MSc Thesis in Geomatics
Faculty of Architecture and the Built Environment

Examining the influence of urban design on cyclist route choice in different weather conditions

Erik van der Wal
January 2020

TU Delft
Delft University of Technology

EXAMINING THE INFLUENCE OF URBAN DESIGN ON CYCLIST ROUTE CHOICE IN DIFFERENT WEATHER CONDITIONS

A thesis submitted to the Delft University of Technology in partial fulfillment
of the requirements for the degree of

Master of Science in Geomatics for the Built Environment

by

Erik van der Wal

January 2020

Erik van der Wal: *Examining the influence of urban design on cyclist route choice in different weather conditions* (2020)

🅭🅯 This work is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit
http://creativecommons.org/licenses/by/4.0/.

The work in this thesis was carried out in the:

Faculty of Architecture & the Built Environment
Delft University of Technology

Supervisors:	Dr. Kees Maat
	Dr. ir. Martijn Meijers
Co-reader:	Dr. ir. Stefan van der Spek

ABSTRACT

With regard to climate change and air pollution within cities, interest in sustainable modes of transportation for regular use has taken a rise. Utilitarian cycling is being seen as a frontrunner for replacing everyday motorized travels within and between cities, supported by the rapid emergence of the electric bicycle. Governments are trying to use the increasing opportunities involving bicycle transportation to reduce car traffic and the related air polution, by stimulating the use of bicycles. In this light, interest is drawn to cyclist travel behavior to uncover preferences of cyclists.

Existing literature shows a significant impact of weather conditions on cyclist travel behavior in terms of tranportation mode choices. Especially adverse weather condtions leave their mark on the use of bicycles as a means of transportation, as it is recognized by many studies as a main deterrent for cycling. On the other hand, the relation between weather conditions and cyclist route choice is an underexplored topic in existing literature. Consequently, it has remained unclear to what extent cyclists attempt to mitigate the influence of weather condtions through choice of route, and based on which determinants. Insights in ways to mitigate unchangeable external circumstances like weather conditions could be another step forward in stimulating utilitarian use of bicycles in the search for transportation modes that can replace motorized trips.

This thesis made an attempt to partially address the research gap in existing literature, by departing from findings in the field of pedestrian mobility. In these studies, pedestrians have been found to adapt their choice of route to the degree of shelter that is offered by the built environment as a measure to change level of weather exposure. These findings were projected on cyclist route choice, to evaluate to what extent cyclist alter their choice of route based on weather conditions and the degree of shelter that can be found within a built environment. An elaborate methodology was proposed in which observed routes throughout the study area of Tilburg (the Netherlands), comprising trips made with conventional and electric bicycles, were compared with shortest and fastest alternatives. The weather conditions under which a route was conducted were modelled through a set of individual meteorological factors, spatially related to the location of an observed route. To operationalize the degree of shelter provided by the built environment, a new method was developed using aspects from existing theories on street climate design and spatial openness in order to provide a detailed description of the potential shelter along a route. Three different shelter factors were established, describing the degree of mean building shelter, maximum building shelter, and vegetational shelter in the form of tree density along a route. Through estimation of a set of linear regression models, independent and combined effects of the meteorological and shelter factors on cyclist route choice were modelled.

Initial moderate influences of windspeed, temperature, and cycling under twilight conditions on the choice of route were found, while cyclists generally chose routes with a lower degree of building and vegetational shelter compared to alternative shortest and fastest routes. Interactions between the effects of meteorological and shelter factors showed very limited additional effects, suggesting that utilitarian cyclists in the study area did not value the degree of built environment shelter along a route sufficiently as a mitigator of weather conditions to diverge from the shortest or fastest route. These findings imply that built environment shelter does not have to be accounted for in policy design to stimulate utilitarian cycling.

ACKNOWLEDGEMENTS

This thesis report is the sole tangible product of a process that lasted one month short of a year, and is also my final submission as a Geomatics student. Over the course of these eleven months, I have received plenty of feedback, advice, and support. So, eventough my name is written on the frontpage, I could not have reached this point without the help of a number of people.

From the academic perspective of this thesis, I would like to express special thanks my supervisors, Kees Maat and Martijn Meijers. The expertise of Kees on bicycling problems has provided me with a lot of guidance in defining the scope of this thesis, and helped me in focussing on the global lines in between all technical details. Besides this, I always felt great enthousiasm about the topic and process from Kees, which gave me extra motivation after every meeting we had. Martijn approached this thesis with the same enthousiasm, and has often put me on the right track with new ideas and discussions about technical developments. In every discussion he reserved the time to analyze problems that I was facing, and elaborate on possible solutions. Having two supervisors with different backgrounds and proficiencies resulted in many different perspectives in discussions we had, making it possible and enjoyable to address this gap in existing research. Finally, I would like to thank Stefan van der Spek for being involved as the co-reader of this thesis. During feedback moments he provided me with novel insights and useful suggestions for the remainder of the process.

Then I would like to thank all my fellow Geomatics students. It has been a great journey together, and I will cherish the memories we made on the travels abroad, when drinking at the Bouwpub, and even those when working in the famous Geolab. It has definitely been the best period in my student life, and I sincerely thank you all for this.

During the course of this thesis, I started a part-time job to gain experience in the professional field. In the last months I have learned a lot of additional skills, some of which turned out to be very useful for this thesis. Therefore, I would like to express my thanks to my colleagues at Krinkels. For providing me this opportunity of putting my academic knowledge in practice, but also for giving me all the space I needed to finish my studies. I am very grateful that we can continue this cooperation on a full-time basis.

As it has been the case for many things throughout my life, I owe a huge thanks to my parents and sisters. Whether it were moments where they tried to sparkle my imagination with discussions about the topic, or times where they reminded me to go out of the house and enjoy other activities, it all helped me to reach this point. To have my family by my side at any time has been a true blessing.

Last, but the opposite of least, I deeply thank Farah for her unconditional love, inspiration, and support. You were always there to encourage me at the moments that I needed it, but were also at any time willing and interested to think along about the content of this thesis. And due to your graphical skills, this report has a wonderful cover. You have made my life a lot easier, when writing the thesis, and at all other times. Thank you for everything.

Erik van der Wal, January 2020

CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Algorithms

ACRONYMS

3D BAG	3D Basisregistratie Adressen en Gebouwen	41
AHN	Actueel Hoogtebestand Nederland	41
BAG	Basisregistratie Adressen en Gebouwen	41
DTM	Digital Terrain Model	41
IDW	Inverse Distance Weighted	24
KIM	Netherlands Institute for Transport Policy Analysis	49
KNMI	Royal Dutch Meteorological Institute	5
SVF	Sky View Factor	13

1 INTRODUCTION

1.1 MOTIVATION

Over the last decades, active transportation modes have rapidly gained interest within the transport sector [Böcker, 2014]. Most dominantly with regard to climate change and air pollution within cities, interest in sustainable modes of transportation for regular use has taken a rise. Utilitarian cycling is being seen as a frontrunner for replacing everyday motorized travels within a city [Amiri and Sadeghpour, 2015], fueled by the fact that around 40% of all trips made within cities are shorter than 2.5 kilometres, whereas 50% of all car trips do not reach more than 5 kilometres [Mertens et al., 2017]. However, since the last decade, the use of bicycles can be stretched beyond short-distance trips. The recognized potential of bicycle transportation is emphasized by the rapid development of electric bicycles, increasing the speed and range of cycling trips, and decreasing the effort for cyclists [Gojanovic et al., 2011; Lin et al., 2008]. With the emergence of electric bicycles, the suitability of the bicycle as a means of transport is not limited anymore to short-distance trips, but enables inter-city transportation in more densely built environments. Consequently, governments operating on different scales are trying to use the increasing opportunities involving bicycle transportation to reduce car traffic and the related air pollution, by promoting the use of (electric) bicycles [Strauss et al., 2015]. Apart from the transport sector and governments, the expanded range of cycling possibilities has also drawn great interest from the academic world. Although cycling has been a widely researched phenomenon since many years, the increasing awareness of the need to switch to more environmental-friendly transportation modes, and the development of electric bicycles, resulted in travel behavior of cyclists being researched more intensively and in relation to a wider scale of influencing factors [Heinen et al., 2010], including the changing climate. Whereas cycling is seen as a means to counter air pollution in cities and global warming in general, changing weather conditions due to climate change could potentially affect cyclist travel behavior [Böcker et al., 2013; Böcker, 2014; Helbich et al., 2014].

Existing literature on transportation mode choices shows a significant impact of weather conditions on the decision whether to travel by bicycle or not [Amiri and Sadeghpour, 2015; Böcker, 2014; Flynn et al., 2012; Heinen et al., 2010; Sears et al., 2013; Zhao et al., 2018]. Among others, Heinen et al. [2010] state that favorable weather conditions positively correlate with the frequency that people travel by bicycle. Yet, the negative influence of unfavorable weather conditions is believed to be more significant, resulting in a lower likelihood to choose the bicycle as means of transportation [Amiri and Sadeghpour, 2015; Spencer et al., 2013; Zhao et al., 2018]. Only a few studies have focused on the impact of weather conditions on actual route choice by cyclists. Cyclists tend to minimize trip lengths when conducted under adverse weather conditions, while favored conditions motivate cyclists to make longer trips [Bergström and Magnusson, 2003; Böcker and Thorsson, 2014; Liu et al., 2017]. Although this indicates that cyclists try to adjust the travel distance or time according to the weather conditions, it does not clarify whether cyclists change their route in such circumstances, and for what reasons. This thesis attempts to identify potential determinants for cyclist route choice in different weather conditions by departing from findings outside the field of cyclist travel behavior.

Whereas cyclist route choice in different weather conditions has remained rather unexposed in existing literature, studies regarding the mobility of pedestrians have shown that the experience of weather does not stand alone, but is heavily influenced by the design of the built environment [Helbich et al., 2014; Lenzholzer and van der Wulp, 2010; Nikolopoulou and Lykoudis, 2007]. As people desire different levels of weather exposure during different types of weather conditions, the degree of shelter that is provided by the surrounding built environment came forward as factor that mitigates the influence of weather conditions on pedestrian mobility. These findings can be related to researches on street climate design [Oke, 1988; Theurer, 1999]. These studies suggest that the experience of weather conditions is directly affected by the degree of shelter that is provided in an urban area.

Although the findings from other research disciplines are not directly applicable on cyclist route choice, they can form a base for the expansion of knowledge about preferences of cyclists. By analyzing observed cyclist travel data and relating these to the weather conditions and urban design at the time of cycling, this thesis built on the findings from the fields of pedestrian mobility and street climate design, and attempted to examine whether the degree of shelter that is provided by the built environment can be considered as a mitigating factor of weather conditions while cycling. Understanding of the effect of urban design on cyclist route choice in different weather conditions may lead to policy designs that incorporate the degree of shelter that is provided on a route to stimulate utilitarian cycling.

1.2 PROBLEM STATEMENT

As has been found in existing literature on cyclist travel behavior, weather conditions have a significant influence on the frequency and duration of cycling trips. Insights in ways to mitigate unchangeable external circumstances like weather conditions could be another step forward in stimulating utilitarian use of bicycles in the search for transportation modes that can replace motorized trips. To identify the suitability of shelter provided by the built environment as a mitigating factor, knowledge on how cyclists behave in different weather conditions once the decision has been made to travel by bicycle is required, as well as on how the observed behavior relates to the design of the built environment. However, the current literature is not sufficient to draw conclusions on this aspect.

As a response to this problem, this thesis attempted to partially fill this gap by relating the degree of shelter provided by the built environment to observed travel data in different weather conditions. A theory-driven methodology was developed and employed to reveal cyclist route choice preferences based on observed travel data. A central aspect of the developed methodology was a method to quantify and spatially model the design of the built environment in such a way that it described the degree of shelter that is provided on a certain location in the study area. Furthermore, the weather conditions under which an observed trip was conducted were decomposed into a set of meteorological factors. For each of those factors, the independent influence on cyclist route choice was modelled, as well as the extent to which the degree of shelter provided by the built environment is an explanatory factor in this.

1.3 CASE STUDY

Within the Netherlands, cycling is historically a widely used transportation mode within urban areas. However, with the emergence of the electric bicicyless, intercity transportation can more easily be conducted by bicycle as well. The Province of Noord-Brabant has recognized the reinforced potential of cycling, and has therefore established policies to stimulate the use of bicycles and electric bicycles as a substitute for less sustainable transportation modes [Provincie Noord-Brabant, 2009]. One of the main actions specified in this document is the realization of a network of 'fast bike lanes' through and between the larger cities in the province by 2030, as is shown in Figure 1.1. This network should enable fast traveling between these cities, potentially providing the opportunity to use bicycles and electrical bicycles for longer distance travelling. As these will be significant investments for the Province, the bicycle lanes should be designed in such a way that they meet the demands of potential users to ensure sufficient usage. Better understanding on determinants of route choice in different weather conditions is therefore highly relevant for the Province, with the Netherlands being a country where weather circumstances can change very rapidly.

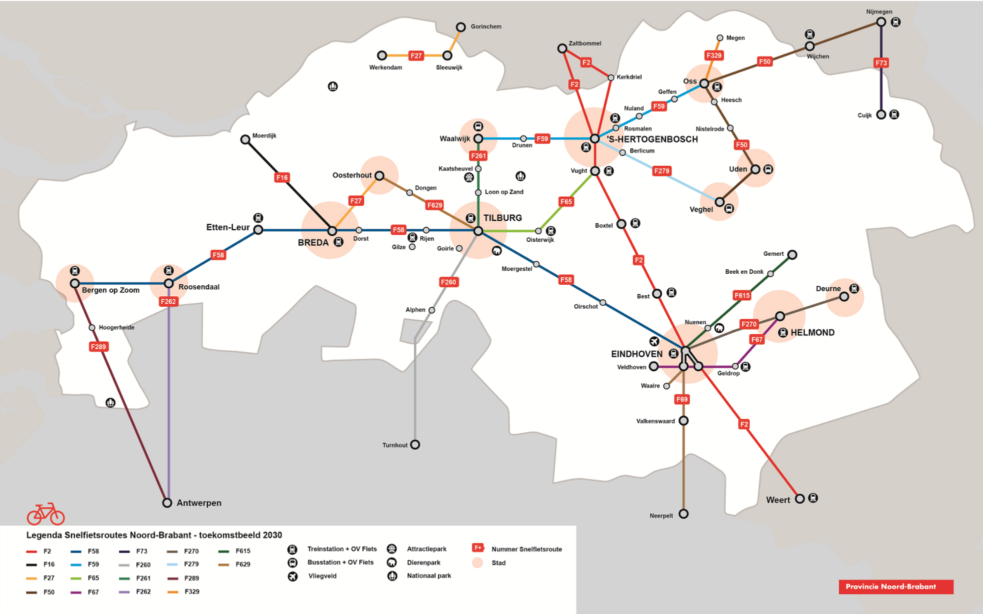


Figure 1.1: Network of fast bike lanes in the province of Noord-Brabant as desired by 2030 [Source: Provincie Noord-Brabant [2019]]

Apart from the real-life 'problem' that exists around the development of the fast bike lanes, the selection of Noord-Brabant as a case study is also fueled by the availability of observed travel data, collected by cyclists that take part in the B-Riders project. B-Riders is a combined initiative from the Province of Noord-Brabant and the Ministry of Infrastructure and Water Management, with the aim of enhancing accessibility of cities in Noord-Brabant through stimulation of sustainable travel behavior [B-Riders, 2016]. Participants of this project are people that commute to work primarily by electric bicycle, or otherwise conventional bicycle, in the province of Noord-Brabant. Therefore, the observed travel data most dominantly comprise utilitarian trips. Through registration of made trips, participants receive financial or related rewards. Trip and user information is registered together with GPS measurements showing the spatial aspect of a trip, which together allows to model a route for a specific trip. As the provided dataset comprises relatively large numbers of travel data, it functioned as a sufficient fundament for this thesis.

1.4 RESEARCH OBJECTIVE, QUESTIONS AND SCOPE

1.4.1 Research objective & research questions

The main objective of this thesis was to examine whether the degree of shelter along a route can be considered a factor that mitigates the influence of weather conditions on cyclist route choice, with the province of Noord-Brabant as case study. Based on the specified research objective, the main research question for this thesis was defined as follows:

To what extent does the degree of shelter provided by the built environment explain cyclist route choice in different weather conditions?

To be able to answer the main research question, a set of sub-questions has been established. The sub-questions are categorized according to the aspect of the research they are relevant for:

1. Development of the theoretical framework
    - *What meteorological factors influence cyclist travel behavior according to existing literature?*
    - *How does urban design affect people's experience of weather?*
2. Operationalization
    - *How can the observed travel data be examined in terms of route choice?*
    - *How can the different meteorological factors be quantified and spatially modelled?*
    - *How can the degree of shelter provided by the built environment be quantified and spatially modelled?*
3. Modelling
    - *To what extent do the different meteorological factors influence cyclist route choice?*
    - *To what extent does the degree of shelter provided by the built environment influence cyclist route choice?*
    - *To what extent are the weather conditions a reason for cyclists to adapt their route choice to the degree of shelter offered by the built environment?*
4. Validation
    - *How can the results generated by implementation of the developed methodology be validated?*

1.4.2 Scope of the research

In order to shape this thesis to meet the defined research objectives, and accomplish this within the desired time span, the following points indicate the scope of the research:

- Since the province of Noord-Brabant provided a clear case study for the problem that this thesis attempted to tackle, the geographical scope limited itself to the extent of the province. For the sake of optimization and testing of the methodology, a smaller geographical extent within the province was used for the development. The test area gave a proper representation of the entire province, as it comprised sufficient observed travel data and variation in the design of the built environment. In the end, the developed methodology can be applied on any location with sufficient availability of observed travel data, built environment data, and weather data.

- Regarding the type of cycling, this thesis focused on utilitarian travelling by bicycle. Any observed travel data potentially classified as leisure trips were disregarded.
- The observed travel data provided for this thesis holds data for the first nine months of the year 2014. Therefore, the temporal scope of this thesis is bounded by the start of the year 2014 and the end of September of that year.
- With regard to factors describing the built environment, this thesis focused solely on quantifying and spatially modelling the degree of shelter provided by the built environment. Operationalization of other spatial factors than shelter was disregarded from this thesis.
- The selection of weather factors that were included in this thesis was formed by findings from the existing literature framework on the influence of weather factors on cyclist travel behavior, and the open data that is provided by the Royal Dutch Meteorological Institute (KNMI).
- A set of other theoretically underpinned influencing factors regarding cyclist route choice was used as control variables to put the results into the appropriate context. These factors included individual and infrastructural indicators, as well as measures of slope.

1.5 THESIS OUTLINE

This thesis report is structured as follows:

***Chapter* 2** provides an overview of the existing literature related to this thesis topic. The studies discussed in this chapter were used to establish the theoretical framework in which the methodology is build.

***Chapter* 3** presents the methodology that was developed for this thesis. In this chapter, a conceptual take on the different aspects of the methodology is provided.

***Chapter* 4** elaborates on the implementation of the methodology described in Chapter 3. This chapter describes the practical execution of the developed methods.

***Chapter* 5** presents the results of the statistical testing of the assumed relationships between cyclist route choice, weather conditions, and the degree of shelter offered by the built environment. An overview is provided of the outcome of a set of statistical analyses which formed the fundament to answer the main research question.

***Chapter* 6** provides conclusions based on the results presented in Chapter 5. In this chapter, the main research question of this thesis is answered, as well as the set of sub-questions. Finally, this chapter includes a discussion on the developed methodology and recommendations for future work.

2 RELATED WORK

This thesis built on the existing literature framework on cyclist travel behavior, the experience of weather conditions, and modelling the potential shelter offered by the built environment. Researches regarding each of those subjects are reviewed in this chapter. At first, Section 2.1 provides a general overview on what factors affect cyclist travel behavior, with the emphasis placed on how cyclists are influenced by weather conditions. Section 2.2 will elaborate on how the experience of weather is affected by the design of the built environment, followed by a recap of existing methods to quantify and model urban design in Section 2.3. Finally, conclusions and a brief discussion about the research gap in the existing literature is presented in Section 2.4, together with subsequent hypotheses that will be tested in this thesis.

2.1 DETERMINING CYCLIST TRAVEL BEHAVIOR

2.1.1 Influencing factors transportation mode choices

Within the field of cyclist travel behavior, two main study topics can generally be identified: transportation mode choices and route choice studies. The first category covers research on the probability that the bicycle will be selected over other modes of transportation, and what factors influence this decision. Studies that fall within the second category add a spatial dimension to the behavior of cyclists, and research where cyclists ride and how the choice of route is determined. Although a chosen route initially follows from the decision to cycle or not, previous studies have found that the determinants for deciding upon transportation mode or route can differ.

Over the last decades, many studies have investigated factors that potentially influence the decision to travel by bicycle [Amiri and Sadeghpour, 2015; Handy and Xing, 2011; Heinen et al., 2010; Moudon et al., 2005]. Whereas a wide range of determinants has been found for choosing the bicycle as transportation mode, a large share agrees upon travel distance as one of the most significant influential factors [Amiri and Sadeghpour, 2015; Heinen et al., 2010, 2011; Winters et al., 2010]. In their overview of the existing literature on bicycle commuting, Heinen et al. [2010] state that an increase in trip distance results in a decrease in the frequency of cycling for commuting, as a form of utilitarian cycling. Especially utilitarian cyclists, using the bicycle merely as a means of transportation rather than for leisure activities [Yeboah and Alvanides, 2015], regard travel distance as the main decisive factor for transportation mode choices [Heinen et al., 2011]. The increased effort and travel time, as a consequence of travelling a larger distance, is often found to be the main reason for choosing other transportation modes over the bicycle. As the travel time and effort related to trip distance take a central place in deciding upon travelling by bicycle, a broad set of studies has identified factors that are not directly related to the trip distance, but to the time of travelling and the effort it takes [Heinen et al., 2010; Moudon et al., 2005; Stinson and Bhat, 2003]. When approaching this from a natural point of view, the factor of slope has a negative influence on bicycle use, as it increases the effort of cycling [Rietveld and Daniel, 2004]. This is confirmed by low cycling shares in cities that are characterized by a hilly surface [Heinen et al., 2010]. In terms of urban environment, a denser road network and a mixture of functions in neighborhoods affect the odds of travelling by bicycle, as it usually results in smaller distances that have to be travelled to desired destinations [Moudon et al., 2005; Stinson and Bhat, 2003]. Apart from factors that are related to urban design, the infrastructural characteristics of bicycle networks matter when people are deciding to travel by bicycle or not. Heinen et al. [2010] confirm that people get discouraged to opt for the bicycle as a transportation mode when the bicycle network that can be used mainly consists of non-separated lanes for cyclists. In this regard, Heinen et al. [2010] make a distinction between bicycle paths, bicycle lanes and roads shared with other forms of traffic, where a clear preference is found for the fully separated bicycle paths. The predilection for separated bicycle paths is mainly caused by the feeling of safety, as it ensures that cyclists are not involved with motorized traffic [Klobucar and Fricker, 2007]. Although in the same research Klobucar and Fricker [2007] state that influence of separated bicycle paths on the objectively measured safety (number of incidents with cyclists) remains unclear, they argue that the subjective safety experienced by cyclists are found to be more decisive with respect to transportation mode choices. This argument is confirmed by Winters et al. [2011], as ”the risk from motorists who don’t know how to drive safely near bicycles” comes forward as one of the main deterring influences on cycling. However, Heinen et al. [2010] have found that the perception of safety varies between men and women. They state that women assign a higher value to safety than men. With regard to safety, the surveyed crowd of cyclists in the research of Winters et al. [2011] have indicated the absence of proper lighting in a dark environment as a major determinant for not using the bicycle as a means of transportation. Besides travelling in poor daylight conditions, well-lit roads are mostly desired when ’bad’ weather conditions are potentially affecting the visibility of cyclists.

In general, weather conditions have been found as a main influence on transportation mode choices [Amiri and Sadeghpour, 2015; Flynn et al., 2012; Heinen et al., 2010; Sears et al., 2013; Spencer et al., 2013; Winters et al., 2011]. Several studies show fluctuations in cycling shares between different season, favoring the summer because of its warmer and dryer days [Heinen et al., 2010]. However, declines in cycling shares throughout the winter vary per region, as locations with milder winters show relatively lower decreases in winter cycling than colder regions [Stinson and Bhat, 2004]. While approaching weather conditions in terms of seasons and climate can be seen as an aggregate of different weather characteristics, existing studies have also researched the influence of individual meteorological factors on cycling. Out of all researched meteorological factors, temperature has been found as an important influencing factor on cyclists, evaluated from different perspectives. Generally, temperature and cycling are related in a positive way. Singled out from other weather factors, temperature is found to be a decisive factor in cases of doubt on whether to cycle or not [Flynn et al., 2012; Sears et al., 2013]. According to those studies, an increase in temperature will raise the odds of using the bicycle as a means of transport. Whereas these findings are supported by other studies, the positive relationship between temperature and the choice to cycle is bound by a maximum temperature, after which the relationship becomes negative [Amiri and Sadeghpour, 2015; Spencer et al., 2013]. This indicates that besides cold, also hot weather conditions negatively affect the experience of cycling. Böcker et al. [2016] confirm this statement in their research on the emotional travel experiences that follow from the relationship between weather and transportation mode choices. They have found that temperatures above 25 degrees Celsius no longer have a positive effect on happiness during a travel, while it increases feelings of tiredness and irritation. In combination with humidity, the aversion of very warm conditions is strengthened even more [Winters et al., 2011].

A second meteorological factor that influences transportation mode choices is precipitation [Flynn et al., 2012; Sears et al., 2013; Winters et al., 2011], however with a magnitude depending on the gradation and type of precipitation [Spencer et al., 2013]. Where light rain does not prevent people from travelling by bicycle, heavy rain and snow result in changes in transportation mode [Spencer et al., 2013;

Zhao et al., 2018]. Especially ice and snow are a main deterrent for travelling by bicycle, since the feeling of safety and comfort will be affected [Winters et al., 2011]. Even tough the experience of temperature and precipitation can be heavily affected by factors describing wind, Spencer et al. [2013] states it is not considered a main determinant for deciding whether to cycle or not. This is in line with the literature review of Heinen et al. [2010], where they confirm that the effect of wind remains unclear. With regard to the significance of minimizing travel time and effort, this is relatively surprising as wind has a clear influence on the effort and resulting travel time while cycling.

Although not a conventional weather factor, environmental darkness has been recognized as an influencing condition for cyclist travel behavior by a set of studies. Spencer et al. [2013] state that lighting conditions have been a determinant for the decision to cycle or not in Vermont, USA, and therefore contributes to a significant decrease of cycling trips during the winter months. In relation to this they confirm that urban areas with a high degree of artificial lighting are more desirable for cyclists, mainly because of better visibility and consequently safer travel conditions. Sears et al. [2013] support the findings by Spencer et al. by stating that a lack of daylight in the morning or evening has frequently been given as a reason for not commuting to work by bicycle. However, the actual minutes or hours of daylight during a day have not been found decisive [Flynn et al., 2012; Sears et al., 2013].

In general, utilitarian cyclists are found to be less sensitive to weather conditions than non-utilitarian cyclists regarding transportation mode choices [Bergström and Magnusson, 2003; Böcker and Thorsson, 2014; Heinen et al., 2010; Liu et al., 2015]. This difference in effect could be caused by high dependency on travelling by bicycle, as utilitarian cyclists might have little choice but to cycle [Heinen et al., 2010].

2.1.2 Influencing factors cyclist route choice

Once the decision to cycle has been made, a route from the current to the desired location needs to be chosen. Similar to transportation mode choices, the length and duration are the main determinants when deciding upon a route for utilitarian cyclists [Broach et al., 2012; Dill and Gliebe, 2008; Heinen et al., 2011; Winters et al., 2011]. Initially, utilitarian cyclists seek for the shortest route to minimize the effort, or the fastest route to minimize the travel time. Where the travel distance is a predefined characteristic of a route, the travel time might be influenced by other factors than can be encountered on a route. A natural characteristic that has been found as an influence on cyclist route choice is slope [Rietveld and Daniel, 2004; Stinson and Bhat, 2003]. Where Section 2.1.1 mentions hilly terrain as a main reason not to cycle, Stinson and Bhat [2003] state that cyclists try to avoid routes with a significant amount of slope. The importance of slope in cyclist route choice is emphasized by Krenn et al. [2014], in their comparison between observed and shortest cycling routes. They have found that cyclists make detours to avoid hilly terrain. It should be noted that both studies mentioned were conducted using mountainous locations as study area.

Based on revealed preference GPS data, Broach et al. [2012] have identified a set of influential infrastructural characteristics related to the potential duration of a route. They have found that cyclists select routes with a minimal level of turn frequencies, traffic lights, and crossings. However, the number of crossings on a route have also been found of influence with regard to the safety of a route, since a high number of crossings increases the interaction with other traffic [Krenn et al., 2014; Stinson and Bhat, 2003]. In general, cyclists try to avoid the interaction with other (motorized) traffic when selecting a route. As it is a main determinant for deciding whether to cycle or not, also route choice is heavily influenced by the presence of separated bicycle paths [Krenn et al., 2014; Winters et al., 2011; Yeboah and Alvanides, 2015]. Krenn et al. [2014] show that cyclists typically opt for routes with a higher share of separated bicycle paths and bicycle lanes over the shortest route. In attempts to avoid interaction with other traffic, Krenn et al. [2014] have also found that cyclists select routes that minimize the share of main roads without separated bicycle paths or lanes, as these roads usually contain larger volumes of traffic. According to Broach et al. [2012], especially commuters are sensitive to traffic volume because of a tighter time schedule than non-commuters.

Whereas the existing literature framework on cyclist route choice state clear preferences with regard to the form of the natural, urban, and infrastructural environment, little is known about where cyclists ride in different weather conditions. However, Böcker and Thorsson [2014] have confirmed that weather conditions lead to minimization or expansion of cycling in terms of frequency and duration. These effects were found to be stronger for leisure cycling than for utilitarian cycling. Although this confirmation merely states that (primarily leisure) cyclists adjust the travel time and frequency based on weather conditions, it provides a base for indications about how people experience cycling in certain circumstances. In this perspective, generalized weather conditions can again be decomposed into the individual weather factors that have been mentioned in Section 2.1.1.

Besides the influence on transportation mode choices, Böcker and Thorsson [2014] conclude that temperature not only affects cycling frequencies, but also the duration of cycling trips. They state that thermal conditions, decomposed into maximum daily air temperature, mean radiant temperature, and physiological equivalent temperature, have a bell-shaped effect on cycling durations, with its optimum around a maximum daily air temperature of 24 degrees Celsius. In the same study, Böcker and Thorsson claim that the factor wind affects both cycling frequency and duration according to a negative relationship. However, the duration of cycling trips is more significantly affected by wind than cycling frequencies. This indicates that people still conduct trips by bicycle on days with more severe wind, but that the length of those trips is minimized. The idea that wind is a factor that influences the behavior of cyclists rather than the choice for transportation mode is supported by other work. Significantly large wind speeds are stated to have a negative effect on emotional travel experiences [Böcker et al., 2016; Helbich et al., 2014], and the way cyclists value their environment [Böcker et al., 2015]. In a similar fashion, the factor precipitation plays an important role in cyclist route choice. Both Helbich et al. [2014] and Böcker et al. [2016] have found that precipitation have a negative effect on the way cyclists experience a trip, as well as how the en-route environment is valued [Böcker et al., 2015]. Furthermore, Böcker and Thorsson [2014] note that precipitation in the form of heavy rain and snow is not only negatively related to the choice to travel by bicycle, but also to the duration of cycling trips. Finally, the effect of daylight conditions on the duration of cycling trips has remained underexposed so far. However, while travelling the absence of light has a negative influence on the value that cyclists give to their environment [Böcker et al., 2015]. In their research on en-route weather and place valuation, Böcker et al. confirm that cyclists assign less value to route surroundings most dominantly because of darkness and subsequent lack of visibility.

In order to assess actual route choice of cyclists under different weather conditions, the geographical context of a route has to be taken into account. Helbich et al. [2014] have touched this aspect by distinguishing variations in cyclist travel behavior in different weather conditions, controlled for the spatial configuration in which the surveyed cyclists usually ride. They have found that the effect of weather on the decision to cycle differs across the extent of the greater Rotterdam area, the case study of their research. Even though Helbich et al. [2014] do not discuss the route choice of cyclists, their findings may be indicative for the relevance of the design of the surrounding environment when cyclists decide upon a route under certain weather conditions.

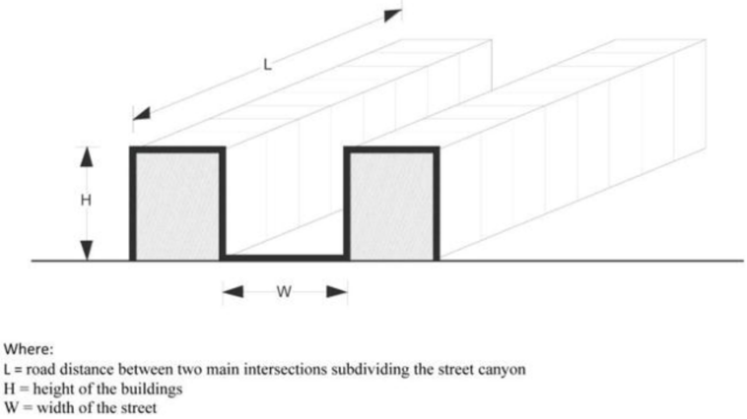
2.2 URBAN DESIGN AND EXPERIENCE OF WEATHER

The framework of the studies mentioned in Section 2.1 confirms the significant influence of weather factors on cyclist travel behavior. However, the experience of different weather conditions prior to, and during a cycle trip does not stand alone, but is influenced by other factors. As Section 2.1.2 mentions, Helbich et al. [2014] have taken this aspect into account by approaching weather conditions in a geographical context, where urban design plays an important role in the experience of weather by cyclists on a certain location. Based on daily travel surveys, Helbich et al. concluded that cyclists seemed to be more heavily affected in their travel behavior by precipitation and wind in remote areas with relatively more open (weather-exposed) areas, compared to cyclists in central areas. Furthermore, they found that openness of space also affects thermal experiences, with differences between day- and nighttime. Where other researches draw conflicting conclusions regarding the influence of daylight on travel behavior [Flynn et al., 2012; Spencer et al., 2013], Helbich et al. have found differences in behavior during day- and nighttime between open and more densely built-up areas.

A few studies in the field of pedestrian mobility patterns have related spatially differentiated weather effects to the design of a built environment. Nikolopoulou and Lykoudis [2007] have investigated the attendance at a seashore boulevard and a central square in greater Athens during different weather conditions, seasons, and during different times of the day. Their findings indicate that people prefer to use less sheltered spaces as the boulevard on sunlit days during colder seasons to get a better exposure to the sun. However, in warmer days people prefer to attend more sheltered areas as the central square, and come to open spaces after sunset. Based on those findings, Nikolopoulou and Lykoudis state that air temperature and solar radiation have the largest influence on the attendance of open or sheltered space. However, they also state that people seek more opportunities to be in the sun when wind speeds increase, although the overall attendance on both the boulevard and the central square drop with higher wind speeds. Finally, the findings of Nikolopoulou and Lykoudis indicate that influence of weather conditions appears to be stronger on the highly exposed seashore boulevard than the sheltered central square. Whereas the results of Nikolopoulou and Lykoudis are based on the warmer Mediterranean climate in Greece, Lenzholzer and van der Wulp [2010] have conducted a similar research on a location with a climate more closely related to the study area of this thesis. Based on interviews with people attending three squares in Den Haag, Eindhoven, and Groningen respectively, they state that the perception of thermal comfort is mainly influenced by wind parameters, and the thermal (dis-) comfort these create. According to their findings this relates to the spatial openness of a square, where spaces that are too open are often experienced as uncomfortable, since the shelter from wind is relatively low.

The relevance of spatial openness in relation to the exposure to weather is further emphasized by a set of studies on street climate design. In an attempt to optimize shelter from wind, dispersion of pollutants, urban warmth, and solar access, Oke [1988] approaches the built environment as a set of urban canyons formed by streets and the surrounding buildings and other built objects. Figure 2.1 shows that an urban canyon is defined by three parameters: the length and width of a street, and the height of the buildings enclosing the canyon. Although the length of the street is a taken into account when describing an urban canyon, the ratio between the height and width of a canyon is the main descriptive parameter in relation to the weather exposure within an urban canyon [Oke, 1988]. An increase in this ratio generally results in a higher degree of shelter from wind, while it limits the solar access to a canyon, as well as the dispersion of pollutants. The findings of Oke [1988] are confirmed by Anselm Akubue [2019] and Shishegar [2013]. Both studies conclude that the flow of wind and solar access are increased in wider canyons, but simultaneously this results in a lower degree of shelter for pedestrians. Using similar parameters to define an urban canyon, Theurer [1999] stresses the negative correlation between the height/width ratio and the flow of wind, and therefore dispersion of pollutants. Although Theurer did find variations in pollutant dispersion for different canyon lengths, he concluded that the height/width ratio is more relevant when designing street environments.

The findings in the field of street climate design are in line with the studies of Nikolopoulou and Lykoudis [2007] and Lenzholzer and van der Wulp [2010], where the associations with weather effects have been found significantly smaller in more sheltered areas. A limitation of the discussed literature framework is that it almost exclusively focusses on the effect of urban design on thermal experiences, and the flow of wind that influence the feeling of thermal comfort. Therefore, the relationship between the shelter provided in an urban design and other weather factors remains unclear.



Where:
L = road distance between two main intersections subdividing the street canyon
H = height of the buildings
W = width of the street

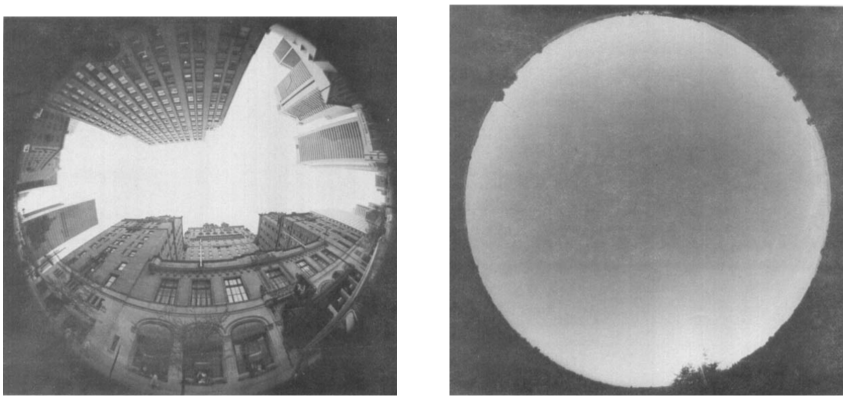
Figure 2.1: Cross section of an urban canyon [Source: Anselm Akubue [2019]]

Whereas the studies discussed in this section take a one-sided stand in approaching urban design as an explanatory factor for experience of weather by only including buildings and built-up objects, the built environment generally also holds natural factors as trees and vegetation. Heisler [1990] has recognized the significance of tree density as an influencing factor on wind speeds. According to the findings of Heisler, wind speeds are heavily reduced by high-density tree arrangements. As a complementary factor to wind speed reductions by buildings, even trees in more scattered arrangements additionally reduce the wind speed.

Besides the effects on wind propagation, Gillner et al. [2015] found that trees play a major role in mitigating the thermal conditions below the canopy layer. Presence of trees results in cooling effects on a street level, with a positive correlation between tree density and the magnitude of the cooling effect. Specifically, the findings of Gillner et al. indicate that the shading provided by the tree canopy, and the capability to absorb and transpire humidity from the air gives trees the ability to mitigate thermal conditions on a street-level.

2.3 SHELTER BY THE BUILT ENVIRONMENT

With potential shelter provided by the built environment, mainly expressed in terms of spatial openness, being recognized as a factor that influences the experience of weather conditions, this section focusses on existing methods for defining and quantifying this built environment characteristic. Spatial openness has been approached in various ways throughout existing literature, from the perspective of multiple different disciplines. In a research centered around design of the built environment, Oke [1981] introduced the Sky View Factor (SVF) as a measure of openness: the amount of sky that is visible from a certain point in the middle of an urban canyon. To compute the SVF for a certain location, the ratio of the canyon height and the street width at the location of measurement is used. Figure 2.2 depicts how Oke visualized the principle of the SVF by picturing the visible area in the middle of an urban canyon as being projected through a so-called fisheye lens. Using this technique, the SVF is then determined by the portion of visible sky in the image. Oke developed the SVF to describe the geometry of urban canyons for simulations of nocturnal cooling rates for urban and rural built environments. However, Oke used this measurement technique of urban canyon geometry in later research on street climate design [Oke, 1988], as has been discussed in Section 2.2.



(a)

(b)

Figure 2.2: Fisheye lens representation of Sky View Factor for an (a) urban and (b) rural environment [Source: Oke [1981]]

Visibility is also the central term in the Isovist method, developed by Benedikt [1979]. The Isovist is the set of all visible points from a certain point in space and in relation to the surrounding environment, as shown in Figure 2.3. In specific terms, the Isovist is formed through a set of line segments that are casted from a vantage point and will intersect with boundary surfaces, and quantified by accounting for the coordinates of the vantage points, the coordinates of the boundary points, and the direction of the radial. Using this method, multiple visibility indicators can be determined: the area of the visible space and the perimeter of the environmental surface that can be seen from the vantage point. As the Isovist is dependent on the location of the vantage point in relation to surfaces that form the surrounding environment, movement of the vantage point within space will lead to a different Isovist. Accumulative understanding of the spatial configuration can be derived from a set of Isovists that belong to a path of vantage points [Benedikt, 1979]. This allows to model the variation in spatial openness along a path, for example a chosen route by a cyclist. Van Rijn [2009] has validated the Isovist method as a measure of openness by comparing it with measured landscape openness. Through comparison with results from a field visit, Van Rijn concluded that the Isovist method gives comparable values as landscape openness measured in the field.

In an approach to model the aesthetic effect of the openness of a built environment on cyclist route choice, Anastasiadou et al. [2018] (a research in which the author of this thesis participated) have included facets of the Isovist method by Benedikt [1979] and the urban canyon method by Oke [1988]. The aim of the method is to identify buildings that interfere with the field of view from a given vantage point, to assess the visual openness at that point. In general, the methods consists of two parts: creating a field of view around the vantage point and identifying which buildings intersect with that field (derived from [Benedikt, 1979]), and evaluating the gradation of interference by accounting for the height of an intersecting building and the distance to the vantage point (derived from [Oke, 1988]).

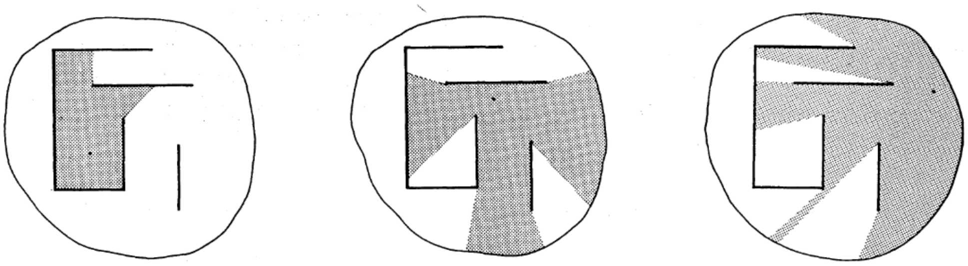
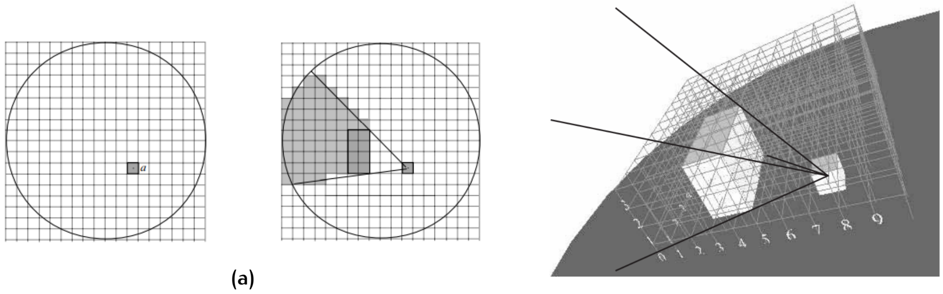


Figure 2.3: Isovists for three vantage points in the same spatial configuration [Source: Benedikt [1979]]

Eventhough both the approaches of Benedikt and Oke are proven methods by themselves, Anastasiadou et al. aimed to integrate the simplified way of computing the openness by Oke in the more detailed method to model the spatial configuration around a vantage point by Benedikt. Oke simplified the built environment by resembling the urban geometry in a set of urban canyons, and measure the openness of the design by the ratio of the height and width of such a canyon. However, this generalization leads to a loss of detail when describing the shape of the urban geometry. By integrating the simplified way of measuring openness in an approximation of the method by Benedikt, Anastasiadou et al. increased the level of detail in describing the urban geometry but maintained the simplified way of measuring the openness around a vantage point.

Fisher-Gewirtzman and Wagner [2003] also approach spatial openness as the volume of open space that can be seen from a given point, however in a more computerized manner. In their Spatial Openness Index (SOI), the world is part of a 3D integer grid and the SOI is determined by taking into account the open space and built volumes. The actual result for the SOI, the spatial openness for a certain location, is given by the number of grid points in the open space that can be seen from a certain location. Figure 2.4a shows a 2D representation of this method, where the left image depicts reduced visibility of space as a building blocks part of the line of sight. Figure 2.4b shows a simplified 3D representation of reduced visbility due to the presence of an obstacle in the direct environment of the vantage point.



(a)

(b)

Figure 2.4: 2D representation of the Spatial Openness Index [Source: Fisher-Gewirtzman and Wagner [2003]]

Following the existing literature framework, openness of the built environment is mainly described in terms of uninterrupted visibility of space with regard to the surrounding (built) environment. The studies of Oke [1988] and Shishegar [2013] show that a visibility analysis to describe openness does not only have an esthetic aspect to it. In their studies they used measurement methods based on the Sky View Factor to assess the potential solar access and wind flow within an urban canyon, as the amount of visible sky gives an indication about the compactness of the built environment. Assessment of the visual and spatial openness of an urban area can therefore be employed to determine the degree of shelter that is provided.

2.4 CONCLUSION

2.4.1 Theoretical fundament derived from literature

The aim of reviewing the literature mentioned in the chapter is the establishment of a theoretical framework to build on in order to develop the methodology for this thesis. In order to form the theoretical framework, two sub-questions have been defined. Based on the reviewed literature in Section 2.1.1 and Section 2.1.2, the first sub-question can be answered:

What meteorological factors influence cyclist travel behavior according to existing literature?

Cyclist travel behavior is mainly influenced by weather conditions with respect to the decision whether to cycle or not. When deciding about a transportation mode, cyclists have found to be most heavily affected by temperature and precipitation. A positive correlation is found between the temperature and the odds to cycle, however bounded by a maximum temperature. On the contrary, precipitation is found to negatively affect the odds of cycling, where the strength of this relationship depends on the gradation and type of precipitation. As a more unconventional meteorological factor, daylight conditions are found as a significant influencing factor with respect to the perceived safety, usually in conjunction with lighting conditions and the overall weather conditions.

The effect of weather conditions on cyclist route choice remains rather underexposed in the existing literature. However, when expressed in cycling durations, cyclist route choice is positively related with the factor temperature. Similar to transportation mode choices, this positive correlation is bound by a maximum temperature. In contrast to transportation mode choices, the wind speed is found as a main influencing factor on cycling durations. Cyclists tend to decrease the duration of a cycling trip when wind speeds are increasing. The effect of precipitation on cycling durations is less pronounced than its effect on transportation mode choices. However, in the form of heavy rain and snowfall, precipitation negatively influences the duration of cycling trips. In general, the effects on cycling durations are weaker for utilitarian cycling, as these type of cyclists are more strongly driven by minimization of travel time and effort.

When combining the findings with regard to transportation mode choices and cycling durations, four main categories of influencing meteorological factors can be distinguished: *temperature*, *wind*, *precipitation*, and *daylight conditions*. These four main parameters can however be decomposed into more detailed descriptors of weather condtions.

To expand the context of cyclist travel behavior in different weather condtions, findings from different study disciplines have been addressed in Sections 2.2 and 2.3. Findings presented in these sections form the base to answer the second sub-question:

How does urban design affect people's experience of weather?

Findings from the field of pedestrian mobility indicate that the experience of weather conditions is influenced by the degree of shelter that the built environment provides, mainly expressed in terms of spatial openness. Attempts to define spatial openness generally come from two different fields: visibility studies and street climate design. The former defines openness from a more aesthetic approach, by considering the magnitude of interruption of the field of view. However, the latter approaches openness as a ratio of height and width of an urban canyon in order to define the exposure to different types of climatological indicators. In the perspective of street climates, shelter by the built environment can be expressed both in terms of building configuration, as well as the density of trees in an urban environment. Both the configuration of buildings and tree density are found to mitigate the experience of temperature and wind when the provided degree of shelter is considered as high. However, the effect that the degree of shelter has on the other two main meteorological factors, precipitation and daylight conditions, has remained underexposed.

2.4.2 Research gap

Whereas the influence of weather conditions on the decision to cycle or not has been reviewed extensively, the choice of route by cyclists under different weather circumstances has remained underexplored so far. Several studies have revealed that cyclists adjust the travel time of a trip according to the weather conditions, but little is known about the spatial context of the chosen routes. Findings from pedestrian mobility studies indicate that the shelter provided by the surrounding built environment plays a significant role in where people travel, and how they experience different weather conditions. Although cyclist route choice behavior differs from that of pedestrians, these findings may form a base for explaining cyclist route choice in different weather conditions.

2.4.3 Hypotheses

As this thesis addressed an underexplored field in existing research, findings from the literature framework presented in this chapter formed the fundament for hypotheses on the effect of shelter by the built environment on cyclist route choice in different weather conditions. Since the weather conditions in this thesis were decomposed into individual weather factors, a set of hypotheses was formulated.

Based on the mild climate characteristics of the Netherlands, general moderate effects of thermal factors as temperature and solar radiation were expected, resulting in limited possibilities to mitigate these effects by the degree of shelter provided by the built environment. However, on hot summer days, cyclists were expected to look for more shelter from the built environment.

Increase of windspeed was expected to result in a higher share of routes through sheltered areas. As an increase in wind affects the effort that cyclists have to make on a trip, attempts to limit this effect of wind were expected to be made when selecting a route. Lower windspeeds were not expected to have a high influence on where cyclists ride, except for very warm days. In those cases, cyclists are expected to travel through more open areas in order to experience the cooling effect of wind. Furthermore, it was expected that cyclists seek for more sheltered routes when the direction of the wind in relation to the general direction of a route is known. The direction of wind in relation to cyclist travel behavior has rarely been discussed in existing literature, but may influence on-the-spot decision on a route while cycling.

A rise in the degree of shelter was expected when the gradation of precipitation increased. Where lighter forms of precipitation will not show a clear preference for open or sheltered routes, heavier precipitation will lead to an increase in the share of sheltered areas on a route. Presence of snow and ice was expected to lead to an increase in routes that go through more sheltered areas, in order to minimize the snow and ice that may form on the road.

When daylight conditions darken, it is expected that cyclists will actually seek the shelter of the built environment, however mostly explained by the presence of artificial lighting in such areas. Since a potential lack of visibility has found to be the main concern of cyclists in poor daylight conditions, a strengthened increase in sheltered routes is expected in combination with fog and precipitation, as these factors influence the visibility of a cyclist.

As this thesis focused on utilitarian cycling, the strength of the hypothesized relationships was not expected to be high since utilitarian cyclists are generally less sensitive to weather influences in the first place. Adjusted for the effects of individual, infrastructural, and environmental characteristics other than shelter, a conceptual view of the influence of weather conditions and shelter by the built environment on cyclist route choice is presented in Figure 2.5.

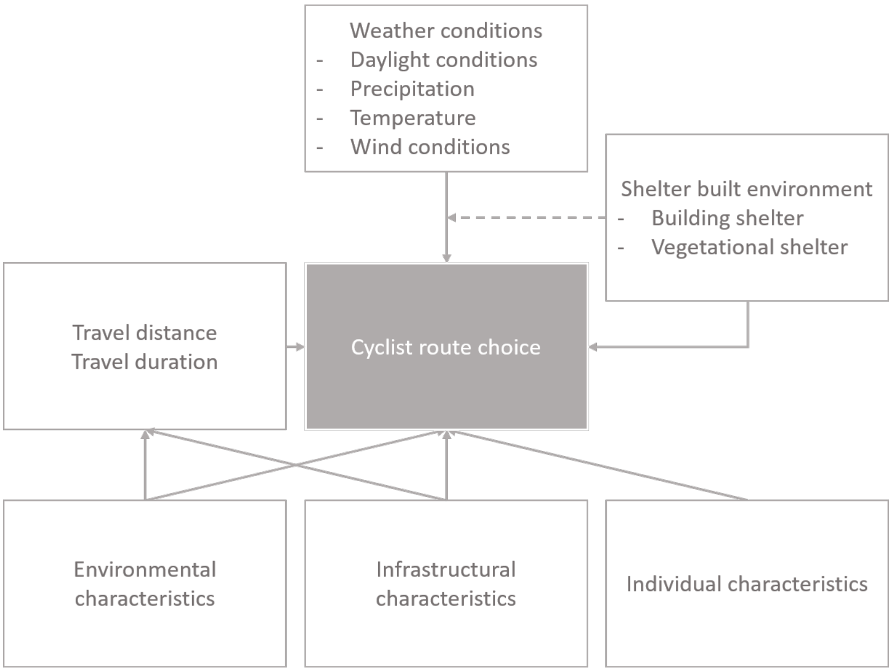


Figure 2.5: Conceptual model of influences on cyclist route choice

3 METHODOLOGY

In this chapter a description is presented of the methodology that has been developed and applied on the case study of this thesis. Figure 3.1 shows a schematic overview of the different aspects of the methodology, which form the base for the outline of this chapter. First of all, in Section 3.1 the definition of cyclist route choice that this thesis departs from will be discussed, as well as the theoretical fundaments that the definition is built on. Secondly, a study area has been used for the development of the methodology, which is presented in Section 3.2. The following three sections will elaborate on the operationalization part of this methodology: the generation of a route model in Section 3.3, establishment of meteorological factors in Section 3.4, and the definition and quantification of shelter provided by the built environment in Section 3.5. Finally, Section 3.6 will explain how the operationalized variables will be modelled with respect to cyclist route choice.

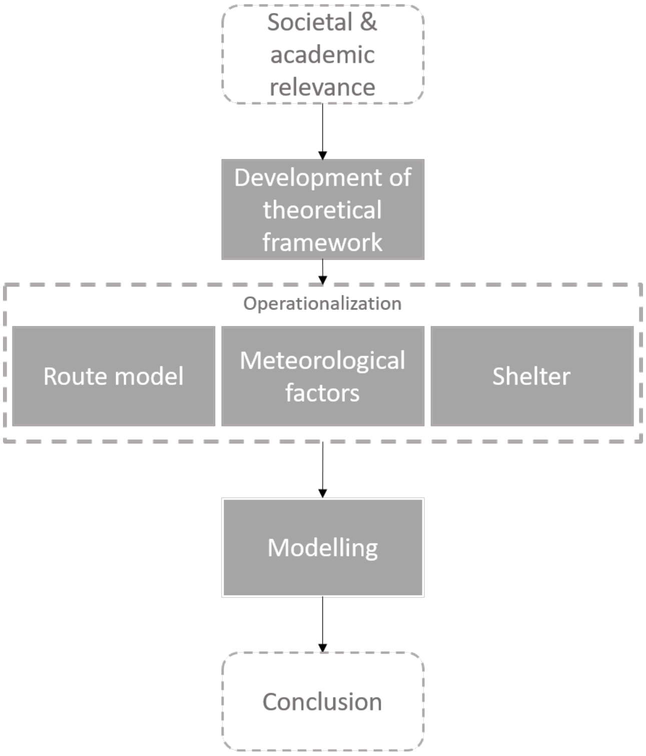


Figure 3.1: Schematic overview of the developed methodology

3.1 DEFINING CYCLIST ROUTE CHOICE

Initiated by both a research gap in the existing literature on cyclist route choice and a real-life case study in the province of Noord-Brabant, the developed methodology finds its fundament in the theoretical framework discussed in Chapter 2. The theoretical fundament not only concerns the design of the methodology, but also provides a base to define and assess cyclist route choice.

In order to draw conclusions on determinants for cyclist route choice, an observed route has to be compared to a theoretically ’optimal’ route. According to the studies on cyclist route choice mentioned in Section 2.1.2, trip distance and duration are the main determinants for utilitarian cyclists when deciding upon a route. For the sake of minimization of effort or travel time, cyclists have been found to initially seek for the shortest or fastest route. As this thesis focused primarily on utilitarian cycling, cyclist route choice was defined as the extent of divergence from the shortest or fastest route. Therefore, this thesis studied the effect of the relationship between weather conditions and the degree of shelter that is provided by the built environment on the extent of divergence from the shortest or fastest route.

3.2 STUDY AREA

Although the idea behind this thesis was to develop a methodology that is applicable on a large geographical extent like the province of Noord-Brabant, the methods were tested on a smaller area within the province: Tilburg. Generally, the selection of a small geographical extent for the development of the methods was done for the purpose of data manageability during the operationalization phase, as a smaller study area limited the size of the necessary input data and subsequently the computation time of the operationalization elements. The reason for using Tilburg as study area was twofold: first of all, the greater Tilburg area has a large variety in spatial configuration as it is formed by a mix of urban and rural areas. The city of Tilburg itself is one of the larger cities in Noord-Brabant, and contains a densely built up central area, as well as more spaciously designed suburbs. Furthermore, inclusion of surrounding towns and villages in the study area allows to model routes between urban and rural areas. Figure 3.2 displays the greater Tilburg area in terms of building configuration and tree density. Secondly, since Tilburg is one of the larger cities in the province of Noord-Brabant, a considerable number of observed travels by a wider variety of cyclists were expected.

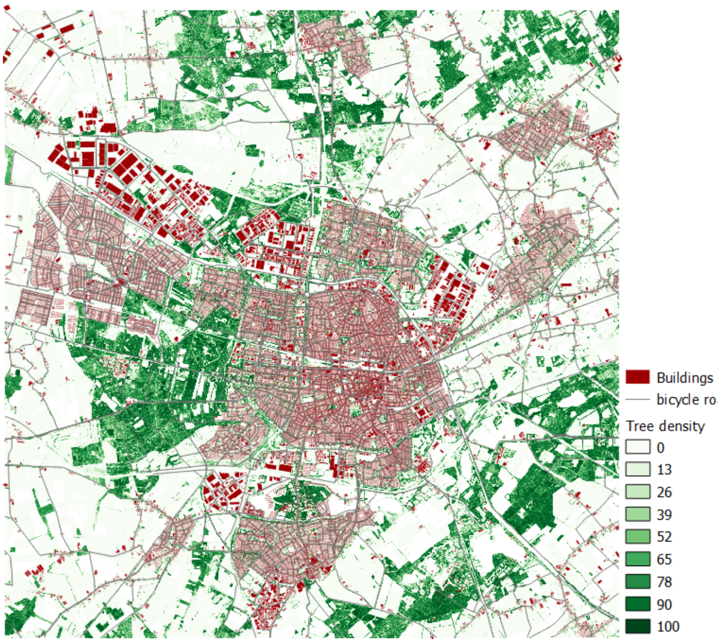


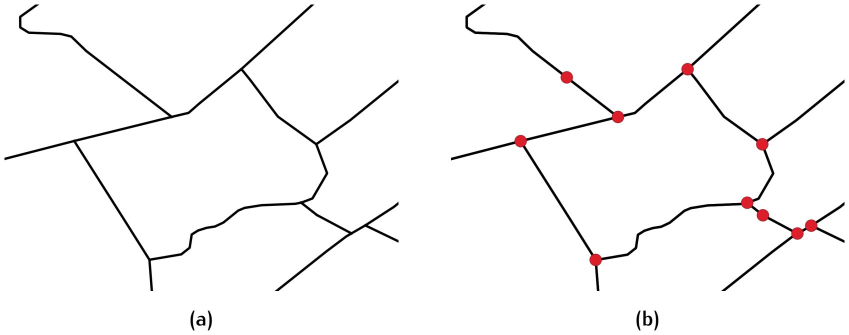
Figure 3.2: Overview of the greater Tilburg area

3.3 OPERATIONALIZATION: ROUTE MODEL

3.3.1 Bicycle road network

The operationalization phase of the route model consists of three parts: *the establishment of a bicycle road network*, *simplification of observed routes*, and *the generation of shortest and fastest routes*. As the first aspect, the establishment of a bicycle road network suitable for route operationalization was a key component of the methodology. Not only did it serve as the base for analyzing the observed routes and generating shortest/fastest routes, but also the operationalization of the degree of shelter provided by the built environment along the road network was done from the perspective of the different segments of the road network.

To generate a route between two points on a network, understanding of connectivity between different segments of that network is required. Where representation of a network in a set of lines lacks the capability of storing spatial relations between the different lines, placement of the bicycle road network in a graph-theoretic perspective enables storage of network topology. This provides insight in the relation between the different components of the network. Urban and Keitt [2001] describe a general graph as a set of unique edges and vertices, where every edge is connected to two vertices. The vertices that share an edge are always adjacent and incident to the edge they share, but vertices can be incident to an unlimited amount of edges. Figure 3.3 displays a visual comparison between a network of lines and the same network stored as a graph. Whereas the lines in Figure 3.3a merely contain information about location and shape, the edges in Figure 3.3b are actually connected through vertices. Connectivity information between edges can subsequently be derived by analysis of edges that share the same vertex.


(a)

(b)

Figure 3.3: Network as a set of lines (a) vs. graph represented by edges and vertices (b)

Storing connectivity information allowed for validation of a topology network, by identifying isolated and falsely unconnected edges. An edge can be considered as isolated if neither of the two incident vertices is shared by another edge. Falsely unconnected edges through digitization of vector lines can be deducted from analysis of vertices with only one connected edge. Although incidence to only one edge is acceptable in the case if dead-ends in the road network, errors concerning incorrectly unconnected edges and vertices should be repaired. Eliminating such topological errors from the generated graph would enable the establishment of a bicycle road network suitable for the purpose of route modelling. As information about adjacency and incidence between edges and vertices is stored, paths between two vertices in a graph could be generated. The ability to model routes was used for the following two aspects of the generation of the route model.

3.3.2 Simplification of observed routes

With the establishment of a bicycle road network with routing abilities, a base was formed for the analysis of the mobility patterns contained in the observed cyclist travel data. To enable compatibility between observed routes and theoretical shortest/fastest routes, the second aspect of the development of the route model was the generalization of the observed travel data to bicycle road network.

The spatial component of the travel data were stored in raw GPS measurements, which by itself did not have a relationship with the bicycle road network, but had been map-matched to the bicycle road network. Therefore, spatial knowledge about which edges of the bicycle road network were used for a route, and the sequence of those edges, already existed prior to the analysis of the observed travel data. However, the observed information regarding trip distance and duration were based on raw GPS measurements and therefore not compatible with trip distances and durations of theoretical shortest/fastest routes that were generated based on length and travel time values of edges in the road network. Consequently, the observed travel data had to be simplified to the match the structure of the bicycle road network through the following two steps:

1. The starting and ending point of a route were re-located to match a vertex in the network, since paths over a graph can only be generated between vertices. As the final route model should contain a shortest/fastest equivalent for every observed route, the source and destination of a generated route should exactly match those of the observed route.

2. The total trip distance and duration of an observed route should be re-calculated based on distance and travel time attributes attached to the edges in the network. Leaving out this step will result in an unfair comparison between an observed route and its shortest/fastest equivalent, as the values regarding trip distance and duration are based on different ways of measuring.

The downside of the simplification of observed travel data was the loss of 'pure' information about trip distance and duration. However, compatibility with a network was required for the comparison with non-observed routes.

3.3.3 Shortest/fastest route generation

Based on the starting and ending vertex of an observed route, hypothetical shortest and fastest equivalents can be generated over the bicycle road network by minimizing the 'travel cost' to go from starting to ending vertex. For a shortest and fastest route, the cost to travel over an edge was defined by the length of an edge in terms of distance and in terms of travel time respectively. To generate the shortest/fastest routes, the A* algorithm is used.

A* is a heuristic search algorithm that evaluates the potential vertices that can be visited from a certain vertex by accounting for the cost to go from one vertex to another, combined with an estimated cost from a potential next vertex to the final destination [Zhang and Zhao, 2014]. Using heuristics enables the algorithm to not having to visit all vertices in the network, and therefore limit computation time compared to algorithms that do not use heuristics. However, compared to non-heuristic algorithms, the output of the A* algorithm might be less accurate as it depends on cost estimations.

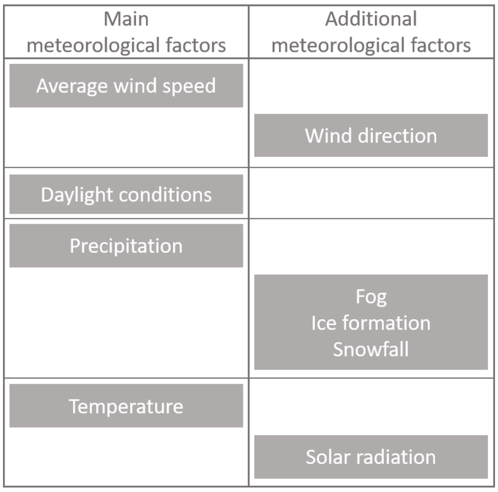
3.4 OPERATIONALIZATION: METEOROLOGICAL FACTORS

For the second element of the operationalization phase, meteorological factors were operationalized to model the weather conditions under which an observed route was conducted. Two operationalization decisions formed the base for the selection and quantification of the set of meteorological factors.

Within the existing literature on cyclist travel behavior variation occurs in the temporal scale in which weather conditions are operationalized. Two approaches can be distinguished: measurements on a daily scale and measurements on an hourly scale [Amiri and Sadeghpour, 2015; Böcker and Thorsson, 2014; Sears et al., 2013]. In order to approach the temporal detail of the observed travel data, the weather conditions were modelled on an hourly scale. To model the operationalized conditions to an observed route, the moment of departure was selected, hereby assuming that cyclists decide upon a route no later than the moment of departure.

A second operationalization decision that can be found in the theoretical framework concerns integration of data into 'weather types' or research the independent effect of each meteorological factor [Böcker et al., 2013; Böcker and Thorsson, 2014]. Although the experience of weather is usually formed through the co-occurrence of different weather factors [Böcker et al., 2013], aggregating meteorological data into integrated weather types will lead to a loss of detail on individual influence of meteorological factors. Therefore, a set of independent meteorological factors was operationalized.

The core of the meteorological factors that were included within this thesis consisted of the four main parameters that have been researched in the theoretical framework discussed in Chapter 2: temperature, wind speed, precipitation, and daylight conditions. However, a further selection of meteorological parameters to model weather conditons was based on the availability in hourly data provided by the KNMI. Ultimately, the aim of the initial weather model was to be as complete as possible. Figure 3.4 presents an overview of the included meteorological factors.



Main meteorological factors	Additional meteorological factors
Average wind speed	Wind direction
Daylight conditions	
Precipitation	Fog Ice formation Snowfall
Temperature	Solar radiation

Figure 3.4: Overview of the included meteorological factors

The meteorological factor describing daylight conditions was formed through a combination of two different types of input: sunrise and sunset times, and durations of the civil twilight. The civil twilight is the time period between the moment that the geometric centre of the sun is 6°below the horizon and sunrise or sunset [Bowditch, 2002]. Within this time period before sunrise and after sunset, the limits of human capabilities to distinguish terrestrial objects without artificial lighting are approximated [Bowditch, 2002]. The daylight conditions were therefore operationalized based on the temporal aspect of a route with respect to sunrise or sunset and the civil twilight period.

Apart from the temporal component, the meteorological data carried a spatial component. The hourly data was collected over a set of official weather stations of the KNMI, spread over the entire country of the Netherlands. As data for the different meteorological factors was only measured at the locations of the weather stations, initially the actual weather conditions at the location of an observed route remained unknown. In terms of accuracy, the collected data from the nearest weather station to an observed route is likely to give the best approximation of the actual weather condtions at the location of the route. However, in cases of a negligible difference in distance between an observed route and multiple weather stations, using only one weather station as a reference might lead to a loss of accuracy. In that light, the value for a meteorological factor was estimated based on input from the three closest weather stations for each observed route. Figure 3.5 provides an overview of the weather stations that are closest to the study area (visualized by the red square).

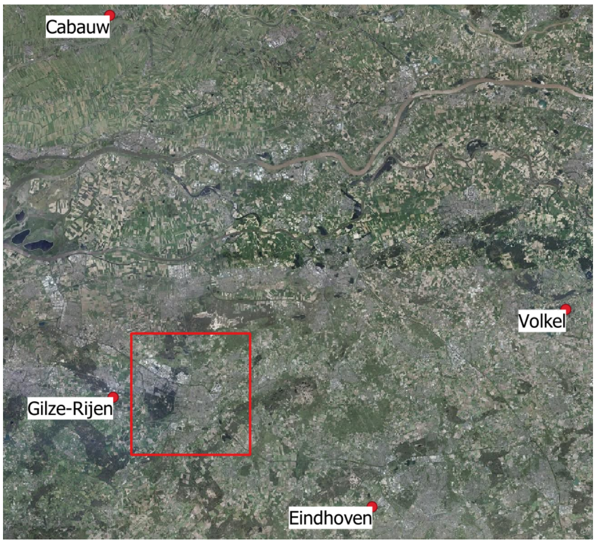


Figure 3.5: Weather stations closest to the study area [Source aerial image: Dutch Cadastre]

To account for the negative correlation between the distance to a weather station and the representativity of the data coming from this station for an observed route, the value for a meteorological factor was determined through an Inverse Distance Weighted (IDW) interpolation over the three selected weather stations. The principle behind this interpolation method is that attribute values of two objects in space are related to each other, but the strength of this relationship is negatively correlated with the distance between the two objects [Lu and Wong, 2008]. More specifically, a closer weather station has a higher influence on the estimated value for a meteorological factor assigned to an observed route than a weather station further away. The formula to perform the IDW interpolation for the application in this method is presented in Equation 3.1. To estimate the value for a meteorological factor at the location of an observed route (V_{route}), the value for this meteorological factor measured at a given weather station ($V_{weatherstation}$) was multiplied by the inverse of the distance between the observed route and the given weather station ($W_{distance}$).

$$V_{route} = \frac{\sum_{i=1}^{n=3} W_{distance} V_{weatherstation}}{\sum_{i=1}^{n=3} w_{distance}} \tag{3.1}$$

Where:

$$W_{distance} = \frac{1}{distance_{\text{weather station-observed route}}} \tag{3.2}$$

3.5 OPERATIONALIZATION: SHELTER

3.5.1 Fundament

The third element of the operationalization phase focused on the quantification and spatial modelling of the degree of shelter provided by the built environment. The fundament for the operationalization of the degree of shelter in this methodology was found in the method that was developed by Anastasiadou et al. [2018]. This thesis built on that method by Anastasiadou et al. by reversing the perspective on openness and operationalize the degree of shelter that is provided by the built environment with respect to weather conditions. Integration of the height/distance ratio as a measure of openness allowed to translate openness to the degree of shelter that is provided by a building in relation to the location of a cyclist. Where Anastasiadou et al. only considered buildings in their approach, this thesis expanded that method by considering the potential shelter that could be offered by vegetation. Besides man-made objects, the built environment generally contains natural objects as trees and vegetation that play a role in the mitigation or emphasis of the experience of weather conditions by a cyclist. As discussed in Section 2.2, existing literature stresses the objective influence of tree density on wind propagation speeds and temperature below the canopy level [Gillner et al., 2015; Heisler, 1990]. Aiming to approximate a complete model that can explain cyclist route choice in different weather conditions, the potential shelter provided by vegetation in the form of tree density was not disregarded for this method.

The main reasoning behind adopting and expanding the more detailed approach of Anastasiadou et al. for this theses, as opposed to the simplified method by Oke, lied in the application on cyclist route choice and the design of the study area. Whereas a large urban area dominated the spatial configuration of the study area, many observed routes were expected to pass through different types of urban design. Therefore, the potential degree of shelter provided by buildings should be modelled on a high level of detail to expose relatively small differences within the urban environment and thus along a route. By including a vegetational shelter indicator, the method of this thesis aimed to examine shelter opportunities along routes passing through rural areas which lack presence of potential shelter by buildings.

3.5.2 Quantification process

To operationalize built environment shelter, the basic workflow structure presented in Figure 3.6 was followed. Whereas this workflow largely complies with the work of Anastasiadou et al., the design of the different stages was adjusted to fit to the objective of this thesis.

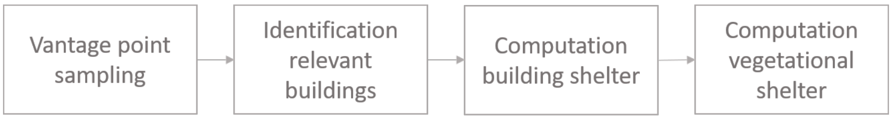


Figure 3.6: Basic workflow structure for the operationalization of shelter

Vantage point sampling

With cyclist route choice being the focal point of this thesis, potential shelter was operationalized in relation to the bicycle road network that cyclists use. To model the movement of cyclists over a road segment, a set of vantage points was sampled over the road segment at an equal distance from each other. Each vantage point resembles the location of a cyclist on the road network, and therefore resembles a location for which the degree of shelter should be determined. In order to place a vantage point in a 3D context of the surrounding built environment, elevation values were attached to each point. The latter is an additional aspect with regard to the method by Anastasiadou et al., where the sampling of vantage points was approached from a 2D perspective. Figure 3.7 provides a simplistic visual representation of a set of sampled vantage points over a road segment.

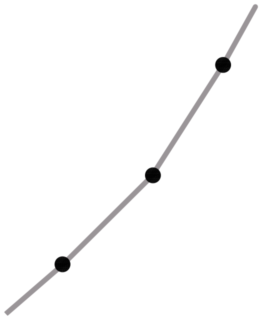


Figure 3.7: A set of sampled vantage points over a road segment

Identification of relevant buildings

The next step in the workflow was determining which buildings can provide shelter for a cyclist on a given vantage point. For this purpose, a method similar to the Isovist method was used. This method is based on the sight and light principle, which aims to quantify a visual field by casting rays from a vantage point and determining whether those rays intersect with surrounding objects [Ncase, 2019]. The method of this thesis used a similar technique, where rays (line segments starting from a vantage point) were casted in a 360°field around a vantage point, as depicted in Figure 3.8

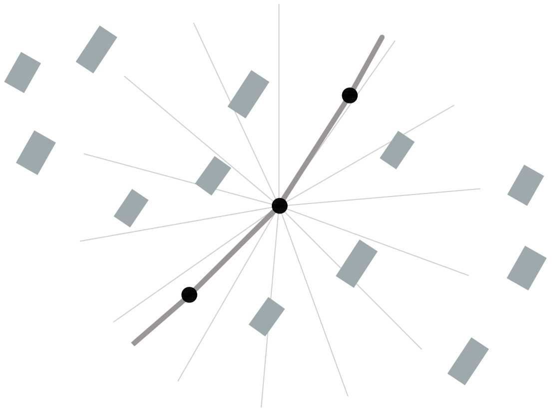


Figure 3.8: Casting of rays from a vantage point

The sight and light principle is based on solving parametric equations to determine intersections between a ray and an object. Parametric equations represent Cartesian coordinates of a line as a function of the same independent variable t [Ncase, 2019]. Describing the x,y component of a line, the set of parametric equations is as follows:

$$\begin{aligned} x &= x_{point} + x_{direction} * t \\ y &= y_{point} + y_{direction} * t \end{aligned} \tag{3.3}$$

Where:

$$\begin{aligned}
x_{point} &= \text{x-coordinate of a point on a line}\\
y_{point} &= \text{y-coordinate of a point on a line}\\
x_{direction} &= \text{direction of motion in x-plane (by increasing t)}\\
y_{direction} &= \text{direction of motion in y-plane (by increasing t)}\\
t &= \text{scalar quantity (parameter)}
\end{aligned}$$

In order to find an intersection between a casted ray and a building, both the ray and the building should be represented as a set of parametric equations. Therefore, a building should be decomposed in a set of line segments that describe the outline of the building. Figure 3.9 shows a simplistic representation of a building footprint, built up of segments ab, bc, cd, de, ef, fa. The start of a segment coincides with the end of the previous segment, and the end of the final segment should coincide with the start of the first segment.

Figure 3.9: Building footprint decomposed into a set of line segments

The sets of parametric equations for both a ray and a building can then be represented as follows:

$$\begin{aligned}
x_{ray} &= ray_{pointx} + ray_{directionx} * t1\\
y_{ray} &= ray_{pointy} + ray_{directiony} * t1\\
x_{segment} &= segment_{pointx} + segment_{directionx} * t2\\
y_{segment} &= segment_{pointy} + segment_{directiony} * t2
\end{aligned} \tag{3.4}$$

In case of an intersection between a ray and a building segment, the x,y components of both parametric equations are the same. Therefore, by equalizing the x,y component, both sets of equations of 3.5 can be solved for the independent variables $t1$ and $t2$. Representing the percent distance between an intersection and the endpoint of a ray or building segment, $t1$ should be bigger than 0 and $t2$ should be between 0 and 1 respectively. If both requirements are met, the intersection lies on an internal point of both the ray and a building segment. In order for a ray and a building segment to intersect, they cannot be parallel, that is, have an equal direction parameter.

$$\begin{aligned}
ray_{pointx} + ray_{directionx} * t1 &= segment_{pointx} + segment_{directionx} * t2\\
ray_{pointy} + ray_{directiony} * t1 &= segment_{pointy} + segment_{directiony} * t2
\end{aligned} \tag{3.5}$$

Since a casted ray is not bound by the first intersection with a building segment, multiple intersections can be found for one ray. As the aim of this method was to operationalize the degree of shelter that a building could bring to a cyclist, only the first intersection with a building segment for each ray was relevant, that is, the intersection with the lowest $t1$ value.

Once a set of intersections has been determined for a vantage point, a subselection is made of the intersecting buildings that provide at least a minimum amount of shelter. This subselection is grounded by a combination of research by Theurer [1999], who developed a more detailed approach at the urban canyon, and Oke [1988]. Oke has found that in order to provide a minimum amount of shelter, the height/width ratio of an urban canyon should not be lower than 0.4. Where Oke only considered urban canyons that are enclosed by buildings on two sides, Theurer also addressed half-open canyons. In this approach, Theurer defined the width of an urban canyon as twice the distance from the middle of a vehicle lane to a building. Integrating this definition into the condition set by Oke, leads to Equation 3.6:

$$\begin{gathered}\frac{height}{2 * distance} > 0.4 \\ distance < \frac{height}{0.8}\end{gathered} \tag{3.6}$$

Translating the distance between a vehicle lane and a building to the distance between a vantage point and an intersection with a building, only intersections where the height delta between the vantage point and the building confirmed to at least 0.8*distance were considered as potential shelter. If an intersection did not meet this condition, it did not contribute to the final building shelter computation for a vantage point. Figure 3.10 visualizes the selection of buildings that provide at least a minimum amount of shelter (buildings in blue).

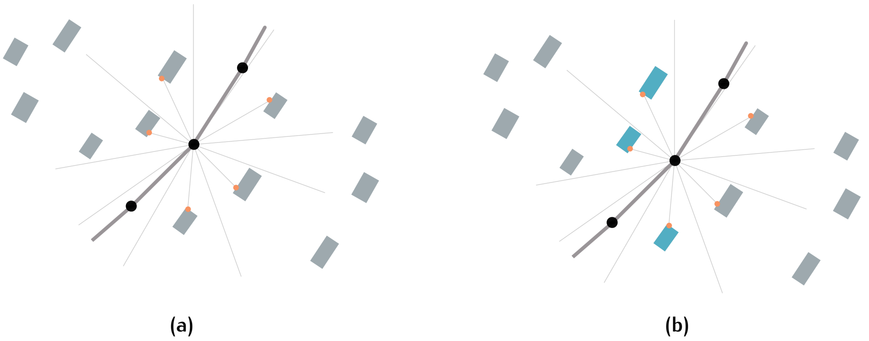

(a) (b)

Figure 3.10: Identification of intersecting buildings (a) and the selection of buildings that can provide at least a minimum amount of shelter (b)

Computation of building shelter

For each ray that intersects with a building, the shelter value of that intersection was determined by the height of the intersecting building h with respect to the height of the vantage point $h_{vantagepoint}$, and the distance d between the vantage point and the intersecting building. In order to compute an average shelter value for a vantage point, an aggregation was made of the set of found intersections, with the mean shelter value being assigned to the vantage point (Equation 3.7). The aim of developing this shelter factor was to quantify an overall degree of shelter, accounting for the entire 360°field around a vantage point. Whereas Equation 3.7 is a derivative from the *height/width* ratio used by Oke [1988], the height delta between the height of the vantage point and a building height was added to the denominator to enable expression of shelter as a percentage.

$$Shelter_{mean} = \frac{\sum_{i=1}^{n} \frac{h - h_{vantagepoint}}{(h - h_{vantagepoint}) + d}}{n} \tag{3.7}$$

Besides the mean shelter value for a vantage point, the maximum degree of shelter that can be found on a vantage point was accounted for by identifying the ray with the highest shelter value (Equation 3.8). The reasoning behind the inclusion of a maximum degree of shelter for a vantage point was mainly based on road segments with buildings only on one side of the road. Whereas the mean degree of shelter for a vantage point on such a road segment would be relatively low, a cyclist might actually experience a high degree of shelter when driving closely to the build-up roadside. Therefore, both a mean and a maximum value for the degree of shelter were established to examine which factor is a better descriptor of cyclist route choice in different weather conditions.

$$Shelter_{maximum} = \max \frac{h - h_{vantagepoint}}{(h - h_{vantagepoint}) + d} \tag{3.8}$$

Computation of vegetational shelter

The computation of the shelter provided by vegetation was approached in a more simplistic manner, mainly based on limitations in available data. With tree density being identified as the main potential source of vegetational shelter, tree density values around each vantage point determined the vegetational shelter at that given location.

3.5.3 Shelter on route-level

Following the workflow presented in Figure 3.6, each vantage point held three values describing the degree of shelter that is provided around that given location. To establish factors describing shelter on a route-level, values for all vantage points were aggregated to the road segments in the bicycle road network. The mean, maximum, and vegetational shelter were then represented as a percentage in distance or travel time over a route where a threshold was met, leading to the following route-level shelter descriptors.

- The percentage of mean building shelter $> 25\%$
- The percentage of maximum building shelter $> 50\%$
- The percentage of tree density $> 50\%$

The defined thresholds for the degree of shelter provided by buildings were an approximate derivative of the conditions for minimum shelter stated in Equation 3.6. For the vegetational shelter, tree densities over 50% were considered as a minimum for sufficient shelter from weather conditions. Figure 3.11 shows a conceptual view on the percentage of mean shelter $> 25\%$, maximum shelter $> 50\%$, and vegetational shelter $> 50\%$ over a route.

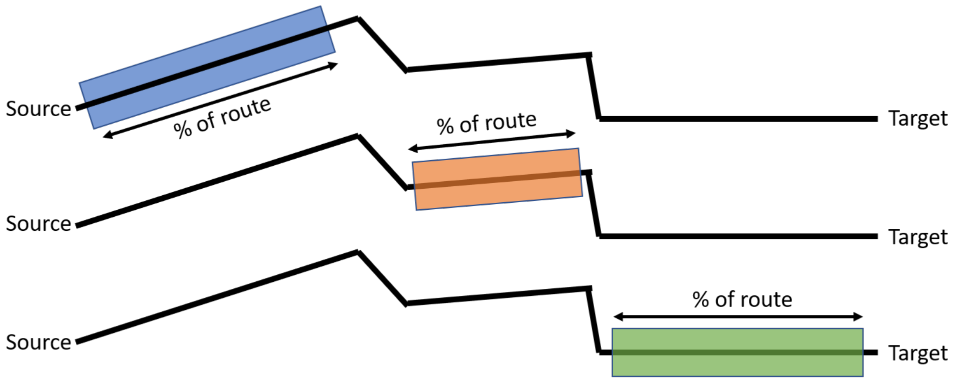


Figure 3.11: Operationalization of mean (blue), maximum (orange), and vegetational (green) shelter factors on route-level

3.6 MODELLING OF CYCLIST ROUTE CHOICE BEHAVIOR

3.6.1 Modelling methods

The aim behind the developed methodology of this thesis was to test the general hypothesis that the degree of shelter that is provided by the built environment is an explanatory factor in the extent of divergence from the shortest or fastest route in different weather conditions. Following the definition of cyclist route choice stated in Section 3.1, the final aspect of the methodology was to model the (combined) effects of the operationalized meteorological and shelter factors to explain two variables: the extent of divergence from the shortest route and the extent of divergence from the fastest route. As these two variables are measured on a ratio scale, two separate sets of multiple linear regression analyses were conducted using each of the two divergence factors as dependent variable. Based on findings in existing literature included in theoretical framework of this thesis (Chapter 2), it was assumed that weather variables and the degree of shelter provided by the built environment are related to each other and that these together determine the degree of detouring. In order to gain a good insight into this, regression models were estimated in which weather and shelter variables were included separately, together as separate variables, and including interactions between both types of variables. Therefore, a set of linear regression analyses held four models of different nature:

1. **Model 1:** in this model, the associations between the dependent variables and the meteorological predictors were estimated. The aim of this regression model was to examine to what extent the set of meteorological predictors explains the divergence from the shortest/fastest route, and to detect co-linearity between the predictors.

2. **Model 2:** in this model, the associations between the dependent variables and the shelter predictors were estimated. Similarly to model 1, regression model aimed to examine to what extent the different shelter predictors explain the divergence from the shortest/fastest route, and to detect co-linearity between the predictors.

3. **Model 3:** this model included the meteorological and shelter predictors that were found to have significant effects on the extent of divergence from the shortest/fastest route. By including both type of predictors in the same model, adjustments for each other's effects on the dependent variable could be examined.

4. **Model 4:** interactions between meteorological and shelter variables were included in this model. Estimating a simple linear regression model did not suffice to model the relationship between the meteorological variables and the shelter variables as a predictor of the extent of divergence from the shortest or fastest route. To test for moderation of the effects of both type of predictors, interactions between meteorological and shelter predictors were included in this model. An interaction between two predictors x and y estimates how the effect of x on the dependent variable z changes when y changes by one unit [Hayes and Matthes, 2009]. This technique allowed to model how a shelter predictor influenced the effect of a meteorological predictor on the extent of divergence from the shortest or fastest route.

Besides the sets of regression models explaining the extent of divergence from the shortest/fastest route, a fifth linear regression model was estimated to explain the adaptation of cyclist route choice to the degree of shelter offered by the built environment. Using each of the shelter factors as independent variable, the aim of this set of models was to examine whether cyclists would adapt the choice of route in different weather conditions to the degree of shelter offered by the built environment.

A final aspect that was accounted for was the fact that multiple routes could be generated by one cyclist, as variations could occur between various cyclists, but also per individual cyclist. Estimating a multilevel linear regression model was the most appropriate way to approach this problem. The aim of conducting a multilevel linear regression analysis to estimate the linear course of a dependent variable, while considering the dependency between data [Sommet and Morselli, 2017]. Using this approach, the observed routes were clustered per cyclist, and the multilevel linear regression model would disentangle the within-cluster effects from the between-cluster effects of the predictors [Sommet and Morselli, 2017]. Application of this modelling technique is not a novelty in research to cyclist travel behavior, as Winters et al. [2007] have used it to estimate the odds of utilitarian cycling as a function of city and individual characteristics. This approach allowed for evaluation of the variance between different cities. In a research more heavily related to this thesis, El-Assi et al. [2017] used a multilevel logistic regression model to estimate the odds of sharing a bicycle as a function of built environment factors and weather factors. By using this method, bicycle sharing activity could be clustered per user.

Within the developed methodology of this thesis, a multilevel linear regression model was exclusively estimated to examine the potential variance in the extent of divergence between cyclists, as well as the effect of individual characteristics of cyclists on the variance.

3.6.2 Modelling context

Although this thesis focused on the interaction between a set of meteorological factors and a set of shelter factors to predict cyclist route choice, valuable conclusions could only be drawn when the output of the different modelling methods was placed into a wider context. The literature framework discussed in Section 2.1.2 confirmed that cyclist route choice can be influenced by a large set of factors. Therefore, the independent predictive power of the interaction between the meteorological and shelter predictors had to be corrected for the influence by a set of control variables. The selection of control variables was based on the established theoretical framework on cyclist route choice, as well as the availability of data. Three categories of control variables were distinguished: individual variables, infrastructural variables, and environmental variables. Addressing the multilevel structure of the model, the individual variables were modelled on cyclist-level, whereas the infrastructural and environmental variables were modelled on route-level.

3.6.3 Model preparation

For each combination of observed and shortest/fastest route in the estimated regression models, variables that could differ between each type of route were measured on route-level (shelter and control variables). Included in the regression models were therefore values for the route-level variables quantified as a difference between an observed route and its shortest/fastest equivalent ($observed - shortest/fastest$).

3.6.4 Validation

Due to time constraints and a limited number of developed factors to include as predictor in the regression models, validation of the statistical models was not included within the methodology. However, a theoretical approach was proposed to validate the results obtained by implementing the developed methodology.

The obtained results for each of the mentioned linear regression models could be validated by predicting the extent of divergence from the shortest or fastest route based on the found effects of each of the included predictors. This validation method would have to be applied on a different study area, or using different observed travel data. In order to apply the proposed validation method, the regression models should have sufficient explanatory strength on values for the dependent variable, usually obtained through inclusion of a substantial number of predictors.

4 IMPLEMENTATION

In order to separate theory and practice, the implementation of the methods from Chapter 3 are described. Divided over the operationalization and modelling methods, the different sections of this chapter elaborate on the used data, preprocessing steps and implementation strategies. For this purpose, the structure of Chapter 3 is maintained. Prior to the elaboration on the implemented methods, a brief overview of the used software and data storage strategies is specified. The three different operationalization methods on the generation of the route model, meteorological factors, and the openness factors are discussed in Section 4.2, Section 4.3, and Section 4.4 respectively. Finally, Section 4.5 treats the quantification of the control variables.

4.1 SOFTWARE SPECIFICATION AND DATA STORAGE

4.1.1 Software

For the implementation of the methodology, different datasets were used in a wide range of formats. Therefore, multiple tools were used for the storage, manipulation, analysis, and visualization of data.

All datasets that were used for the implementation of the different methods were stored in a ***PostgreSQL*** database. *PostgreSQL* is an object-relational database system that allows for both storage and manipulation of the stored data. In conjunction with the *PostGIS* extension, spatial attributes could be stored and operated on.

Python is an open source programming language, mainly used for the implementation of the shelter operationalization method. Use was made of external libraries *pandas*, *psycopg2*, *shapefile*, and *shapely*.

As the main data manipulation platform for the implementation of the meteorological factors method, and preprocessing of data for the openness method, ***FME*** was used. *FME* allows to build a workflow in which data can be loaded from several types of sources, manipulated in terms of content and data model, and converted to different data formats.

Complementary to PostgreSQL/PostGIS and *FME*, ***QGIS*** was mainly used for visualization of the generated output of each method. *QGIS* is an open source geographic information system which allows for analysis, manipulation and visualization of spatial data. *QGIS* offers connection possibilities with a PostgreSQL/PostGIS database, enabling direct visualization from spatial data stored in a database and therefore avoiding unnecessary exports from a database.

To model cyclist route choice, ***SPSS*** was used to statistically analyze the model that follows from the operationalization phase. *SPSS* is a software platform that offers a wide set of statistical analysis methods.

4.1.2 Data storage strategies

In order to reduce computation time when retrieving and manipulating data, two main strategies were applied for more efficient data storage. The distinction in the two methods lies in the difference between general data storage strategies and spatial data storage strategies.

- ***Data storage in B-tree structure:*** is a common method of indexing standard data types like numbers or strings [Elmasri Ramez and Navathe Shamkant, 2010]. This indexing method stores the data in a hierarchical search tree, with the hierarchy defined by the indexed attribute [Elmasri Ramez and Navathe Shamkant, 2010]. Application of this indexing method avoids the traversal of an entire table when requesting a certain value.
- ***Spatial data storage in R-tree structure:*** allows for more efficient access to spatial data by grouping the data in bounding boxes [Van Oosterom, 1999]. By indexing such a bounding box, only spatial data that is grouped within the same bounding box has to be traversed when requesting specific data.

4.2 ROUTE MODEL

4.2.1 Data

The input for the route model can be classified into two main categories: observed travel data and bicycle road network data. With these two categories holding multiple datasets, Table 4.1 presents an overview of the data per category.

Table 4.1: Overview of datasets for the operationalization of the route model

Dataset	Source	Format
Observed travel data		
GPS measurements	B-Riders project	Shapefile (Point)
GPS match	B-Riders project	Table
Bicycle road network data		
Fietsersbond network	Dutch Cyclist Union	Shapefile (Multilinestring)
Links	B-Riders project	Shapefile (Multilinestring)

As described in Section 1.3, the observed travel data consisted of a set of GPS measurements and accompanying route and user information, measured over the year 2014. The GPS measurements held attributes describing the route they belong to (*routeid*), the speed, the heading, and the time of measurement (in UTC). In order to maintain the privacy of the cyclists, the observed travel data were anonymized by removing a part of a route with a random magnitude between 100 and 300 meters at the start and end of a route. To preserve the actual length of a route, the start and end point had been randomly placed at the cut-off distance.

Grouped per observed *routeid*, the GPS measurements were mapmatched to a road network provided by the Dutch Cyclist Union. The mapmatched data resembled an aggregation of the GPS measurements that belong to the same road segment. Along with the spatial aggregation to a road segment, attributes that describe the speed and time of measurement were aggregated for the set of GPS measurements per road segment. Therefore, the mapmatched travel data provided information on what road segments were used for a road, the sequence of the road segments, and the aggregated speed per road segment. As general attributes of an entire route, the mapmatched travel data held values for the total travel distance and travel time.

The road network of the Dutch Cyclist Union consisted of all road segments that were designated as appropriate for cyclist use. These road segments carried attributes describing infrastructural characteristics, as well as indicators of the traffic volume around a road segment. In order to spatially relate the mapmatched GPS measurements with the road network, an intermediate network dataset *Links* was provided with matching geometry and common attributes to join on. In order to structure the relationship between the different datasets, Figure 4.1 presents an overview of the different datasets and how they were connected. As neither of the datasets were openly accessible, both the observed travel data and the road network data were obtained through the main supervisor.

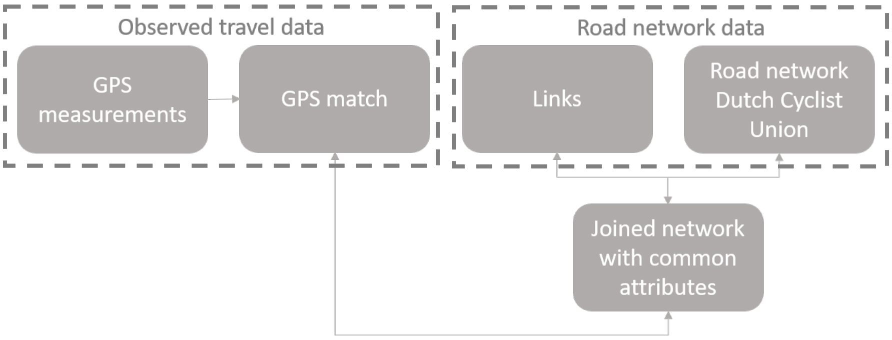


Figure 4.1: Connections between datasets for operationalization route model

4.2.2 Preprocessing

To establish a topologically correct bicycle road network, unconnected edges have been eliminated. As shown in Figure 4.2a, the bicycle road network contains line segments that have falsely been disconnected from a vertex joining multiple line segments in the network. To correctly align the end vertex of such a 'dangling' line segment with another vertex in the network, the vertices are snapped, using a tolerance of one meter. Figure 4.2b visualizes the result of the snapping operation.

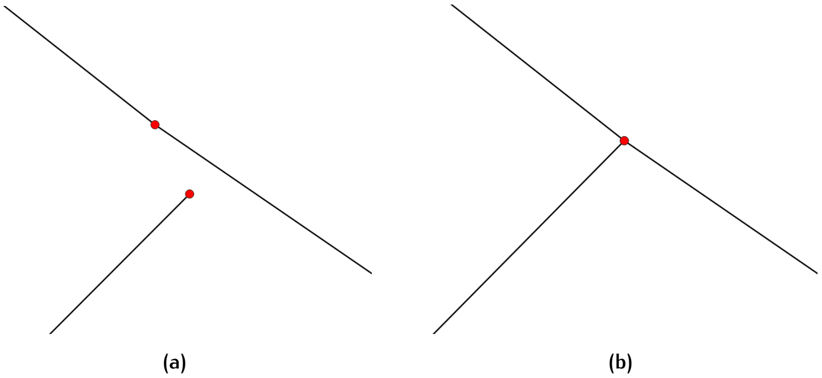


Figure 4.2: Unconnected line segment (a) vs. snapped line segment (b)

4.2.3 Implementation

After completion of the preprocessing steps, a topologically correct bicycle road network is established, with unique identifiers for each road segments that can be linked to the mapmatched travel data. In order to generate a model containing simplified observed routes and accompanying shortest/fastest routes, three implementation steps are defined based on the method presented in Section 3.3: *construct a graph from the bicycle road network*, *simplify and filter observed routes*, and *compute the shortest and fastest routes*. For all implementation steps, *PgRouting* functions are used. *PgRouting* is an extension of *PostgreSQL/PostGIS* that provides spatial routing functionality within a database [pgRouting Community, 2016].

Graph construction

To construct a graph from the topologically correct network, a topology network was created by assigning a *source* and *target* vertex to every edge (road segment) in the network. The network geometry, attributes, and connectivity information were then stored in table format. Since the original network data did not provide information on direction restrictions for a road segment, the constructed graph was bidirectional.

Simplification and filtering of observed routes

After the construction of a graph over the bicycle road network, *source* and *target* vertices could be assigned to the mapmatched observed travel data based on the common attributes to join on. Following the method discussed in Section 3.3.2, the start of an observed route was generalized to the *source* vertex of the first edge used by a route. Similarly, the end of an observed route was generalized to the *target* vertex of the last used edge. Having determined the *source* and *target* vertex of every observed route, shortest/fastest equivalents could be computed between this set of vertices. Prior to the generation of the shortest/fastest routes, the set of observed routes was filtered based on four criteria:

- ***Filter 1:*** disregard observed routes that are not completely contained by the study area. In order to generalize the observed routes to the bicycle road network of the study area, the travel distance and duration was re-calculated over the edges of bicycle road network. Consequently, observed routes that have been mapmatched to edges outside the study area could not be included in the route model.

- ***Filter 2:*** disregard observed routes with average speed measurements over $45km/h$. Although the observed travel data provided by the B-Riders project was meant to be exclusively collected by participants travelling by bicycle or electronic bicycle, the analysis of the observed travel data accounted for potential use of other transportation mode. The filtering of non-bicycle data was based on speed observations contained by the GPS measurements, where routes containing observed average speeds over $45km/h$ were disregarded from the route model. The speed limit of $45km/h$ was defined by the maximum speed that is supported by the fastest type of electronic bicycle: the speed-pedelec [Stelling-Konczak et al., 2017]. A single speed observation above $45km/h$ per route was merely considered as an outlier when the average speed over the entire route was lower than $45km/h$, leading to inclusion of those routes in the route model.

- ***Filter 3:*** disregard observed routes that have the same vertex as the *source* and *target* of the route. The generalization of the observed routes can lead to overlap of the *source* and *target* vertex in case of a round trip. Since *PgRouting* uses the combination of a *source* and *target* vertex as the base for the shortest/-fastest route computation, no route could be computed for overlapping *source* and *target* since the shortest/fastest option is to visit none of the edges.

- ***Filter 4:*** disregard observed routes that are likely to be round trips. As the aim of this thesis is to explain utilitarian cyclist route choice, round trips should not be included in the route model. Therefore, observed routes with shortest/-fastest equivalents shorter than 20% of the observed route and only a small amount of traversed edges by the shortest/fastest route were disregarded.

Shortest/fastest route generation

The final implementation aspect concerned the generation of the shortest/fastest routes over the bicycle road network, using the A* algorithm (with as heuristic function the Manhattan distance, $abs(dx) + abs(dy)$). Where the calculation of the shortest routes was based on the length of an edge, the fastest routes were calculated by using the length and aggregated speed measurements over an edge. The GPS measurements of the observed travel data contained speed values for every measurement, which have been aggregated per edge for the map-matched travel data. Using the length and aggregated speed measurements, a theoretical travel time could be calculated for each edge. In cases where no observed speeds had been recorded for an edge, the average speed over all measurements was used.

4.2.4 Output

Implementation of the steps discussed in the previous section led to a route model consisting of 18424 distinct observed routes and shortest/fastest equivalents, distributed over 322 distinct cyclists. Table 4.2 summarizes the distance and travel time characteristics of the established route model. To verify the generated shortest/fastest routes, a comparison was made with a route model that was generated using Dijkstra's algorithm. As Dijkstra's method uses a more 'greedy' approach to find a shortest path by repeatedly selecting all vertices in every iteration, it produces accurate results but with more computational effort [Reddy, 2013]. By comparing the number of edges and the overall distance/travel time for every shortest/fastest route generated by both algorithms, it was found that the A* and Dijkstra's algorithm produced the same route model.

Table 4.2: Summary of distance and travel time characteristics route model

	Observed distance [m]	**Shortest distance [m]**	**Observed travel time [s]**	**Fastest travel time [s]**
Mean	5311.8	4947.9	947.9	878.4
Median	5417.6	5057.8	969.6	900.4
Maximum	27424.3	16072.9	4893.1	2756.2
Minimum	500.1	177.1	83.2	31.8

Figure 4.3 shows the temporal distribution of the routes within the route model. A rising trend can be observed from the winter months until the midst of spring. The summer months hold the lowest frequency of cycling trips, however with a large peak in the month of September.

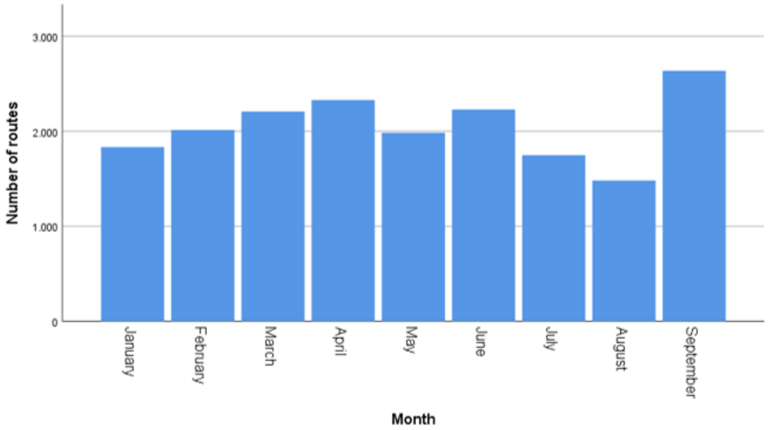


Figure 4.3: Temporal distribution of the routes

4.3 METEOROLOGICAL FACTORS

4.3.1 Data

The operationalization of the meteorological factors was based on open weather data by the KNMI, measured on an hourly scale. These data can be requested per weather station, and for a self-defined timespan. To match the temporal scope of this thesis, this method used a dataset containing hourly measurements from January until September of the year 2014. This dataset includes values for all factors specified in Section 3.4, except values to operationalize the daylight conditions. To quantify the daylight conditions, a dataset containing the sunrise and sunset times per day was combined with a dataset describing the duration of the civil twilight per month, both provided by the Royal Dutch Meteorological Institute. Both datasets are provided with precision to the minute. Due to unavailability of data for the year 2014, data for the year 2019 was used. As the sunset and sunrise times differ slightly between both years, the magnitude of the difference does not exceed five minutes. Therefore, the loss of accuracy was considered as minimal. Table 4.3 provides an overview of the required input to operationalize the meteorological factors.

Table 4.3: Overview of datasets for the operationalization of the meteorological factors

Dataset	Source	Format
Hourly meteorological values (2014)	KNMI	Table
Sunrise/sunset times (2019)	KNMI	Table
Civil twilight duration (2019)	KNMI	Table

4.3.2 Preprocessing

In the first preprocessing step, a spatial dimension was added to the meteorological data by defining the locations of the weather stations from which the data were requested. Included in the hourly datasets were coordinates in *longitude* and *latitude* for every weather station, allowing for representation in point geometry. By creating geometry from the set of coordinates, the observed meteorological data could be spatially related to the observed travel data.

The second preprocessing step concerned the definition of the two civil twilight periods for each day of the year. To define the morning civil twilight period, the civil twilight duration was subtracted from the time of sunrise. Similarly, the evening civil twilight period was defined by adding the civil twilight duration to the time of sunset.

4.3.3 Implementation

In order to assign values for each meteorological factor to an observed route, the general workflow presented in Figure 4.4 was followed. The workflow can be broken down into four aspects: *route preparation*, *identification of three closest weather stations*, *inverse distance weighted interpolation*, and *determination of final values*.

Route preparation

Prior to any involvement of meteorological data, the observed routes were prepared to enable a link based on time, as well as a spatial link with the meteorological data measured at the weather stations. As described in Section 3.4, the meteorological data was modelled to an observed route based on the moment of departure. Following this principle, the temporal and spatial component of an observed route were represented by the timestamp and location of the first GPS measurement of a route.

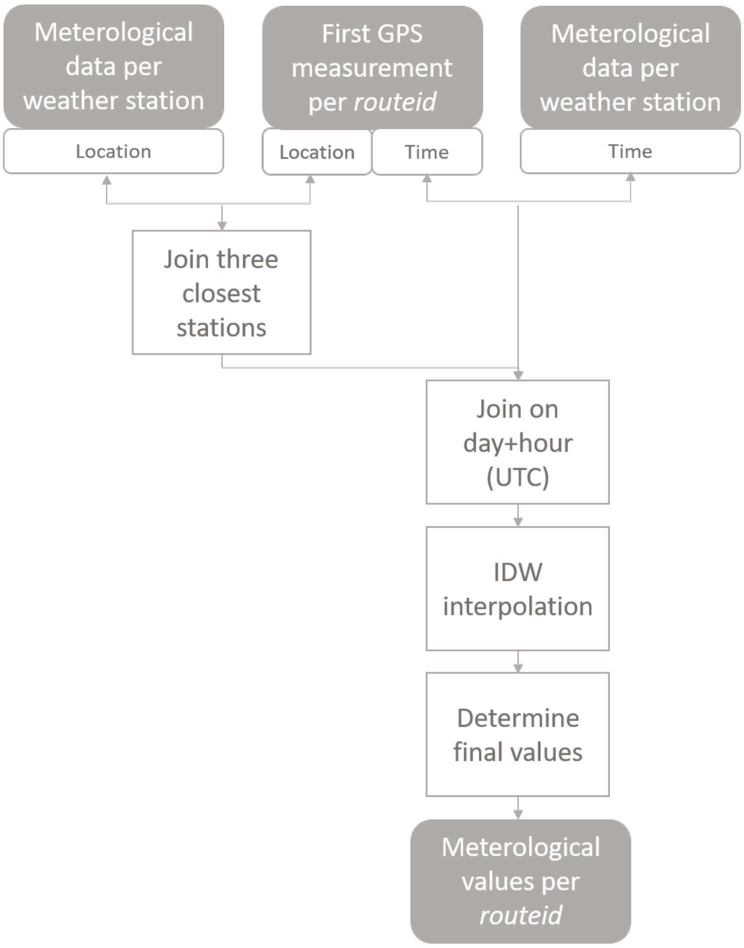


Figure 4.4: General workflow for implementation meteorological factors method

The second aspect of the preparation of the route data is the determination of the average route direction, in order to compute the relative wind direction in a later stage of the implementation. For this step, the average heading was taken for all GPS measurements belonging to an observed route.

The final preparation step involved meteorological data, as the observed routes were assigned an attribute value describing the daylight conditions at the time of departure. Distributed over three categories, the assignment of daylight conditions to a route is as follows:

- **Full daylight:** when the moment of departure is between sunrise and sunset
- **Twilight:** when the moment of departure is within the civil twilight period
- **No daylight:** when the moment of departure is between the end of the current, and the start of the next civil twilight period

Identification of three closest weather stations

With the location of an observed route represented by its first GPS measurement, the next implementation step was the identification of the three closest weather stations. For each point representing an observed route, the distance to each of the weather stations was calculated, whereafter the three closest stations are selected. The meteorological data for each of the three closest weather stations was joined with an observed route based on the departure time of the route. As the meteorological data is scaled on an hourly level, departure time of the observed route is generalized from minute level to hour level.

Inverse distance weighted interpolation

After completion of the previous implementation step, the observed routes hold meteorological data coming from three weather stations, as well as one general value describing the daylight conditions. In order to estimate a value for each of the meteorological factors (measured at a continuous scale) at the location of an observed route, the data from the three closest weather stations is aggregated based on the inverse distance weighted interpolation method described in Section 3.4.

Determination of final values

By implementing the inverse distance weighted interpolation method, the observed routes hold an interpolated value for each of the continuous meteorological factors. However, the input data for the factors *snowfall*, *ice formation*, *fog*, and winddirection is measured on a nominal scale. Values for those factors could therefore not be estimated through the IDW interpolation method. In order to estimate a value for the *snowfall*, *ice formation*, *fog* and *wind direction* at the location of an observed route, a nearest neighbor interpolation was applied. For each of the four factors, the value that was measured at the closest weather station was assigned to the observed route.

Finally, the operationalization of the variable wind direction required an extra step. In order to place the direction of wind in the context of an observed route, the direction of a route was considered. Figure 4.5 presents a conceptual visualization of the applied method. Approaching the context of wind direction on the level of an entire route, the general route direction between start and destination of a route was compared to the wind direction. The difference between both directions, measured in degrees, was then aggregated into a more meaningful category of wind direction. Table 4.4 shows how the categories regarding wind direction were quantified. Cases without wind, or with rapidly changing wind directions have been left out of this table as those were not identified through the same principle. In both cases the direction of the observed route did not have to be considered, as the influence of either changing wind conditions or absence of wind is independent from the route direction. Cases without wind have been classified as *wind in the back*, as the cyclist should experience no resistance in such conditions.

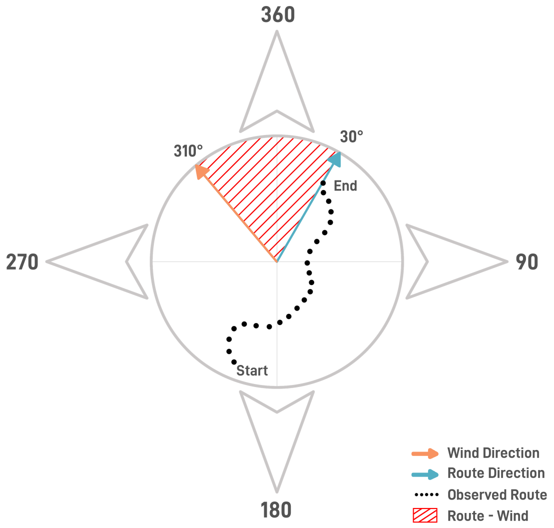


Figure 4.5: Determination of the difference between route direction and wind directions

Table 4.4: Categorization of wind direction in relation to the general route direction

Difference route-wind direction	Range	Category
0	+/- 45	
360	- 45	wind in the back
180	+/- 45	wind in the face
90	+/- 45	
270	+/- 45	wind from the side

4.3.4 Output

By implementing the operationalization method for the meteorological factors the route model was extended with values that represent the weather conditions under which a route was conducted. Table 4.5 provides an overview of the meteorological factors that were assigned to the route model, and how each of the factors was measured. Since the shortest and fastest routes were computed from the same departure time as the observed routes, a shortest or fastest route were assigned the same values for each of the meteorological factors as an observed route.

Table 4.5: Overview of meteorological factors

Meteorological factor	Code	Scale
Average windspeed	[m/s]	Ratio
Daylight conditions	0 = no daylight 1 = full daylight 2 = twilight conditions	Nominal
Fog	0 = no fog 1 = fog	Nominal
Ice formation	0 = no ice formation 1 = ice formation	Nominal
Precipitation sum	[mm]	Ratio
Snowfall	0 = no snowfall 1 = snowfall	Nominal
Solar radiation	[J/cm^2]	Ratio
Temperature	[$^\circ C$]	Ratio
Wind direction	1 = wind in the back 2 = wind in the face 3 = wind from the side 4 = changing direction	Nominal

4.4 SHELTER

4.4.1 Data

The input for the operationalization of the shelter variables consisted of three parts: data describing the configuration of buildings and the height of those buildings, data describing the tree density at a certain location, and a Digital Terrain Model (DTM). Table 4.6 specifies each of the three datasets.

To quantify the degree of shelter provided by buildings, the 3D Basisregistratie Adressen en Gebouwen (3D BAG) is used. The 3D BAG dataset is a project of the 3D Geoinformation Group of the TU Delft [Dukai, 2018], in which the building footprints of the Basisregistratie Adressen en Gebouwen (BAG) dataset are enriched with height values based on the point cloud of the Dutch height model, Actueel

Table 4.6: Overview of datasets for the operationalization of shelter variables

Dataset	Source	Format
3D BAG	3D Geoinformation Group TU Delft	Shapefile (Multipolygon)
DTM	Dutch Cadastre	Raster (0.5x0.5m)
Tree density	RIVM	Raster (10x10m)

Hoogtebestand Nederland (AHN). For the generation of the 3D BAG dataset, the building footprint was elevated based on the height values of points located above the footprint that are classified as building. For each building footprint, several height values were included to provide multiple options to model the roof height. As the difference in height values generally did not exceed centimeter level, the potential effect on the output of this method was considered negligible. Therefore, the highest roof height was selected as building height. Furthermore, the 3D BAG dataset held the following original BAG attributes that are used for the implementation of the shelter method: *identificatie* (a unique identifier for each building) and *bouwjaar* (the date of construction). In contrast with the general BAG data, the 3D BAG dataset was already filtered from buildings that were designated as demolished, or yet to be built in the year 2019.

The other two datasets that were used to operationalize the shelter variables concern raster data describing the elevation of the terrain of the study area, and the density of trees throughout the study area. The former of the two datasets was used complementary to the 3D BAG dataset to determine the height of a vantage point relative to the height of a building. The raster file describing the tree density for a given cell was used to quantify the vegetational shelter. In this dataset only trees higher than 2.5 meters above the terrain surface were included.

4.4.2 Preprocessing

In order for the 3D BAG dataset to match the temporal scope of this thesis, all buildings that were not yet constructed by the beginning of 2014 should be disregarded from the dataset. Therefore, the following filter was applied:

- *bouwjaar* $\leq$ 01-01-2014

4.4.3 Implementation

After filtering the 3D BAG dataset to match the temporal scope of this thesis, a fitting dataset describing the building configuration in the study area was established. Combined with raster data describing the tree density, the set of three shelter factors could be operationalized. To spatially relate the shelter factors with the routes in the route model, all three factors were operationalized from the perspective of the bicycle road network. For the operationalization, four implementation steps were distinguished: *sampling vantage points*, *subselecting vantage points*, *identification of relevant buildings*, and *computation of building and vegetational shelter*.

Sampling vantage points

As the first implementation step a set of vantage points was sampled over the bicycle road network at a regular interval of 20 meters, where each vantage point carried a unique identifier and *linknummer* to identify over which road segment it was sampled. An interval of 20 meters was selected in order to generate sufficient coverage of the study area. Since shelter characteristics for each unique road segment in the bicycle road network had to be computed, the sampling method accounts for road segments shorter than 20 meters by assigning at least one vantage point to each road segment. The height of a vantage point was estimated by extracting elevation values from the generated DTM of the study area. The estimated elevation value for a vantage point was an estimated average over the four nearest raster cells. In cases of missing elevation on the location of a given vantage point, the elevation value from the previous vantage point was assigned to that vantage point. For the study area, a set of 72845 vantage points was sampled.

Subselecting vantage points

As a measure to reduce computational effort, a subselection of the sampled vantage points was made in the second implementation step. As is described in Section 3.5.2, in order for a building to provide at least a minimum amount of shelter, the height of the building cannot be lower $0.8 * distance$. Therefore, the vantage points that did not fall within this distance threshold of any building in the study area could be left out of the process to compute the degree of shelter provided by buildings.

To determine which vantage points did not fall within the distance threshold of any building, variable buffers were created around each building using the following equation: buildingheight/0.8. Whenever a vantage point did not intersect with at least one buffer, it was not used as input to compute the degree of shelter provided by buildings. Figure 4.6a and Figure 4.6b depict cases where a vantage point was included and was not included for further computations of building shelter respectively. By implementing the subselection step, the number of vantage points included for analyses was reduced to 49901.

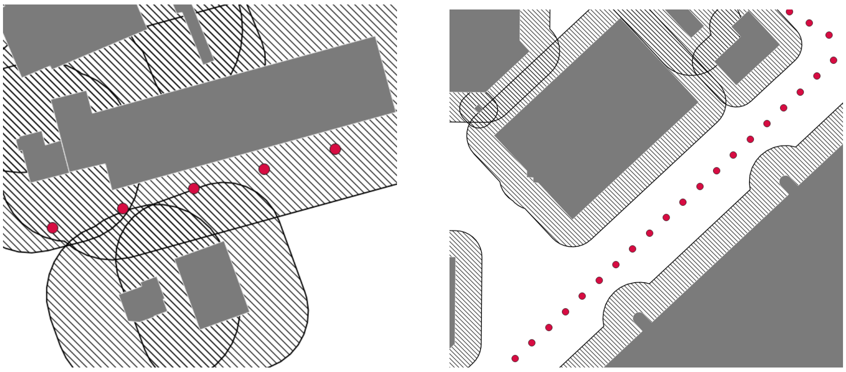

(a) (b)

Figure 4.6: Intersecting vantage points (a) vs. non-intersecting vantage points (b)

Identification of relevant buildings

Combining the output of the previous two implementation steps, buildings that provide shelter on a given vantage point could be identified. To implement this part of the method, the subselection of vantage points and the building polygons of the 3D BAG were used as input. To identify the buildings that could provide shelter for a given vantage point, the following steps were implemented:

1. Subdivision of the building polygons into segments based on the combination of two sets of x,y coordinates for two sequential vertices, belonging to the same building. The set of segments was then stored together with the unique identifier of the building that the segments belong to, and the height value of that building.

2. For each vantage point, a subselection of the segments was made based on the building buffers with which it intersected in the previous implementation step. As only those buildings could potentially provide shelter to the given vantage point, segments of other buildings were disregarded for the given vantage point.

3. In order to maintain a reasonable probability of incidence between a ray and a building, rays were casted every 10°for each vantage point.

4. For each casted ray, the algorithm iterates over the building segments that were linked to the vantage point. By writing both the ray and the segments in parametric form, intersections between a ray and segment could be found by solving the set of parametric equations as described in Equation 3.5.

5. For each intersection that was found per ray, a check was performed to determine whether the intersection lied on an internal point of the building segment, and which intersection was closest to the vantage point. Only the closest internal intersection was maintained. Algorithm 4.1 presents an algorithmic overview of steps 3 to 5.

6. For each closest intersection per ray, a check was performed whether the distance to the point of intersection is smaller than $(height_{building} - height_{\text{vantage point}})/0.8$. If this condition was not met, the intersection is disregarded.

The output of the six-step process was for each vantage point a set of casted rays holding information on the intersected building, or holding no additional information in case the ray did not intersect with any building.

Algorithm 4.1: Finding intersections

```
Input: vantage points, building segments
Output: a set of intersections, grouped per vantage point
initialization;
for each vantage point do
    cast rays;
    for each ray do
        solve ray.x = building segment.x;
        solve ray.y = building segment.y;
        if intersection is internal then
            if intersection is closest intersection then
                set as intersection point;
            else
                disregard intersection;
```

Computation of building and vegetational shelter

Based on the generated output of the previous implementation step, building shelter values could be computed for each intersecting ray by accounting for the height delta between the intersected building and the vantage point, and the distance between the vantage point and the intersected building. Whenever a ray did not intersect with a building, a building shelter value of 0 was assigned. For each vantage point, two building shelter values were then computed: a mean shelter value by averaging over the values of all the rays per vantage point, and a maximum shelter value by considering the maximum shelter value of all rays per vantage point. The vantage points that were not included in the subselection made in an earlier stage of the operationalization were assigned a value of 0 for both the mean and maximum degree of shelter.

To determine the vegetational shelter at a given vantage point, a tree density value was extracted from the raster dataset provided by the RIVM by estimating over the four nearest cells. As the input data only includes trees higher than 2.5 meters above the height of the terrain, the elevation value of the vantage point did not have to be considered. For this step, all 72845 vantage points were considered.

Finally, the three shelter values that were determined for each vantage point were aggregated per road segment. An average value for each of the three shelter factors was taken over all vantage points belonging to the same road segment.

4.4.4 Performance

The implemented method to operationalize the three shelter factors proved to be costly in terms of computation time. Initial runs including all 72845 vantage points resulted in computation times around 8 hours. Reduction of the number of included vantage points (49901) through the described subselection decreased computation times to ± 4.5 hours. Due to the limited time extent for this thesis, no further attempts to improve the computation time were implemented. However, a set of steps to theoretically decrease the computation time are provided within this section.

The most substantial computational load was caused by the six-step process to identify relevant buildings for a given vantage point. The current implementation method used an inefficient approach by segmenting all building polygons in every run, and by iterating over every building segment that could potentially provide the closest intersection for the vantage point. Improvements in the computational efficiency of the implementation method could be made by separation of the segmentation part, storage of segments in an R-tree structure, and usage of parallel processing strategies.

Separating the segmentation aspect of the process from further steps would avoid unnecessary generation of building segments, as the building segments from a previous run could be re-used. The computational effort would then only have to be made once. Additionally, the generated segments could be stored in an R-tree structure for more efficient traversing when identifying the closest intersection for a ray. Figure 4.7 shows a conceptual take on this aspect. By storing the segments (gray solid lines), only segments in bounding boxes that intersect with a ray would have to be traversed. Initially this means that only bounding box R_1 would be considered, whereafter a subselection of bounding boxes R_{11} and R_{12} would be made. Therefore, none of the other segments would have to be traversed. The final step would then be to identify intersecting segments (orange), and find the closest intersection (green).

A final decrease in computation time could be obtained by subdividing the set of vantage points, and running the operationalization process for each subset of vantage points simultaneously. In the current implementation method, the process for each vantage point is run in a linear manner. Since initiation of the operationalization process for one vantage point does not depend on output for another vantage point, parallel processing techniques could be applied.

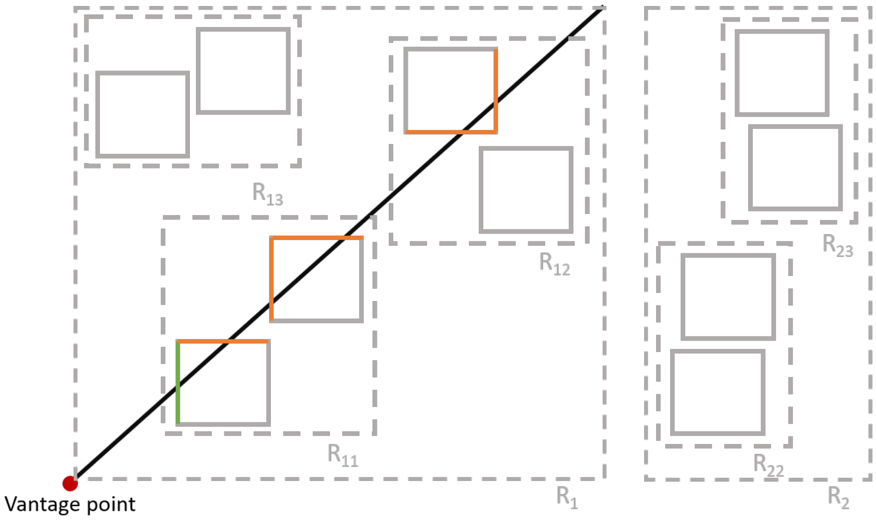


Figure 4.7: Conceptual view on the storage of segments in R-tree structure

4.4.5 Output & visualization

With all road segments of the bicycle road network holding values for the mean shelter, maximum shelter, and tree density, shelter values on route-level were computed by considering the length/travel time of road segments that meet the criteria described in Section 3.5. Consequently, the route model was expanded with three attributes describing the extent of shelter that can be found on found along a route.

In the absence of reference data, visualization of the distribution of the mean shelter, maximum shelter, and tree density over the study area was used as a form of validation. Following the concept behind the method described in Section 3.5, values for both factors describing the building shelter are expected to be higher in densely built up areas. Figure 4.8 and Figure 4.9 show the distribution of the mean shelter and maximum shelter respectively, averaged per road segment, over the study area.

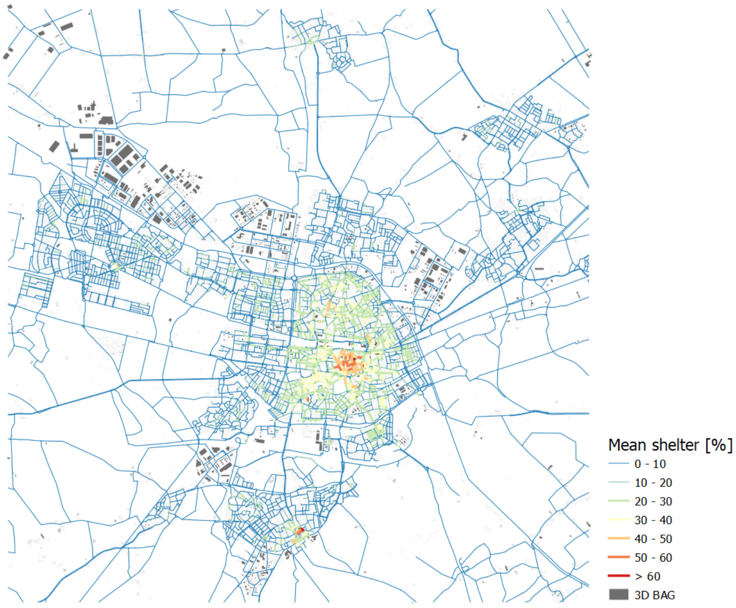


Figure 4.8: Distribution of mean shelter value per road segment over the study area

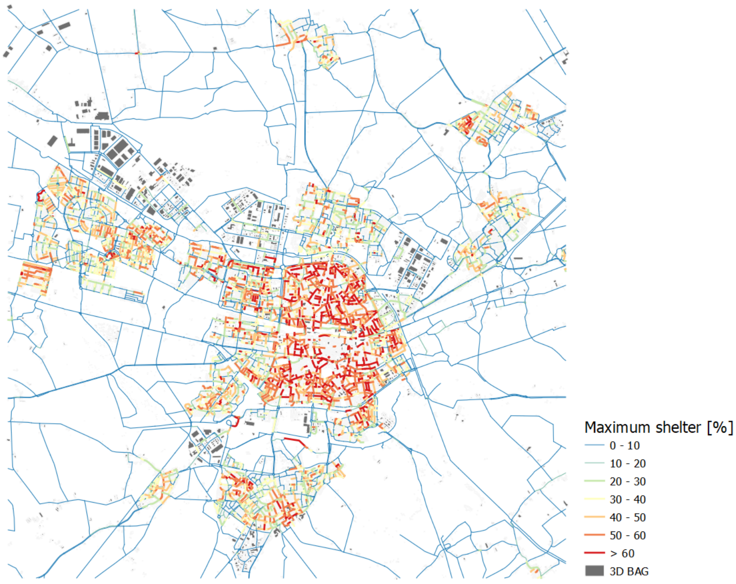


Figure 4.9: Distribution of maximum shelter value per road segment over the study area

Both Figure 4.8 and Figure 4.9 show higher degrees of mean and maximum shelter respectively in more densely built up areas like the city centre, and suburbs with denser street patterns, conforming with the previously stated expectations. However, occurrence of high degrees of mean shelter in central areas were more limited compared to high degrees of maximum shelter.

As described in Section 3.5.2, the difference between considering the mean shelter at a location and considering the maximum shelter at a location was expected to be most pronounced for road segments with buildings only on one side of the road. Figure 4.10a and Figure 4.10b depict the same set of road segments in Tilburg assigned with an averaged mean and maximum shelter value respectively. Comparison between both figures indicates that a significantly higher maximum degree of shelter (orange and red segments) were experienced on road segments close to buildings, while road segments in further away from buildings share low degrees of mean and maximum shelter (blue segments). This output corresponded with the expected outcome.

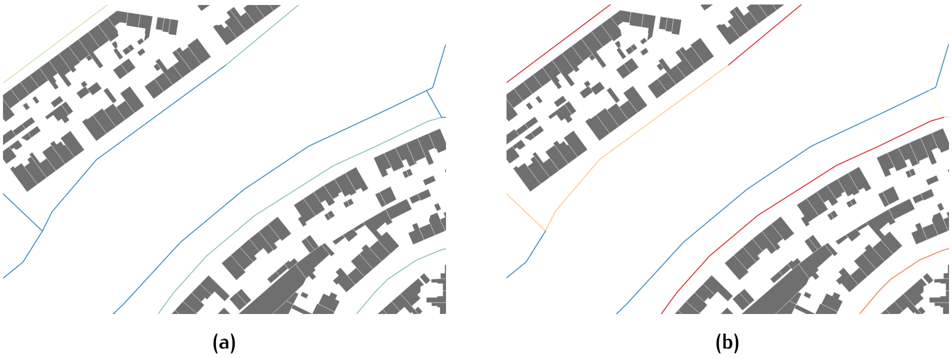

(a)
(b)

Figure 4.10: Mean shelter for road segment with buildings on one side (a) vs. maximum shelter for road segment with buildings on one side (b)

As high tree densities could both occur inside and outside urban area, no clear expected distribution was stated. Since the most obvious form of reference data was used as input for the generation of tree density values per road segment, a visual comparison between aerial imagery and the modelled tree density was made as validation method. For this purpose, small samples of the study area on which the aerial imagery showed relatively high tree densities along the road were selected for visual inspection. Figure 4.11 shows an example of a sample. The majority of the road segments held high values for tree density, whereas most of the road segments bordering this area held lower tree density values. As this complied with the underlying aerial image, the output of the operationalized vegetational shelter approached the expected values to a considerable level.

Figure 4.11: Visual comparison of reality vs. operationalized averaged tree density per road segment [Source aerial image: Dutch Cadastre]

4.5 OPERATIONALIZATION OF CONTROL VARIABLES

As a final implementation aspect, the set of route-level control variables was quantified. For the quantification of the control variables, attributes assigned to the bicycle road network provided by the Dutch Cyclist Union were used. Each road segment in the bicycle road network held attributes describing the infrastructural characteristics, environmental characteristics, and descriptors of the traffic volume surrounding a road segment. Table 4.7 presents an overview of the attributes provided by the Dutch Cyclist Union.

Table 4.7: Overview of attributes provided by the Dutch Cyclist Union

Attribute Dutch Cyclist Union	Scale
Infrastructural characteristics	
Lighting conditions	Nominal
Road quality	Nominal
Type of bicycle road	Nominal
Environmental characteristics	
Slope	Nominal
Traffic volume	
Traffic volume around bicycle road	Nominal

All of the attributes presented in Table 4.7 held nominal values describing each road segment for that given attribute. In case of missing data for a certain attribute, values from neighboring road segments were selected to represent the attribute value for that given road segment. In order to assign an aggregated value of the route-level characteristics to an observed route and its shortest/fastest equivalent, the control variables were operationalized as a proportion over an entire route. Appendix A elaborates on the exact formation of the control variables based on attributes of the Dutch Cyclist Union. Not based on attributes provided by the Dutch Cyclist Union is the variable describing the number of crossings per route. Identification of crossings in the bicycle road network was based on the number of incident edges for a vertex. To avoid side streets being identified as crossings, a vertex was considered a crossing when it had more than three incident edges. Table 4.8 provides an overview of the complete set of route-level control variables that was used in the modelling phase.

Table 4.8: Overview of the developed control variables

Control variable	Code
Individual characteristics	
Gender	[0 = female, 1 = male]
Age	[number]
Infrastructural characteristics	
Number of crossings	[number]
Proportion of good lighting	[%]
Proportion of good road quality	[%]
Proportion of shared road	[%]
Environmental characteristics	
Proportion of slope $> 1\%$	[%]
Traffic volume	
Proportion of high traffic volume	[%]

5 RESULTS

This chapter provides an overview of the results that were generated by implementing the developed methodology. Descriptive statistics of the study sample are presented in Section 5.1, whereafter Section 5.2 addresses the multilevel character of the sample. Section 5.3 elaborates on the set of regression analyses that were conducted to model the effect of weather condtions and the degree of built environment shelter on cyclist route choice. At the end of the chapter, the results of the regression analyses are synthesized.

5.1 DESCRIPTIVE RESULTS

5.1.1 Demographic characteristics

Implementation of the developed methodology on the study area of Tilburg resulted in a study population of 322 cyclists. Table 5.1 shows the demographic characteristics of the study population. A slight majority of the cyclists were women with a percentage of 54%, compared to men covering 46% of the population. With regard to the distribution of age over the population, more than 80% of the cyclists were between 40 and 69 years old.

Table 5.1: Demographic characteristics of the study population (N = 322)

	Share [%]
Gender	
Female	54%
Male	46%
Age	
24-29	4.3%
30-39	14.9%
40-49	38.4%
50-59	38.1%
60-69	6.1%

The study population has been compared to research by the Netherlands Institute for Transport Policy Analysis (KiM), which are representative for the entire country [Kennisinstituut voor Mobiliteitsbeleid, 2019]. The findings of KiM state that the larger part of the users of electric bicycles consists of people of 50 years or older, and women make more use of electric bicycles than men. As the observed travel data provided by the B-Riders projects comprise a substantial number of users of electric bicycles, the study population of this thesis corresponded to the findings of KiM.

5.1.2 Meteorological characteristics

The meteorological factors that were assigned to the route model can be divided into two categories: continuous and nominal/categorical variables. Table 5.2 presents an overview of the descriptive characteristics of the meteorological factors that are measured on a continuous scale. It can be concluded that there was considerable variation in the values for each of the four variables. Regarding the variable temperature, the route model lacks routes that were conducted under substantial freezing temperatures, with the minimum observed temperature just below $0°C$.

Table 5.2: Descriptive characteristics of continuous meteorological factors

Variable	Mean	Maximum	Minimum
Average windspeed $[m/s]$	3.7	11.2	0.0
Precipitation sum $[mm]$	1.0	71.6	0.0
Solar radiation $[J/cm^2]$	81.0	325.8	0.0
Temperature [°C]	14.3	33.2	-0.1

Table 5.3 shows the frequencies of occurrence for each of the categories of the nominal/categorical meteorological variables (total number of routes; $N = 18424$). The majority of the routes were conducted during daytime, with full daylight. However, a considerable number of routes took place during the twilight period, or under conditions without any daylight. The distribution of routes over the different categories of the variable wind direction was more equal. For the majority of the routes cyclists experienced side wind, while the categories back wind and face wind account for almost a third and fifth of the routes respectively. The number of routes conducted under changing wind conditions was considerably lower.

The distribution of routes for the variables fog and ice formation was heavily skewed, although the route model still held a considerable number of routes were fog was measured. As could be concluded from Table 5.2, the route model did not include observations of substantial freezing temperatures. Consequently, the number of routes conducted under conditions where ice formation was measured were limited. Due to similar reasons, the route model does not include any observations of snowfall. Since both variables lack a reasonable amount of variation, ice formation and snowfall were not considered for further analyses.

Table 5.3: Descriptive characteristics of nominal/categorical meteorological factors

Variable	Category	N	%
Daylight conditions	No daylight	1684	9.14
	Full daylight	15708	85.26
	Twilight	1032	5.60
Fog	No fog	18346	99.58
	Fog	78	0.42
Ice formation	No ice formation	18419	99.97
	Ice formation	5	0.03
Snowfall	No snowfall	18424	100.00
	Snowfall	0	0.00
Wind direction	Back wind	5024	27.27
	Face wind	3664	19.89
	Side wind	9243	50.17
	Changing wind	493	2.67

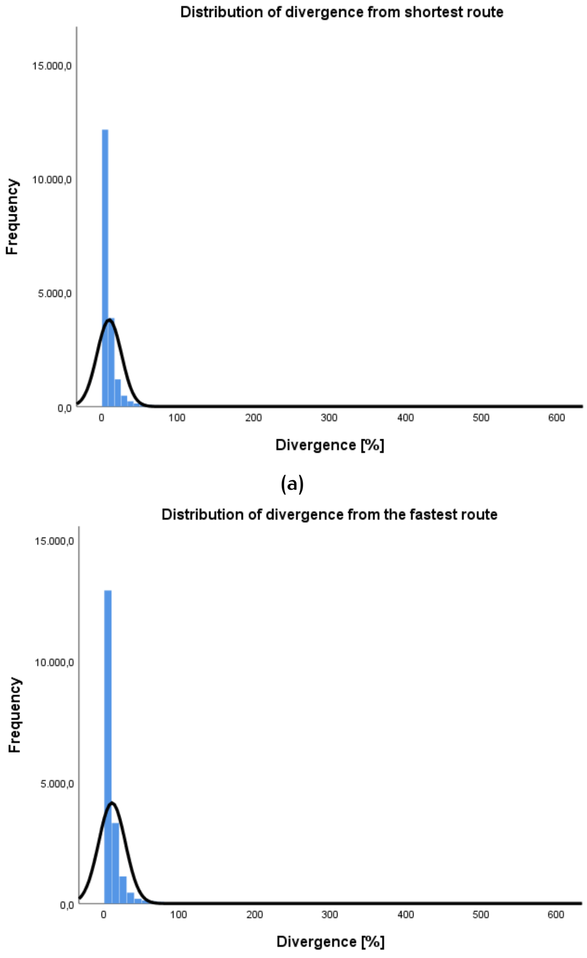
5.1.3 Route characteristics

In contrast to the meteorological characteristics of the route model, observed and shortest/fastest routes did differ in route characteristics. Differences between observed and shortest/fastest routes were initially described by the extent of divergence in terms of distance and travel time respectively, the dependent variables for the majority of the further regression analysis. Table 5.4 shows the descriptive characteristics of the extent of divergence from the shortest/fastest route over the study sample.

Table 5.4: Descriptive characteristics of the extent of divergence from the shortest/fastest route

	Divergence from shortest route [%]	**Divergence from fastest route [%]**
Mean	9.50	10.43
Maximum	539.27	556.63
Minimum	0.00	0.00

On average the divergence from the fastest routes was slightly larger than the divergence from the shortest routes, with an accompanying higher maximum value. Figure 5.1 shows that both types of divergence variables follow a similar distribution. The large majority of the observed routes did not diverge more than 10% from the shortest or fastest route, and approximately 90% of the observed routes did not diverge more than 20%. These findings confirmed the utilitarian character of the study sample, where minimization of distance and travel time was expected to be highly influential.



(a)

(b)

Figure 5.1: Distribution of divergence from shortest route (a) and fastest route (b)

As a first indication of the variation of the shelter and control variables over the route model, a set of Paired Samples t-test was executed to assess the differences between the variable mean of an observed route and the variable mean of its shortest/fastest equivalent. Table 5.5 presents a summary of the output of the different tests, where the insignificant differences are highlighted in bold.

Table 5.5: Route characteristics of the study sample (N = 18424)

Variable	Mean observed (% of length)	Difference shortest (sig.)	Mean observed (% of travel time)	Difference fastest (sig.)
Shelter variables				
% mean shelter > 25%	8.57	-1.04 (0.00)	8.66	-0.85 (0.00)
% max shelter > 50%	23.31	-1.63 (0.00)	23.83	-1.11 (0.00)
% tree density > 50%	12.18	1.24 (0.00)	11.71	**0.07 (0.13)**
Control variables				
% good lighting	93.27	-0.15 (0.01)	93.48	-0.23 (0.00)
% good road quality	74.94	2.64 (0.00)	74.08	**0.001 (0.98)**
% high traffic volume	38.93	4.46 (0.00)	35.47	1.07 (0.00)
% shared road	48.71	-1.60 (0.00)	48.40	0.44 (0.00)
% slope > 1%	2.68	-0.31 (0.00)	2.89	**0.02 (0.27)**
No. of crossings	11.97	0.57 (0.00)	11.97	0.55 (0.00)

The differences between the observed and shortest routes regarding the shelter variables showed that on average cyclists seek for less shelter by buildings in comparison to the shortest routes. For both the percentage of mean shelter > 25% and maximum shelter > 50% the differences between the mean values were negative. On the contrary, cyclists seeked on average more vegetational shelter as the difference between the mean values for the percentage of tree density > 50% is positive. When comparing the observed routes with the fastest equivalents similar, but less pronounced, results were found. However, the difference between the mean values for the percentage of tree density > 50% was statistically insignificant.

The limited average differences in shelter variables between observed and shortest or fastest routes were confirmed by the observed distributions of the differences. Figures 5.2 and 5.3 present the distributions of the differences in maximum shelter > 50% between observed and shortest routes, and observed and fastest routes respectively. The normally distributed values have high peaks around the zero point. Whereas the distributions for the other two shelter variables show similar characteristics, relatively little variation in the degree of shelter between observed and shortest/fastest was observed for the large majority of the route model.

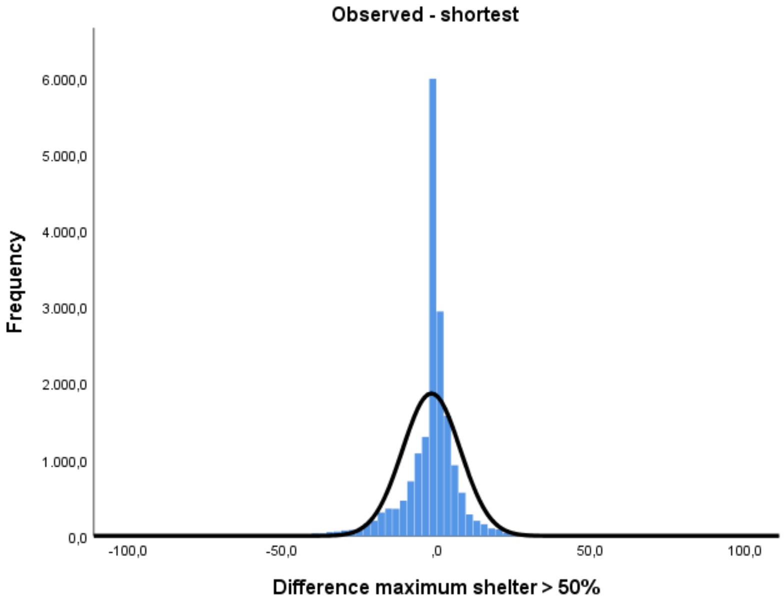


Figure 5.2: Distribution of differences in maximum shelter between observed and shortest routes

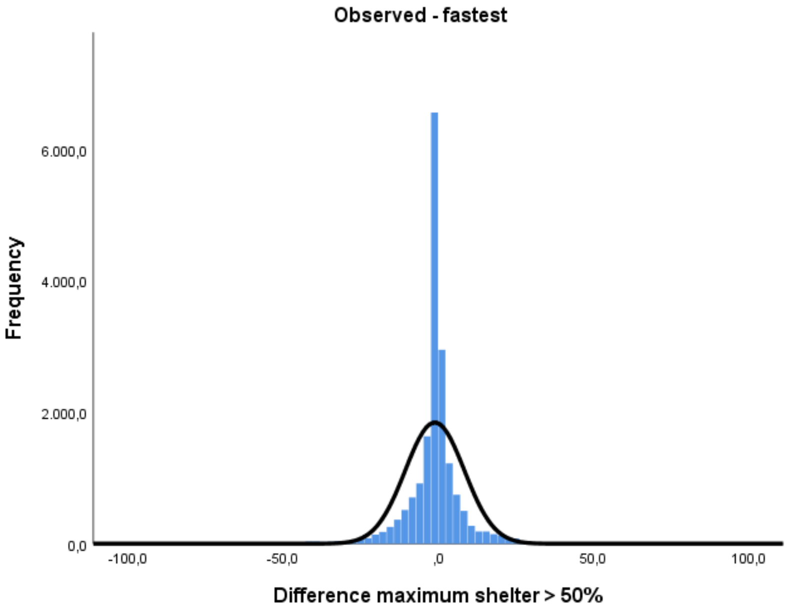


Figure 5.3: Distribution of differences in maximum shelter between observed and fastest routes

With regard to the control variables, the observed routes held on average a higher percentage of good road quality and high traffic volume then their shortest equivalents, as well as more crossings. These results suggest that cyclists on average did not felt disturbance from a high traffic volume or crossings along a route. Significant negative differences were noted for the variables percentage of good lighting, percentage of shared road, and percentage of slope $> 1\%$. Tests comparing the observed and fastest routes gave different results than the comparison between the observed and shortest routes. First of all, the differences for the variables percentage of good road quality and slope $> 1\%$ were found statistically insignificant. Furthermore, the positive difference for the variable percentage of shared road indicated that on average cyclists seeked for routes with a higher percentage of shared road.

In general, the differences between the mean values for the observed routes and the fastest routes were less pronounced then the differences between the observed routes and shortest routes. These results suggest that on average cyclists selected routes that were approaching the configuration of fastest routes.

5.2 MULTILEVEL ANALYSES

To address the nested character of the route model, a set of parsimonious multilevel linear regression analyses was conducted to assess the variance in the extent of divergence from the shortest/fastest route between cyclists. The aim of these analyses was to examine whether a multilevel approach was required to model the effects of the meteorological and shelter factors on the dependent variable, as well as the potential effect of the individual-level variables age and gender on the variance in the extent of divergence. For both age and gender, fixed and random effects on the extent of divergence were estimated to get examine the average effect over the entire study population and the variance in effects between cyclists respectively.

Table 5.6 presents the estimates of the fixed and random effects of age and gender on the extent of divergence from both the shortest and fastest routes. The intercept of the fixed effects indicated that the average divergence from the shortest route over the entire sample of routes is nearly identical to the average divergence from the fastest route (8.38 vs. 8.51). Overall the effect of both age and gender on the divergence from the shortest/fastest route was found statistically insignificant. For the

Table 5.6: Fixed and random effects of variables age and gender

Parameter	Estimate shortest (sig.)	Estimate fastest (sig.)
Fixed effects		
Intercept	8.38 (0.00)	8.51 (0.00)
Age	0.05 (0.13)	0.07 (0.08)
Gender	-0.56 (0.40)	-0.39 (0.61)
Random effects		
Intercept	2.75 (0.69)	8.62 (0.46)
Age	0.01 (0.07)	0.01 (0.24)
Gender	*	*

* Estimate could not be made

general population, the variances between cyclists in the divergence from the shortest and fastest path were 16.7 (standard error (SE) = 2.50) and 23.48 ($SE = 3.45$) respectively in the null model (not included in Table 5.6), which indicates considerable heterogeneity in the extent of divergence between cyclists. Adjusting for age and gender, the random intercepts presented in Table 5.6 indicated insignificant between-cyclist variances for the extent of divergence from the shortest and fastest route. No significant variance between cyclists in the effect of age on diverging from the shortest/fastest route was found, while the effect gender remains unexplained.

From the estimated null model, it was concluded that the extent of divergence from the shortest/fastest route varied significantly between cyclists, meaning that a multilevel approach would be appropriate for further regression analyses. However, as adjustment for the individual-level variables age and gender resulted in insignificance of the between-cyclist divergence, further multilevel approaches were not carried out. As the effects of age and gender on the extent of divergence were not significant over the general population, both variables were disregarded for the further regression analyses described in Section 5.3.

5.3 MODEL ESTIMATION

5.3.1 Model 1: meteorological variables

To model the effect of the established meteorological, shelter, and control variables on cyclist route choice, a set of four linear regression models was estimated having the divergence from the shortest/fastest route as the dependent variable. In the first model, meteorological and control variables were included as predictors.

Prior to the estimation of the regression model, a set of one-way ANOVAs were performed to analyze differences in mean divergence from the shortest/fastest route for the categorical meteorological variables daylight conditions and wind direction. Table 5.7 shows a summary of the output generated by the ANOVAs.

Table 5.7: Output summary of the one-way ANOVAs for the variables daylight conditions and wind direction

Variable	F-value shortest (sig.)	F-value fastest (sig.))
Daylight conditions	13.76 (0.00)	22.61 (0.00)
Wind direction	0.57 (0.64)	0.61 (0.61)

The one-way ANOVA to test the differences in mean divergence from the shortest/fastest route between the three types of daylight conditions, showed that there was significant variation between the different categories in terms of divergence from shortest route ($F(2, 18421) = 13.76, p = 0.00$) and the fastest route ($F(2, 18421) = 22.61, p = 0.00$). Post-hoc comparisons using the Tukey HSD test indicated that the mean divergence from the shortest routes were significantly smaller during twilight conditions ($Mean = 7.14$, standard deviation (SD) $= 11.21$), compared to full daylight ($Mean = 9.72, SD = 16.18$) and no daylight conditions ($Mean = 8.88, SD = 17.54$). However, the mean divergence did not significantly differ between full daylight conditions and no daylight conditions ($p = 0.10$). Using the divergence from the fastest route as dependent variable, the mean divergence differed significantly between full daylight conditions ($Mean = 7.11, SD = 11, 10$), no daylight conditions ($Mean = 9.56, SD = 18.61$), and twilight conditions ($Mean = 10.74, SD = 17.93$).

The F-values for the variable wind direction presented in Table 5.7 indicate that no significant variation was found in terms of divergence from shortest route ($F(3, 18420) = 0.57, p = 0.00$) and the fastest route ($F(3, 18420) = 0.61, p = 0.61$) for the different categories of wind direction. Therefore, the variable wind direction was not included as independent variable in further regression models.

Table 5.8: Output summary for regression model 1

Variable	**B shortest (sig.)**[1]	**B fastest (sig.)**[2]
Average windspeed	0.13 (0.04)	0.20 (0.01)
Fog	-0.33 (0.86)	-1.23 (0.54)
Precipitation sum	0.03 (0.25)	0.02 (0.52)
Solar radiation	0.01 (0.00)	0.01 (0.00)
Temperature	0.09 (0.00)	0.11 (0.00)
Daylight conditions		
No daylight	-0.06 (0.89)	-0.09 (0.85)
Twilight	-1.21 (0.02)	-1.88 (0.00)
Control variables		
% good lighting	-0.26 (0.00)	-0.23 (0.00)
% good road quality	0.04 (0.00)	-0.05 (0.00)
% high traffic volume	0.08 (0.00)	0.03 (0.00)
% shared road	0.05 (0.00)	0.05 (0.00)
% slope $> 1\%$	-0.18 (0.00)	-0.27 (0.00)
No. of crossings	0.21 (0.00)	0.39 (0.00)

[1] $R^2 = 4.2\%$
[2] $R^2 = 2.8\%$

After conducting the one-way ANOVAs, the first regression model was estimated. Table 5.8 presents an overview of the unstandardized B-coefficients and significance values. Adjusted for the effects of the control variables, Table 5.8 indicates that the variables average windspeed, temperature, solar radiation, and dummy variable twilight (reference category: full daylight) were significant predictors of the extent of divergence from the shortest route. An increase in average windspeed of 1 m/s would result in 0.13 % extra divergence from the shortest route, while an increase of 1 $^\circ C$ in temperature would result in 0.09 % extra divergence. Although significant, the effect of both meteorological variables were therefore not substantial considering mean values of 11.2 m/s and 14.3 $^\circ C$ for the average windspeed and temperature respectively (Table 5.2). Similarly, with a B-coefficient of 0.01 the effect of solar radiation on the divergence from the shortest route is relatively small ($Mean = 81.0$). Finally, cycling under twilight conditions would lead to a decrease of 1.21% in the divergence from the shortest route compared to cycling under full daylight conditions. Overall, 4.2% of the variation in the extent of divergence from the shortest route could be explained by the meteorological and control variables in regression model 1 (R^2-value of 4.2%).

The regression model explaining the divergence from the fastest route gave comparable results regarding the significance of the meteorological variables. However, the positive effects of the average windspeed and temperature were slightly more substantial than in the model explaining the divergence from the shortest route. Additionally, the negative effect of cycling under twilight conditions is larger when explaining the divergence from the fastest route ($B = -1.88$). However, the overall explanatory strength of this model was lower ($R^2 = 2.8\%$). For both models, the effects of the precipitation sum, and cycling under fog or no daylight conditions were not statistically significant. Therefore, these variables were not included for further analyses.

5.3.2 Model 2: shelter variables

The second set of linear regression models that were estimated included the three shelter variables and the control variables. The aim of this analysis was to model the associations between the shelter variables and the extent of divergence from the shortest/fastest route, adjusted for the effects of the control variables. Table 5.9 presents an overview of the unstandardized B-coefficients and significance values for the second regression model.

Table 5.9: Output summary for regression model 2

Variable	**B shortest (sig.)**[1]	**B fastest (sig.)**[2]
% mean shelter > 25%	-0.06 (0.04)	-0.16 (0.00)
% max shelter > 50%	-0.16 (0.00)	-0.06 (0.00)
% tree density > 50%	0.01 (0.41)	-0.05 (0.01)
Control variables		
% good lighting	-0.24 (0.00)	-0.21 (0.00)
% good road quality	0.04 (0.00)	-0.05 (0.00)
% high traffic volume	0.08 (0.00)	0.01 (0.27)
% shared road	0.07 (0.00)	0.05 (0.00)
% slope > 1%	-0.16 (0.00)	-0.26 (0.00)
No. of crossings	0.21 (0.00)	0.45 (0.00)

[1] $R^2 = 4.6\%$
[2] $R^2 = 2.8\%$

Regarding the model explaining the divergence from the shortest route, the strongest effect came from the percentage of maximum shelter > 50%. An increase of 1% in the percentage of maximum shelter > 50% for an observed route compared to the shortest route, would lead to a decrease in divergence of 0.16%. While the percentage of mean shelter > 25% also had a significant negative effect on the dependent variable, the association was weaker. The effect of the percentage of tree density > 50% on the divergence from the shortest route was found insignificant. Therefore, this variable was not included in further regression models explaining the extent of divergence from the shortest route.

When explaining the divergence from the fastest route, the percentage of mean shelter > 25% was found to be a stronger predictor. An increase of 1% in the percentage of mean shelter > 25% for an observed route compared to the fastest route, would lead to a decrease in divergence of 0.16%. The effect of the percentage of maximum shelter > 50% was also found significant, but less substantial. In contrast with the model explaining the extent of divergence from the shortest route, the percentage of tree density > 50% had a significant effect on the divergence from the fastest route. However, the effect could not be considered as substantial: an increase of 1% in the percentage of tree density > 50% for an observed route compared to the fastest route, would lead to a decrease in divergence of 0.05%. Overall, regression model 2 explained the variation in the extent of divergence from the shortest and fastest route for 4.6% and 2.8% respectively.

The negative associations between the significant shelter variables and the extent of divergence in both regression models route suggested that cyclists did not diverge from the shortest or fastest route to obtain a higher degree of shelter. Only adjusted for the effect of the control variables, this would indicate a preference for generally more open routes when diverging from the shortest/fastest route. However, the overall results show limited effects for each of the shelter variables on the extent of divergence from the shortest or fastest route. Based on the distributions presented in Figures 5.2 and 5.3, the found effects were considered as reasonable.

5.3.3 Model 3: meteorological & shelter variables

After estimating the individual effects of the meteorological and shelter variables on the extent of divergence from the shortest/fastest route, a third set of linear regression models was estimated in which both types of predictors were included. The aim of this set of models was to examine whether adjustment for each other's effects, as well as for the effects of the control variables, would lead to different associations for the included meteorological and shelter variables with the dependent variable. Table 5.10 summarizes the output of both regression analyses.

Table 5.10: Output summary for regression model 3

Variable	B shortest (sig.)[1]	B fastest (sig.)[2]
Meteorological variables		
Average windspeed	0.13 (0.05)	0.21 (0.01)
Solar radiation	0.01 (0.00)	0.01 (0.00)
Temperature	0.08 (0.00)	0.11 (0.00)
Daylight conditions		
Twilight	-1.19 (0.02)	-1.85 (0.00)
Shelter variables		
% mean shelter > 25%	-0.06 (0.00)	-0.16 (0.00)
% max shelter > 50%	-0.15 (0.00)	-0.06 (0.00)
% tree density > 50%	*	-0.05 (0.01)
Control variables		
% good lighting	-0.24 (0.00)	-0.21 (0.00)
% good road quality	0.04 (0.00)	-0.05 (0.00)
% high traffic volume	0.08 (0.00)	0.01 (0.23)
% shared road	0.07 (0.00)	0.05 (0.00)
% slope > 1%	-0.16 (0.00)	-0.27 (0.00)
No. of crossings	0.21 (0.00)	0.39 (0.00)

[1] $R^2 = 5.1\%$
[2] $R^2 = 3.4\%$
* Effect was not estimated

The third set of regression models showed similar magnitudes of associations when including all types of variables into one model, compared to the regression models that were estimated in Section 5.3.1 and Section 5.3.2. The minimal differences in effect strengths on the divergence from the shortest/fastest route indicate that the meteorological and shelter predictors were very weakly associated with each other. Including all types of predictors in a single regression model increased the predictive power compared to the previously estimated regression models ($R^2 = 5.1\%$ and $R^2 = 3.4\%$).

5.3.4 Model 4: interaction effects

For the final set of linear regression models that were estimated to explain the extent of divergence from the shortest/fastest route, interaction effects between each meteorological and shelter variable were included as independent variables. Separate models were estimated in which the effect of one interaction was adjusted for the effects of the meteorological, shelter, and control variables. Table 5.11 and 5.12 show the effects of each interaction on the divergence from the shortest and fastest route respectively, as well as the change in explanatory strength of the models caused by adding an interaction effect. It should be noted that for the model explaining the extent of divergence from the shortest routes, the shelter variable percentage of tree density $> 50\%$ was not included, nor were any interactions with this variable due to found insignificant effects in previous models.

Table 5.11: Output summary for regression model 4.1: observed vs. shortest

Interaction	**B-value (sig.)**	R^2 **change**[1]
% mean shelter $> 25\%$		
x average windspeed	0.02 (0.08)	0.00%
x solar radiation	0.00 (0.65)	0.00%
x temperature	0.00 (0.76)	0.00%
x twilight	**-0.23 (0.03)**	0.00%
% max shelter $> 50\%$		
x average windspeed	0.01 (0.12)	0.00%
x solar radiation	0.00 (0.13)	0.00%
x temperature	0.00 (0.53)	0.00%
x twilight	-0.08 (0.16)	0.00%

[1] Compared to R^2 in regression model 3 where $R^2 = 5.1\%$

Table 5.11 indicates that only the interaction effect between the percentage of mean shelter $> 25\%$ and cycling under twilight conditions had a significant effect on the extent of divergence from the shortest route. The B-coefficient of -0.23 suggests that when cycling in twilight conditions, the divergence from the shortest route would decrease with an additional 0.23% for every increase of 1% in the percentage of mean shelter $> 25\%$, compared to cycling under full daylight conditions (B-value for twilight = -1.36, $p = 0.01$). The found effect suggested that cyclists preferred a lower degree of mean shelter $> 25\%$ along a route when cycling during the twilight period. Although substantial, the found effect did not add explanatory strength to the regression model compared to regression model 3. All other interaction effects between the meteorological and shelter variables were found insignificant.

For the model explaining the extent of divergence from the fastest route (Table 5.12), the interaction effects between the percentage of mean shelter $> 25\%$ and temperature, and between the percentage of maximum shelter $> 50\%$ and average windspeed were found significant. An increase of 1% for the mean shelter $> 25\%$ along an observed route, controlling for the fact that the temperature cannot differ between an observed or shortest/fastest route, would lead to an additional 0.01% decrease in the divergence from the fastest route. Therefore, the interaction effect was considered as relatively weak. Similarly, a B-coefficient of 0.02 suggests a small influence of the interaction between the percentage of maximum shelter $> 50\%$ and average windspeed on the divergence from the fastest route. Consequently, both significant interaction effects did not lead to an increase in the overall explanatory strengths of the models. Finally, neither of the interaction effects between a meteorological factor and the percentage of tree density $> 50\%$ was found significant.

In general, inclusion of interaction effects between meteorological and shelter variables did not add substantial value to the explanatory strengths of the regression models. As the general magnitude of the significant interaction effects were relatively low, the regression models suggest that the shelter variables could not be considered as good predictors of the effect of meteorological variables on the extent of divergence from the shortest/fastest route. The sole exception was the interaction effect between the variables twilight and the percentage of mean shelter

Table 5.12: Output summary for regression model 4.2: observed vs. fastest

Interaction	B-value (sig.)	R^2 change[1]
% mean shelter > 25%		
x average windspeed	0.01 (0.57)	0.00%
x solar radiation	0.00 (0.61)	0.00%
x temperature	**-0.01 (0.03)**	0.00%
x twilight	-0.18 (0.09)	0.00%
% max shelter > 50%		
x average windspeed	**0.02 (0.05)**	0.00%
x solar radiation	0.00 (0.06)	0.00%
x temperature	-0.00 (0.06)	0.00%
x twilight	-0.04 (0.53)	0.00%
% tree density > 50%		
x average windspeed	-0.01 (0.33)	0.00%
x solar radiation	0.00 (0.40)	0.00%
x temperature	0.00 (0.35)	0.00%
x twilight	0.03 (0.68)	0.00%

[1] Compared to R^2 in regression model 3 where $R^2 = 3.4\%$

> 25%, which added substantially to the individual effect of twilight on the extent of divergence. In other words, cyclists in the study area generally do not chose between shorter or longer routes based on the degree of shelter from weather conditions offered by the built environment. However, the found effects do not fully exclude the possibility of route choice based on the degree of shelter: cyclists could have routes similar in length or travel duration compared to a shortest or fastest alternative, but with different shelter characteristics. This perspective on route choice is further elaborated on in Section 5.3.5.

5.3.5 Route choice in terms of shelter

The final set of linear regression models that was estimated aimed to examine whether cyclists adapted their route choice to the degree of shelter offered by the built environment, based on the weather conditions under which a route was conducted. In this regard, three separate linear regression models were estimated using the percentage of mean shelter > 25%, maximum shelter > 50%, and tree density > 50% as dependent variable, and the meteorological variables as predictors. The effects of the meteorological variables on the dependent variables were adjusted for the effects of the control variables in each of the models.

To analyze variation in the mean values for each shelter variable over the categorical meteorological variables daylight conditions and wind direction, a set of one-way ANOVAs were carried out prior to the estimation of the regression models. Table 5.13 summarizes the generated output of the different one-way ANOVAs.

When comparing observed routes with shortest routes, the output of the one-way ANOVAs indicated significant variation in mean values for the variables percentage of mean shelter > 25% ($F(2, 18421) = 3.64, p = 0.02$) and percentage of tree density > 50% ($F(2, 18421) = 8.45, p = 0.00$). Post-hoc comparisons using the Tukey HSD test showed that the variation in mean values was primarily caused by significant differences between the categories full daylight ($Mean = -1.09, SD = 6.86$) and no daylight ($Mean = -0.69, SD = 5.97$) for the variable percentage of mean shelter > 25%. For the variable percentage of tree density > 50%, significantly higher mean values were found for the category twilight ($Mean = 2.09, SD = 8.29$), compared to

Table 5.13: Output summary of the one-way ANOVAs for the variables daylight conditions and wind direction in relation to each shelter variable

Variable	F-value shortest (sig.)	F-value fastest (sig.)
Mean shelter $> 25\%$		
Daylight conditions	**3.64 (0.02)**	0.27 (0.76)
Wind direction	0.88 (0.45)	0.97 (0.41)
Maximum shelter $> 50\%$		
Daylight conditions	2.93 (0.06)	0.85 (0.43)
Wind direction	**3.15 (0.02)**	1.06 (0.37)
Tree density $> 25\%$		
Daylight conditions	**8.45 (0.00)**	0.09 (0.91)
Wind direction	0.66 (0.58)	1.39 (0.24)

full daylight ($Mean = 1.17, SD = 7.13$) and no daylight ($Mean = 1.40, SD = 7.23$). These findings suggest that on average, cyclists opted for a lower degree of mean building shelter $> 25\%$ along a route during conditions without daylight, while a higher degree of tree density $> 50\%$ was desired when a route was conducted under twilight conditions. In the ANOVAs examining the observed vs. fastest route model, no significant variation in mean values was found for any of the shelter variables.

Regarding the variation in mean values between the different wind direction categories, significant variation was found for the shelter variable percentage of maximum shelter $> 50\%$. Output from the post-hoc Tukey HSD test showed significant differences in variation in mean values between the categories face wind ($Mean = -1.24, SD = 9.32$) and side wind ($Mean = -1.75, SD = 9.47$). Variation in mean values between the other categories were found insignificant. These findings suggest that cyclists on average seeked for a lower percentage of maximum shelter $> 50\%$ in an observed route compared to the shortest route when cycling with side wind, then when cycling under face wind conditions. Similar to the findings for the variable daylight conditions, no significant variation was found in mean values for any shelter variable describing the difference between observed and fastest routes.

Based on the output of the one-way ANOVAs, the meteorological variables daylight conditions and wind direction were only included in further regression models explaining the difference in shelter variables between the observed and shortest routes where significant variation in mean values was found between the different categories of the variables. For regression models explaining the difference in shelter variables between the observed and fastest routes, neither of the two variables were included.

Table 5.14 summarizes the output of the regression models analyzing the effects on differences between the observed and shortest route for each shelter variable. For the model explaining the difference in the percentage of mean shelter $> 25\%$ between observed and shortest routes, a significant effect was found for the average windspeed. According to the B-coefficient of -0.09, an increase in average windspeed of $1 m/s$ would lead to a decrease of 0.09% in the percentage of mean shelter $> 25\%$ along an observed route compared to the shortest route. Considering that the minimum and maximum average windspeeds varied between 0 and 11.2 (Table 5.2), the association was considered as weak. The effects of the other meteorological variables were insignificant. Consequently, the overall explanatory strength of the model was relatively low with an R^2-value of 2.1%.

On the contrary, the overall explanatory strength of the model explaining the percentage of maximum shelter $> 50\%$ was relatively high with an R^2-value of 17.5%. However, a substantial part of the explanatory strength was caused by strong associations with the control variables percentage of good lighting, percentage of slope $> 1\%$, and the number of crossings along a route.

Table 5.14: Summary shelter regression model 5.1: observed vs. shortest

Variable	**Mean shelter $>$ 25%[1]**	**Max shelter $>$ 50%[2]**	**Tree density $>$ 50%[3]**
	B-value (sig.)	**B-value (sig.)**	**B-value (sig.)**
Meteorological variables			
Average windspeed	**-0.09 (0.00)**	-0.06 (0.11)	0.01 (0.81)
Solar radiation	0.00 (0.59)	0.00 (0.54)	**-0.00 (0.01)**
Temperature	-0.02 (0.12)	**-0.05 (0.00)**	-0.01 (0.19)
Daylight conditions			
No daylight	0.31 (0.09)	*	-0.04 (0.83)
Twilight	0.15 (0.51)	*	0.41 (0.07)
Wind direction			
Face wind	*	0.32 (0.09)	*
Side wind	*	-0.21 (0.17)	*
Changing wind	*	-0.36 (0.38)	*
Control variables			
% good lighting	0.01 (0.19)	0.21 (0.10)	0.04 (0.00)
% good road quality	0.02 (0.00)	-0.02 (0.00)	0.06 (0.00)
% high traffic volume	-0.09 (0.00)	0.00 (0.11)	-0.01 (0.00)
% shared road	0.02 (0.00)	0.10 (0.00)	0.00 (0.91)
% slope $>$ 1%	-0.04 (0.01)	0.17 (0.00)	-0.76 (0.00)
No. of crossings	0.17 (0.00)	0.79 (0.00)	-0.12 (0.00)

[1] $R^2 = 2.1\%$
[2] $R^2 = 17.5\%$
[3] $R^2 = 10.6\%$
* Not included as predictor based on output of one-way ANOVA

Regarding the meteorological variables, only the effect of temperature was found significant. As an increase of 1 °C would lead to a decrease of 0.05% in the percentage of maximum shelter $>$ 50%, the effect was not considered as substantial.

Finally, the model explain the percentage of tree density $>$ 50% only showed a significant influence of the variable solar radiation. Nevertheless, with a B-coefficient approximating 0, the effect was negligible. The explanatory strength of the model was substantial ($R^2 = 10.6\%$), mainly caused by a strong association between the percentage of slope $>$ 1% and the dependent variable.

Table 5.15 provides a summary of the output of the regression models analyzing the effects on differences between the observed and fastest route for each shelter variable. For the model explaining the percentage of mean shelter $>$ 25%, significant effects were found for the meteorological variables temperature and solar radiation. However, with a B-coefficient of -0.03 for temperature, and a B-coefficient approximating 0 for solar radiation, both effects were considered relatively small. Combined with the associations for the control variables, the explanatory strength of the model just exceeded 5%.

A much higher explanatory strength was found for the model explaining the percentage of maximum shelter $>$ 50% with an R^2-value of 12.4%, most dominantly caused by strong associations for the control variables percentage of slope $>$ 1% and number of crossings along a route. The only significant meteorological effect was found for the variable solar radiation. As the B-coefficient approximated 0, the effect was considered negligible.

Table 5.15: Summary shelter regression model 5.2: observed vs. fastest

Variable	Mean shelter $> 25\%$[1]	Max shelter $> 50\%$[2]	Tree density $> 50\%$[3]
	B-value (sig.)	**B-value (sig.)**	**B-value (sig.)**
Meteorological variables			
Average windspeed	-0.04 (0.11)	0.04 (0.28)	0.07 (0.02)
Solar radiation	**0.00 (0.01)**	**0.00 (0.03)**	-0.00 (0.44)
Temperature	**-0.03 (0.00)**	-0.02 (0.20)	0.01 (0.43)
Control variables			
% good lighting	0.04 (0.00)	0.11 (0.00)	0.08 (0.00)
% good road quality	0.02 (0.00)	0.01 (0.30)	0.03 (0.00)
% high traffic volume	-0.09 (0.00)	-0.08 (0.00)	-0.06 (0.00)
% shared road	-0.03 (0.00)	0.04 (0.00)	0.02 (0.00)
% slope $> 1\%$	0.07 (0.00)	0.37 (0.00)	-0.38 (0.00)
No. of crossings	0.12 (0.00)	0.64 (0.00)	-0.05 (0.00)

[1] $R^2 = 5.3\%$
[2] $R^2 = 12.4\%$
[3] $R^2 = 4.7\%$

The output for the final model explaining the percentage of tree density $> 50\%$ showed a significant effect for the average windspeed, where an increase in average windspeed of $1m/s$ would lead to a decrease of 0.07% in the percentage of tree density $> 50\%$ along an observed route compared to the fastest route. Based on the minimum and maximum average windspeeds presented in Table 5.2, the association was considered as weak. With an R^2-value of 4.7%, the model explaining the percentage of tree density $> 50\%$ had the lowest explanatory strength of the three models.

The results of the set of regression models explaining the difference in shelter variables between the observed and shortest/fastest routes suggested that for the study sample, adaptation of route choice to the degree of shelter provided by the built environment was generally not substantially influenced by the weather conditions under which a route was conducted. Whereas significant effects of meteorological factors were limited, generally found substantial effects for the control variables percentage of slope $> 1\%$ and number of crossings suggests that cyclists rather account for infrastructural and environmental characteristics when adapting the choice of route to the degree of shelter offered by the built environment.

5.4 SYNTHESIS

The aim of estimating the set of statistical models presented in this chapter was to test the general hypothesis that the degree of shelter that is provided by the built environment is an explanatory factor in the extent of divergence from the shortest/fastest route in different weather conditions, as well as to examine the set of hypotheses stated in Section 2.4.3. Estimations of the initial associations between the set of meteorological variables and the extent of divergence from the shortest/fastest route indicated that the cyclists of the study population were most dominantly influenced by average windspeed, temperature, and cycling under twilight conditions. Cyclists would diverge more with higher average windspeeds and temperature, but diverge less during the twilight period.

The estimated associations between the three shelter variables and the extent of divergence from the shortest/fastest route suggested preferences for less sheltered routes when diverging. In relation to the divergence from the shortest route, the percentage of mean shelter $> 25\%$ along a route was found to be the strongest descriptor. However, when diverging from the fastest route, cyclists were more heavily influenced by the percentage of maximum shelter $> 25\%$. Inclusion of interaction effects between the meteorological and shelter variables in general did not show substantial relationships between the effects of both type of variables on the extent of divergence from the shortest/fastest route. Finally, the effects of meteorological factors on adaptation of route choice to the degree of shelter offered by the built environment were limited for all three shelter variables.

To synthesize the results of the estimated statistical models, in general, the cyclists in the study population only diverged from the shortest or fastest route to a limited extent. When diverging, cyclists did not seem to adjust the length of their route to find shelter from weather conditions, nor were weather conditions a substantial reason to adapt the choice of route to the degree of shelter offered by the built environment.

6 CONCLUSION

In this chapter conclusions are drawn based on the results that were produced by implementing the developed methodology on the study area. First of all, the main research question and sub-questions are answered in Section 6.1. Secondly, the obtained results are discussed in Section 6.2. Thirdly, Section 6.3 elaborates on implications of the found results for policy design regarding stimulation of cycling. Finally, the contributions of this thesis to existing literature are mentioned in Section 6.4, whereafter recommendations for future work are provided in Section 6.5.

6.1 ANSWER RESEARCH QUESTIONS

6.1.1 Main research question

The main objective of this thesis was to examine whether the degree of shelter along a route can be considered a factor that mitigates the influence of weather conditions on cyclist route choice. Based on this objective, the main research question for this thesis was defined as follows:

To what extent does the degree of shelter provided by the built environment explain cyclist route choice in different weather conditions?

An elaborate methodology was proposed in which observed travel data, comprising trips made with conventional and electric bicycles, were compared with theoretically optimal alternative routes. Due to the focus on utilitarian cycling, optimal routes were operationalized based on minimization of travel distance and travel time. To model the weather conditions under which a route was conducted, a set of meteorological factors were quantified and spatially modelled to the location of an observed route. As protection from weather has been an underexplored topic in existing cyclist route choice studies, a new method to operationalize the degree of shelter provided by the built environment was developed. This method used aspects from existing theories on street climate design and spatial openness in order to provide a detailed description of the potential shelter along a route. Three different shelter factors were developed, describing the degree of mean building shelter, maximum building shelter, and vegetational shelter in the form of tree density along a route. For each of the meteorological and shelter factors, independent and combined effects on the extent of divergence from the shortest or fastest route were modelled, adjusted for potential effects of other factors related to cyclist route choice. Additionally, a different approach was applied where cyclist route choice was examined in terms of the degree of shelter based on weather conditions.

Application of the developed methodology on the study area of Tilburg indicated initial moderate influences of windspeed, temperature, and cycling under twilight conditions on the choice of route. Regarding the effects of shelter along a route, cyclists in the study sample seemed to seek for a lower degree of shelter by buildings and trees compared to alternative shortest and fastest routes. Whereas the magnitude of the effects differed slightly between the different shelter factors, a general moderate but negative trend was observed. Combining the effects of meteorological and shelter factors showed very limited additional effects on the extent of divergence from the shortest or fastest route. Complementary to these findings, no substantial influences of meteorological factors were found when explaining the adaptation of route choice to the degree of shelter.

The findings in this thesis generally suggest that the degree of shelter along a route cannot be considered a factor that mitigates the influence of weather conditions on cyclist route choice. Therefore, the route choice behavior of (utilitarian) cyclists seems to differ from that of pedestrians, as existing literature used as a base for this thesis indicated high influences of built environment shelter on pedestrian mobility patterns. However, as the results of the implemented methodology were not validated, the found effects within this thesis should not be generalized outside the study area.

6.1.2 Sub-questions

For the development and implementation of the methodology to answer the main research question, a set of sub-questions was established. These sub-questions could be categorized in the following phases of the methodology: *development of the theoretical framework*, *operationalization*, *modelling*, and *validation*. Since the sub-questions that were relevant for the development of the theoretical framework have already been answered in Chapter 2, they are not treated in this section. The answers to each of the sub-questions provide a more elaborate insight in the different aspects that contributed to answering the main research question.

Operationalization: *how can the observed travel data be examined in terms of route choice?*

Route choice by cyclists was examined by comparing observed GPS measurements describing routes with theoretically optimal alternatives for those routes. As this thesis focused on utilitarian cycling, the assumption was made that cyclists initially seek for a minimization of effort or travel duration based on findings in existing literature. Therefore, cyclist route choice was defined as the extent of divergence from the shortest or fastest route.

In order to enable a fair comparison between observed and theoretical routes, the observed travel data was generalized to a bicycle road network. Initially, the GPS measurements had already been mapmatched to the bicycle road network. By constructing a graph over the bicycle road network consisting of edges and connecting vertices, the start and end of the mapmatched observed routes were re-located to match a vertex in the graph. The generalization of the start and end of an observed route allowed for the generation of shortest and fastest equivalents for each observed route. As a result, the route model enabled examination of the observed travel data through comparison of generalized observed routes with their shortest and fastest alternatives.

Operationalization: *how can the different meteorological factors be quantified and spatially modelled?*

To avoid a loss of detail in the representativity of a meteorological factor, an initial quantification decision was made to model the weather conditions under which an observed route was conducted through a decomposed set of individual meteorological factors over integrated weather conditions. The input data for the quantification process came from official weather stations of the KNMI, measured on an hourly scale. By using this temporal scale to quantify the meteorological factors, the assumption was made that cyclists in the study sample decided upon a route based on the weather conditions at the start of a trip.

In order to approximate the weather conditions under which a route was conducted, an IDW interpolation method was applied to estimate the value for a meteorological factor on the start location of route, based on input from the three closest weather stations. For factors measured on a nominal or categorical scale, a nearest neighbor interpolation was implemented.

Operationalization: *how can the degree of shelter provided by the built environment be quantified and spatially modelled?*

Shelter provided by the built environment was quantified and spatially modelled by expansion of a method developed by Anastasiadou et al. [2018]. Their method used aspects from theories to quantify street climate design [Oke, 1988] and spatial openness [Benedikt, 1979] to measure visual openness along a route conducted by cyclists. By reversing the perspective of this method, the degree of shelter from surrounding buildings at a certain location could be computed by accounting for the distance between a cyclist and a building, and the height delta between the cyclist and the height of that building. By determining the height/distance ratio for several points around a given location, two shelter indicators were established: the mean building shelter and the maximum building shelter at that location. The distinguishment between two building shelter factors was based on the idea that fully enclosed streets give were expected to give different mean shelter values than streets having buildings only on one side. However, as streets with buildings on one side could potentially provide enough shelter to cyclists, both factors were developed to determine which would be a better descriptor regarding cyclist route choice. As an additional component, a vegetational shelter factor based on the tree density around a location was included to account for the potential shelter provided by vegetation inside and outside urban areas in the study area.

Whereas each of the shelter factors were quantified on a set of static locations over the study area, route-level values were quantified as a percentage of distance/travel time over a route in which predefined thresholds were met. Consequently, this led to the following factors indicating the degree of shelter provided by the built environment along a route:

- The percentage of mean building shelter $> 25\%$
- The percentage of maximum building shelter $> 50\%$
- The percentage of tree density $> 50\%$

Based on the moderate found effects within this thesis, the suitability of the method for application on cyclist route choice problems remains questionable. The developed method to quantify and spatially model the degree of shelter used aspects from existing literature in which spatial openness and potential shelter are approached from a static point of view. Due to the highly dynamic character of cycling, the degree of shelter might be experienced differently by cyclists.

Modelling: *to what extent do the different meteorological factors influence cyclist route choice?*

Using the extent of divergence from the shortest or fastest route as dependent variable, a set of linear regression models were estimated that included the meteorological factors as predictors. Adjusted for the effect of infrastructural and environmental control variables, moderate associations were found for average windspeed, temperature, and cycling under twilight conditions. Where increases in average windspeed and temperature would lead to a higher extent of divergence from both the shortest or fastest route, a negative effect was found for cycling under twilight conditions compared to cycling in full daylight. Effects of precipitation, fog, wind direction relative to the route direction, were found to be insignificant in each of the estimated models.

Comparing the found results with existing literature, the lack of impact of the factor precipitation is not in line with researches stating negative influences of precipitation on trip distances and durations [Böcker et al., 2016; Helbich et al., 2014]. These findings could be caused by transportation mode choices in case of rain, an aspect that fell outside the scope of this thesis.

The positive effect of temperature does comply with existing literature [Böcker and Thorsson, 2014], suggesting that cyclists in the study population were willing to make longer trips and spend more time on the bicycle on warmer days. However, the found positive attitude towards higher windspeeds conflicts with studies stating that windspeeds cause a decrease in trip lengths and duration [Böcker and Thorsson, 2014; Böcker et al., 2016; Helbich et al., 2014]. The difference in behavior compared to earlier studies could imply that cyclists in the study population were generally indifferent to increases in windspeed when it comes to route choice, as trips were barely conducted under severe wind conditions.

The negative association with cycling under twilight does match findings in existing literature where darkness was identified as a deterring factor for cycling [Böcker et al., 2015]. On the contrary, conditions of no daylight did not influence the route choice of cyclists in the study population, potentially caused by differences in artificial lighting conditions during twilight and nighttime.

Modelling: *to what extent does the degree of shelter provided by the built environment influence cyclist route choice?*

The influence of the degree of shelter provided by the built environment on cyclist route choice was examined through a similar approach used to model the effect of the different meteorological factors. For each of the three developed shelter variables, negative associations were found with the extent of divergence from the shortest or fastest route. When diverging from the shortest route, the percentage of maximum shelter $> 50\%$ along a route was the strongest descriptor. On the other hand, the percentage of mean shelter $> 25\%$ had the strongest effect on the extent of divergence from the fastest route. To compare the suitability of the two variables describing building shelter as a descriptor of the degree of shelter provided by the built environment, the varying strengths in effects for the two variables provided inconclusive evidence.

The found effects suggest that cyclists in the study sample actually seeked for less sheltered areas when diverging from the shortest or fastest route, avoiding characteristics attached to more densely buil-up areas like heavier traffic. Regarding the vegetational shelter, insubstantial negative effects were found on the extent of divergence from the fastest route, while the association with diverging from the shortest route was insignificant. In other words, these findings suggest that vegetational shelter was not considered as a decisive factor for route choice by cyclists in the study sample.

Modelling: *to what extent are the weather conditions a reason for cyclists to adapt their route choice to the degree of shelter offered by the built environment?*

In order to address this sub-question, a set of linear regression models was estimated in which differences in the degree of shelter along an observed route compared to a shortest/fastest route were explained using the set of meteorological and control variables as predictors. The output of each of the models showed a general tendency where none of the meteorological factors substantially influenced variation for each of the three shelter variables between observed routes and shortest/fastest equivalents. Heavier influences were found from infrastructural and environmental route characteristics as the number of crossings and higher slope, included in the regression models as control variables. The findings indicate that for the study sample, weather conditions cannot be considered as a substantial reason for cyclists to adapt their route choice to the degree of shelter offered by the built environment.

Validation: *how can the results generated by implementation of the developed methodology be validated?*

Although the results that were generated by implementing the developed methodology were not validated within this thesis, a design of a potential validation method was provided in which cyclist route choice in a different study area would be predicted using the found effects of meteorological and shelter variables. As the estimated statistical models showed relatively low explanatory power, application of this validation method was expected to lead to highly varying predictions regarding cyclist route choice, leading to indefinite conclusions about the validity of the obtained results within this thesis. As a result, the results of implementing the developed methodology of this thesis stayed invalidated.

6.2 DISCUSSION

Implementation of the developed methodology on the study area of Tilburg showed that the degree of shelter provided along a route, decomposed into three different factors, did not substantially influence cyclist route choice in different weather conditions for the study sample. Several methodological aspects and other reasons could have underlain those findings.

A first methodological aspect regards the generation of the route model, where the set of observed routes was filtered based on completeness in terms of the road segments that were used. None of the included routes travelled over a road segment that is not included in the bicycle road network for the study area. This spatial filtering of the observed travel data was done to ensure a fair generation of an alternative shortest/fastest route for an observed route, since the applied shortest path algorithm was able to choose over the same network as used for the generalized observed route model. However, disregarding routes that go over the boundary of the study area had several implications. First of all, a boundary problem was created in the sense that the majority of the observed routes traversed through the center of the study area. This implies that a higher degree of observed routes went through areas with a higher degree of building shelter, relative to more exposed routes in the rural areas on the boundary of the study area. Similarly, as higher tree densities could be found in areas surrounding Tilburg, relatively more routes passed through areas with low tree densities. Therefore, the imbalance in the spatial distribution of the observed routes might have led to a degree of bias in the results.

A second implication of only including complete routes in the observed route model is the fact that observed routes could exclusively span over the extent of the study area. In that light, disregarding boundary crossing routes might have resulted in a relatively higher loss of 'longer' routes compared to 'shorter' routes, leaving less space for variation in the route-level degree of shelter provided by the built environment.

Both implications mentioned would suggest that cyclists in the study sample only had a limited choice in diverging from the shortest or fastest route to seek for more or less shelter from weather conditions. Whereas the study area of Tilburg was specifically chosen due to significant variation in spatial configuration, the methodological decisions regarding the route model might have restricted the potential effects of the variation. Application of the methodology on a larger study area could eliminate the issue in spatial variation that arose within this thesis.

For the quantification of the control variables, missing input data related to the bicycle road network was dealt with by assuming similarity with values of neighboring road segments. Although the implementation of this assumption enabled the generation of a more complete model regarding factors that potentially influence cyclist route choice, the validity of the results with respect to reality might have been influenced. Consequently, the adjustments of the (combined) effects of meteorological and shelter variables could potentially have been affected.

Taking a more general approach on the reasoning behind the found effects, existing literature indicates that in general utilitarian cyclists are less sensitive to external influences when it comes to route choice. The moderate and insignificant effects of the meteorological factors underlined this tendency. With only a limited effect of the weather conditions on utilitarian cyclist route choice by itself, the question could be posed whether potential mitigating effects of the shelter provided by the built environment would be desired by cyclists. As an additional factor that could strengthen the low sensitivity to weather conditions, the study population comprised cyclists that used electric bicycles for utilitarian trips. Providing the possibility to generate higher speeds through less effort compared to conventional bicycles, usage of electric bicycles can shorten trip durations and therefore decrease the time of exposure to external influences.

Finally, a second reason for the selection of Tilburg as the study area was an expected sufficient amount of observed travel data. Although the study population closely represented general trends within the national cycling community, the obtained results of implementing the developed methodology on this study area could be characteristic for this area only. Validation of applied methodology and results would enable conclusive statements on this aspect.

6.3 IMPLICATIONS FOR POLICY DESIGN

With a main motivation behind this thesis being to provide a base for policy design to stimulate utilitarian cycling, the obtained results suggest that there is no need to account for the degree of shelter provided by the built environment along a route in policies regarding mitigation of the effect of weather conditions in the study area. Utilitarian cyclists in the study area primarily opted for routes approaching the shortest or fastest route. This route choice strategy leads to minimization of effort and travel duration, but, from the perspective of mitigating weather conditions, also minimizes exposure time to external influences like the weather. Higher degrees of built environment shelter by itself did in that light not seem to provide sufficient extra protection from weather conditions to make cyclists increase the exposure time through detouring from the shortest or fastest route.

The strong preference for short and fast routes by utilitarian cyclists in the study area implies that for the case study of this thesis, the fast bike lane network in the province of Noord-Brabant, policy design should evolve primarily around enabling fast travelling rather than mitigating the effect of weather conditions by providing shelter. However, the found route choice preferences were based on a study sample mainly including routes through urban areas, whereas a significant number of the desired fast bike lanes are supposed to traverse through rural areas. As urban areas by definition offer a larger variety of sheltering options than rural areas, the provided guidelines for policy design are therefore mainly applicable for parts of the fast bike lane network within or close to urban areas.

6.4 CONTRIBUTION TO THEORETICAL FRAMEWORK

This thesis addressed a research gap in existing literature on cyclist travel behavior by examining whether the degree of shelter provided by the built environment could be considered a mitigating factor for the effect of weather conditions on route choice by utilitarian cyclists. A brief indication on the placement of the developed methods and found results from this thesis within the established theoretical framework is provided in the section.

Regarding the field of cyclist travel behavior, the contribution of this thesis is twofold. Firstly, a method has been developed in which cyclist route choice can be tested for the combined effects of meteorological factors under which a route was conducted, and the degree of shelter that was provided by the built environment. Estimating effects of individual meteorological factors on actual cyclist route choice is an underexplored aspect in existing literature, but not a novelty [Böcker and Thorsson, 2014; Helbich et al., 2014]. However, the addition of shelter by the built environment as an explaining factor was a first attempt to identify determinants for route choice by cyclists in different weather conditions. Therefore, a novel methodology was developed which could be adopted, modified, and applied on other study areas. Secondly, the obtained results give a first indication of the relevance of built environment shelter to cyclists when deciding upon a route in different weather conditions. The limited results might be a motivation to research other potential determinants of cyclist travel behavior, especially in relation to weather conditions.

The inclusion of the degree of shelter provided by the built environment as a possible mitigating factor of weather conditions on cyclist route choice found its fundament in the field of pedestrian mobility studies [Lenzholzer and van der Wulp, 2010; Nikolopoulou and Lykoudis, 2007]. The findings of this thesis strengthen the idea that cyclist travel behavior is differs significantly from travel behavior from pedestrians. Where pedestrians were found to be highly affected by the degree of shelter provided against certain weather conditions, cyclists in the study area showed no substantial changes in behavior. As the dynamic character of pedestrian movement is not as strong as for cyclists, the experience of shelter on a given location could therefore be heavier for pedestrians. This assumption could open the door for application of (aspects of) the developed methodology of this thesis on a pedestrian mobility problem.

Finally, the developed method to operationalize the degree of shelter provided by the built environment used aspects from studies on street climate design [Oke, 1988] and spatial openness [Benedikt, 1979]. Combination of both study fields led to a detailed approach to quantify urban geometries, enabling application on street climate design studies.

6.5 RECOMMENDATIONS FOR FUTURE WORK

As the developed methodology addressed an underexplored topic regarding cyclist route choice, shortcomings and new ideas came forward over the course of this thesis. Therefore, several recommendations for future work are presented:

- In order to increase the odds of significant variation in the degree of shelter between observed routes and theoretically optimal routes, a larger study area could be selected for further research. Using a larger study area would lead to inclusion of a more considerable share of 'longer' routes and more opportunities to vary in spatial configuration around a route.

- A potential cause of the lack of found substantial effects could be that the degree of shelter provided by the built environment was operationalized in a higher level of detail than required. Due to the highly dynamic character of cycling, shelter provided by the built environment might be experienced differently. To operationalize the degree of shelter along a route in a more fitting manner in relation to cyclist route choice, further qualitative research could study the perception of environmental shelter by cyclists.

- Within this thesis, independent meteorological factors were used to model the weather conditions under which a route was conducted. However, co-occurrence of the different factors can change the experience of each independent meteorological factor. Therefore, future research could use integrated weather effects to model the effects of weather conditions on cyclist route choice.

- As a measure of the degree of vegetational shelter, tree densities surrounding a road segment were used as indicator. Since this method was primarily used due to absence of alternative data, future research could employ different ways of operationalizing vegetational shelter when different data is available.

- As the obtained results of the implemented methodology within this thesis have not been validated, no conclusions can be drawn about their reproducibility. Inclusion of the proposed validation method in further applications of the developed method would increase the credibility of the obtained results.

BIBLIOGRAPHY

Amiri, M. and Sadeghpour, F. (2015). Cycling characteristics in cities with cold weather. *Sustainable Cities and Society*, 14(1):397–403.

Anastasiadou, A., Herrera Herrera, M., Oldenburg, D., Tsakalakidou, I., and van der Wal, E. (2018). Cyclist Route Choice Modeling: A research on influence of Openness & Monotony on cyclist route choice.

Anselm Akubue, J. (2019). Effects of Street Geometry on Airflow Regimes for Natural Ventilation in Three Different Street Configurations in Enugu City. In *Different Strategies of Housing Design [Working Title]*.

B-Riders (2016). B-Riders — Fietsen naar het werk.

Benedikt, M. L. (1979). To take hold of space: isovists and isovist fields. *Environment and Planning B: Planning and Design*, 6(1):47–65.

Bergström, A. and Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and Practice*, 37(8):649–666.

Böcker, L. (2014). *Climate, Weather and Daily Mobility Transport Mode Choices and Travel Experiences in the Randstad Holland*. PhD thesis.

Böcker, L., Dijst, M., and Faber, J. (2016). Weather, transport mode choices and emotional travel experiences. *Transportation Research Part A: Policy and Practice*, 94:360–373.

Böcker, L., Dijst, M., Faber, J., and Helbich, M. (2015). En-route weather and place valuations for different transport mode users. *Journal of Transport Geography*, 47:128–138.

Böcker, L., Prillwitz, J., and Dijst, M. (2013). Climate change impacts on mode choices and travelled distances: a comparison of present with 2050 weather conditions for the Randstad Holland. *Journal of Transport Geography*, 28:176–185.

Böcker, L. and Thorsson, S. (2014). Integrated Weather Effects on Cycling Shares, Frequencies, and Durations in Rotterdam, the Netherlands. *Weather, Climate, and Society*, 6(4):468–481.

Bowditch, N. (2002). *The American practical navigator : an epitome of navigation*.

Broach, J., Dill, J., and Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46(10):1730–1740.

Dill, J. and Gliebe, J. (2008). Understanding and Measuring Bicycling Behavior: A Focus on Travel Time and Route choice. Technical Report December, Portland State University.

Dukai, B. (2018). 3D Registration of Buildings and Addresses (BAG).

El-Assi, W., Salah Mahmoud, M., and Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation*, 44(3):589–613.

Elmasri Ramez and Navathe Shamkant (2010). Fundamentals of Database Systems 6th. In *6*.

Fisher-Gewirtzman, D. and Wagner, I. A. (2003). Spatial openness as a practical metric for evaluating built-up environments. *Environment and Planning B: Planning and Design*, 30(1):37–49.

Flynn, B. S., Dana, G. S., Sears, J., and Aultman-Hall, L. (2012). Weather factor impacts on commuting to work by bicycle. *Preventive Medicine*, 54(2):122–124.

Gillner, S., Vogt, J., Tharang, A., Dettmann, S., and Roloff, A. (2015). Role of street trees in mitigating effects of heat and drought at highly sealed urban sites. *Landscape and Urban Planning*, 143:33–42.

Gojanovic, B., Welker, J., Iglesias, K., Daucourt, C., and Gremion, G. (2011). Electric bicycles as a new active transportation modality to promote health. *Medicine and Science in Sports and Exercise*, 43(11):2204–2210.

Handy, S. L. and Xing, Y. (2011). Factors correlated with bicycle commuting: A study in six small U.S. cities. *International Journal of Sustainable Transportation*, 5(2):91–110.

Hayes, A. F. and Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*, 41(3):924–936.

Heinen, E., Maat, K., and Van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. *Transportation Research Part D: Transport and Environment*, 16(2):102–109.

Heinen, E., van Wee, B., and Maat, K. (2010). Commuting by bicycle: An overview of the literature. *Transport Reviews*, 30(1):59–96.

Heisler, G. (1990). Mean wind speed below building height in residential neighborhoods with different tree different tree densities. *ASHRAE Transactions*, 96(1):1389–1396.

Helbich, M., Böcker, L., and Dijst, M. (2014). Geographic heterogeneity in cycling under various weather conditions : evidence from Greater Rotterdam. *Journal of Transport Geography*, 38:38–47.

Kennisinstituut voor Mobiliteitsbeleid (2019). Mobiliteitsbeeld 2019. Technical report.

Klobucar, M. S. and Fricker, J. D. (2007). Network evaluation tool to improve real and perceived bicycle safety. *Transportation Research Record*, 2031(2031):25–33.

Krenn, P. J., Oja, P., and Titze, S. (2014). Route choices of transport bicyclists: A comparison of actually used and shortest routes. *International Journal of Behavioral Nutrition and Physical Activity*, 11(1):31.

Lenzholzer, S. and van der Wulp, N. Y. (2010). Thermal experience and perception of the built environment in Dutch urban squares. *Journal of Urban Design*, 15(3):375–401.

Lin, S., He, M., Tan, Y., and He, M. (2008). Comparison study on operating speeds of electric bicycles and bicycles experience from field investigation in Kunming, China. *Transportation Research Record*, 2048(2048):52–59.

Liu, C., Susilo, Y. O., and Karlström, A. (2015). Investigating the impacts of weather variability on individual's daily activity-travel patterns: A comparison between commuters and non-commuters in Sweden. *Transportation Research Part A: Policy and Practice*, 82:47–64.

Liu, C., Susilo, Y. O., and Karlström, A. (2017). Weather variability and travel behaviour–what we know and what we do not know. *Transport Reviews*, 37(6):715–741.

Lu, G. Y. and Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers and Geosciences*, 34(9):1044–1055.

Mertens, L., Compernolle, S., Deforche, B., Mackenbach, J. D., Lakerveld, J., Brug, J., Roda, C., Feuillet, T., Oppert, J. M., Glonti, K., Rutter, H., Bardos, H., De Bourdeaudhuij, I., and Van Dyck, D. (2017). Built environmental correlates of cycling for transport across Europe. *Health and Place*, 44:35–42.

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., and Weather, R. D. (2005). Cycling and the built environment, a US perspective. *Transportation Research Part D: Transport and Environment*, 10(3):245–261.

Ncase (2019). SIGHT & LIGHT - How to create 2D visibility/shadow effects for your game.

Nikolopoulou, M. and Lykoudis, S. (2007). Use of outdoor spaces and microclimate in a Mediterranean urban area. *Building and Environment*, 42(10):3691–3707.

Oke, T. R. (1981). Canyon geometry and the nocturnal urban heat island: Comparison of scale model and field observations. *Journal of Climatology*, 1(3):237–254.

Oke, T. R. (1988). Street design and urban canopy layer climate. *Energy and Buildings*, 11(1-3):103–113.

pgRouting Community (2016). pgRouting Project — Open Source Routing Library.

Provincie Noord-Brabant (2009). Fiets in de Versnelling Visie 2020 met actieprogramma. Technical Report december.

Provincie Noord-Brabant (2019). Fietspaden.

Reddy, H. (2013). Pathfinding-Dijkstra's and A* Algorithm's. *International Journal in IT and Engineering*, pages 1–15.

Rietveld, P. and Daniel, V. (2004). Determinants of bicycle use: Do municipal policies matter? *Transportation Research Part A: Policy and Practice*, 38(7):531–550.

Sears, J., Flynn, B. S., Aultman-Hall, L., and Dana, G. S. (2013). To Bike or Not to Bike. *Transportation Research Record: Journal of the Transportation Research Board*, 2314(1):105–111.

Shishegar, N. (2013). Street Design and Urban Microclimate: Analyzing the Effects of Street Geometry and Orientation on Airflowand Solar Access in Urban Canyons. *Journal of Clean Energy Technologies*.

Sommet, N. and Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, 30(1):203–218.

Spencer, P., Watts, R., Vivanco, L., and Flynn, B. (2013). The effect of environmental factors on bicycle commuters in Vermont: Influences of a northern climate. *Journal of Transport Geography*, 31:11–17.

Stelling-Konczak, A., de Groot-Mesken, J., Vlakveld, W. P., Wesseling, S., and SWOV, S. W. O. V. (2017). Speed pedelec on the carriageway : First practice research on behavioral effects. Technical report.

Stinson, M. and Bhat, C. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record*, (1828):p. 107–115.

Stinson, M. A. and Bhat, C. R. (2004). Frequency of bicycle commuting: Internet-based survey analysis. In *Transportation Research Record*, volume 1878, pages 122–130.

Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis and Prevention*, 83:132–142.

Theurer, W. (1999). Typical building arrangements for urban air pollution modelling. *Atmospheric Environment*, 33(24-25):4057–4066.

Urban, D. and Keitt, T. (2001). Landscape Connectivity: A Graph-Theoretic Perspective. *Ecology*, 82(5):1205.

Van Oosterom, P. (1999). Spatial Access Methods. volume 1, pages 385–400. Wiley.

Van Rijn, N. (2009). *Validating the use of the Isovist tool as a means to assess landscape openness in the Netherlands*. PhD thesis, Wageningen University.

Winters, M., Brauer, M., Setton, E. M., and Teschke, K. (2010). Built environment influences on healthy transportation choices: Bicycling versus driving. *Journal of Urban Health*, 87(6):969–993.

Winters, M., Davidson, G., Kao, D., and Teschke, K. (2011). Motivators and deterrents of bicycling: Comparing influences on decisions to ride. *Transportation*, 38(1):153–168.

Winters, M., Friesen, M. C., Koehoorn, M., and Teschke, K. (2007). Utilitarian Bicycling. A Multilevel Analysis of Climate and Personal Influences. *American Journal of Preventive Medicine*, 32(1):52–58.

Yeboah, G. and Alvanides, S. (2015). Route choice analysis of urban cycling behaviors using OpenStreetMap: evidence from a British urban environment. *Lecture Notes in Geoinformation and Cartography*.

Zhang, Z. and Zhao, Z. (2014). A multiple mobile robots path planning algorithm based on a-star and dijkstra algorithm. *International Journal of Smart Home*, 8(3):75–86.

Zhao, J., Wang, J., Xing, Z., Luan, X., and Jiang, Y. (2018). Weather and cycling: Mining big data to have an in-depth understanding of the association of weather variability with cycling on an off-road trail and an on-road bike lane. *Transportation Research Part A: Policy and Practice*, 111:119–135.

A FORMATION OF CONTROL VARIABLES

Table A.1: Grouping of nominal attribute values provided by the Dutch Cyclist Union into control variables

Atttribute Dutch Cyclist Union	**Attribute value**	**Control variable**
Verlichting	'goed verlicht'	Good lighting
Wegkwaliteit	'goed'	Good road quality
Wegniveau	'belangrijke hoofdweg'	High traffic volume
	'langs drukke weg'	
Wegtype	'fietsstraat'	Shared road
	'ventweg'	
	'weg met fiets(suggestie)strook'	
	'normale weg'	
Gem_st_fwd	'1-2%'	Slope $>$ 1%
	'2-4%'	
	'$>$4%'	

COLOPHON

This document was typeset using LaTeX. The document layout was generated using the arsclassica package by Lorenzo Pantieri, which is an adaption of the original classicthesis package from André Miede.

TU Delft

Delft University of Technology