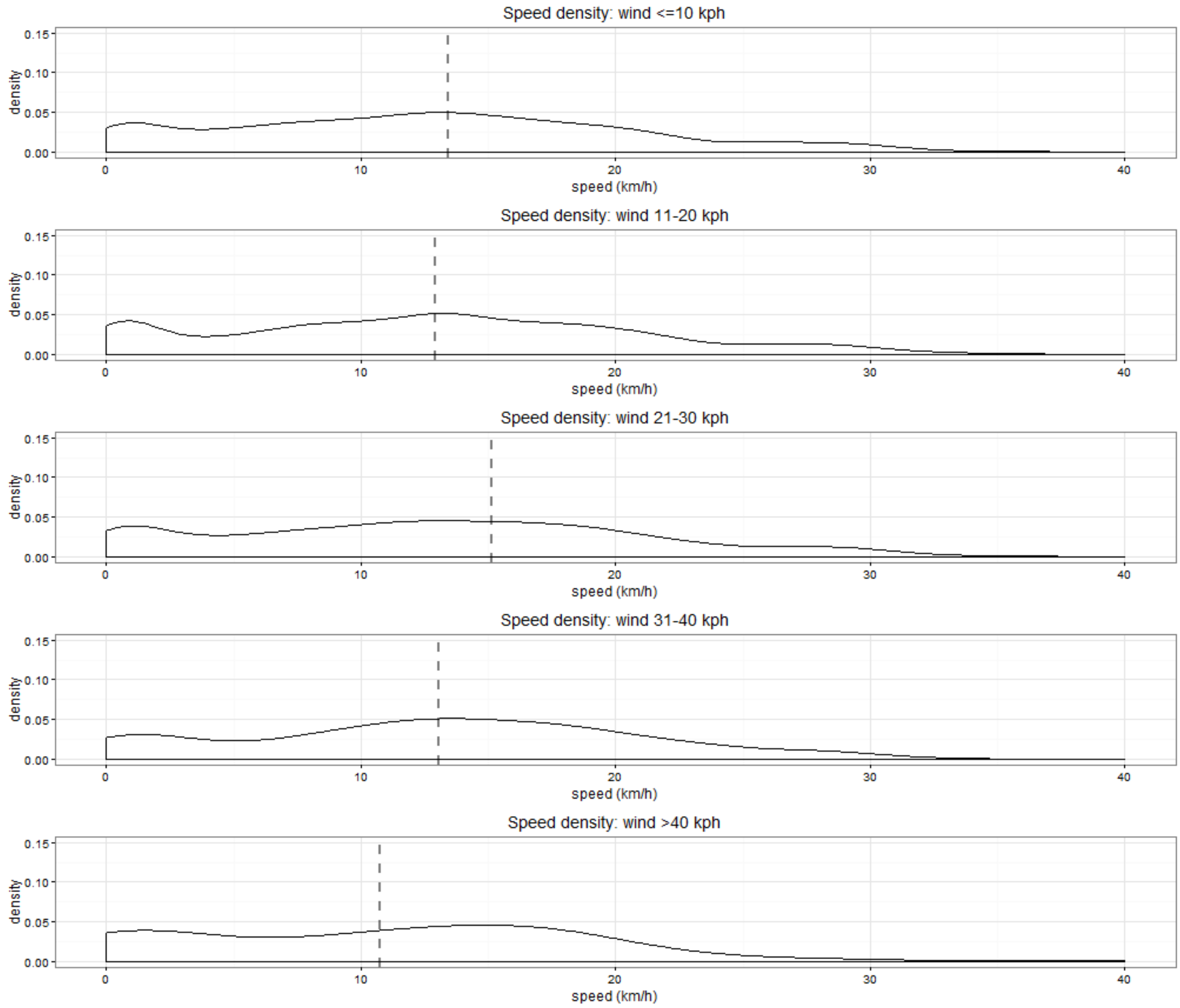
Figuur 21: Probability density of the speed per user's profile

As expectable, the average speed is higher for males and for people with higher experience with an age included between 41-59 years. However, the speed of the shared e-bikes (technically comparable to the average of the sample but for which we don't know the user's profile) is higher than the owned e-bikes.

We also analyzed the distribution of the speed in connection with the weather conditions, these results are summarized in Figure 22 and Figure 23. The dataset employed for this analysis came from the weatherbase.com website and contains observations about the weather conditions for the cities of Leuven, Ghent, Hasselt, Bruges, Kortrijk and Antwerp. For the majority of the trips we were able to get visibility, wind and temperature data in the days and times in which these trips were performed.



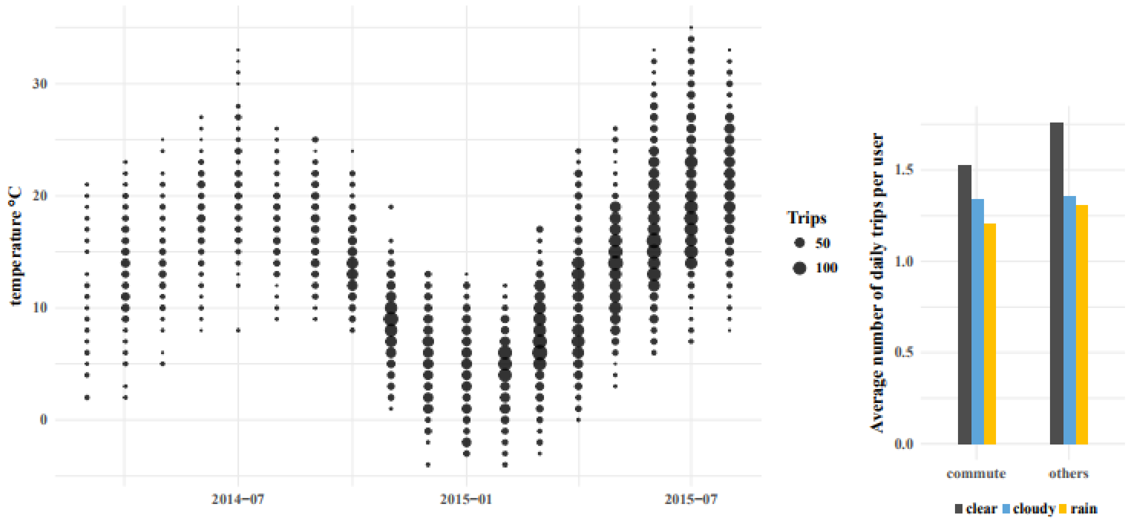
Figuur 22: Probability density of the speed per weather conditions



Figuur 23: Probability density of the speed for different wind conditions

Figure 22 shows how the travel speed is not affected by the scarce visibility (Foggy) or by Rainy conditions. The latter result may seem surprising at first, but could be explained by a selection bias: in adverse weather, the fraction of less fanatic cyclists is more likely to be repelled by the weather than the well-trained cyclist who uses his/her bicycle regardless of the weather. Analysing the influence of the wind in riding an e-bike, we can notice (Figure 23) that when the wind has a velocity higher than 40 kph, the average speed decreases by 2-3 km/h.

We also decided to analyse in detail the correlation between the weather conditions and the use of e-bikes. To the best of our knowledge, this is the first study that provides insights on this issue for e-bikes. The relation between ordinary bikes and weather conditions was addressed in [41]. The authors found that weather conditions influence cycling for both recreational and commuting purposes. Frequencies are higher for fine weather conditions; however, in cooler weather, commuting by bicycle occurs more frequently than recreational cycling. In the case of rainy weather, this difference is more pronounced: there are 10% more of commuting cyclists than recreational cyclists ride during rainfall. Our expectation was that, in the case of e-bikes, this effect could be smaller due to higher cost of e-bikes. In Figure 24, we provide an overview of the weather conditions for the trips included in our sample. In particular, we plot the trip distribution in terms of temperature. The number of trips was larger in year 2015 than in 2014 due to a growth in our sample. For both years, the graph clearly shows that most trips take place in summer months.



(a) (b)

Figuur 24: Weather conditions from sampled trips: (a) number of trips and temperature; (b) commute trips and other types of trips.

We found that 61% of trips are performed in “clear” conditions (by definition of the website) while the 25% and 9% in “mostly cloudy” and “rain” conditions, respectively. Analysing the average number of trips per day (for each person), we obtained two trips per day in clear conditions, 1.5 trip per day in cloudy conditions and 1.4 in rainy conditions. These values show the impact of the weather conditions on the number of trips per day (higher in clear conditions). In order to provide as many details as possible, we also divided our sample into “commuters” and “others”, defining “commuting trips” as all trips performed from Monday to Friday between 6:00 a.m. to 10:00 a.m. and between 4:00 p.m. to 7:00 p.m. The results in Figure 24b show how in rainy and cloudy conditions that the number of trips per day is basically constant, while, in clear conditions, the number of trips with a purpose different from commuting is larger. This means that weather conditions have more influence on recreational trips than on commuting trips.

2.3 Methodology

2.3.1 Investigating route choice decisions

In this section we present a case study that investigates route choice decisions of e-bike owners. First the technicalities needed and the data requirements are illustrated; then, the results of the implemented case study are discussed.

2.3.1.1 Map-Matching

As mentioned earlier, the data used for performing this analysis comes from the AGIV network. In this section, we will first describe the Belgian network and then we will show which type of links result as the most used for our users.

The Belgian network (as represented from AGIV) is composed of 1012672 links and 783564 nodes; of these links, 99% are currently “in use”. The dataset contains information related to the morphology of the links (Table 10), their category (Table 11) and the types of nodes (Table 12). For each item, the total number of links and the number of nodes per class have been computed. The three most represented morphological types of road in the Belgian network are respectively: “weg bestaande uit één rijbaan”, “aardeweg” and “wandel- of fietsweg, niet toegankelijk voor andere voertuigen”. The road category is unfortunately unknown for the 86% of the cases. The “wegknoop” dataset identifies 72% of the nodes as “Echte knoop” (i.e. an intersection where no less than three roads meet).

Tabel 10: Road morphology of the links in the network

LBLMORF	N links
aardeweg	136787
autosnelweg	1403
dienstweg	907
in- of uitrit van een dienst	549
in- of uitrit van een parking	1752
niet gekend	914
op- of afrit, behorende tot een gelijkgrondse verbinding	535
op- of afrit, behorende tot een niet-gelijkgrondse verbinding	2466
parallelweg	321
rotonde	7035
speciale verkeerssituatie	356
tramweg, niet toegankelijk voor andere voertuigen	150
veer	7
ventweg	2881
verkeersplein	103
voetgangerszone	65
wandel- of fietsweg, niet toegankelijk voor andere voertuigen	120481
weg bestaande uit één rijbaan	696897
weg met gescheiden rijbanen die geen autosnelweg is	39063

Tabel 11: Road category of the links in the network

LBLWEGCAT	N links
hoofdweg	1392
lokale weg type 1	35384
lokale weg type 2	58555
lokale weg type 3	4757
niet gekend	870717
niet van toepassing	3280
primaire weg I	1564
primaire weg II	6380
primaire weg II type 4	274

secundaire weg	97
secundaire weg type 1	5142
secundaire weg type 2	13076
secundaire weg type 3	12054

Tabel 12: Type of nodes in the network

LBLTYPE	Nodes
echte knoop	565130
eindknoop	164105
keerlusknoop	3412
minirotonde	29
schijnknoop	50888

As a first step to understand route choice behavior, we performed the analysis shown in Tables 13 and 14 to assess which types of links appear in the collected and segmented trips. The tables show the percentage of coverage that the different link types exhibit, both with respect to the case-study sub-network and to the full Belgian network. In line with the full network, among the most represented road morphologies appear, respectively, “weg bestaande uit een rijbaan”, “aardeweg”, “wandel of fietsweg”. However, a higher share of this sub-network is composed of roads defined as “weg met gescheiden rijbanen die geen autosnelweg is”.

Tabel 13: Road morphology for the used network

LBLMORF	% Coverage Used network	% Coverage Full Network
aardeweg	11.35	13.51
autosnelweg	0.23	0.14
dienstweg	0.09	0.09
in- of uitrit van een dienst	0.21	0.05
in- of uitrit van een parking	0.26	0.17
niet gekend	0.07	0.09
op- of afrit, behorende tot een gelijkgrondse verbinding	0.14	0.05
op- of afrit, behorende tot een niet-gelijkgrondse verbinding	0.48	0.24
parallelweg	0.09	0.03
rotonde	0.74	0.69
speciale verkeerssituatie	0.04	0.03
tramweg, niet toegankelijk voor andere voertuigen	0.20	0.01
veer	0.01	0.0001
ventweg	0.39	0.28
verkeersplein	0.06	0.01
voetgangerszone	0.05	0.01
wandel- of fietsweg, niet toegankelijk voor andere voertuigen	11.11	11.90
weg bestaande uit één rijbaan	69.58	68.82
weg met gescheiden rijbanen die geen autosnelweg is	4.91	3.86

The classification of the network coverage for the subnetwork considered also reflects the properties of the full network.

Tabel 14: Road category of the used network

LBLWEGCAT	% Coverage Used network	% Coverage Full Network
hoofdweg	0.21	0.14
lokale weg type 1	3.77	3.49
lokale weg type 2	7.32	5.78
lokale weg type 3	1.15	0.47
niet gekend	81.84	85.98
niet van toepassing	0.76	0.32
primaire weg I	0.34	0.15
primaire weg II	0.68	0.63
primaire weg II type 4	0.00	0.03
secundaire weg	0.00	0.01
secundaire weg type 1	0.00	0.51
secundaire weg type 2	2.40	1.29
secundaire weg type 3	1.52	1.19

We also analyzed which types of links are used at least once from our sample and which are the most used. In Table 15 we report the links that have been used at least once by our sample, and the number of times each link has been used. Note that each person could contribute multiple times to the computation of this measure: recurrent activities increase the time each link is used. On the other hand, the links rarely used could also be affected by self-correlation: one person living far from the rest of the sample could massively contribute to the number of times a specific link is used. This is of course a limitation of our randomly selected sample. In addition, some links totally not accessible by bikes (e.g. autosnelweg) are observed: these measurements might capture people who carried their bike on a different means of transport (car, train, bus, …). These observations have been removed from the general dataset, and stored for further analysis. In contrast with the subnetwork description included in Table 14, the category "aardeweg" is rarely used: it does not represent a bicycle-friendly type of road. Fortunately (for our analysis), the unclassified roads characterize only a small percentage of the used network.

Tabel 15: Network usage

Used at least Once	%Usage
aardeweg	0.31
autosnelweg	0.04
dienstweg	0.12
in- of uitrit van een parking	0.33
niet gekend	0.02
op- of afrit, behorende tot een gelijkgrondse verbinding	0.24
op- of afrit, behorende tot een niet-gelijkgrondse verbinding	0.47
parallelweg	0.08
rotonde	1.43
speciale verkeerssituatie	0.16

tramweg, niet toegankelijk voor andere voertuigen	0.46
veer	0.01
ventweg	0.40
verkeersplein	0.17
voetgangerszone	0.23
wandel- of fietsweg, niet toegankelijk voor andere voertuigen	12.00
weg bestaande uit één rijbaan	69.75
weg met gescheiden rijbanen die geen autosnelweg is	13.78

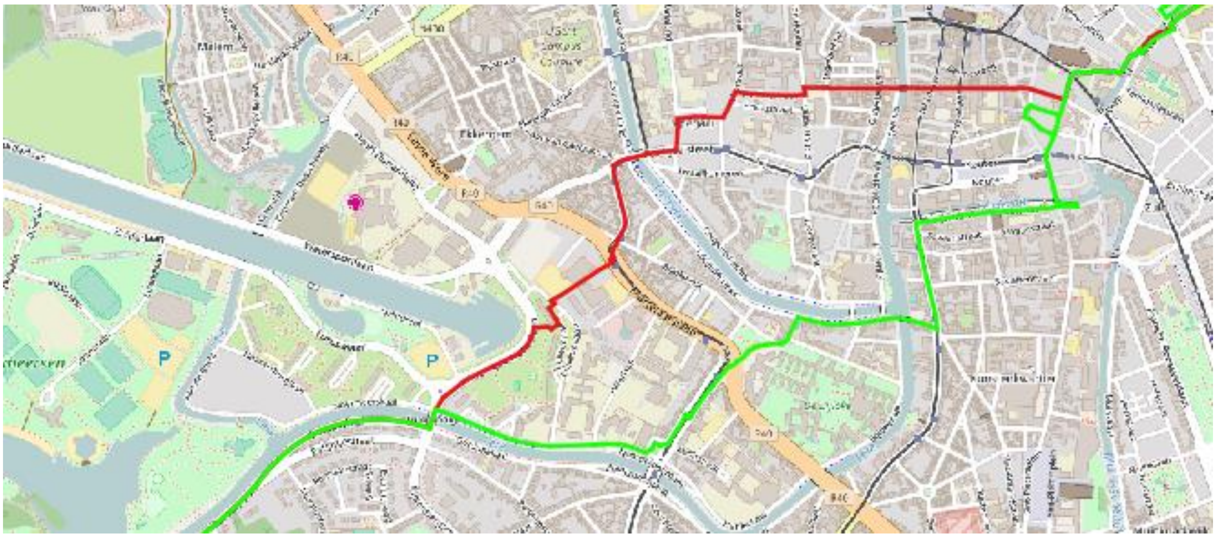
The link properties listed above have been used as input for the route choice analysis detailed in the next section. Extra attributes have also been computed as a combination of the stated information. Further details can be found in section 2.3.1.2.

2.3.1.2 From route set generation to route choice model

The aim of the route set generation step is to identify all the routes that travelers might consider. In literature, different algorithms are proposed [42]. A well-known method, the K-shortest path algorithm [43], generates the first K shortest paths for a given origin-destination pair. Two heuristics are used to achieve this result, called *link penalty* and *link elimination* methods respectively. Both techniques proceed iteratively after identifying the topological shortest path. The *labelling* approach exploits instead the availability of multiple link attributes to formulate different generalized cost functions that produce alternative routes.

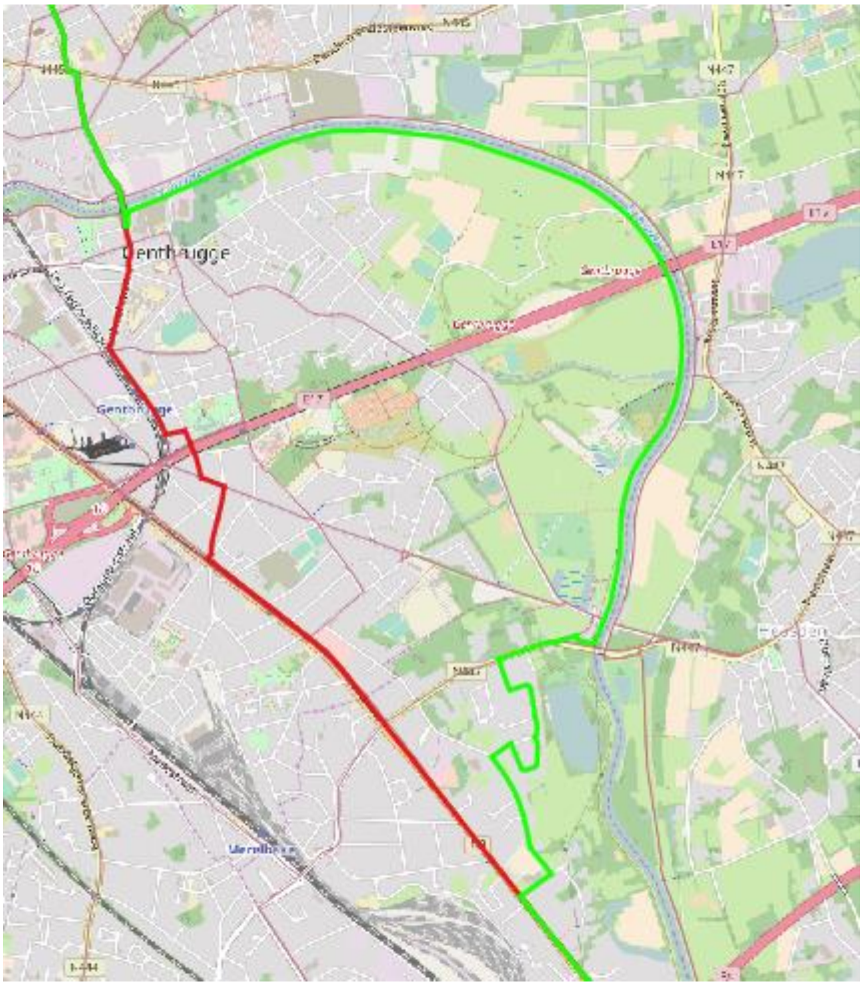
We initially employed a link elimination approach to generate five alternative routes, based on the topological costs of the network. The value of K was chosen following [44], according to whom each individual normally considers not more than 5 alternatives in his route choice decisions. However, as stated in [45] and [44], the different route set generation algorithms do not influence significantly the robustness of the model, while the size of the choice set itself has a rather strong impact: bigger sets consistently lead to more reliable results.

The first attempt of generating routes led however to unsatisfactory results: the originated paths resulted systematically shorter than the chosen path. Figures 25-26-27 show a few instances of said issue. Please note that the beginning and the end of the trips have been removed from the visualization, to respect user privacy. The chosen paths are shown in green, while the first (topologically shortest) path is shown in red. Upon inspection, it’s rather immediate to realize that when a “scenic” (e.g. along a canal) route is present, the latter is apparently preferred, even when incurring longer distances (at least from a topological perspective). In the example shown in Figure 27, both the chosen and the generated paths use mostly links with a bicycle path, meaning that the presence / abundance of dedicated cycling lanes might not be the most representative parameter for this particular observation.



Figuur 25: Trip 1 example

Figuur 26: Trip 2 example



Figuur 27: Trip 3 example

To address this issue and better capture the preferences of e-cyclists, we exploited the AGIV dataset in combination with the trip characteristics with the objective of deriving interesting attributes capable of justifying this uncommon behavior. In particular, we computed four different attributes associated with each trip (in appendix a full description of the tools used for their computation):

- Number of true (echteknoop) intersections crossed during the trip
- Number of Right turns done during the trip
- Number of Left turns done during the trip
- Number of changes in hierarchy of the road (e.g. from a residential street to a primary road) met during the trip

We then computed the number of times these measures are lower for the chosen path compared to the topologically shortest path. The change in hierarchy and the turns are for around 60% of the cases lower for the chosen path, while the number of intersection only for 19%. These results seem to anticipate the important role of the other attributes compared to the standard trip length, which justifies further investigation hereafter.

To further investigate these results, we developed a route choice model that explicitly considers not only the commonly employed attributes such as distance, but also includes these derived measures in its utility function. One of the most critical issues in route choice modeling is correctly representing the significant correlation among alternatives. In literature, two approaches are normally proposed to address it: (i) a deterministic correction of the path utilities in a Multinomial Logit [46] and (ii) an explicit modeling of the correlations through assumptions about the error terms and the use of advanced discrete choice models [39]. The first approach is simpler, and most often applied in practice. We thus decide to follow the same approach, called “Path Size Logit” [47]. This includes a Path Size (PS) attribute in the utility specification, which represents a scale parameter that quantifies path overlap, computed as follows:

$$PS_{in} = \sum_{\alpha \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in C_n} \delta_{aj}}$$

Equation 1

where δ_{aj} is the link-path incidence variable that equals one if link *a* is on path *j* and zero otherwise. We denote by C_n the set of paths considered by individual *n*. L_a and L_i denote the respectively the length of link *a* and path *i*.

The utility function of traveler *n* associated with path *i* in its general form is expressed as follows:

$$U_{in} = V_{in} + \beta_{PS} ln PS_{in} + \varepsilon_{in}$$

Equation 2

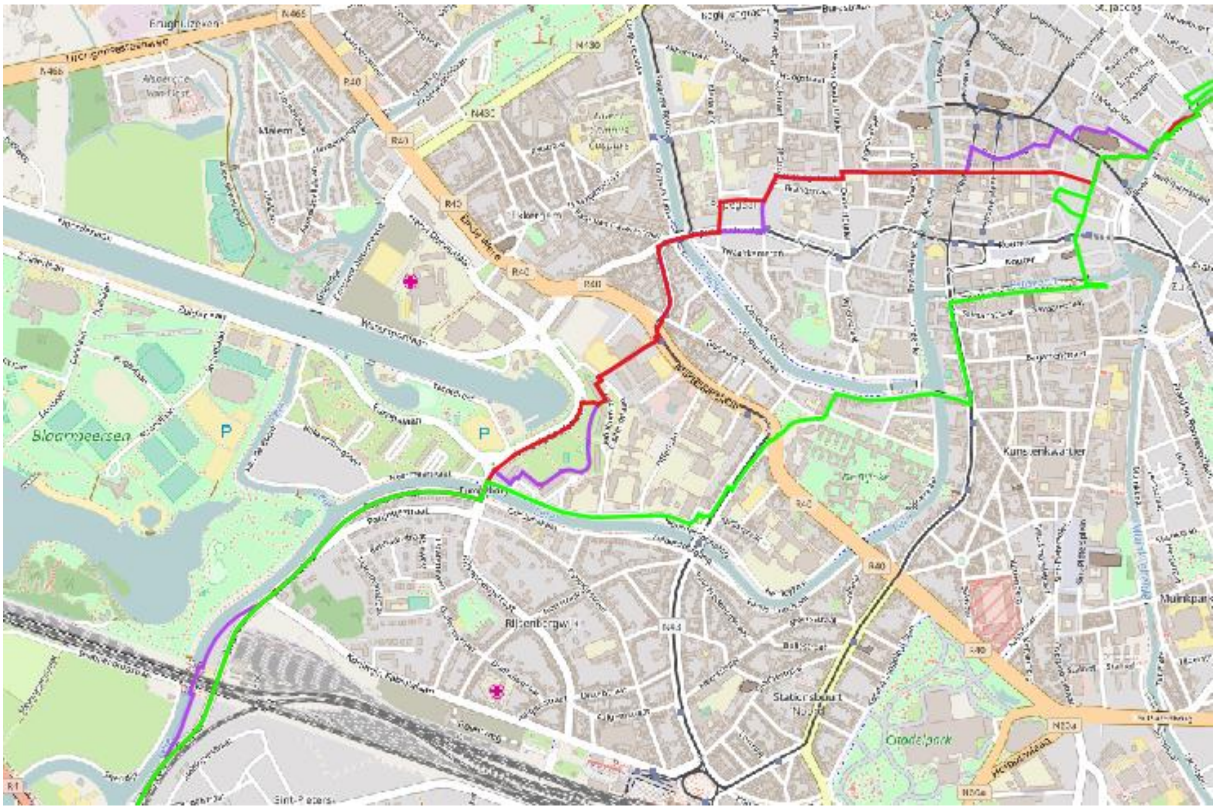
The term V_{in} denotes the systematic utility of path *i* while the random residual ε_{in} is assumed to be identically and independently Gumbel distributed across routes and observations.

The attributes considered in the systematic utility are:

- Type of road: from Table 14, the type of roads with similar characteristics (e.g. speed) have been aggregated in 4 groups: “cycleway road”, “primary road”, “secondary road” and “other”. Specifically, with the term “cycleway” we defined all roads exclusively accessible to bikes and pedestrians (this includes dirt roads etc, so in spite of the name, cycleways are not always the most comfortable option for cyclists). Each of these four classes is explicitly expressed in the utility function as a length-cost. Basically, for each path, the lengths of all links with the same road label have been added.
- Number of intersections: this parameter represents the total number of true (echteknoop) intersections encountered along the path, normalized by the total length of the path (to avoid high penalization of longer routes – which would be double counting as length is already accounted for in the ‘Type of road’ attribute).
- Number of right/left turns: they represent the total number of right and left turns met during the path normalized by the total length of the path.
- Change in hierarchy (HN): this parameter represents the total number of changes in the hierarchy of the network an individual met in performing a journey normalized by the total length of the path. In this specific study the changes in the network hierarchy are associated with a change in the type of road between two consecutive links.

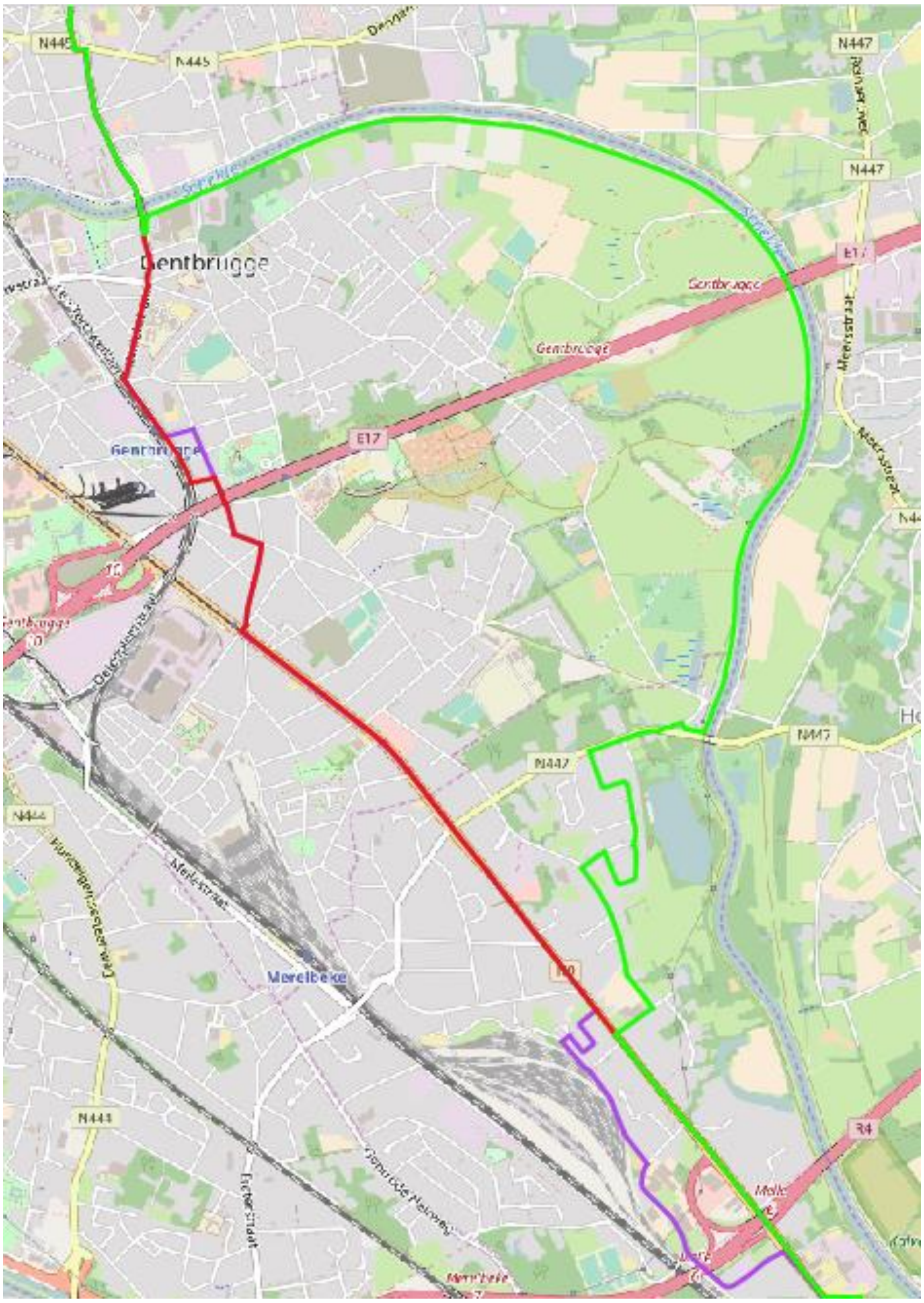
The Multinomial Logit model (MNL) was calibrated through Biogeme. The first attempt, which included the alternative paths generated during the first iteration of the k-shortest paths, was not successful. The length-cost attributes turned out to be positive, due to the high detours exhibited by the chosen routes (see Figures 25,26,27) compared to the topologically shortest paths.

As the generated K-shortest paths only represented topological costs, a generalized link cost attribute was computed for each link, endogenizing the effect that the attributes composing the systematic utility would have on the choice of different paths in the network. This led to the generation of new, hopefully better representative paths, behaviourally closer to the users’ chosen set. In Figures 28,29 and 30 the new generated paths are shown in purple. For trip 1 and 3 the new generated paths are still far from the chosen path. However their total length increase making them close to the length-cost of the chosen route. For trip 2, instead the new generated path totally overlaps the chosen path. Even though the idea to generate new paths seems reasonable, a further refinement of the link costs (or the addition of new attributes) could be necessary.



Figuur 28: Trip 1 second iteration

Figuur 29: Trip 2 second iteration



Figuur 30: Trip 3 second iteration

To verify the validity of these generalized costs, the multinomial logit was then re-estimated. These estimation results are shown in table 16.

Tabel 16: MNL calibration

Coefficient	Value	Std err	t-test	p-value
B_Cycleway	-0.00202	0.000148	-13.69	0
B_HN	-6.46	0.36	-17.94	0
B_Intersection	-1.61	1.32	-1.22	0.22*
B_LeftTurn	-6.53	0.388	-16.85	0
B_Other	-0.00165	0.000304	-5.43	0
B_PS	10.4	0.643	16.1	0

B_Primary	-0.00401	0.000371	-10.8	0
B_RightTurn	-5.42	0.351	-15.45	0
B_Secondary	-0.000642	5.44E-05	-11.79	0
Not significant at 95% level				

All the estimated parameters exhibit the expected signs: the length-cost attributes are negative, as well as the parameters associated with the number of intersections, changes in hierarchy and number of turns. In our assumption, all these attributes negatively influence the attractiveness of a route. As these measures result similar in terms of order of magnitude, the most decisive attributes are the left turns and the changes in hierarchy. Left turns are a well-known deterrent in route choice decisions [48] for other vehicles, so their relative importance in bike route choice isn't in itself surprising [5]. We are instead not aware of other studies that investigate the importance of the changes in the hierarchical network from a bicycle perspective. However, these results highlight the need to further pursue this attribute. Given the full analysis performed on the length of the type of roads, this result shows how length-costs attributes exhibit less influence compared to the other measures in e-bicycle route choice, in contrast with the results obtained for ordinary bike in [5]. This result could anticipate substantial differences in route choice decisions between normal bikes and electric bikes. The preference of secondary roads over cycleways could be for example explained by the type of roads surface: "cycleway" are for the definition employed here, all roads exclusively accessible to bicycles and pedestrians; they then also include unpaved track (e.g. cobblestone, dirt road). People could be then more inclined to share the road with the other vehicles than travelling alone but on an uncomfortable path. The sign of the PS corresponds to a scale parameter, and should consequently be positive. The value of the parameter associated with the number of intersections is the only one which results not significant at 95% level. This may be due to the correlation it could have with the other attributes included in the model: for instance, a change in hierarchy is often present when there a true intersection is encountered. However, the opposite is not true. Finally, the results in Table 16 show how and the model is highly inclined to reproduce the reality ($\rho^2 = 0.8$).

2.3.1.3 Future Research

The case study proposed in the previous section employed a multinomial logit model which incorporated a Path Size Logit term to consider the overlap among the paths included in the choice set. The limitation of this approach is that the results are strongly dependent from the type of the route set generation approach. This also appears from the fact that the type of cost used for generating the paths (topological distance vs generalized cost) was a decisive factor for the model estimation being meaningless or significant. For this reason, future research direction aims to overcome this issue by testing the so-called Recursive Logit model [48]. This model has the advantage to be a random utility model for the choice of a path with no restriction on the choice set. Similar to the Path Size Logit model, the original authors proposed a link-size attribute that corrects utilities of overlapping paths, but unlike the former approach, this attribute is link additive: at each node, the decision maker chooses the utility maximizing the outgoing link with link utilities given by instantaneous cost, the expected maximum utility to the destination and i.i.d. extreme value error term. However, compared to the PS model, the utility specification is application specific and different specification should be tested to investigate the sensitivity of the final estimation results.

From a behavioural point of view much is still unknown when e-bikes are involved. Current studies compare e-bikes and ordinary bikes solely on trip characteristics such as travelled distance, travel time and travel speed. However, route choice decisions are strongly affect by other factors, as we have partially shown in our route choice model estimations. The type of surface for instance could play a different role between e-bikes and ordinary bikes: using unpaved road could reduce the speed of an e-bike to that one of an ordinary bike. In this case, while the ordinary bikes could still prefer ride on difficult path but avoiding other vehicles, the e-bikes could be more available to ride in a vehicle-shared space where they can keep their cruise speed.

Other possible factors to include in the route choice model are personal characteristics, attitudes or purpose of the trip. This would allow to understand whether route choice decisions are influenced by the time and destination of the trip but also by socio-demographic information (e.g. age). Finally, we have treated in this research all subjects as one homogeneous group (since given the small number, it didn't make sense statistically to further subdivide our population). However, it is widely acknowledged that cyclists are a group with heterogeneous preferences, which is likely the case for e-bikers as well. Acknowledging this would probably yield much better model estimation results.

2.3.2 Investigating the influence of daily modal choice on ownership decisions[2]

Motivating people to use sustainable transport modes requires understanding their motivations to use certain modes for their trips. Traditionally, there are two approaches that relate to this question: vehicle ownership models and mode choice models. The former category of models aims at predicting whether people would acquire a certain vehicle (usually focused on cars) and possible the type of vehicle (e.g. sedan, break, SUV,…). The main explanatory variables considered are usually socio-demographic factors like income, household constitution etc. The latter category of mode choice models normally takes vehicle ownership as exogenously given, and examines which mode from those that are available to a person will be selected for a given trip or tour (sequence of successive trips, e.g. home-work-shop-home).

Both approaches have their limitations when trying to understand transitions to sustainable transport modes. Ownership models purely based on socio-demographic explanatory variables are insensitive to changes in the supply; hence unsuited to predict whether an improvement of a sustainable mode, or introduction of a new one such as the (fast) electric bicycle will lead to increased adoption of this mode. Modal choice models may exhibit some sensitivity, but fail to predict structural changes in people's modal choice, like acquiring a new (sustainable) vehicle or abandoning a less sustainable one (as vehicle ownership is exogenous to these models). Also, they consider modal choice as only being determined by the characteristics of the trip or tour under consideration, herewith neglecting aspects like habit or satisficing. As a result, they fail to predict why people whose mobility pattern is best served for 95% by one mode (e.g. car oriented) may refuse to consider any other mode for the remaining 5% (even though for these 5% of trips this may appear a much more rational (cheaper, faster) choice).

The aim of this section is to explore modal choice model structures that combine both aspects of ownership and daily modal choice and their interrelations, considering choices that consider trips as being part of a mobility pattern. Ownership decisions may be understood better when considering whether the cost (monetary, but in principle the framework presented may also consider the mental effort of changing habit and familiarizing oneself with a formerly unknown option) for acquiring a vehicle (or more general: any mobility tool including also subscriptions to services like public transportation or mobility as a service) is justified, considering its impact on the quality of daily travel. Daily modal choice decisions may be better understood when considering that people deliberately constrain their choice set to only those modes that they decided to own after considering their entire mobility pattern.

These considerations apply to any mobility tool or service, whether existing or introduced new into the market. As such it also applies to decisions to acquire and subsequently select for a certain trip/tour a (fast) electric bicycle, and all their related mobility, environmental and safety impacts. Even though it was our initial intention to develop the model methodology and apply it to the survey data acquired specifically during this project, it appeared that the dataset was too limited for such effort. We hence decided to develop the methodology and prove the concept using different data sets (one reported in this document; another similar effort on German data is under review at the time of writing). With the experience reported here, application of the methodology to e-bike related decisions in future research has become feasible, as we now better understand both the technicalities of the modeling, and the data requirements for this purpose.

The case study reported further on thus focuses on proving the concept and showing the explanatory power of daily-mobility-related attributes in explaining vehicle ownership decisions. Compared to the state of the art, in which socio-demographic information is usually employed to explain vehicle ownership decisions, we highlight how the quality of travel offered by a given mode of transportation in serving one's activity pattern influences the ownership of a mode of transport. To implement this approach, we employ a two-level model: first, we estimate the mode choice for performing a set of activities, and second, we use the results of the mode choice calibration to estimate an ownership model. We also show the need to estimate this two-step model, originally modelled as a two-step multinomial logit, with more complex logit structures. With the aim to consider all possible correlations included in each step, a Mixed Logit model and a Cross-nested logit model are employed. Finally, the influence of the socio-

[2] The content of this section corresponds entirely to a journal paper which will be soon published in "Research in Transportation economics"

demographic information on the quality of travel offered by a given transport mode in performing a particular activity is discussed.

2.3.2.1 Literature Framework

Tabel 17: List of terminology

Restricted choice set (RCS)/ Portfolio	A set of mobility resources. The term 'portfolio' is often used in literature, but we preferred this terminology to avoid confusion with financial terms. The term stems from the fact that daily modal choice is restricted to the modes contained in the RCS that is chosen on a longer time scale.
Mobility Resources	Any tool(s) that enable(s) movement (e.g. car, bicycle, public transport ticket, car-sharing subscription, specific modes, etc.)
Mobility-related attributes/Quality of travel	Factors that influence the travelling experience (e.g. out of pocket costs, travel time by mode)
Travel needs	Trips required to perform activities that people consider indispensable (e.g. working, visiting family on Sunday)

This study addresses the interdependency between daily transportation modal choice and ownership decisions: on the one hand, daily modal choice is restricted to those modes contained in one's personal set of vehicles, subscriptions or other resources giving access to mobility; on the other hand, the decision to own or acquire these vehicles or mobility resources depends on one's expectation on how useful they will be in daily use, balanced against the cost (or effort) of acquisition. Where standing literature leans heavily on socio-demographic characteristics to explain vehicle ownership, we turn attention to the explanatory power of mobility-related attributes: can vehicle ownership decisions be explained by their expected impact upon the (aggregate) quality of trips that one intends to make, and to what extent? We thus need to theoretically and empirically investigate the relationship between the choice of a transport mode and the activity-travel chain in order to confirm the following hypothesis: a relationship exists between the different activities that people perform daily and the modes of transport an individual decides to own.

It is common practice to model transport choices using discrete choice theory [49], which allows modeling choice from a set of mutually exclusive and collectively exhaustive alternatives. This choice is based on the principle of utility maximization, wherein a decision maker selects the alternative with the highest utility among those options available at the time of choice. Within this framework, the dependency of daily modal choice on vehicle ownership is taken into account naturally, as the vehicle ownership (or more general: restricted choice set; see terminology defined in Table 1) of the decision maker defines directly the set of options (modes) available for daily trip making. Models that take into account this direction of the mutual dependence between vehicle ownership and daily modal choice are therefore common in literature [50].

Our work hence addresses explicitly the other direction: how do vehicle ownership decisions depend on daily modal choice? It is natural to assume that the quality of travel supplied by the vehicles and mobility resources that one owns, determines the decision to acquire such resources. Yet, in most literature vehicle ownership is explained by socio-demographic characteristics of the decision maker, while for the quality of supply, at best a sort of *general* accessibility quantification is considered [51], [52]. To contribute in filling this gap, we quantify more specifically the supply side's quality for one's *personal* accessibility on the vehicle ownership decision-making process.

In [53], the authors provided an overview of existing ownership models highlighting how the ownership decisions are always given in input in the mode choice decisions. These models are then compared using different criteria (e.g. inclusion of demand and supply side of the car market, inclusion of income, of license holding, etc.), however at the best a measure related with fixed and/or variable car cost are taken in consideration.

Already in [54] the authors included mobility and lifestyle components (long term decisions), for instance auto ownership, and short term decisions, such as travel decisions for different activities, in their modeling approach. One of the example they provided considers the daily activity and travel pattern of one person who drove alone to work, returned at home and stopped to shop on the way to home. After some financial incentives this person decided to buy a transit pass hold which strongly influenced his activity pattern, both in terms of timing and activity. Thus, since years, the quality of travel offered by a given transport mode that, as in this specific example, could influence the decision to acquire a transit pass, has not been taken in consideration.

[55] further elaborated on car ownership models and developed a comprehensive model of mobility tool ownership. They proposed a model to capture the dependencies between the mobility tool choices at the household level by allowing not only the modeling of the presence of a tool, but of the exact number of each type. They demonstrated how people's ownership of cars and of public transport seasonal tickets are linked choices with the possibility of correlated error terms. Their results were also confirmed by [56] that showed a strong interdependence between mobility attributes such as car ownership and transit pass holding. Both these papers tried to include mobility attributes (such as household car ownership, employer-provided transportation for commuting, etc) in their models, nevertheless their concept of mobility attributes did not include a specific measure related with the performances of a specific mode of transport in reaching an activity. On a similar line of thoughts there are explanatory ownership models that employ only socio-demographic information as explanatory variables [57]. The work of [58] tried to expand this typical approach by including extra information about the trips shared by more than one member of the household. In particular, the authors applied a Nested Logit model to investigate household travel behavior with respect to vehicle ownership, mode choice and trip sharing decisions. They developed a two-level model with the upper level representing the ownership choice while the lower level captures the mode choice combination for two-traveler households. More recently, in [59] the authors investigated the correlation in the class-membership model (i.e. it defines prior segment membership probabilities as functions of concomitant variables) among the travel mode choices across multiple trips for any individual.

The idea of this study finds its origins in a recent line of research which states that ownership choices, as well as residential location choices, are decisions that structurally affect accessibility. To the best of our knowledge, Le Vine and co-workers are, to date, the only ones to investigate the two-way relationship between decisions of this nature and personal travel needs [60]. The authors define a *personalized/disaggregate* accessibility measure, rather than a general/aggregate accessibility measure for the zone in which one resides. To this end, they developed the concept of *perceived activity set* (PAS): the subset of activities that each person considers for covering his travel needs. Consequently, the mobility resource(s) a person holds is, in the authors' assumptions, a function of their PAS. Given the PAS, each person will then choose a *Restricted Choice Set* (RCS), defined as a subset of a more general set of mobility resources. For modal choice in daily trip making, the user is assumed to only consider options in his/her restricted set of mobility resources. The individual RCS choice is based on the cost of acquiring each resource and the added value that the mode(s) of transport enabled by these resources would provide in accessing the activities in the person's *perceived activity set*.

Le Vine's approach constrains the daily modal choice to the modes chosen during the vehicle ownership choice phase- the RCS is chosen first, and the resulting daily mode choice is *conditional* on the available RCS. While making this first choice, the traditional approach does not anticipate how the RCS choice will constrain the daily mode choice, while the Le Vine approach, employed in this paper, does this explicitly. It predicts the total disutility of daily mode choices when made under different RCS constraints and then chooses the one with the best balance between this total disutility and fixed costs.

However, as a criticism to the work of Le Vine, we argue that his model disregards two types of correlations. Firstly, correlation exists between the utility perception of multiple trips performed by the same traveler (e.g. a traveler may (dis)like cycling more than average, which will be reflected in similar bias towards cycling in all his trips). Secondly, correlation exists between the utility perception of different RCSs that contain the same mobility resource (e.g. the perceived convenience of car+bike or car+public_transport will be correlated because both contain car). To account for such correlations, here we revise the two step model's structure proposed by Le Vine (who used MNL model for both) using a Mixed Logit approach for the first step and a Cross-Nested Logit model (for whose specification we refer to [61]) for the second step. We also perform a comparison of different model structures in order to statistically justify our modeling choices. Results are obtained with two different software packages, BIOGEME [62] and NLOGIT [63], to confirm the robustness of the estimated parameters.

The contributions of this analysis can be summarized as follows:

1) Independent validation of Le Vine's approach (i.e. PAS) on another dataset;
2) Proposing a modeling structure that accounts better for correlations, and justifying its use empirically;
3) Analyzing the influence of socio-demographic information on the personal accessibility (related to the quality of travel) to PAS activities.

The rest of the chapter is structured as follows. Section 2 details the structure of our proposed two-step approach. In section 3, we present our case study, demonstrating the strength of the approach described in the previous section. Section 4 discusses the influence of the socio-demographic information on the personal accessibility to PAS. Finally, conclusions and future research directions are offered.

2.3.2.2 Methodology

Tabel 18: Glossary of terms

PAS	**Perceived activity set**
RCS	Restricted Choice set
MXL	Mixed multinomial logit
CNL	Cross-nested logit
GMXL	Generalized mixed logit
MNL	Multinomial Logit
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
PT	Public Transport

IIA	Independence from irrelevant alternatives
FCs	Fixed costs
SD	Socio-Demographic information

Tabel 19: List of notations

r,R	**Mobility resources (where 0 means nothing)**
i	Index for individual
d	Index for RCS,portfolio
$\mathbf{U^i_d}$	Utility of RCS,portfolio
$\mathbf{V^{i,travel}_d}$	Systematic component of utility associated with variable costs
$\mathbf{V^{i,non\text{-}travel}_d}$	Systematic component of utility associated with acquiring and/or maintaining the RCS/portfolio
$\boldsymbol{\varepsilon^i_d}$	Error term
m,M	Mode of travel
j,J	journey
λ	Variance of Gumbel distribution
γ	Importance weights
n	Index for generic alternative
k	Index for nest
µ_nest	Parameter associated with the logit tree-structure employed, expressing the degree of correlation among the alternatives that belong to the same nest.
α_nest_n	Parameter associated with the generic alternative *n*, expressing the degree of membership of the alternative *n* to each nest.
$\mathbf{ASC_m}$	Alternative specific constant for each mode of transport m
$\mathbf{B\_purpose_m}$	Calibration parameter associated with the type of purpose and the alternative m

SIGMA_ASC_m/SIGMA_purpose_m	Variance of the distribution of the associated parameter
B_RCS_Gender	Calibration parameter associated with the gender of each individual that own a particular RCS
B_RCS_HighIncome	Calibration parameter associated with people that own a particular RCS and have a high income (>2500 €)
$\mathbf{A^i_d}$	Personal accessibility to PAS activities provided by the set of transport modes enabled by RCS alternative *d* perceived by individual *i*

In order to prove that a measurement related to the aggregate trip cost also has explanatory power in vehicle ownership modeling, we employ the concept of PAS (perceived activity set: *the array of activities which a person views as encompassing their potential travel needs, when making decisions that structurally affect their accessibility*) and, moreover, we propose an improvement of the model's structure defined by Le Vine (Sections 2.3.2.3.1 and 2.3.2.3.2). In [60] it is underlined that each person is required to choose a "Restricted Choice Set" (RCS, portfolio in his terminology) from a choice set of all possible and available combinations of modes, on the basis of both the cost of acquiring each resource and the added value of the modes of transport enabled by it, allowing the users to (better) access activities in their PAS.

In the most general form, where a person can choose anywhere from zero to all resources without restrictions, they face a fully-factorial choice set of 2^R separate restricted choice set options, where R represents the total number of resources.

Based on Le Vine's work, the utility of the RCS *d* perceived by person *i* (U^i_d), is as follows:

$$U_d^i = V_d^{i,non-travel} + V_d^{i,travel} + \varepsilon_d^i$$

Equation 3

which includes a systematic component of utility ($V_d^{i,non-travel}$) associated with acquiring and/or maintaining the RCS, a second systematic component of utility ($V_d^{i,travel}$) which relates to the use of modes of travel enabled by RCS *d* to access activities in their PAS, and an error term ε^i_d.

Note that it is possible to consider further generalizations of eq. (1), where instead of the utility to a single person, one considers RCS at the household level. It is then assumed that a household makes a joint decision on RCS, taking into account a similar utility component as in eq. (1), representing acquisition cost that is shared among the members of the household; however, instead of just one travel utility component, the modeler should then consider separate travel utility components per household member (each with his own PAS) and for trips shared by household members (in the intersection of their PAS's). Although conceptually conceivable, we are unaware of any attempt reported in the literature so far pursuing this research direction. Also in this study, for the sake of simplicity, we take the perspective of individual RCS decisions and leave the household perspective as a topic for future research.

The utility function for individual RCS decision making then takes the following explicit form:

$$U_d^i = \sum_{r_d}^{R} V_r^{i,non-travel} + \sum_{j_i}^{J} (\gamma_{j_i} * \frac{1}{\lambda^{travel}} \ln \sum_{m \in d}^{M} e^{(\lambda^{travel} V_{m_{j_i}}^{i,travel})}) + \varepsilon_d^i$$

Equation 4

where the indices *i, r, d, m, j* represent people, mobility resources included in the RCS *d*, RCSs, modes of travel and journeys, respectively. ε_d^i are the error terms.

The first term shows that the fixed cost associated with owning a personally restricted choice set, composed by different mobility resources, is obtained by adding up the costs of the single resources within the choice set itself.

The second term represents the combined utility related to making all trips j_i in a person *i*'s PAS, employing only the (perceived) best travel mode *m* contained in the restricted choice set *d* . This model assumes that individuals evaluate how well the most suitable mode, enabled by their RCSs, would perform in providing access to a particular activity. These accessibility factors are then summed across all activities (and best modes) composing the person's choice set. The selection of this most suitable mode is reflected through the logsum [60] form included in eq. (2) where the systematic part of the utility of travel ($V^{i,travel}$) contains the attributes reflecting the accessibility to a particular activity (e.g travel time to reach a certain destination). The terms γ represent "importance weights" associated with each journey (e.g. capturing frequency of the trip, when the daily commute is considered more important than an occasional trip; we would like to note that trip importance can also be more subjective than frequency, e.g. a single yearly holiday trip may weigh equally in one's decision than the daily commute).

In this study, the model is estimated in two separate steps: we first estimate parameters of the utility definition in the second term of eq. (2) through a Mixed Logit (MXL) model [64] and then, using the parameters calibrated from this first model, we estimate the RCS's choice through a Cross-Nested Logit (CNL) model [61]. Using the Mixed Logit model for the mode choice step allows to get rid of the independence of irrelevant alternatives (IIA) property and accounts for correlation of unobserved factors over repeated choices by each user (also referred to as 'mixing heterogeneity', [65]). The better fit of the MXL model, compared to the traditional Multinomial Logit (MNL) model, is discussed in section 3.1, using three statistical criteria: the Log-likelihood, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

The Cross-Nested Logit model, instead, allows a given alternative to belong to more than one nest, with different degrees of membership. In our particular case, the structure allows to explicitly consider the correlation among the same mobility resource included in the different RCSs, an effect which we believe far from negligible. Consequently it is not possible to apply the standard MNL.

2.3.2.3 Application

We base our study on the data obtained from the "BMW" project (Behaviour and Mobility within the Week) [25], in which behavioural panel data were collected. The original dataset was composed by 712 users, who filled in a travel survey for two weeks during the winter months of 2008, in the Belgian city of Ghent. This dataset includes journey characteristics (time, cost, purpose, etc.) for 8991 trips as well as the participants' socio-demographic information. The sample is split as follows: 52% males and 48% females. 68% of them live in the city center (zip code 9000 for Ghent) and 32% outside the city center. Originally four different income classes were included: Lower than 1500 euro per month (8%), between 1500 and 2500 (25%), between 2500 and 3500 (29%), higher than 3500 (38%).

The dataset is the result of two questionnaires: the first part contained the traditional questions related to the socio-demographic information; the second part consisted of a travel diary where each person had to fill in the origin, the destination, the purpose, the travel time, the chosen mode and the day of the week of each movement. The questionnaires were both in the local language (Flemish) and are therefore not included in this paper.
For the model estimation, also the travel times of the non-chosen travel modes need to be known. We obtained these by querying through an API interface the online multimodal route planner of Google Maps.

Our first step is to understand which modes of transport were available for this study and with which frequency they were used (Table 4).

Tabel 20: Modal share

Mode Choice per trip	Number of trips	Modal Share
Car	5030	56 %
Bike	1955	22 %
Walk	1405	16 %
Public Transport (Pt)	601	7 %

Initially, we decided to consider four different mobility resources: Bike, Car, Public Transport (PT) (in form of a seasonal ticket subscription) and Walk (for which we assume universal availability). Under this assumption, there are 2^R different restricted choice sets, where R=4 represents the total number of resources. However, since walk is considered to be always available, the final number of combinations is 2^3 (Table 5).

Tabel 21: Ownership share

RCS	Car	Bike	Walk	PT	N users	%
1		x	x		3	0.4
2	x		x		100	14.0
3			x	x	13	1.8
4	x	x	x		144	20.2
5		x	x	x	12	1.7
6	x		x	x	153	21.5
7	x	x	x	x	286	40.2

8			x		1	0.1

Qualitatively, it's easy to notice how the high percentage of car ownership (only 4% doesn't own a car) corresponds to a high usage of the car, while for the public transport this effect is not present. As Table 5 shows, 65.2% of the participants declare to own a Public transport seasonal ticket (Student seasonal ticket is included in this category) but in Table 4 we can see that this subscription is only used for 7% of the total amount of trips. The high share of Public transport seasonal ticket ownership could be due to "business pass": companies provide free subscription to their employees, however they do not use it because they prefer another mode of transport. Unfortunately, we do not have the necessary data to verify this hypothesis.

Given the aforementioned statistics, we decided to focus solely on the RCSs in which car is included: even though the daily modal choice estimation is feasible for all modes (there are enough observations for each of them), with 4 non-car RCSs representing in total only 4% of our sample, any parameter estimation on these rare RCSs will inevitably be statistically insignificant. Our final sample is hence composed by 8574 trips and 682 users (Table 6).

Thus, the final goal of our estimation process is capturing the decision to complement car ownership (and walk) with the ownership of the other resources bike and/or PT seasonal ticket, based on utility of these additional resources for one's PAS as described by model (2).

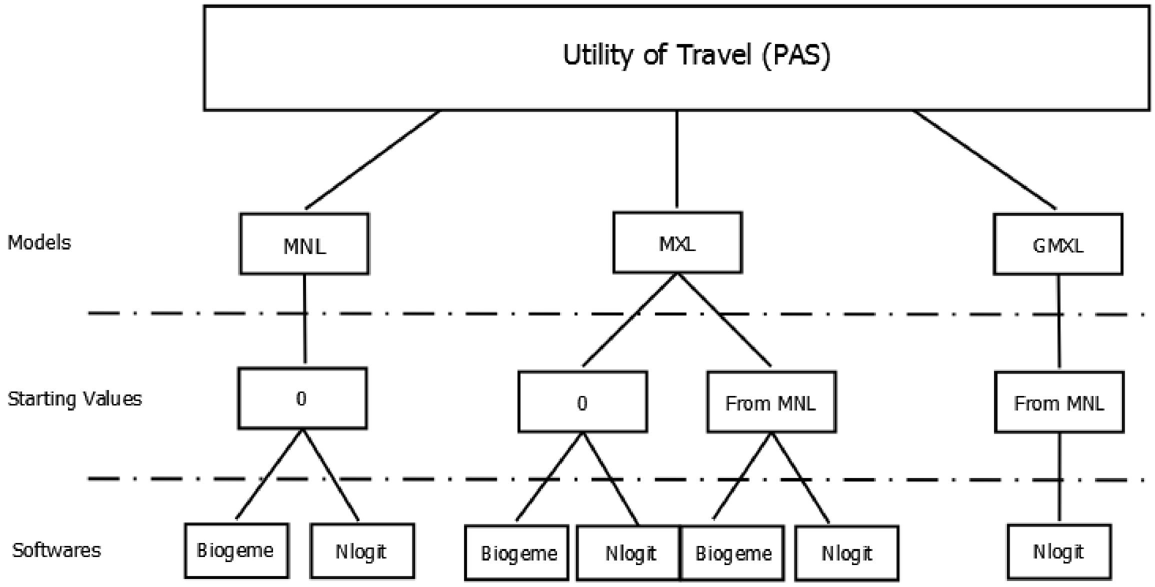
Tabel 22: Modal share for RCS in which car is included

RCS	Mode Choice	Number of trips
RCS_C	Car	985
	Walk	137
RCS_CB	Car	1301
	Bike	604
	Walk	179
RCS_CP	Car	1100
	Pt	226
	Walk	344
RCS_CBP	Car	1644
	Bike	1239
	Pt	293
	Walk	630

2.3.2.3.1 Utility of Travel

In this sub-section the process that led us to decide which type of model structure had to be employed is firstly discussed and summarized. Afterwards, the final estimation results are given and interpreted.

To facilitate the reader in the comprehension of the different steps we performed to estimate the Utility of Travel, Figure 31 has been included.



Figuur 31: Estimation steps for the utility of travel

Three model structures have been compared: we fitted the model using a MNL structure, then a MXL structure and finally a GMXL structure [66]. The latter approach has been included since it introduces a random parameter that accounts for heterogeneity in the scale of the logit error term in the MXL model. The MNL and the MXL were estimated with both BIOGEME and NLOGIT software packages, while the GMXL was fitted solely in NLOGIT. Two possible starting values have been used for estimation of both the MXL and the GMXL models: starting values of zero for all parameters, and starting values set to the output results of the previously obtained MNL fit. The number of draws (required for Monte Carlo sampling of the underlying correlation structures) for both models has been set to 600 in order to reach convergence (Note that GMXL could not reach convergence using starting values equal to zero).

In Table 23 we compare the different model/starting value/software combinations and results through three statistical criteria: the Log-Likelihood, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). AIC is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model, consequently a lower AIC means a model closer to the truth. BIC is an estimate of a function of the posterior probability of model being true, under certain Bayesian setup, therefore a lower BIC means a model to be more likely true. Despite various subtle theoretical differences, their main practical difference is the size of the penalty: BIC penalizes model complexity more heavily. For more details over these criteria we refer the reader to [67] and to [68].

Tabel 23: Model comparison

	MNL		MXL (no initial)		MXL (MNL initial)		GMXL
	BIOGEME	**NLOGIT**	**BIOGEME**	**NLOGIT**	**BIOGEME**	**NLOGIT**	**NLOGIT**
Log-Likelihood	-7561.47	-7561.47	-7084.97	-7087.46	-7084.97	-7085.62	-6955.68
AIC	15148.94	15148.94	14197.94	14202.92	14197.94	14199.24	14203.36
BIC	140996.9	140996.9	14297.4	14302.38	14297.4	14298.7	14317.02

Table 23 shows how the MXL models exhibit a better fit compared to the MNL structure. Moreover, using the MNL results as initial values for the MXL estimation procedure decreases the computation time in both software packages, especially in BIOGEME. Focusing on the different criteria, BIC provides guides (the lower the better) for whether scale heterogeneity is present in order to choose the most proper model between the MXL and the GMXL. In our case the MXL model is the best one as it has the lowest AIC and BIC values.

Summarizing, the reliability of the parameters estimated with a MXL has been proved through the employment of two types of software and different starting values. In the rest of the paper we will therefore show only the results obtained through the MXL.

The computation of the utility of travel focuses on the second term of eq. (2). As anticipated earlier, we introduce a different choice model structure compared with Le Vine's original work: the Mixed Logit (MXL).

Following the definition of PAS, we collected information related to the purpose of the trips. In this case the PAS was approximated by the set of out-of-home activities the users undertook during the survey's period. This pattern will, of course, be an incomplete description of their full PAS and might consequently introduce bias; e.g. less frequent trips like holiday or family visits are less likely to be observed during a two (non-holiday) weeks' survey even though they may be important for one's RCS decision.

Four different categories of purpose have been finally considered: "work", "shopping", "leisure" and "other".

Summarizing, for each trip we know the employed mode, the corresponding travel time, the travel time for the same trip if undertaken by alternative modes and the purpose of the journey.

The expression of the systematic utility function for each person *i* and each alternative *n* takes the following form:

$$V_n^i = \beta_i X_{in}$$

Equation 5

Where X represents a vector of attributes and β a vector composed by the associated weights. The explicit form of eq. 3 is:

$$V^i_{n,j} = ASC^i_{n,j} + \beta^i_{time,j} * Tt_{n,j} + \sum_p \beta^i_{P,j} * P_{n,j}$$

Equation 6

For this study, the following variables were included in the model: the travel time (Tt_n) associated with each mobility resource (*n*: car, bike, walk, public transport seasonal ticket) and each journey (*j*); three different purposes (*P*): *Work, Shop and Other*, which take a value of 1 if they are the purpose of the given trip and 0 otherwise. The *Leisure* activity is considered as reference (if all three purpose variables are 0) and is therefore not explicitly included in the model.

In the work of Le Vine [60] one of the key underlying assumptions is that the influence of travel time is equal for every user. We relax this assumption by estimating the mean and variance of the distribution in the population of all the considered parameters (betas in eq.4). In Table 8, we denote by "B_" the mean of the distribution and by "SIGMA" the variance of the distribution, assumed to be normal. Thus, we will estimate a Mixed Logit (MXL) model which generalizes the Multinomial Logit model (MNL) where the variance of this distribution is assumed to be zero.

Tabel 24: Mixed Logit parameters for the utility of travel

Coefficients	Values	Std error	p value
ASC_{car}	-1.23	0.127	0.00
ASC_{bike}	-0.883	0.124	0.00
ASC_{walk}	0.00 Fixed	…	…
ASC_{Pt}	-3.31	0.329	0.00
B_Other_{Car}	-0.533	0.138	0.00
B_Other_{Bike}	-1.41	0.202	0.00
B_Other_{Walk}	0.00 Fixed	…	…
B_Other_{Pt}	-3.86	0.607	0.00
B_Shop_{Car}	-0.993	0.113	0.00
B_Shop_{Bike}	-1.29	0.155	0.00
B_Shop_{Walk}	0.00 Fixed	…	…
B_Shop_{Pt}	-3.46	0.343	0.00
B_Work_{Car}	-1.94	0.204	0.00

B_Work$_{Bike}$	-1.06	0.227	0.00
B_Work$_{Walk}$	0.00 Fixed	…	…
B_Work$_{Pt}$	-3.76	0.459	0.00
B_time	-0.148	0.00807	0.00
SIGMA_ASC$_{Car}$	1.15	0.166	0.00
SIGMA_ASC$_{Bike}$	0.973	0.166	0.00
SIGMA_ASC$_{Pt}$	2.84	0.386	0.00
SIGMA_Other$_{Car}$	1.04	0.194	0.00
SIGMA_Other$_{Bike}$	1.91	0.221	0.00
SIGMA_Other$_{Pt}$	2.33	0.599	0.00
SIGMA_Shop$_{Car}$	0.912	0.189	0.00
SIGMA_Shop$_{Bike}$	1.50	0.181	0.00
SIGMA_Shop$_{Pt}$	1.93	0.321	0.00
SIGMA_Work$_{Car}$	2.32	0.287	0.00
SIGMA_Work$_{Bike}$	2.70	0.279	0.00
SIGMA_Work$_{Pt}$	3.70	0.401	0.00
SIGMA_Time	0.111	0.00606	0.00

Table 24 summarizes the results of the Mixed Logit model estimation obtained through BIOGEME [69]. The ρ^2 is 0.3 and all the calibrated standard deviations of the random parameters are significantly distinct from zero. The mean for the travel time coefficient (B_time) is negative, as expected, and the standard deviation (SIGMA_Time) is significantly different from zero. Its numerical value indicates that the probability that the parameter has a negative value is very close to 1. Similar considerations can be made for the other random coefficients. The random parameters associated with the different purposes of the trip show the highest heterogeneity for the activities performed by public transport. Moreover, the utility of the public transport in performing the different activities is lower than that one of the other modes of transports, as expected [70]. The utility of bike is higher compared to that of car for working and leisure (represented by the alternative specific constants) activities (ceteris paribus), while for shopping and other activities the opposite is true. This is plausible in a bicycle friendly city like Ghent: if it is equally fast, the bike is preferred for work commute and for leisure trips; for shopping and other activities (like escorting kids, social motives,…) the need to carry luggage or move other people makes the car more preferred (note that these findings correspond to the analyses on the same data by [71]).

2.3.2.3.2 Utility of non-Travel

This second step focuses on the estimation of the utility of the RCS for which the computation of the first term of eq. (2) is necessary.

As we introduced earlier, rather than employing the common MNL approach, we prefer to use the Cross-Nested Logit choice modelling (CNL) framework. This choice stems from the fact that a relation clearly exists between the RCSs sharing common mobility resources, a fact which cannot be modelled with the standard MNL model and that cannot be neglected theoretically.

In order to estimate the model correctly, we compute the fixed costs for each mode of transport taking into consideration the following:

- for the Car: price, insurance, registration tax, road tax, maintenance;
- for the Bike: price of acquiring, maintenance and cost of replacing it in case of theft;
- for Public Transport (Pt) we consider the cost of the seasonal ticket and the cost of replacing it in case of theft.

Unfortunately, this information was not included in the original surveys and no official statistics are available. Therefore, we retrieved this data through different online governmental databases (at regional, provincial and federal levels). Through the available resources, we computed an average cost for each mobility resource, scaling them in monthly units.

Furthermore, in order to compute the utility of travel for each person, we needed to (i) take the output of the calibration results of the first step, (ii) calculate the utility of each trip for each person, (iii) sum the values of the utilities for the different mode combinations, (iv) aggregate them per user, (v) finally collect four different values of personal accessibility, each of which considers the different mode combinations (RCSs).

Finally, given the utilities of travel (in the remainder of this paper written as: $V_d^{i,travel} = A_d^i$) and the fixed costs, it is possible to estimate the final utility for each RCS alternative through equation (1).

As example, let us consider one individual with two modes enabled by his RCS and two trips in his PAS, one with shopping as purpose and the other with work. Let's assume that the first trip takes 10 minutes with the first mode and 12 minutes with the second mode. While the second trip takes 5 minutes with the first mode and 15 minutes with the second one. His utility of travel is computed as follows: the output parameters of the first step multiply the associated attributes (i.e. travel times (10-12; 5-15) and the dummies related to shopping (1,0) and work (0,1)) for all modes enabled by the RCS. A personal accessibility measure (A^i_d) is then computed as a sum of the utilities derived by each trip performed with the modes of transport permitted by the RCS alternatives considered. This aggregate measure will feed the utility function in which the ownership choices are examined.

Each utility function (eq. 1) was computed as a sum of three different terms where the systematic part ($V_d^{i,non-travel} + V_d^{i,travel}$) is given by:

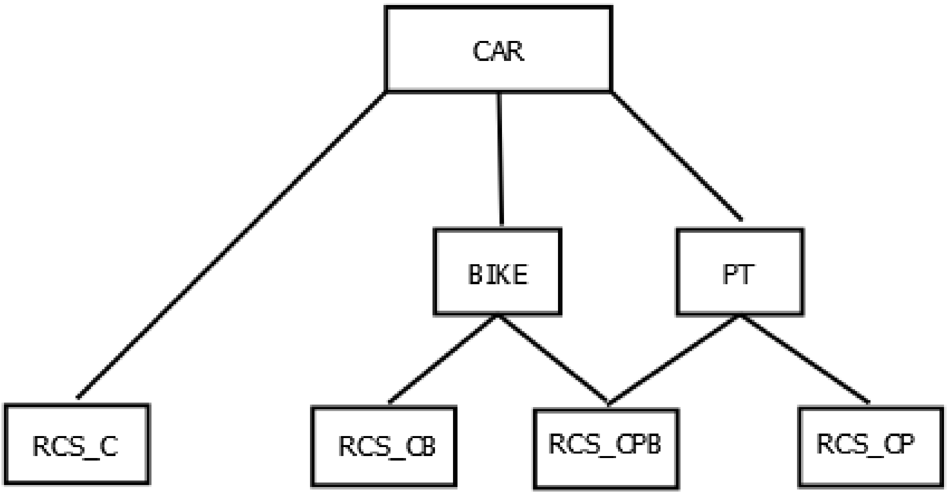
$$V_d^i = ASC_d + \beta_{FC} * FC_d + \beta_{PAS} * A_d^i$$

Equation 7

where *d* takes values C (car), CB (car+bike), CP (car+public transport), CBP (car+bike+public transport) representing the available alternatives (RCSs); *i* is the user; ASC is the "Alternative Specific Constant". *FCs* are the fixed costs associated with acquiring/maintaining the RCS d and *A* is the personal accessibility to PAS activities derived from the previous level. For each user, we computed the utility (with the individually estimated parameters from the previous section) associated with performing their set of activities considering the modes available inside the personal restricted choice set. The “importance weights” (γ) included in eq. (2) were chosen based upon the frequency of an activity over all trips done by each user, scaled to “trips/week”.

In contrast to the previous step, where we had panel data with multiple observations for each user, at this stage we have one single observation for each user which identifies his RCS.

The cross-nested correlation structure used to estimate the model takes the form shown in Figure 32.



Figuur 32: CNL correlation structure

We estimated the model through the software package BIOGEME. To help the reader in the comprehension of Table 9 which contains the parameters explicitly mentioned in eq. (5) but also the correlation parameters typical of the CNL specification, the general formulation of the choice probability for the CNL is reported [72] in eq. (6).

$$p[n] = \frac{\sum_k [\alpha_{nk}^{\frac{\mu_0}{\mu_k}} e^{\frac{V_n}{\mu_k}} (\sum_{i\in I_k} \alpha^{\frac{\mu_0}{\mu_k}}{}_{ki} e^{\frac{V_i}{\mu_k}})^{\frac{\mu_k}{\mu_0}-1}]}{\sum_h (\sum_{i\in k} \alpha^{\frac{\mu_0}{\mu_k}}{}_{ki} e^{\frac{V_i}{\mu_k}})^{\frac{\mu_h}{\mu_0}}}$$

Equation 8

Where the degree of membership of an alternative *n* to nest *k* is represented by α_{nk} and by definition is included in the interval [0,1]; μ_k is a parameter associated with nest *k* and capture the correlation among the error terms of the alternatives belonging to the nest and μ_0 is the parameter associated with the first level of choice (root); I_k is the set of alternatives belonging to nest k.

Tabel 25: CNL estimation for RCS choice

Coefficients	Values	Std error	p-value

ASC_RCS_C	*-6.71*	1.65	0.00
ASC_RCS_CB	-9.90	1.49	0.00
ASC_RCS_CP	2.18	0.612	0.00
ASC_RCS_CBP	0.00 Fixed	…	…
B_FC	-0.407	0.0634	0.00
B_PAS	12.6	1.20	0.00
µ_Bike	9.52	2.02	0.00
µ_Pt	9.88	2.02	0.00
α_Bike_RCS_CBP	0.999	0.240	0.00
α_Pt_RCS_CBP	0.00114	0.240	0.00

Table 9 shows the results of the estimation procedure for the RCS choice. The ρ^2 is 0.8 and all the parameters have the expected sign: the Fixed Cost and the personal accessibility both represent costs at the individual level, but since in the input data file the FCs costs are given as positive and the personal accessibility costs as negative, the output must show inverted signs for both. The µ and α parameters play an important role in the computation of the choice probability (eq.6). The α parameters, which express the degree of membership of an alternative to a particular nest of the CNL, show the preference of owning the bike compared to the public transport: the value of α_Bike_RCS_CBP assigns a higher influence to the bike compared to the public transport in the RCS in which both of them are included. The µ parameters show a clear correlation among the RCSs that share the same mobility resources (significantly larger than 1). The latter proves our initial assumption: the correlation among the same mobility resource included in the different RCSs cannot be neglected.

To conclude this section, we investigate whether specific commonalities characterize the people choosing the same RCS. To achieve this objective, we introduce in the model some socio-demographic information. We added them one after the other to investigate both the single and the combined effect. For the income parameter we considered two classes: low income (<2500 euro per month) and high income (>2500 euro per month). Gender is represented by a dummy variable that takes a value 1 for male and 0 for female. Residential locations are also represented by a dummy variable, which takes value 1 for the city centre and 0 otherwise.

For these tests, the systematic portion of the utility function ($V_d^{i,non-travel} + V_d^{i,travel}$) is computed as follows (only the third iteration contains the complete specification):

$$V_d^i = ASC_d + \beta_{FC} * FC_d + \beta_{PAS} * A_d^i + \beta_{income}^d * INCOME^i + \\ \beta_{gender}^d * GENDER^i + \beta_{res\_location}^d * RES\_LOCATION^i$$

Equation 9

We again estimated the model through BIOGEME.

Tabel 26: CNL parameters with Socio-demographic information

Coefficients	**Model A** **Value (Std error)**	**Model B** **Value (Std error)**	**Model C** **Value (Std error)**
ASC_RCS_C	-6.29 (1.73)	-6.89(11.0)*	-0.871(22.7)*
ASC_RCS_CB	-10.0(1.54)	-10.0(10.5)*	-9.99(0.516)
ASC_RCS_CP	3.00(0.665)	2.18(0.765)	7.63(22.7)*
ASC_RCS_CBP	0.00 Fixed	0.00 Fixed	0.00 Fixed
B_RCS_C_Gender	*Not included*	0.304(0.807)*	-1.69(3.94)*
B_RCSCB_Gender	*Not included*	0.00 Fixed	0.00 Fixed
B_RCSCP_Gender	*Not included*	0.619(0.781)*	-1.89(3.94)*
B_RCSCBP_Gender	*Not included*	0.0394(0.0748)*	0.213(0.291)*
B_RCS_C_HighIncome	0.00 Fixed	0.00 Fixed	0.00 Fixed
B_RCS_CB_HighIncome	0.334(0.777)*	-1.40(0.918)*	-0.307(4.17)*
B_RCS_CP_HighIncome	-0.191(0.368)*	-0.237(0.379)*	-0.333(0.417)
B_RCS_CBP_HighIncome	0.459(0.778)*	-1.34(0.916)*	-0.360(4.17)*
B_RCS_C_CityCentre	*Not included*	*Not included*	-2.83(22.7)*
B_RCS_CB_CityCentre	*Not included*	*Not included*	0.263(0.267)*
B_RCS_CP_CityCentre	*Not included*	*Not included*	-1.37(22.7)*
B_RCS_CBP_CityCentre	*Not included*	*Not included*	0.00 Fixed
B_FC	-0.415(0.0651)	-0.415(0.418)*	-0.413(0.0281)
B_PAS	13.0(1.23)	14.5(1.45)	29.5(5.59)
μ_Bike	8.34(1.82)	8.00(1.57)	7.72(0.614)
μ_Pt	9.98(1.82)	10.00(1.57)	1.00(0.604)*
α_Bike_RCS_CBP	0.997*	1.00(1.80e+308)*	0.984(0.0544)*
α_Pt_RCS_CBP	0.00292	0.000100(1.80e+308)*	0.0165(0.0544)*

Not significant			

Tabel 27: Goodness of fit statistics

Model Statistics	Base	Model A	Model B	Model C
ρ^2	0.8	0.8	0.8	0.8
Final Log-Likelihood	-112.736	-113.385	-111.264	-91.855

Tables 10 shows the significant influence of the personal accessibility attribute (associated with B_PAS parameter). Comparing the parameters in the model without (Table 9) and with socio-demographic information leads to the conclusion that the value of B_PAS is rather stable (except for Model C). The same applies to both µ and α parameters. From Table 10 we can also conclude how people with higher income tend to prefer RCS_CBP, in which all the mobility resources are indeed included. Interestingly, RCS_CP, in which Car and PT are included, is less preferred than RCS_C, in which only Car is included.

Table 11 reports on the goodness of fit of the models, which all have a ρ^2 equal to 0.8. A Chi^2 test was carried out in order to check whether two models are significantly different from each other. This test fails when gender and residential location are included.

Being the models equal in terms of ρ^2, checking the significance of the parameters, the best model specification is the basic one even though the value of the final log-likelihood is slightly higher compared to Model B and Model C. This means that adding socio-demographic information do not improve the fitting of the model. This issue could also be related with the structure of our dataset: for each Restricted Choice Set the percentage of people is balanced among groups that compose the same attribute (Table 12).

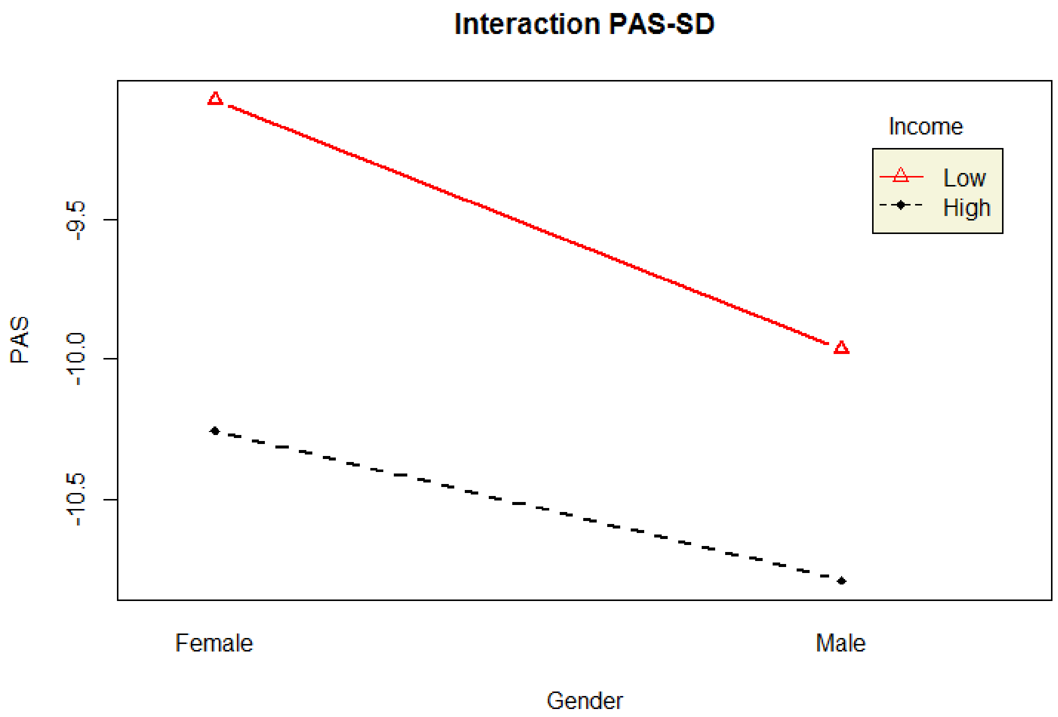
Tabel 28: Percentage of people for each SD group

	% People Outskirts	% People City Centre	% People Male	% People Female	% People Low Inc	% People High Inc
RCS_C	13	15	13	16	14	16
RCS_CB	20	22	19	23	21	21
RCS_CP	26	21	23	21	25	19
RCS_CBP	41	42	44	39	41	43

2.3.2.4 Discussion

In order to further understand the role of a measure related with the quality of travel and its importance, we used t-test statistics to investigate how socio-demographic information influences it. These tests allow to discover whether there are any differences between the means of two groups. In general, if the t-test is significant this indicates that the two groups have means that are significantly different from one another.

We firstly performed this analysis to check how variables such as income and gender could affect personal accessibility value. Considering the income, the p-value is equal to 0.02, showing therefore evidence of differences between Low Income and High Income groups. This difference is instead not significant for gender where the p-value is equal to 0.07. This hints to the fact that the average value of the personal accessibility is influenced by the value of income but not from gender. To check whether the variable gender could play a relevant role in combination with another attribute, an interaction effect has further been investigated. Interaction effects represent the combined effects of factors on the dependent measure. When an interaction effect is present, the impact of one factor depends on the level of the other factor. This moreover indicates that the interpretation of the main effect might be incomplete, or misleading. In Figure 33, the interaction between gender and income on personal accessibility measure is shown: the mean for the female group is higher (less negative) than the mean of the males, however the role of the income in this context is unclear. Interaction results whose lines do not cross are called "ordinal" interactions. If the slope of lines is not parallel in an ordinal interaction, the interaction effect will be significant, given enough statistical power. If the lines are parallel, then there is no interaction effect. In this case, a difference in level between the two lines would indicate a main effect of income; a difference in level for both lines between male and female would indicate a main effect of gender. Therefore, the gender variable plays a role if interacting with the income.



Figuur 33: Interaction between PAS and socio-demographic information

Furthermore, the t-test shows also differences according to the residential location (p-value equal to 0.03). In particular, people who live in the outskirts exhibit a higher average value of personal accessibility which corresponds to a lower disutility. Since, by definition, its value is influenced by the number of trips and the length of the trips, we explored whether these two attributes could be affected by the residential locations. The t-test showed no influence of the residential location on the number of trips. However, a significant difference is present for trip lengths: for people who live in the city centre, the average duration of the trips is higher, no matter the travel mode.

2.3.2.5 Results and conclusion

Given a restricted set of travel modes, this study investigates the relationship between people's daily modal choices and the choice of owning mobility resources that enable the individuals to access their activities. Through a two-level model, we show that ownership of a mode of transport is indeed influenced by the quality of travel offered by the selected mode for serving one's activity pattern (*A* attribute). This work is relevant not only for better prediction of vehicle ownership in classical conditions, it is even more important in case where changes in the supply side are done with the intention of convincing people to structurally change their modal choice. For instance, this process allows us to discover the magnitude of the influence that individual activity patterns have in medium/long term decision processes, such as membership or ownership of different modes of transport. This is particularly useful when modeling the impact of the introduction of a new alternative in the choice set. Focusing on an activity set's influence on the decision to own certain modes of transport offers the potential to model, for example, the transition towards sustainable modes of transport (e.g. e-bike, etc.); from a behavioral standpoint, it also holds the potential to reveal factors that make users less car dependent. The idea of combining different transport options, aiming at increasing sustainable transport usage, is gaining popularity- as an example, the MaaS (Mobility as a Service) concept [73] combines options from different transport providers into a single mobile service, decreasing the stress of planning and single payments.

Considering the three main mobility resources (Car, Bike, PT subscription), the results show that Car is the preferred mode for *commuting* (working trip), while the modal share is equally divided between bike and car for *shopping*. For the purpose *other*, bike is the favourite mode. The results of the estimation of the utility of non-travel show the importance of employing a Cross Nested Logit. This approach properly captures the relationship between RCSs that share the same mobility resources.

The model parameters show that the personal accessibility has a large effect, especially compared to the effect of the socio-demographic variables. From the results, we can conclude that when making a RCS ownership choice, the main parameters taken into account by the users are (i) that the choice of the transport modes enables all of the trips in their activity patterns, and (ii) the fixed costs associated with each mobility resource. From a policy perspective, this kind of approach could help the stakeholders to understand which factors determine whether a person would consider adopting a new mobility option, whether a new service (e.g. shared car) or resource (e.g. electric bicycle). First and foremost, this depends on the person's activity pattern and whether the new option brings sufficient reduction in travel cost over one's perceived activity set. In addition, the widely recognized socio-demographic characteristics impact only indirectly the ownership choice. Mind however that models built solely on socio-demographic explanatory factors, would be totally insensitive to improvements in supply (e.g. better public transport, improved bicycle infrastructure, improved MaaS package,…), and would be totally incapable of estimating market shares after the introduction of new mobility options (like e-bikes or MaaS). Disaggregate accessibility measures like personal accessibility costs would in these conditions at least be capable of providing reasonable estimates even before the novel mobility options are introduced. Market explorations for private mobility service providers, as well as sustainable policy plans could benefit from application of such models.

The model presented could also be translated to studies where vehicle type choices are discussed. For instance, in [74] they explicitly consider the purpose of the trip associated with car type choice. Our model would collapse in something similar if travel time would not have been included in the utility specification (i.e. sum of the utilities over activities consisting of how suitable a vehicle type is for purpose). However, the vehicle type could also be intrinsically included in our model, incorporating different costs for different types of cars without changing the mode alternative labelled "car" as mobility resource (e.g. RCS_CB would still be CAR+BIKE). Compared to Baltas and Saridakis, this model is able to simultaneously consider different mobility resources (car, PT, bike, etc.) but also different types of vehicles within each mode.

2.3.2.6 Future Research

Future research could analyse these socio-demographic attributes in order to predict the personal accessibility value [75] and could include household-level decisions, such as to combine multiple PASs (each member of the household will have his or her own PAS). This sets the stage for answering our key objective: predicting the impact of introducing a new mode of transport (e-bike) on a user's behavior [76]. In particular, the question remains, how could the introduction of the e-bike change the daily mobility habits of the users and consequently influence their ownership decisions? While this goal is not reachable using the dataset employed herein, the idea is to test this modelling approach on a longitudinal dataset that includes observations measuring individuals' mobility habits before and after the acquisition of an e-bike. In this way, we could analyse how people reconstruct their PAS considering this new mobility resource.

2.3.3 Hierarchical network

This section introduces an ongoing research. It then includes only a literature framework and methodological suggestions.

2.3.3.1 Introduction and literature review

Cyclists are a heterogeneous group of road users. They differ in speed, purpose of the trip, preferences of the surroundings, cycling equipment, age, skills, and experience with (complex) traffic situations. The emergence of cargo-bikes, bike trailers, e-bike and speed pedelec only amplifies this variety. How should road authorities respond to this variety, as different classes of cyclists may have different expectations and needs for experiencing an efficient, comfortable and safe ride? Even the same cyclist may behave differently depending on the trip (s)he is making, and with this may also have different preferences. To some extent, the same holds for motorized traffic, where motorcycles, cars, vans, buses, and trucks may have different infrastructural needs. The traditional solution of transportation network planners has been to develop networks with functional differentiation, yielding a hierarchical road network. It can be expected that cyclists have similar needs, which is recognized for instance in cycling policy in Flanders with the development of BFF ("Bovenlokaal functioneel fietsnetwerk" – interlocal functional cycling network) and "fiets-o-strades" (cycling highways).

With its different reach, weight and speed, the (fast) electric bicycle reinforces the need to reflect on the functional differentiation and corresponding hierarchy of cycling network links and nodes, maybe in the future even separation of different forms of cycling. Traditionally, hierarchical functional differentiation and network design have been developed following a top-down approach, where the spatial structure of traffic-generating activities determines the need for connectivity with different qualities and for different quantities of traffic, which is then projected onto existing (or newly planned) physical infrastructure [88]. In principle, this approach is generically applicable, hence independent of the transport mode. However there is far less experience with application of such methodologies for cycling than for motorized traffic and public transportation. Moreover, since cyclists fit very well on smaller roads, which usually form a dense, fine meshed network, there are (too) many possible ways to project a theoretically desired network function onto links of the physical infrastructure. Moreover, since our understanding of cyclists' needs and preferences is relatively limited, it is likely that such top-down exercise would yield networks that do not correspond to the preferences of users. They would simply use other links and intersections than the ones that we planned for them. A mismatch of planned hierarchical networks and behavioural preferences may thus lead to unforeseen and undesired behavior, unnecessary conflicts between different types of road users, dissatisfaction of cyclists, and – foremost – unsafe situations.

In this section we therefore explore an inverse approach: bottom-up. We hypothesize that – even though the currently existing cycling infrastructure may or may not have some (possibly incomplete or imperfect) hierarchical structure, emergent behavior of cyclists may reveal an underlying desired functionally differentiated hierarchical network. With data sets like the continuously tracked gps data of our e-cyclists, a quite complete picture of the variety of trips of different distances over space is obtained. In this section, we try to interpret from this, an implicitly desired hierarchy of the network. The idea is that we interpret various stages of each trip as having a certain function (e.g. giving access to the origin or destination, or covering the main distance between origin and destination) and we observe which links and nodes in the infrastructure are used for this implicit function. In this way, each network link can be related to a mix of functions which it is apparently performing for the e-cyclists, herewith revealing what could be considered an implicit 'empirical network hierarchy'.

This gives rise to a rich variety of possible analyses from the network or user perspective: is the infrastructure of a link suited for the mix of functions it performs in practice? does the mix of functions lead to functional, comfort or safety problems that could be mitigated by re-design of the network? do links with important functions for many cyclists form 'backbone' routes or is any structure missing, or are there missing links? how is the match between how an individual uses the network and how the rest of the cyclist population uses it? what is the correspondence or difference between planned hierarchical cycling infrastructure and the implicit empirical version? how does the empirical hierarchy relate to the planned hierarchy of motorized traffic? are there important conflicts between these? can known problems with cycling comfort or safety be (partially) explained by tension between its empirically detected function and other functions of the infrastructure or surrounding space? etcetera.

The aim of this section is hence to explore a methodology which could help policy makers in identifying a hierarchical network for (e-)cyclists. The term "hierarchical" anticipates the idea to represent and/or organize the physical space (the bicycle network in this specific case) in different levels with different orders of aggregation and/or different functions. For instance, longer trips may set different requirements to the infrastructure and its management than shorter trips. This may reflect eventually in different design parameters (such as lane width, speed, priority at junctions, need for segregation of traffic types, etc.). In this context, the individual perception plays a fundamental role in perceiving lower ranked levels of the bicycle network with higher degree of detail. In order to clarify the origin of this idea and its foundation, this section is organized as follows: first the origins of the term "hierarchy" in people's minds is introduced; consequently, how this term is applied in transportation studies is shown; in addition a subsection specifically focused on bicycling is included; finally, the hypotheses and contributions of the study are discussed.

2.3.3.1.1 The role of the cognitive maps on the definition of the network hierarchy

Early studies in which the term hierarchy, in the context of land use, has appeared are positioned in the seventies when [77] started to introduce the idea of a hierarchical structure in the framework of travel behavior and land use studies. In their book, the authors link for the first time this concept to the observer's location: people tend to perceive the area closer to them in greater detail than areas further away. Each area belongs to a specific "class" which has its own properties. These characteristics derived from people's perception of different locations may be more or less accurate in terms of the real world characteristics of the places constituting the classes. The authors further claimed the presence of certain similarities among individuals within identifiable groups sharing certain sensory and/or verbal experience of the environment. Particularly relevant with this respect seem to be groups based on life cycle, social class and ethic status. Some years later [78], using the same line of thoughts of [77], stressed the idea that people acquire spatial knowledge of the environment and other information with the aim to build a multilevel structure of the space. This structure includes hierarchies, reference points, distance points and semantic information about landmarks in the space. Through the analysis of land use data, they also showed how mental representations of the actual spaces are composed of both spatial and non-spatial information. The latter being hierarchical in nature even for spaces exhibiting non objectively predefined hierarchies.

Continuing in the framework of mental representation, in the same decade the concept of cognitive maps was established [79]. Thanks to their experiments, the authors proved the presence of a hierarchical structure in cognitive maps. They observed systematic errors in people's behavior when directional or spatial judgment was required. They concluded that a superordinate structure systematically distorts how spatial relationships are remembered and therefore it is evidently stored hierarchically. On this same line but twenty-five years later, Voicu [80] described a computational model of spatial navigation based on experimental studies conducted with human participants. Specifically, the model employed a hierarchical cognitive map of large environments to show that (i) the network builds a hierarchical representation of space based on the associations between landmarks and the environment; (ii) the information stored in the network is used by the people to plan their paths in the environment; (iii) experimental data brings evidence for the hierarchical organization of the space.

In the nineties, the group of A. Car [81],[82] started to apply the concept of hierarchy to road network data with the aim of investigating the relationship between hierarchy of data and wayfinding tasks. In [81], the authors investigated this link analyzing the human reasoning process when searching the fastest (not the shortest) path in the multilevel road network. The validation process determines properties of the street network in different hierarchical levels in order to find the fastest path. This conceptualization enables the determination of the relations between speed in different levels of the hierarchy and the size of the meshes per hierarchy level. In their successive work [82], the authors explicitly defined the level hierarchy based on the type of road and following these assumptions: (i) humans may divide a large road network hierarchically; (ii) the amount of detail increases from highest to lowest level; (iii) hierarchical levels are formed according to expected travel speed; (iv) wayfinding occurs in small sub-networks.

Finally [83], proposed a totally different way to look at the hierarchy of a network even though in the same domain of A. Car. The authors proposed three levels of abstraction to clarify how people do employ mental maps when planning and navigating, and explained them with a simple example "one starts planning a trip by asserting that there is a way to drive by car from Boston to Baltimore, by estimating its approximate duration and by determining the Interstate highways to be used (Planning Level). Next, one determines entry and exit points for each leg of the trip (Instructional level) and finally the driver makes decisions about which lanes to use (Driver Level)".

In this context, we aim at understanding how the individual perception of e-bikers reflects on the usage of the physical network. Does the hierarchy revealed by the observed paths coincide with the network hierarchy (if present) designed by policy makers?

2.3.3.1.2 Hierarchy in transportation network

One of the pioneers who introduced the concept of hierarchy in traffic-related issue was Hans Monderman. In the 80's he was recognized for radically challenging the criteria used to evaluate engineering solutions for street design. He introduced the idea of "shared space", an urban design approach that seeks to minimize demarcations between vehicles and pedestrians, often by removing features such as kerbs, road surface markings, etc. In this context he also introduced the concept of "het trappetje van Monderman" [84] (only available in original language) in which he represented a travelers' movement from an origin to a destination as a "staircase": travelers search for low road hierarchy in proximity of the origin/destination but for a higher hierarchy in the middle of their trips.

Years later, P. Bovy [85] applied this concept to analyze and compare individual routes in an urban road network with the objective to assess the predictive quality of the shortest time route choice model for car drivers.

More recently, several studies [86] [87] [88] investigated the role of a hierarchical approach in transportation network design. Specifically, in [86] the author examines the consequences of multimodal travelling for designing multimodal transport networks. The analysis focuses on the way transport networks are organized in hierarchical network structures and determines the main mechanism leading to them. Even though the main focus of the author's thesis was on multimodal transportation, he disserted considerably also on the meaning of hierarchy in case of unimodal networks. Hierarchy in private transport was defined as a network bearing different levels, each having their own transport function in terms of serving specific types of settlements or travel distances, while also providing access to higher network levels. Moreover, each level is suited for specific trip types, especially with respect to trip length: higher levels are suited for long-distance travel and have low access densities, low network densities and high network speeds; lower levels are meant for short-distance travel and thus have high network densities and low speeds. In addition, the number of levels is usually limited due to the costs of building/maintaining/operating related to the investors but also because of the costs of transferring between network levels from a traveler's point of view.

In [88] the authors explained the network design process and the variables and perspectives to be taken into account. The network design problem is beyond the scope of this paper, however some interesting concepts included in the book can be exploited to different applications, e.g. when the identification of a hierarchical structure of a network is involved, as it is the case of this study. For instance, travelers might prefer specific routes, even though all routes are equal in time and length from an objective perspective. Such preference might be because of habit or perception. As a consequence, some routes appear to be more attractive than others from the point of view of traffic demand. This naturally influences the supply side, because the most intensively used routes will be developed so to offer better facilities, thus becoming more attractive. Therefore, the supply side of the transport system will strengthen the hierarchy implicitly developed at the demand side.

Another fundamental concept included in [88] is the idea of "integrating functions": in urbanized areas there is a strong tendency to integrate functionally different network levels within a single physical network. On one hand this reduces the investments, but on the other hand the quality of the transport network might be affected: medium and short trips experience a higher quality because of the higher accessibility and higher speed of the higher level network. This leads to an overuse of the higher level which, in case of cars, is reflected in increased congestion levels.

In this framework the lack of hierarchical network analysis which involve cyclists is evident. Starting from the assumption that also cyclists search for higher hierarchy in the middle of their trips (e.g. search straight connections), our scope is to understand if the existing bicycle network does reflect this need. In addition, the impact of a functionally mixed network on user's behavior is so far totally underestimated: is a link with a mixed function (i.e. highly used both at the beginning/end of the trip and in the main leg) preferred over a link with a homogenous function but characterized by a non-friendly bicycle structure? Or on the other hand a link with a homogenous function is preferred because decrease the number of interaction with pedestrians (for instance)?

2.3.3.1.3 How cyclists perceive the network hierarchy

In the last decades authorities have developed different policies to increase the use of sustainable modes of transport, which resulted in an effective growth of the cycling levels in many cities. The most common policies promoted both physical expansion of the bicycle network (e.g. bicycle lanes, bicycle facilities) and specific strategies such as traffic calming zones. However, with this increase of bicycle usage, policy makers have started to think in terms of a comprehensive "bicycle network", rather than focusing solely on the design of single lanes or facilities. This type of approach examines not only the bicycle infrastructure as a whole but determines metrics useful to characterize the features (e.g. nodes and links) of the network itself.

The majority of the studies [89] focus on examining the individual features of the network instead of investigating the link between its characteristics and the amount of cycling. One exception is [90], where the authors propose a standard methodology for measuring bicycle network quality at the macroscopic level and testing its performance in association with bicycle commuting. They consider four different measures (density, connectivity, directness, fragmentation) and employ linear regression models to predict influence on bicycle commuting. Their findings revealed how connectivity and directness are the most influential factors.

The most common urban bicycle planning studies focus on: (i) the relation between built environment and mode choice [91],[92]; (ii) exploring the match or mismatch between the objective and the perceived bicycling environment [93],[94]; (iii) computing different scores useful for urban planners [95],[96]. In [91], the authors investigate the effect of the built environment on healthy transportation mode choices (bicycle vs car) for trips made by current and potential cyclists. They characterized the built environment at origins, destinations and long routes, hypothesizing that within each of the three spatial zones different built environment features would influence decisions to travel by bicycle instead of by car. They reported an increase of the bicycle share due to reduced hilliness, reduced use of arterial roads, presence of bicycle specific infrastructure, vicinity to commercial, educational and industrial areas, higher population density.

The work performed in [92] contributes to this line of research and partially to (ii), analyzing the correspondence between perceptions of people's residence and the objectively measured spatial characteristics of that residence. Through an internet-based survey submitted from a population sample from the region of Flanders (Belgium), the authors found that people tend to overrate the urbanized character of their residence in terms of constructed density. Among urbanities, (mis)matched spatial perceptions do not influence mode choice, which remains mainly influenced by urban characteristics and not by personal perceptions as such. However, the travel consequences of (mis)matched spatial perceptions depend on the residential neighborhood type.

More recently, [93] explored the match or mismatch between the objective and perceived bicycling environment and its influence on people's bicycle behavior. They firstly employed a factor analysis for both the perceived and the objective bikeability, and then clustered their sample assigning each participant who shared similar characteristics in perceptions or who lived in similar bicycling environments to a cluster. Using a scale of "low", "moderate", "relatively good", "good", "high", their results showed how 44% of the participants perceived their environment at the same level as the objective metric of bikeable environment, while about 7% perceived their relatively good cycling environment as good; 11% perceived the moderate-bikeability environment as high, while 10% as low.

In [94] the authors introduced also the impact of individual attitudes on bicycle commuting. The authors found that residential preference for a good environment for bicycling is associated with bicycle commuting. Moreover, changes to physical environment alone are likely to have little impact since the most important factors are the attitudes of the commuters themselves: their comfort level with bicycling and how much they subjectively enjoy bicycling.

The last line of research (iii) focuses on determining a score (or overall metric) to explain cycling behavior. For instance, in [95], the authors computed a bike score made by four components: bike lane score, hill score, destination and connectivity, bicycle mode share. They then investigated how this score was associated with between and within-city variability in cycling behavior. Employing a linear regression to model associations between bike score and journey to work, they found a significant association across all cities. However, their within-city analysis identified important nuances in city specific models. Therefore, the authors concluded that their bike score shows utility for national or multicity studies but closer inspection may be needed prior to its application for city-specific analysis and planning at a microscopic level.

Merkuria et al. [96], investigated which kind of measures of low-stress connectivity can be computed to guide bicycle network planning. They propose a set of criteria by which road segments can be classified into four levels of traffic stress. They then demonstrated that two measures of connectivity can be applied for a certain level of traffic stress: percent of trips connected (defined as the fraction of trips in the regional trip table that can be made without exceeding a specified level of stress) and percent of nodes connected (which counts the fraction of nodes in the street network that are connected to each other).

To conclude this overview the PRESTO [97] guidelines deserve a particular mention. This fact sheets represents the first effort to bundle state of the art European knowledge and experience on urban cycling policy in an easily accessible format. They indeed serve as European reference guides.

Firstly they define five quality requirements for cycling infrastructure starting from user needs:

Safety: it can be provided reducing traffic intensities and lowering speeds below 30 km/h, separating cyclists in space and in time from fast and heavy motorized traffic; when conflicts points between motorized traffic and cyclists cannot be avoided, these should be presented as clearly as possible;

Directness: cyclists can use a route as direct as possible to reach their destination. Detours must be kept small and overall travel time for cyclists needs to be minimized;

Cohesion: it is about the extent to which cyclists can go from any origin to any destination without interruption;

Attractiveness: it means that the bicycle infrastructure is well integrated into agreeable surroundings. This is a matter of perceptions, which can strongly encourage or discourage cyclists.

Comfort: it is about creating an enjoyable, smooth and relaxed cycling experience. For smooth driving irregular efforts should be avoided: having to stop and start repeatedly is stressful.

Often these requirements can be contradictory among each other so it is important to find the right balance among them (Figure 34): safety must always be the top priority; functional routes and recreational routes have different set of priorities.

Utility cycle network	Recreational cycle network
Safety	Safety
Directness	Attractiveness
Cohesion	Cohesion
Comfort	Comfort
Attractiveness	Directness

Figuur 34: Cycling requirements

They also provide some design requirements in order to take into account the physical space needed for cycling:

Stability: bicycles are unstable vehicles. To maintain balance, a speed of at least 12 km/h is required

Zigzagging: when riding, cyclists constantly have to maintain their balance. At normal cycling speeds the resulting zigzagging requires about 0.20m of space. For speed lower than 12 km/h more free space is required (around 0.80 m)

Fear distance from obstacles: cyclists want to keep their distance from kerbs, edges and walls (0.25-0.63 m)

Section of free space: an absolute minimum pavement width of 0.9 m is required. When possible this width should be at least 1.5 to allow adults to ride next to children

With the advent of the electric bikes is important to understand if they seek the same requirements than ordinary bike. We then aim for example at understanding if e-cyclists, due to their high speed, are more willing to share the road with the other modes of transport without necessary looking for a safer road (with bicycle path) or if they prefer longer routes but less used over the more direct paths.

2.3.3.2 Objective of the present study

The goal of this section is to explore a methodology to identify a hierarchical network for (e-)cyclists by means of their revealed behavior. To the authors' knowledge, no one has investigated this problem before based solely on users' perspectives. As a consequence, the aim of this approach is to check whether the infrastructural bicycle hierarchy currently available corresponds to the implicit needs of the people. This will allow policy makers to have a clear overview of the area in which investments are needed/superfluous.

Moreover, for the first time, the characterization of a bicycle network is seen from the point of view of electric bike owners and not from ordinary bike riders. This could lead to identify a new market to be satisfied, and whose needs might not necessarily coincide with those of ordinary bike users.

Our hypothesis related to the possible structure of an (e-) bicycle hierarchical network, are the followings:

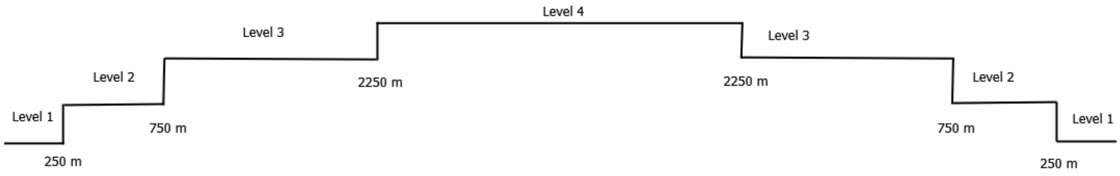
- (e-)bikers use certain parts of the network for different functions (providing access to origin/destinations, accessing higher order connections, longer distance connection) but the existing infrastructural hierarchy may not match these implicit needs
- Parts of the network have a dominant function (from the perspective of the e-bikers) and other have more mixed functions. Design may not always correspond to these functions
- For the e-bikers, factors such as directness and cohesion could have a higher impact compared to ordinary bikes

At the moment of the submission we are only able to propose a methodology without providing its results.

2.3.3.3 Methodology

The methodological approach we are going to discuss requires in input two types of data: GPS data and network data (links and nodes). After the map matching step (illustrated in section 2.3.1.1), which links are used by each trip is a known variable. This allows to assign a desired amount of hierarchical levels to each trip based on certain thresholds.

For instance in Figure 35 the representation of a trip characterized by three levels is presented. The hierarchical levels are presented in "staircase" shapes, where the hierarchy decreases in proximity of the origin and of the destination. The length of the levels is built in a way in which the sum of two segments of one level is equal to the length of one segment at the successive upper level. The number of levels and their length can of course have alternative specifications.

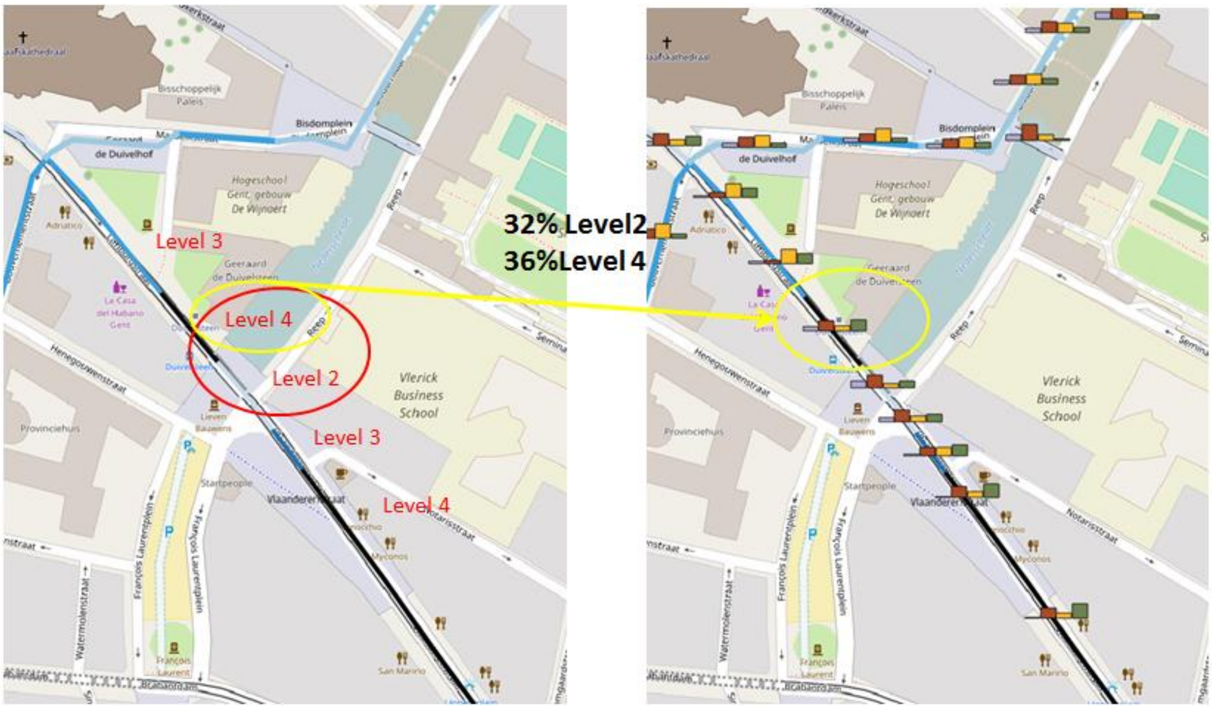


Figuur 35: Hierarchical levels in staircase

After having described the whole trip sample in hierarchical levels it is possible to have an overview of their network location with the aim to understand how the network is perceived by e-bikers. Mapping the hierarchy allows to also functionally describing the network itself: which part of the network is more homogenous (i.e. the type of levels exhibit little mix/overlap) and where instead the links are used both for having access to origin/destination (level 1-2) and to ride the main leg of the trip (level 3-4). The latter case is obviously potentially the more problematic: the interactions among the different interpreters of the road increase. This may influence both travel speed and safety aspects.

Due to the small size of our sample (we hardly have a reasonable amount of individuals that travel in the same area, e.g. contributions for determining a link function come from at most 37 different persons (other participants are Leuven-based), which is obviously not representative for the Ghent e-cycling population) we are not able to provide a proper case study. We will therefore only show some analysis enabled by our methodological approach. At the time of writing, we are co-operating with research groups governing more elaborate data sets in order to more extensively test the methodology. These efforts will be published in due time.

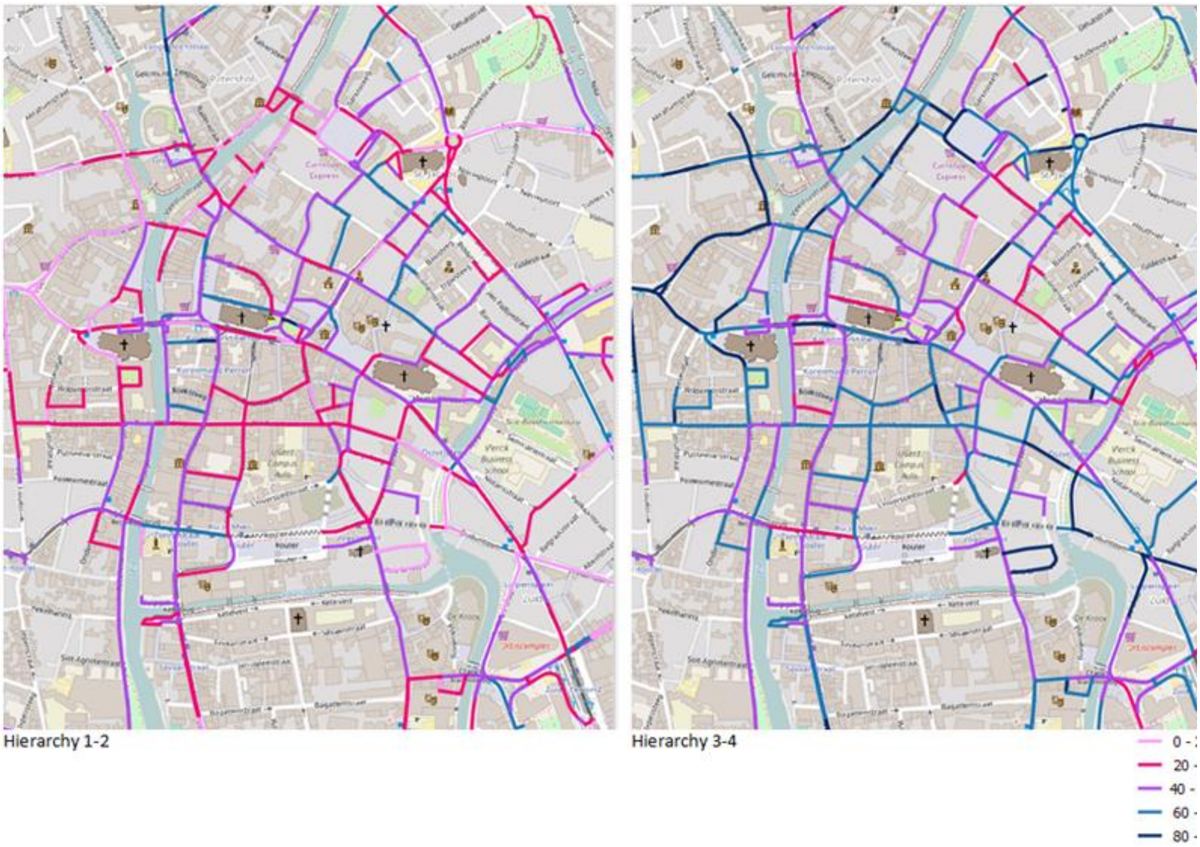
Yet, our partial data set of Ghent allows for some preliminary analyses that give the reader a first idea of what can be further elaborated. We discuss some examples below. From a policy's perspective, characterizing the network in hierarchical levels could be important from two different points of view. Let us consider the example in Figure 36, on the left side the transition between two consecutive links with very different functional mix is shown (red circle, Level 2 and Level 4). However, analyzing the bar plots on the right side of the picture, where the functional mix for each link is shown, it is possible to notice how the brown bar (Level 2) and the green bar (Level 4) have approximately the same height: 32% for Level 2, 36% for Level 4. Therefore, the transition of the hierarchy among consecutive links is not as sharp as appeared on the left side: the light blue link, which is located immediately before the link we are examining, has an evident profile (dominant Level 2), then on the dark blue link Level 2 and Level 4 are mixed, finally on the blue link, Level 3, located immediately after the link we are examining, probably incorporates Level 2 making this transition smoother. Unfortunately, it is not possible to ensure that Level 3 has indeed incorporated Level 2 because it is not possible to recognize the inflow and the outflow of each node: the links are bi-directional for cyclists in the AGIV network and therefore it is not possible to investigate the node continuity (e.g. understanding the direction of the demand that is moving in a certain hierarchical level). Policy makers could then use this map in with a twofold scope: (i) identifying the backbone of the network; in this case they will simply consider the part of the network which has a relevant percentage in level 4; (ii) highlighting the links with major conflicts and which thus may be more critical.



Figuur 36: Transition between consecutive links

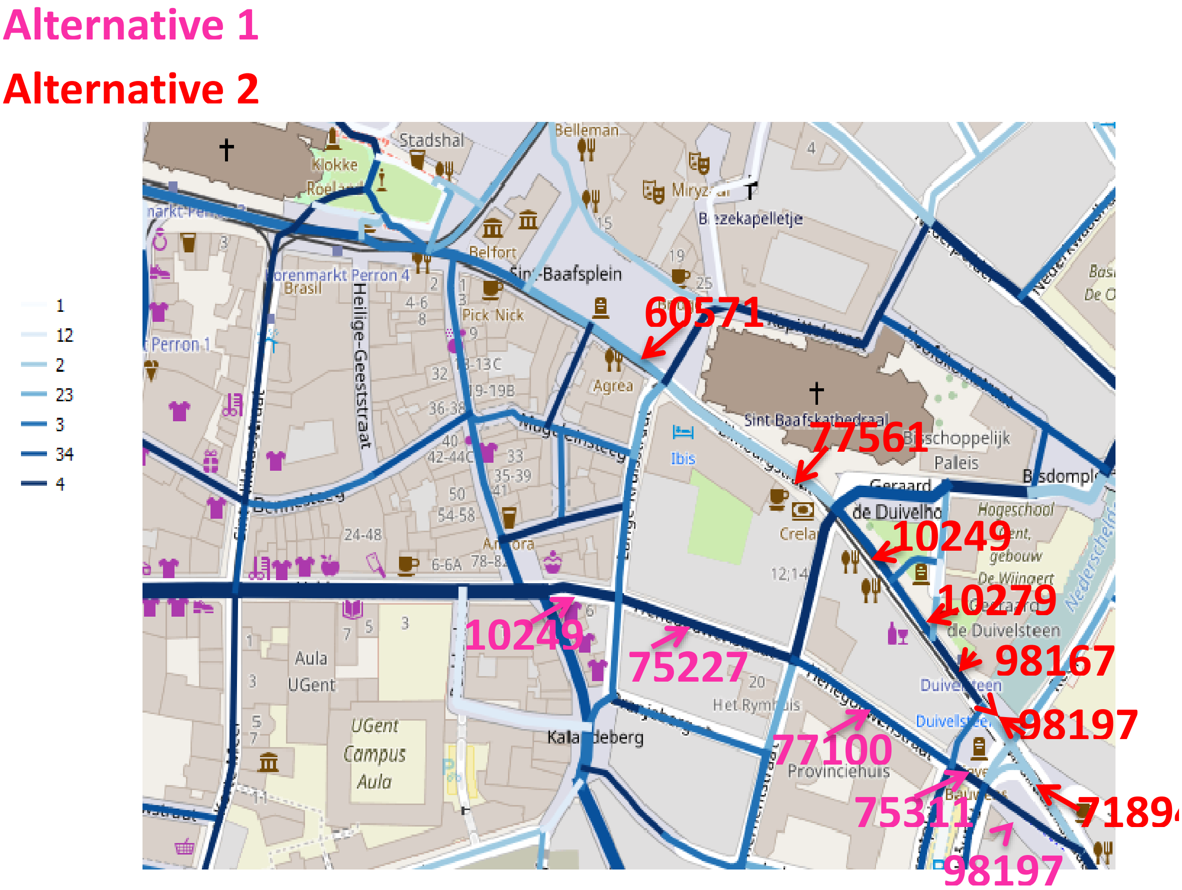
For the kind of analysis we are presenting, it is important to recognize which parts of the network have a mixed function (e.g. can be used both for having access to an origin/destination or as main leg of a longer trip). As visual support, it is useful to consider only four levels of aggregation

(1,2,3,4) and further aggregate them in two groups: low hierarchy that includes Level 1 and Level 2 and high hierarchy that aggregates Level 3 and Level 4. Figure 37 shows for each link in which range the sum of the percentages between level 1-2 and 3-4 falls. This means that in both maps the purple links are potentially problematic: they have a highly mixed function; the sum of level 1-2 and 3-4 lies between 40% and 60%.



Figuur 37: Mixed function vs homogenous function

We will now show an example of these potentially problematic events in case of two alternatives paths with an origin in common (Figure 38). However, due to privacy issues we will not include in the picture the full trip but only part of it. The chosen individuals use both the alternatives contributing differently to the aggregated hierarchy. From a policy perspective, the most difficult alternative to recognize is “Alternative 2” where levels are consistently mixed: transition between Level 2 and Level 4 for example. The causes underlying this behavior can be multiple: the purpose of the trip, the day and the hour in which the trip is performed, etc. In this situation, with two parallel alternatives, it could be discussed whether it may be beneficial to concentrate the demand with similar characteristics on one of the two alternatives in order to decrease conflicts between too heterogeneous cyclists.



Figuur 38: Parallel alternatives

As a consequence of the diverse behaviour, the links with a mixed function have a higher variance in terms of speed as well as in terms of the length of the trips that cross these links (ongoing analyses not shown here).

The examples showed so far investigate how the network infrastructure reflects the hierarchy of the trips/users. But what is the role of the individual perception in this context? How does the aggregated hierarchy (multiple users and multiple trips) influence the individual satisfaction?

Our methodology could also help in answering these questions. For instance, in one of our ongoing analyses, we introduce a “score” of a person’s conformity, assuming that conformity may be correlated to subjective satisfaction with cycling. Conformity objectively measures the correspondence between the function of the link in one’s current trip and that of the general cyclist population on the same link. E.g. a cyclist cruising a link at level 4 (i.e. as part of the connecting stage of a long trip) would be 100% conforming if all other cyclists are also using it as level 4, and only 10% if 90% of users are using it in levels 1, 2, or 3. Quantifying the level of conformity opens several avenues for interesting analyses. Aggregating conformity over a route may be a measure that could (partially) explain route choice. Aggregating conformity over several trips of the same individual into an overall conformity of one’s cycling pattern may be correlated to one’s satisfaction and thus positive or negative attitude towards (e-)cycling in general. If such correlation could indeed be validated, then infrastructural measures could be aimed at increasing the level of conformity in the cycling population, as it would be a predictor for overall satisfaction with the measure. Also, it could be an aspect of (e-)bike promotion campaigns (like projects offering commuters a trial period with an e-bike): such campaigns could be targeted at people whose (non-cycling) mobility pattern, when routed over the cycle network, would exhibit high levels of conformity, as such people would be more likely to experience the infrastructure in their environment as “(e-)bike friendly”.

2.3.3.4 Future Research

As this section just presented a concept without full elaboration nor results, there are several future research directions:

Methodological:

- Parameter values: The parameters defining the hierarchical levels are so far arbitrary: what is a rational ground for determining an appropriate number of levels and the thresholds separating them? This question may be answered based on behavioral arguments (which separations make sense in the minds of cyclists?) and on pragmatic arguments (which separations lead to an outcome that allows meaningful interpretations?).

- Time or distance? So far we suggested separating tracks based on distance from origin or destination. It may also make sense to separate based on time since origin/to destination. Arguments to choose either approach need to be developed.

- Data selection: It may be that the implicit function of the network changes depending on conditions and travel patterns. What is the influence of the time of day (morning/evening peak, low), days of the week, work zones with cyclists rerouting, events and other temporary situations. Not filtering for non-recurrent situations may lead to seemingly mixed functions that are in reality more homogeneous; omitting such situations may yield an incomplete picture and may be not fully representative.

- Visualization: A functional mix is a multidimensional characteristic that belongs to a 2D object (link or intersection). Visualization in 2 dimensions can hence be challenging as it inevitably requires simplified views on the data: either one loses overview, or one loses information. No experience exists as yet with visualizations that are meaningful and can be interpreted by both experts and non-experts.

- Validation: Application of the empirical hierarchy to a sufficiently large dataset is required to prove the concept and further refine and exploit it.

- Privacy issues: In dense areas, many tracks overlap and it will be impossible to refine individual data. However as density of tracks decreases, tracking data typically becomes more discrete until individuals could be recognized. The methodology should in that case protect privacy by design.
  Inversely, source data for the analysis may have constraints that originate from privacy concerns. For instance, data of the first/last X meters may have been deleted to protect privacy. Also, the data may contain multiple trips from the same person, but the key allowing to link these trips may have been removed. These and similar limitations of the data may compromise certain analyses in the context of empirical hierarchy and conformity. More experience is needed on the consequences of this for the value of the analyses.

Behavioural:

- Behavioural relevance of hierarchy: So far it has not been validated that the empirical hierarchy would have any meaning to cyclists: are they aware of it? is the basic assumption that bicycle trips would preferably form 'staircases' justified? (Note that for cars, hierarchy correlates with speed regimes and most cars can cover the entire relevant range of speeds. Cyclists may exhibit far less difference in speeds over the hierarchical levels; then is the concept relevant at all?)

- Relation conformity – satisfaction: Is there indeed a correlation between conformity and satisfaction as discussed in earlier sections?

- Homogeneity: Can the hierarchy be determined in general for all cyclists, or would the implicit hierarchy be dependent on gender, age, type of bike used (regular, electric, speed pedelec, cargo bike,…)? Especially potential difference in needs between regular cyclists and (fast) e-bikes may be of interest as little is known about how to integrate the latter into the multimodal transportation infrastructure.

- Impact on behavior: Could hierarchy or conformity indeed be related to route choice, or satisfaction with certain types of bicycle or cycling in general, as suggested before?

- Speed (variance) vs. Function: Is there a (positive) correlation between speed variance within a link and the mixture of functions? Is this perceived as natural, (un)comfortable, (subjectively) unsafe,…? Are there links that – within their category – are outliers in the sense that speeds are exceptionally high or low, or variance is higher or lower? If so, are they perceived as more or less attractive/comfortable/safe? What is the relation and perception of speed within a link of a certain functional mix, relative to the desired speed over the track? The assumption here is that each traveler has a desired speed for a trip (which may vary depending on his purpose, time constraints, company, bicycle,…), which may be observed as some high percentile of speed over the track. Is it acceptable that lower level stages of the trip allow lower speeds, and if so, what are acceptable reduction levels?

Transportation planning:

- Mixed functions vs. Monofunctionality: Which mixes of functions do exist and should some be encouraged or rather avoided? How can networks be designed that comply with these guidelines?What is the relation between functional mix and land use, specific attractors,…? Which interactions are desired or need to be avoided?Do links with similar function or functional mix form routes/corridors or even networks? How are they connected? what are the network characteristics like network shape (grid (if so, rectangular, triangular, random?), concentric,…), maze width,…?What is the link between implicit empirical functions and the planned hierarchy of the infrastructure? What do we learn from correspondences or differences? Are there missing links or other modifications needed to the existing bicycle network planning?What is the relation between the functional mix and how the physical infrastructure is designed? Does the empirical function follow physical characteristics (e.g. concentration on high-standard cycling infrastructure)? Are there consequences for design guidelines of cross sections or of intersection types, priority at intersections?What is the relation between the empirical hierarchy and the planned hierarchy of the motorized road traffic infrastructure? Where do they overlap/cross? Is this desired or not? What is the relation between the empirical hierarchy and public transportation planning? and bicycle parking facilities? shared bicycle infrastructure?Is there a relation between empirical functional hierarchy, functional mix, intersections with other infrastructures, and safety statistics (black spots)? This may reveal structural causes for safety problems, as well as structural ways of how to avoid or solve them.

3 Safety Aspects

This section aims at understanding which areas are perceived as potentially dangerous from our sample. During the period of tracking we have asked our users to spontaneously fill in a survey in each time they perceived a situation as dangerous. The outcome of this survey is very limited: in total 18 dangerous situations were reported. In a scale of risk from 1 (very low) to 5 (very high) all the responses were between 3 and 5. Moreover, when we asked to better describe the situation (“crash alone”, “crash with another (e-)bike”, “crash with pedestrian”, “crash with a vehicle”) all uses reported a crash with a vehicle. Asking instead the reason for which the accident could happen, the most frequent answers were “traffic conditions” and “absence of bicycle path”. Analyzing additional notes and remarks (which could freely be added in the survey to better specify the answers), we noticed how the majority of the situations are related with missing give way from the other road users and with the complicate junctions where the bicycle path is not even clearly painted.

Among the 18 different locations, we show below the most critical situations reported by our users.

Figuur 39: Dangerous situation

Figuur 40: Dangerous situation

Figuur 41: Dangerous situation

Figuur 42: Dangerous situation

Figure 39 and 41 represent two typical situations where road user distraction might cause accidents. In the first case (Figure 39) there is a right turn with a highly visible bicycle path, therefore the accident is totally attributable to the vehicle approaching the right turn; in the second case (Figure 41), instead, there is both a clear insufficiency in the cycling infrastructure (totally absent) and in the attention of drivers while opening the door. Figure 40 and 39 show again the lack of cycling infrastructure. Moreover in Figure 42 we have a roundabout with many interactions: 4 accesses and parking on the sides that increase the risk at the intersection.

4 Conclusions and Policy Implications

In this section we extrapolate the results of our analyses to propose some policies concerning the functional use of e-bikes. This includes some suggestions derived from the literature and not directly (yet) observable from the new analyses we have done.

In this project we analysed survey data and GPS data coming from 60 e-bike owners.
The initial survey helped us in: defining the user's profile; understanding the mobility habits of e-bikers; investigating the satisfaction of the users related with the e-bike infrastructure (e.g. storage, re-charge spots). The sample is well distributed in terms of gender and age. The income is proper of the Belgian middle class and the majority of the users are employed (Table 2). The e-bike is mostly used for commuting trips (to work/school) and during free time ("habitual leisure"). Instead for "general shopping" and "non-recurrent activities" the car remained the preferred mode of transport. This may be associated among other factors (that were not investigated) with users' dissatisfaction with the presence of re-charge spots and safe public bicycle storage. While these two issues can easily be solved at home/work location, this is not the case for other type of trips. Given the high acquisition cost of the e-bike, and in order to support its usage also for non-work related trips, future investment should be focusing on improving at least the presence of safe storage spots which, in case of short trips, is considered more important than the presence of re-charge spots.

The GPS data analysis allowed us to get an overview of the preferred routes of our users and to better know their profiles in terms of speed, distance and time. From the density estimation (Figure 14) we identify at least two "speed-groups": one distributed around an average of 15 km/h and the other between 25 and 30 km/h. As expectable, the average speed is higher for males and for people with higher experience with an age included between 41-59 years. However, the speed of the shared e-bikes (technically comparable to the average of the sample but for which we don't know the user's profile) is higher than that of owned e-bikes. An interesting result is to see how the average speed increases with trip length. This result shows one of the interesting potentials of the electric bicycle: the hypothesis seems empirically confirmed that electric support enables longer trips at reasonable speeds (probably higher on average than traditional cycling), making cycling an attractive alternative to the car for more trips.

The route choice model developed in Section 2.3.1 shows the importance of attributes such as turns and changes in hierarchy in route choice decisions of e-bikers. They prefer routes with less intersections, especially when they cannot go straight and need to maneuver (right, and even more so left turns). This suggests (understandably) that e-cycling could be promoted if more uninterrupted cycling routes would be provided. This encourages the current policy of developing cycling highways, but extends also beyond highways alone: preference for routes with less interruptions and heterogeneity in hierarchy is common to all e-bike trips, also shorter ones. Whereas this probably also holds for regular bicycles, our results suggest that e-bikers are somewhat more prepared to make detours if this additional distance is compensated by fewer interruptions for intersections or turning maneuvers. Analyzing the length rode using different types of road, the most preferred roads are the "secondary". It is not entirely clear why secondary roads seem to be preferred over local roads and cycleways. Some hypotheses are that this might be related with the number of interaction the e-bikers have with the traditional bikes using cycleway roads, as well as with the fact that this category includes also less comfortable paths like dirt tracks, single tracks (e.g. in parks) and cobblestones.

Our observations on route choice preferences highlight the importance of a well-structured hierarchical cycling network. Efforts along this line are already existing in Flanders (BFF, fiets-o-strade) but may need to be reinforced. In order to strengthen our knowledge on the true needs of the cyclist population, we propose a new methodology to infer from gps-data an implicit hierarchical cycling network as it emerges from the user's perspective. Although the data available in this project did not allow validating this approach, we anticipate that this can be a highly valuable instrument in cycling transport planning but also in warranting structural safety in the network.

The case study shown in Section 2.3.2 showed that not only the traditionally considered socio demographic attributes (income, household composition, gender,…) matter when people decide which vehicles and mobility tools they will use in their daily life. We give empirical evidence that the convenience that a vehicle or mobility service brings in the personal daily mobility pattern, balanced against the cost of acquisition, has a strong influence on such decisions. Therefore from a policy perspective, this knowledge helps the stakeholders in understanding which factors determine whether a person would consider adopting a new mobility option (a new service (e.g. shared car) or resource (e.g. electric bicycle)): this depends in the first place on the person's travel pattern (not just on single trips or tours, as many policy support models assume!) and whether the new option brings sufficient reduction in travel cost over this travel pattern for a reasonable price. Moreover, we propose a new mathematical model that quantifies the sensitivity of such decisions to changes in the travel cost over one's travel pattern; whereas state-of-the-art models so far disregarded this and predicted ownership decisions purely based on socio-demographic arguments. In contrast to these existing models, the new model can thus be used by policy makers to predict how market shares (modal shares, adoption of sustainable modes like e-bike,…) evolve in response to policy measures aimed at making certain mobility options more (or less) attractive, whether by affecting the acquisition cost or by improving travel costs of these modes. The socio-demographic information is instead useful for understanding the commonalities among the people that perform the same ownership choice. In this project however, we were able to develop the methodology and models for such analyses, and to validate the concept based on a limited dataset only. The data contained ownership information of only traditional modes: car, public transport, and bicycle. Moreover the number of subjects on which information was available was too low to split the population into homogeneous groups with different sensitivities (which would be interesting policy information). Although the current model offers insight into the important role of travel cost over a full set of trips in a travel pattern, the true potential to policy can be much larger still when more data could be used. Such data should contain multiday trip patterns of a relevant fraction of the population, tracked over a longer period such that changes in one's mobility tool portfolio can be monitored. Traditional survey methods base on mobility panels could be useful for this purpose (ongoing work in our group exploits for instance the Germany mobility panel travel survey where information from 1994 to 2015 are included). But in addition to that, data transferability techniques (allowing reuse of data collected elsewhere – of course after appropriate transformations) may be used to combine multiple smaller surveys of different regions and years. And finally the continuous tracking through all kinds of big data techniques may offer new opportunities for this kind of policy-relevant modeling.

Focusing on safety, in literature little is known about the differences between safety aspects related with an ordinary bike and an e-bike. In [98], for example, the authors cite weight of the bicycle, speed and width of the bicycle path as three possible factors. However, these differences appear of little significance when needing to justify different/dedicated e-cycling infrastructures. These results are also confirmed by our own study: the potentially dangerous situations in which the e-bikers are involved don't present any differences with those of ordinary bikes when the cause is the total lack or the low quality of the bicycle infrastructure. However, we are strongly convinced that a misperception of the higher speed compared to the ordinary bike increases the risk of accidents at intersections, as was also showed in [20]. Moreover, as stated in [99], one of the most important requirements for sustainable safe road used by cyclists is the prevention of crashes between fast and slow traffic.

Literatuurlijst

[1] S. L. Handy, Y. Xing, and T. J. Buehler, "Factors associated with bicycle ownership and use: a study of six small U.S. cities," *Transportation*, vol. 37, no. 6, pp. 967–985, Nov. 2010.

[2] M. A. Stinson and C. R. Bhat, "Frequency of bicycle commuting: internet-based survey analysis," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1878, no. 1, pp. 122–130, 2004.

[3] J. Pucher and R. Buehler, “Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany,” *Transp. Rev.*, vol. 28, no. 4, pp. 495–528, Jul. 2008.

[4] B. Caulfield, E. Brick, and O. T. McCarthy, “Determining bicycle infrastructure preferences – A case study of Dublin,” *Transp. Res. Part Transp. Environ.*, vol. 17, no. 5, pp. 413–417, Jul. 2012.

[5] J. Broach, J. Dill, and J. Gliebe, “Where do cyclists ride? A route choice model developed with revealed preference GPS data,” *Transp. Res. Part Policy Pract.*, vol. 46, no. 10, pp. 1730–1740, Dec. 2012.

[6] E. Heinen, B. van Wee, and K. Maat, “Commuting by Bicycle: An Overview of the Literature,” *Transp. Rev.*, vol. 30, no. 1, pp. 59–96, Jan. 2010.

[7] E. Heinen, K. Maat, and B. van Wee, “The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances,” *Transp. Res. Part Transp. Environ.*, vol. 16, no. 2, pp. 102–109, Mar. 2011.

[8] J. Weinert, C. Ma, and C. Cherry, “The transition to electric bikes in China: history and key reasons for rapid growth,” *Transportation*, vol. 34, no. 3, pp. 301–318, May 2007.

[9] C. Cherry and R. Cervero, “Use characteristics and mode choice behavior of electric bike users in China,” *Transp. Policy*, vol. 14, no. 3, pp. 247–257, May 2007.

[10] C. Cherry, “Electric Bike Use in China and Their Impacts on the Environment, Safety, Mobility and Accessibility.” Working Paper UCB-ITS-VWP-20017-03, 2007.

[11] J. MacArthur, J. Dill, and M. Person, “Electric Bikes in North America,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2468, pp. 123–130, Dec. 2014.

[12] J. Dill and G. Rose, “Electric Bikes and Transportation Policy: Insights from Early Adopters,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2314, no. 1, pp. 1–6, Dec. 2012.

[13] N. Popovich, E. Gordon, Z. Shao, Y. Xing, Y. Wang, and S. Handy, “Experiences of electric bicycle users in the Sacramento, California area,” *Travel Behav. Soc.*, vol. 1, no. 2, pp. 37–44, May 2014.

[14] E. Fishman and C. Cherry, “E-bikes in the Mainstream: Reviewing a Decade of Research,” *Transp. Rev.*, pp. 1–20, Jul. 2015.

[15] “http://www.bike-eu.com.” .

[16] “http://www.epomm.eu.” .

[17] “The Global E-bike market.” 2014.

[18] “http://www.xpats.com/electric-bicycle-gaining-popularity.” .

[19] L. Bai, P. Liu, Y. Chen, X. Zhang, and W. Wang, “Comparative analysis of the safety effects of electric bikes at signalized intersections,” *Transp. Res. Part Transp. Environ.*, vol. 20, pp. 48–54, May 2013.

[20] T. Petzoldt, K. Schleinitz, J. F. Krems, and T. Gehlert, “Drivers’ gap acceptance in front of approaching bicycles – Effects of bicycle speed and bicycle type,” *Saf. Sci.*, Aug. 2015.

[21] B. C. Langford, J. Chen, and C. R. Cherry, “Risky riding: Naturalistic methods comparing safety behavior from conventional bicycle riders and electric bike riders,” *Accid. Anal. Prev.*, vol. 82, pp. 220–226, Sep. 2015.

[22] E. Fishman and C. Cherry, “E-bikes in the mainstream: reviewing a decade of research.pdf,” *Transport Reviews*, Jul. 2015.

[23] M. Dozza, G. F. Bianchi Piccinini, and J. Werneke, “Using naturalistic data to assess e-cyclist behavior,” *Transp. Res. Part F Traffic Psychol. Behav.*, May 2015.

[24] “http://www.mech.kuleuven.be/sprint/.” .

[25] F. Viti, C. Tampère, R. Frederix, M. Castaigne, F. Walle, and E. Cornelis, “Analyzing weekly activity-travel behaviour from behavioural survey and traffic data.,” Jul-2010. [Online]. Available: https://lirias.kuleuven.be/handle/123456789/294554. [Accessed: 15-Jul-2014].

[26] T. Streit, C. Weiss, B. Chlond, and P. Vortisch, “MORE CYCLING OR MORE CYCLISTS– DYNAMICS AND POTENTIALS OF 1 BICYCLE USE IN GERMANY BY A CLUSTER BASED APPROACH 2,” in *Transportation Research Board 93rd Annual Meeting*, 2014.

[27] Y. Yue, T. Lan, A. G. O. Yeh, and Q.-Q. Li, “Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies,” *Travel Behav. Soc.*, vol. 1, no. 2, pp. 69–78, May 2014.

[28] J. Wolf, M. Oliveira, and M. Thompson, “Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey,” *Transp. Res. Rec.*, vol. 1854, no. 03, pp. 189–198, 2003.

[29] L. Yalamanchili, R. Pendyala, N. Prabaharan, and P. Chakravarthy, “Analysis of Global Positioning System-Based Data Collection Methods for Capturing Multistop Trip-Chaining Behavior,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1660, pp. 58–65, Jan. 1999.

[30] G. Draijer, N. Kalfs, and J. Perdok, “Global Positioning System as Data Collection Method for Travel Research,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1719, pp. 147–153, Jan. 2000.

[31] J. Wolf, M. Oliveira, and M. Thompson, “Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1854, pp. 189–198, Jan. 2003.

[32] A. Bolbol, T. Cheng, I. Tsapakis, and J. Haworth, “Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification,” *Comput. Environ. Urban Syst.*, vol. 36, no. 6, pp. 526–537, Nov. 2012.

[33] N. Schuessler, K. Axhausen, and N. Schüssler, “Processing Raw Data from Global Positioning Systems Without Additional Information,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2105, pp. 28–36, Dec. 2009.

[34] M. Bierlaire, J. Chen, and J. Newman, “A probabilistic map matching method for smartphone GPS data,” *Transp. Res. Part C Emerg. Technol.*, vol. 26, pp. 78–98, Jan. 2013.

[35] Y. Zheng, Y. Chen, Q. Li, X. Xie, and W.-Y. Ma, “Understanding transportation modes based on GPS data for web applications,” *ACM Trans. Web*, vol. 4, no. 1, pp. 1–36, 2010.

[36] N. Schuessler and K. Axhausen, “Identifying trips and activities and their characteristics from GPS raw data without further information,” ETH, Eidgenössische Technische Hochschule Zürich, IVT, 2008.

[37] S. T. Doherty, N. Noël, M. L. Gosselin, C. Sirois, and M. Ueno, “Moving beyond observed outcomes: integrating global positioning systems and interactive computer-based travel behavior surveys,” 2001.

[38] A. J. Lopez, D. Ochoa, and S. Gautama, “Detecting Changes of Transportation-Mode by Using Classification Data,” in *18th International Conference on Information Fusion (Fusion 2015)*, 2015, pp. 2078–2083.

[39] E. Frejinger, “Route choice analysis: data, models, algorithms and applications,” École Polytechnique Federale de Lausanne, 2008.

[40] J. H. Ward Jr., “Hierarchical Grouping to Optimize an Objective Function,” *J. Am. Stat. Assoc.*, vol. 58, no. 301, pp. 236–244, Mar. 1963.

[41] C. Brandenburg, A. Matzarakis, and A. Arnberger, “Weather and cycling—a first approach to the effects of weather conditions on cycling,” *Meteorol. Appl.*, vol. 14, no. 1, pp. 61–67, Mar. 2007.

[42] S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming, “Evaluation of choice set generation algorithms for route choice models,” *Ann. Oper. Res.*, vol. 144, no. 1, pp. 235–247, May 2006.

[43] J. Y. Yen, “Finding the K Shortest Loopless Paths in a Network,” *Manag. Sci.*, vol. 17, no. 11, pp. 712–716, Jul. 1971.

[44] C. Prato and S. Bekhor, “Modeling Route Choice Behavior: How Relevant Is the Composition of Choice Set?,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2003, pp. 64–73, Jan. 2007.

[45] “Meta-analysis of choice set generation effects on route choice model estimates and predictions - Transport - Volume 27, Issue 3.” [Online]. Available:

http://www.tandfonline.com/doi/abs/10.3846/16484142.2012.719840#.V1FexNl96M8. [Accessed: 03-Jun-2016].

[46] E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta, “A MODIFIED LOGIT ROUTE CHOICE MODEL OVERCOMING PATH OVERLAPPING PROBLEMS. SPECIFICATION AND SOME CALIBRATION RESULTS FOR INTERURBAN NETWORKS,” presented at the TRANSPORTATION AND TRAFFIC THEORY. PROCEEDINGS OF THE 13TH INTERNATIONAL SYMPOSIUM ON TRANSPORTATION AND TRAFFIC THEORY, LYON, FRANCE, 24-26 JULY 1996, 1996.

[47] M. Ben-Akiva and M. Bierlaire, “Discrete Choice Methods and their Applications to Short Term Travel Decisions,” in *Handbook of Transportation Science*, R. W. Hall, Ed. Springer US, 1999, pp. 5–33.

[48] M. Fosgerau, E. Frejinger, and A. Karlstrom, “A link based network route choice model with unrestricted choice set,” *Transp. Res. Part B Methodol.*, vol. 56, pp. 70–80, Oct. 2013.

[49] M. E. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, 1985.

[50] K. Train, “A Structured Logit Model of Auto Ownership and Mode Choice,” *Rev. Econ. Stud.*, vol. 47, no. 2, pp. 357–370, Jan. 1980.

[51] K. Kockelman, “Travel Behavior as Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from San Francisco Bay Area,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1607, pp. 116–125, Jan. 1997.

[52] C. Zegras, “The Built Environment and Motor Vehicle Ownership and Use: Evidence from Santiago de Chile,” *Urban Stud.*, vol. 47, no. 8, pp. 1793–1817, Jul. 2010.

[53] G. D. Jong, J. Fox, A. Daly, M. Pieters, and R. Smit, “Comparison of car ownership models,” *Transp. Rev.*, vol. 24, no. 4, pp. 379–408, 2004.

[54] J. L. Bowman, “ACTIVITY BASED TRAVEL FORECASTING1 JOHN. L. BOWMAN AND MOSHE BEN-AKIVA,” 1996.

[55] D. M. Scott and K. W. Axhausen, “Household Mobility Tool Ownership: Modeling Interactions between Cars and Season Tickets,” *Transportation*, vol. 33, no. 4, pp. 311–328, Jul. 2006.

[56] P. Vovsha and E. Petersen, “Model for Person and Household Mobility Attributes,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2132, no. 1, pp. 95–105, Dec. 2009.

[57] C. R. Bhat and J. Y. Guo, “A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels,” *Transp. Res. Part B Methodol.*, vol. 41, no. 5, pp. 506–526, Jun. 2007.

[58] D. Dissanayake and T. Morikawa, “Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed preference/stated preference Nested Logit model: case study in Bangkok Metropolitan Region,” *J. Transp. Geogr.*, vol. 18, no. 3, pp. 402–410, May 2010.

[59] A. Vij, A. Carrel, and J. L. Walker, “Incorporating the influence of latent modal preferences on travel mode choice behavior,” *Transp. Res. Part Policy Pract.*, vol. 54, pp. 164–178, Aug. 2013.

[60] S. Le Vine, M. Lee-Gosselin, A. Sivakumar, and J. Polak, “A new concept of accessibility to personal activities: development of theory and application to an empirical study of mobility resource holdings,” *J. Transp. Geogr.*, vol. 31, pp. 1–10, Jul. 2013.

[61] E. Abbe, M. Bierlaire, and T. Toledo, “Normalization and correlation of cross-nested logit models,” *Transp. Res. Part B Methodol.*, vol. 41, no. 7, pp. 795–808, Aug. 2007.

[62] M. Bierlaire, “BIOGEME: a free package for the estimation of discrete choice models,” presented at the Swiss Transport Research Conference, 2003.

[63] W. H. Greene, *Nlogit*. Version, 2002.

[64] J. L. Walker, “Extended discrete choice models: integrated framework, flexible error structures, and latent variables,” Massachusetts Institute of Technology, 2001.

[65] M. F. Yáñez, E. Cherchi, B. G. Heydecker, and J. de D. Ortúzar, “On the Treatment of Repeated Observations in Panel Data: Efficiency of Mixed Logit Parameter Estimates,” *Netw. Spat. Econ.*, vol. 11, no. 3, pp. 393–418, Aug. 2010.

[66] D. G. Fiebig, M. P. Keane, J. Louviere, and N. Wasi, “The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity,” *Mark. Sci.*, vol. 29, no. 3, pp. 393–421, Jul. 2009.

[67] S. I. Vrieze, “Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).,” *Psychol. Methods*, vol. 17, no. 2, pp. 228–243, 2012.

[68] Y. Yang, “Can the strengths of AIC and BIC be shared? A conflict between model indentification and regression estimation,” *Biometrika*, vol. 92, no. 4, pp. 937–950, Dec. 2005.

[69] Michel Bierlaire, “BIOGEME: A free package for the estimation of discrete choice models ,” *3rd Swiss Transportation Research Conference*, Ascona,Switzerland, 2003.

[70] C. Domarchi, A. Tudela, and A. González, “Effect of attitudes, habit and affective appraisal on mode choice: an application to university workers,” *Transportation*, vol. 35, no. 5, pp. 585–599, Aug. 2008.

[71] V. Plevka, P. Segaert, C. M. J. Tampère, and M. Hubert, “Analysis of travel activity determinants using robust statistics.,” *Transportation*, no. 43(6), pp. 979–996, 2016.

[72] E. Cascetta, *Transportation systems analysis: models and applications*, vol. 29. Springer, 2009.

[73] R. Srp, *Mobility as a service–Czech approach to MaaS*. WoCoMoCo, 2016.

[74] G. Baltas and C. Saridakis, “An empirical investigation of the impact of behavioural and psychographic consumer characteristics on car preferences: An integrated model of car type choice,” *Transp. Res. Part Policy Pract.*, vol. 54, pp. 92–110, Aug. 2013.

[75] V. Plevka and C. M. J. Tampère, “Modelling framework for optimal mobility services considering modal choice and vehicle ownership decisions,” presented at the Transport Research Board, 2016.

[76] A. J. Lopez, P. Astegiano, D. Ochoa, C. Tampere, C. Beckx, and S. Gautama, “Unveiling E-Bike Potential for Commuting Trips from GPS Traces,” *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 7, p. 190, Jun. 2017.

[77] R. M. Downs and D. Stea, *Image and Environment: Cognitive Mapping and Spatial Behavior*. Transaction Publishers, 1973.

[78] S. C. Hirtle and J. Jonides, “Evidence of hierarchies in cognitive maps,” *Mem. Cognit.*, vol. 13, no. 3, pp. 208–217, May 1985.

[79] A. Stevens and P. Coupe, “Distortions in judged spatial relations,” *Cognit. Psychol.*, vol. 10, no. 4, pp. 422–437, Oct. 1978.

[80] H. Voicu, “Hierarchical cognitive maps,” *Neural Netw.*, vol. 16, no. 5–6, pp. 569–576, Jun. 2003.

[81] A. Car and A. U. Frank, “Hierarchical street networks as a conceptual model for efficient way finding,” in *Proceedings of the EGIS*, 1993, vol. 93, pp. 134–139.

[82] A. Car and A. Frank, “General principles of hierarchical spatial reasoning-the case of wayfinding,” in *the Proceedings of the 6th International Symposium on Spatial Data Handling*, 1994, vol. 2, pp. 646–664.

[83] S. Timpf, G. S. Volta, D. W. Pollock, and M. J. Egenhofer, “A conceptual model of wayfinding using multiple levels of abstraction,” in *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, vol. 639, A. U. Frank, I. Campari, and U. Formentini, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 348–367.

[84] B. Immers and J. Diepens, “Achtergrounddocument 1b: Het ‘trappetje van Monderman.’” 2015.

[85] P. H. L. Bovy, “Het korste-tijd routekeuzecriterium: een empirische toetsing.” .

[86] R. van Nes, *Design of multimodal transport networks: a hierarchial approach*. Delft: Eigenverl, 2002.

[87] M. K. Y. Bhanu, “The emergence of hierarchy in transportation networks,” 2003.

[88] B. Immers, B. Egeter, and R. van Nes, “TRANSPORT NETWORK PLANNING: THEORETICAL NOTIONS. IN: HANDBOOK OF TRANSPORTATION ENGINEERING,” 2004.

[89] R. Buehler and J. Dill, “Bikeway Networks: A Review of Effects on Cycling,” *Transp. Rev.*, vol. 36, no. 1, pp. 9–27, Jan. 2016.

[90] J. E. Schoner and D. M. Levinson, “The missing link: bicycle infrastructure networks and ridership in 74 US cities,” *Transportation*, vol. 41, no. 6, pp. 1187–1204, Jul. 2014.

[91] M. Winters, M. Brauer, E. M. Setton, and K. Teschke, “Built Environment Influences on Healthy Transportation Choices: Bicycling versus Driving,” *J. Urban Health*, vol. 87, no. 6, pp. 969–993, Dec. 2010.

[92] V. V. Acker, B. Derudder, and F. Witlox, “Why people use their cars while the built environment imposes cycling,” *J. Transp. Land Use*, vol. 6, no. 1, pp. 53–62, Apr. 2013.

[93] L. Ma and J. Dill, “Do people’s perceptions of neighborhood bikeability match ‘Reality’?,” *J. Transp. Land Use*, vol. 0, no. 0, Jan. 2016.

[94] S. L. Handy and Y. Xing, “Factors Correlated with Bicycle Commuting: A Study in Six Small U.S. Cities,” *Int. J. Sustain. Transp.*, vol. 5, no. 2, pp. 91–110, Jan. 2011.

[95] M. Winters, K. Teschke, M. Brauer, and D. Fuller, “Bike Score®: Associations between urban bikeability and cycling behavior in 24 cities,” *Int. J. Behav. Nutr. Phys. Act.*, vol. 13, no. 1, Dec. 2016.

[96] M. Mekuria, P. Furth, and H. Nixon, “Low-Stress Bicycling and Network Connectivity,” *Mineta Transp. Inst. Publ.*, May 2012.

[97] “presto_policy_guide_cycling_infrastructure_en.pdf.” [Online]. Available: https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/presto_policy_guide_cycling_infrastructure_en.pdf. [Accessed: 11-Aug-2016].

[98] J. P. Schepers, E. Fishman, P. den Hertog, K. K. Wolt, and A. L. Schwab, “The safety of electrically assisted bicycles compared to classic bicycles,” *Accid. Anal. Prev.*, vol. 73, pp. 174–180, Dec. 2014.

[99] F. Wegman, F. Zhang, and A. Dijkstra, “How to make more cycling good for road safety?,” *Accid. Anal. Prev.*, vol. 44, no. 1, pp. 19–29, Jan. 2012.

[100] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numer. Math.*, vol. 1, no. 1, pp. 269–271, Dec. 1959.

[101] RStudio Team, *RStudio: Integrated Development for R. RStudio, Inc., Boston, MA*. 2015.

[102] Quantum GIS Development Team (YEAR), *Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project*. .

APPENDIX A

In this annex the softwares used during the whole analysis with their related input/output will be described. This annex is structured in three parts related to the three main methodological topics discussed in section 2.3.

1) Tools used for the analysis in section 2.3.1

The analysis done to model route choice decisions covered three main themes:

a. Map-Matching

it required two sources of data: GPS data and network (links and nodes) data. Managing such big datasest was challenging and forced us to use an object-relational database management system: PostgresSQL. It allowed to safely store the data but also to return them in response to requests of other applications. In the first step of the map matching process, the gps points were aligned to the OSM (© OpenStreetMap contributors) network. To do so, the OpenStreetMap routing engine has been used. It assigned each GPS point to a link of the OSM network. The limitation of this approach is that the points are aligned to the link axis: two points coming from trips in opposite direction but on the same road will overlap. Since we were interested in matching the trips on the AGIV network, the following step was to intersect the results of the first step with this network. This was done using PostGIS. This is a spatial database extender for PostgresSQL which allows location queries. The output of this phase is a dataset in which, for each GPS coodinate the associated link is included.

b. Route set generation

To handle the large scale of the case study network, a python implementation of the igraph (http://igraph.org) library has been employed both for storing the graph representation of the network itself and for computing the large amount of shortest paths necessary. Receiving as input i) a descriptor of the network's topology in terms of links and nodes, ii) a set of origin-destination node pairs and iii) a set of (generalized) link costs, the developed algorithm computes a given K (in our case, 5) shortest paths connecting each origin-destination pairs. Shortest paths are computed utilizing the well-known Dijkstra algorithm [100], which is generalized to compute the successively k-shortest routes by iteratively removing one arc from the network's graph (*link elimination* approach) and re-running the algorithm for the same origin-destination pair. Results are stored in comma separated value files, including the original (measured) path, the 5 algorithmically generated alternatives, the costs for each link composing each path and the Path Size indicators computed per each couple of paths (Eq 5). Computationally, the proposed algorithm handles, in average, about 100 trips per hour.

c. Model calibration

This step required in input the results of the output of the map-matching and of the route set generation phase. Before to proceed with the proper calibration, some data-processing activities were necessary. These were done in Rstudio [101] which is an integrated development environment and support both code execution and plots analysis. At this stage the attributes mentioned in 2.3.1 were computed. Deriving the type of road and the type of intersection was a relatively simple activity. While the computation of the changes in hierarchy and above all the turns was demanding. Specifically, the computation of the number of turns required to compute the angle among consecutive points which was calculated with a query in Qgis [102]. The computation of the different attributes were done both for the chosen route and for the five generated routes. The measures were then aggregate per trips. In the final dataset each row included the information related with one trip: all the attributes of the chosen path and of the five generated routes. This dataset has been then given in input to Biogeme [69] which produced the estimated parameters associated with each attribute.

2) Tools used for the analysis in section 2.3.2

The analysis done to model the relation between daily modal choice and owenrship decisions covered two main themes:

a. The model calibration

The calibration of the first step (Section 2.3.2.3.1) needed in input a txt file included all the trips performed by the sample. Specifically, each row represent a trip with its attributes (i.e. time, purpose). Multiple trips (rows) can belong to the same individual. This file, together with the model specification assumed in Equation 10 are then given in input to Biogeme which produced in output the estimated parameters. These allowed to compute the personal accessibility measure which represents one of the attributes of the second step.

The calibration of the second step(Section 2.3.2.3.2), even though similar in the practical steps, required a different strutcure of the input file. In this case each row coincides with a different individual and the attributes considered are the ones mentioned in Equation 11. On the same line of the previous step the model is then calibrated with Biogeme.

b. The interaction analysis

To investigate the influence of the socio-demographic information on personal accessibility an interaction plot has been drawn. The latter required a dataset which included individual characteristics: gender, income, zip code. With the support of Rstudio and its function “interaction.plot” Figure 30 has been produced.

3) Tools used for the analysis in section 2.3.3

The analysis suggested for the hierarchical network analysis required two main steps:

a. Map Matching

We refer the reader to section 1.a of this annex.

b. Trips characterization

This step required in input the output of the map-matching. Rstudio with its package “dplyr” allowed to (i) describe each trip with a certain number of levels; (ii) assign to each link in the network its hierarchical level: each link could also be characterized by more than one level; (iii) compute statistics per link and per individual. The output generated with Rstudio has been then given in input to Qgis in order to produce some maps to immediately verify the structure of the tested network (e.g. links with homogenous profile vs links with a mixed profile).

Het Steunpunt Verkeersveiligheid 2012-2015 is een samenwerkingsverband tussen de volgende partners: