Pedestrian-Cyclist Interactions

Analysing Pedestrian Movements to Cyclists Nearing Bike Paths Using Trajectory Data

Hidde Vincken

Pedestrian-Cyclist Interactions

Analysing Pedestrian Movements to Cyclists Nearing Bike Paths Using Trajectory Data

by

Hidde Vincken

Student:	H.L.E.Vincken
Chair:	W. (Winnie) Daamen
Daily supervisor:	Y. (Yufei) Yuan
Project duration:	November, 2024 - July, 2025
Arup supervisors:	Johan Krom, Sara Zein, Martin van Oosten
Faculty:	Faculty of Civil Engineering and Geosciences, Delft
Department:	Traffic and Transport Engineering

Cover: Ossip Architectuur Fotografie

TUDelft

Preface

This thesis is the result of my last eight months of the master of Civil Engineering. I look back at my seven years at the TU Delft as the most enriching period of my life. The experience of intensive study days, various study trips, and many hours of collaborative work with my friends have shaped me. This thesis hopefully reflects some of those experiences I have gained over the years. The topic of this thesis has intrigued me from the very beginning and that feeling has never left me to this day.

My gratitude goes out to all my supervisors who have not only helped me through providing feedback, but also through an overall nice cooperation between the university and the company. You have always considered each other's interests with care, which made for a pleasant way of working. Winnie, you kept me sharp and gave me guidance, and I learned to shape this thesis to be something I am very proud of. Yufei, I have thoroughly enjoyed our meetings in which we could delve into specific technical matters for a long time, to really get to the bottom of things. Johan, our weekly conversations have not only assisted me with structuring my thesis, but it has also shown me the nice working culture that lives within Arup. Sara, your kind words have helped me a lot in times where I encountered more difficulties and it reminded me that I can look back positively at what I had already achieved.

I want to thank all other colleagues at Arup who have always supported me, whether that was through a presentation, short update or just a regular conversation. You made me feel truly part of the team. Thank you to the students in the thesis room. It has always been a source of motivation for me to talk about common pitfalls, and day-to-day struggles, but also our achievements. Furthermore, I would like to thank my friends and family for always showing interest in my progress on my thesis and study as a whole, this really convinced me that my commitment to it mattered. Lastly, I want to thank my girlfriend, Irene, for always being there to listen, whether I came home after a good or a bad study day. Even though you beat me by graduating five days before me, I look forward to our period of free time together.

And to you, the reader, I sincerely hope you enjoy reading this thesis as much as I have while studying this topic.

Hidde Vincken
Delft, July 2025

Summary

Active mobility becomes a more prominent transportation option in the urban environment because of its benefits in terms of sustainability and population health (Donaire-Gonzalez et al., 2015). An increase in volumes of cyclists and pedestrians could result in more frequent and severe conflicts, especially when they share narrow designed spaces (Huyghebaert, 2021). For accommodating pedestrians and cyclists, proper infrastructure needs to be designed, but this requires knowledge on the behaviours that these two modes perform. This study gives a description of the interactive behaviour between pedestrians and cyclists, providing input for the social forces model MassMotion, developed by the company Arup. More specifically, the focus is on the movement changes a pedestrian performs when confronted with a cyclist on bike paths, because often times the pedestrian is at a disadvantage in these cases considering its vulnerability (Letsel & Schade, 2025). The study aims to address the research question:

In what way do pedestrians change their movements when approaching and crossing bike paths with oncoming cyclists?

Previous studies on interactive behaviour between pedestrians and cyclists focus on head-on or rear-end conflicts, but not many studies focus on sideways crossing scenarios. Yet these are often the most critical for pedestrians, as crossing a dedicated bike path sideways typically offers the quickest route to exit the area designated for cyclists. Many studies describe conflict severance either by time-to-collision (TTC) or by post-encroachment time (PET). While TTC quantifies how soon two traffic participants would collide if they maintain their current trajectories, PET is used in this study because it captures a broader range of crossing scenarios, including situations where paths intersect but are not strictly on a collision course. Furthermore, the literature shows that simulation of movements for both pedestrians and cyclists operates best if a prediction of movement is taken into account.

This study uses trajectory data collected by smart sensors of two intersections at the TU Delft campus for a behavioural analysis. In total, 289216 trajectories were detected in a full month. Speeds, location with respect to the environment, and origins and destinations are determined for all trajectories to determine the type of mode and to distinguish approach cases. Sideways crossing cyclists and pedestrians within a time frame (PET) of 0 to 5 seconds are studied. The movements of a total of 7310 pedestrians have been detected of which 4780 crossed with a single cyclist. A classification is made between pedestrians crossing before and after the cyclist to compare the outcomes of the behaviours in both scenarios, both with each other and with reference trajectories of non-crossing pedestrians. Conflicts that are in close range (0 - 3 seconds of PET) in general show more stopping movements for pedestrians than interactions that happen further away from each other (3 - 5 seconds of PET). This second classification from a PET of 3 seconds threshold allows further analysis to distinguish whether pedestrians behave differently when the interaction is near compared to when there is more time in between the two modes.

Stopping behaviour is analysed for pedestrians that cross with a cyclist and compared with pedestrians that do not cross with a cyclist. From the four conflicting scenarios that were analysed, three considered pedestrians crossing straight and one scenario, the pedestrian crossed the path after walking along the cycling path. In the straight walking scenarios, the pedestrians in the category of 0 to 3 seconds of PET that cross after the cyclist stop around 35% of the time. The stopping percentage decreases with an increasing PET, indicating that the closer in time a pedestrian crosses after a cyclist, the more likely it is that the pedestrian will stop. The average distance that pedestrians tend to keep measured from the instant the pedestrian stops to the eventual crossing point in this close range scenario is slightly more than 3 meters. The average distance in the scenario where the pedestrian walked along the cycling path was 2.5 meters. The stopping distance where the pedestrian crosses after the cyclist within a PET of 0 to 3 seconds follows a normal distribution.

Apart from the actual PET value of a conflict, this study presents a method to derive a predicted PET

value per time step. The predicted PET provides insight into how an approach towards a potential conflicting situation is handled. From these predicted PET values, it is found that when a collision or near collision is predicted (PET between 0 and 1 seconds), the pedestrian is more likely to stop or slow down to let the cyclist pass. In many other cases where no collision is predicted, the pedestrian often maintained a steady course.

Furthermore, deviating behaviour is analysed for pedestrians that need to cross straight and that have a conflict with cyclists. This is determined based on the overall and maximum deviation of a trajectory from the straightest path. Whether a trajectory shows notable deviation is determined based on the common difference with non-crossing trajectories. The findings suggest that pedestrians that cross behind a cyclist that approaches from the side, deviate about 20% of the time. This deviation is often directed towards the cyclist, because the pedestrian can cross the cyclist earlier.

This study shows that pedestrians adapt their behaviour in response to approaching cyclists by stopping, deviating, slowing down, or combining these strategies, when yielding is needed. These adaptations are anticipatory and depend on the relative timing, approaching side, and available space of the interaction. The findings suggest pedestrians often act more cautiously, implicitly recognising cyclists' speed advantage. This behavioural distinction highlights the need for infrastructure that reflects these dynamics and supports safe, intuitive interactions at active-mode crossings.

The findings of this study can fit into the simulation model MassMotion by implementing a combination of adaptation to the social forces for the direction of movement and overruling these forces when stopping or slowing down behaviour needs to be performed.

The findings could refine the behavioural assumptions that are made in future research on pedestrians stopping and deviating movements towards approaching cyclists, but are also required as input for behavioural models. This study furthermore shows that the implemented method for a prediction of PET can capture more nuanced pedestrian avoidance behaviours outside stopping and deviating behaviour, thereby broadening the scientific understanding of conflict measures beyond traditional PET.

More testing of the MassMotion model is necessary to eventually implement a proper cyclist agent into the model by calibrating the parameters and implementing terms with stochasticity. Further research could be done in the movement changes that the cyclists tend to make when interacting with the pedestrian, especially by focussing on longer ranges.

The results of this study can inform urban planners to develop evidence-based decisions on designs, implementing the knowledge on pedestrians hesitancy when crossing a bike path. The findings can be used by policymakers to develop intervention methods at busy pedestrian-cyclist crossings.

Contents

List of Figures

List of Tables

Nomenclature

Terminology

Term	Definition
Agent	A simulated pedestrian or cyclist in a virtual environment
Approach	Coming near a situation from a certain origin
Conflict	The state where two or more traffic participants' intended paths overlap
Crossing	Two paths of traffic participants overlapping **or** traffic participant passing over a path assigned to another mode of transport **or** an intersection (see intersection)
Interaction	The process of two or more traffic participants influencing each others movements
Intersection	Site where more than two roads come together
Shared road	Road where the paths of multiple modes of traffic are intertwined in longitudinal direction
Shared space	Space where multiple modes of traffic can cross each other in all directions
Trajectory	A sequence of points in time belonging to the path of a traffic participant

Abbreviations

Abbreviation	Definition
FPS	Frames Per Second
PET	Post-encroachment time
RQ	Research Question
SDK	Software Development Kit
SFM	Social Forces Model
SRQ	Sub-Research Question
TTC	Time To Collision
UI	User Interface

Symbols

Symbol	Definition	Unit
d	Distance	[m]
p	Position	[m]
PET	Post-Encroachment Time	[s]
$RMSD$	Root Mean Square Deviation	[m]
t	Time	[s]
v	Velocity	[m/s]
w	Weight	[-]
φ	Angle	[Degrees]

1 Introduction

With the need for more sustainable transportation options, the urban environment will undergo a modal shift from the car to cycling and walking, among other modalities (Ng et al., 2024). An additional benefit of this shift towards active mobility is the increase in physical activity of a population accompanied by improved health (Donaire-Gonzalez et al., 2015). This transition to active modes comes with the challenge of facilitating sufficient and high-quality infrastructure that meets the needs of pedestrians and cyclists. While active modes are associated with significantly lower external costs compared to motorized transport, the remaining costs are largely linked to traffic accidents (Pisoni et al., 2022). An increase in the number of cyclists and pedestrians may lead to more frequent and severe accidents, particularly in areas where they share narrow spaces (Huyghebaert, 2021).

Consequently, infrastructure should be designed properly, so it not only encourages walking and cycling but also promotes behaviours that are safe and consistent with the intended use of the space. However, the movements of cyclists and pedestrians differ significantly in terms of speed and agility. While Dutch legislations make a clear distinction between motor vehicle drivers and active modalities, labelling the latter group as vulnerable road users which affords them greater legal protection in traffic incidents (ARAG, 2025), no legal distinction is made between cyclists and pedestrians (Letsel Hulp Service, n.d.). At the same time, the general higher speed of the cyclist might suggest that the pedestrian is a more vulnerable road user (Letsel & Schade, 2025) and behaves accordingly by a more cautious attitude (Huyghebaert, 2021 and voetgangersvereniging Nederland, 2021), yet it is unclear in what way exactly this conflict avoidance behaviour between pedestrians and cyclists takes place. In addition, transport planners will need insight into this behaviour when creating policies that should contribute to active mobility needs. Therefore, there is a need for knowledge about this behaviour, especially considering that the limited number of studies that focus on this interactive behaviour do not specify all the typical encounters that take place in areas dedicated to active modes.

1.1. Research Problem

Despite the growing need for sustainable transportation, there is a significant gap in understanding the interactions between cyclists and pedestrians, which this research aims to address. Predominantly, studies have been conducted on the interaction in pedestrian crowds (Shahhoseini and Sarvi, 2019), which is an enormous challenge for festivals and events Klüpfel (2014). However, only a handful of studies have been performed to simulate the interactive behaviour between cyclists and pedestrians, of which the interactive cases have been outlined by W. Wang et al. (2024), Dias et al. (2018), Beitel et al. (2018), Yuan et al. (2019), and Afghari et al. (2014). These studies consider different approaching cases between a cyclist and a pedestrian. The studies of W. Wang et al. (2024) and Dias et al. (2018) focus on the trajectories of pedestrians and cyclists in a head-on conflict or an approach from behind. And while Beitel et al. (2018) also considers the approaching scenario of a sideways interaction, it only makes an assessment of the safety of different interactions and does not analyse movements or movement changes based on the interactive cases. The study of Yuan et al. (2019) performs an analysis on cyclist and pedestrian interactions based on data collected from a shared space. Although shared space is a concept that is widely applied and tested within the urban area, general intersection cases occur even more frequently. Afghari et al. (2014) performs a study at a bike path with a pedestrian crossing, yet after filtering only 17 interactions were analysed to detect general movement patterns.

Thus, there is a knowledge gap on the general behaviours that cyclists and pedestrians show, when in conflict with each other at crossings. Not only could the general knowledge on cyclist-pedestrian dynamics be advanced, also the particular behaviours of individuals and its variance are of interest both for scientists and practitioners. If insights on the behaviours of pedestrians and cyclists are gathered, these could form a fundament for studies on safety, for example by developing methods to assess the risks involved in the interactions. Urban planners or policy makers could use the behavioural information on interactions to create design interventions that reduce conflicts or improve flows. A suitable way to develop design ideas, would be to implement behavioural rules into a simulation model to be able to test these designs. This is a preferred method rather than applying designs directly in real-life situations, because several scenarios can be tested in simulation software with relatively low costs, while it is also possible to detect severe risk issues at an early stage.

The company Arup and their software development branch Oasys have a simulation model, MassMotion, in which the movements of pedestrians and cyclists can be analysed. MassMotion operates under the conditions of social forces that direct agents to their respective goals. Although the development of the pedestrian is already in a far stage of development, the cyclist is only at the beginning. Some general modifications have been made to the pedestrian agent to resemble a cyclist agent, yet when interacting with pedestrians, this new cyclist and subsequently the pedestrians still show behaviour that is not realistic. Transforming a pedestrian agent into a cyclist involves rethinking how the agent behaves and interacts with its environment. Cyclists move and respond differently and occupy more space compared to pedestrians. Modelling these differences presents a challenge, as it requires adapting both the shape and functional characteristics of the original agent to reflect cyclists. Moreover, when cyclists are correctly adapted into the model, the responses of pedestrians need to be tuned to these cyclists as well to mimic real-life responses. These responses are pivotal for the implementation when the pedestrian is considered a more vulnerable party in the the interaction with a cyclist.

1.2. Research Objective

The aim of this research is to investigate how pedestrians adapt their movement when approaching cyclists in crossing scenarios, with a focus on identifying behavioural patterns that contribute to collision avoidance. With a data driven approach, this study analyses pedestrian responses of stopping and deviating to a cyclist approaching on a bike path. The resulting insights are intended to provide a foundation for future development of cyclists and pedestrians in simulation models, such as MassMotion.

The first objective is to conduct a literature review by identifying distinct types of cyclist-pedestrian interactions. The studies should differentiate in approaching scenarios and what consequences these approaches have on the behaviours that are performed. Furthermore, it should be clarified how pedestrian and cyclist behaviours have been incorporated into social force models.

The second objective is to define a structured list of data requirements that is needed to study pedestrian-cyclist interaction. This list should follow from a developed framework that clarifies the factors influencing cyclist and pedestrian behaviour.

The third objective is to implement a data preprocessing method to prepare for the analysis on behaviours, while retaining a significant amount of the pedestrians and cyclists. The purpose of this method is to extract the parameters that explain behavioural traits and distinguish types of interactions.

The fourth objective is to analyse significant pedestrian-cyclist interaction trajectories and qualitatively and quantitatively assess the stopping and deviating behaviour that pedestrians tend to perform in different conflicting situations.

The last objective is to detect possible deviations when implementing cyclists in a social forces model, focussing on the movement changes that pedestrians perform towards these cyclists. These movements should be compared in a test case to the movements that have been detected in the data, to determine what resulting behaviour needs to be adapted in the simulation model.

1.3. Research Scope

This research focusses on the interaction between a pedestrian and a cyclist, specifically examining how pedestrians react to the presence of cyclists approaching from cycling paths with no other modes of traffic involved. There should be a pavement separation between the two modes in the form of distinction by either kerbs, different (coloured) pavement or ribboned pavement.

The individual characteristics of the participants are not considered as influencing factor of the behaviour. This study only distinguishes between a cyclist and pedestrian, without further distinctions into different types, e.g. runners or e-bikes, and different demographics such as gender, age and fitness. This is because focussing on a broad categorisation enables the findings to be more generalisable to pedestrian-cyclist interactions, rather than limiting them to specific subgroups. Additionally, this research prioritises the quantity of data over detailed individual profiling, as the aim is to provide a first insight into pedestrian behavioural adaptations during encounters with cyclists. As a result, nuances in how different subgroups might react to conflicts remains undetected and the conclusions drawn may overlook the influence of individual characteristics.

The study will analyse individual interactions between a single pedestrian and a single cyclist, excluding any group dynamics. While general cyclist characteristics will be considered for modelling purposes, the primary emphasis remains on pedestrian behaviour. This focus is chosen because accurately modelling cyclist adaptation in MassMotion would require significant adjustments, which is beyond the scope of this study.

The interactive cases that this study focusses on are pedestrians and cyclists crossing each other sideways, because these are often the most critical for pedestrians, as crossing a dedicated bike path sideways typically offers the shortest and therefore fastest path to exit the area designated for cyclists. Therefore, no head-on or rear-end conflicting situations are considered.

Lastly, weather circumstances and lighting conditions are not taken into account to generalise behaviour of pedestrians and cyclists in all weather circumstances and lighting conditions. Although the study does not include lighting conditions, the visual range of pedestrians and cyclists on site is sufficient due to the application of lampposts.

1.4. Research Questions

To achieve the aim and the corresponding objectives presented in this thesis, it is necessary to explore how the behaviour towards cyclists affects pedestrian movements. Therefore, this research seeks to address the following question:

In what way do pedestrians change their movements when approaching and crossing bike paths with oncoming cyclists?

First, the different possible interactive behaviours in existing research will need to be studied and the main takeaways for social force-based modelling of the studies need to be described. Second, the interactions should be chosen that could contribute to describing general movement cases of cyclists and pedestrians and then data requirements need to be determined. Third, the data needs to be processed to make an analysis of the behaviour possible. Fourth, the behavioural patterns need to be detected that explain the movement of pedestrians when interacting with cyclists. Lastly, this study aims to translate some of the findings into a social forces model. Therefore, the sub-research questions become:

1. What types of movement patterns between cyclists and pedestrians have been identified in existing research, and how have these been implemented in social force-based models?
2. What data is needed to describe the movements of pedestrians towards cyclists?
3. Which data processing and enrichment methods are required to prepare pedestrian and cyclist data for a behavioural analysis?
4. What movement changes does a pedestrian perform when approaching a cyclist in a sideways conflict?

5. What adaptations can be done to the social forces model to mimic the behaviour of pedestrians towards approaching cyclists?

These sub-questions will be answered in their respective chapters. The main research question will be answered in the Conclusion.

1.5. Contribution to Science and Practice

This study contributes to both scientific understanding and practical application of cyclist–pedestrian interactions by extracting and analysing data-driven behavioural patterns from trajectory data. Scientific contributions are provided on the movement adaptations pedestrians make towards cyclists through a framework that explains this concept, a method for deriving these adaptations by making a prediction of paths, and a quantification of their various types. This methodology can find more use cases not only limited to cyclist-pedestrian interactions. Furthermore, data-based classification methods are presented that illustrate ways of organising interaction patterns. These findings are translated into an empirically grounded behavioural description that can complement social force-based models to demonstrate the feasibility of incorporating such behaviour. These insights provide a foundation for further model calibration. This can eventually inform practitioners for developing designs that better accommodate the needs of active modes. In addition, the behavioural insights that this study presents could already provide practitioners with making substantiated design or policy choices

1.6. Outline of the Thesis

In this thesis, cyclist-pedestrian interactions will be analysed with the following approach. The Literature Review examines existing research on various types of cyclist-pedestrian interactions and identifies which behaviours have been studied. The Methodology outlines the complete approach to study the movement changes of pedestrians towards cyclists. The Data Selection chapter motivates the choice of data collection method and site selection. The Data Preparation details the enriching processes taken to prepare the data for a behavioural analysis. The Results chapter presents the findings from the data analysis and evaluates the outcomes of simulation tests. The Discussion addresses the limitations of the study and reflects on the relation with other literature. Finally, the Conclusion gives an interpretation of the most relevant findings, discussing pedestrian behaviours towards cyclists and areas for future research.

2

Literature Review

This chapter outlines the existing knowledge regarding cyclist-pedestrian interaction and discusses what previous studies have done to analyse pedestrian and cyclist behaviour that could contribute to microscopic modelling. The chapter aims to address sub-research question one: *What types of movement patterns between cyclists and pedestrians have been identified in existing research, and how have these been implemented in social force-based models?* Therefore, literature is studied that explains the situations where interactions arise and describes what movements can be expected in these situations.

First, studies on possible interactions between traffic participants are discussed. Second, the social forces model utilised in MassMotion is examined in the literature by following the modelling approach of each of the studies. Finally, a conclusion is provided that highlights the gaps in the literature and how this research will build upon the existing knowledge.

2.1. Interactions

The interactions between cyclists and pedestrians are strongly influenced by the design of the infrastructure, as it shapes the space in which encounters occur. Certain types of infrastructure actively facilitate these interactions, making them valuable settings for studying the behaviours of both modes. Understanding how infrastructure influences encounters can help identify the conditions under which pedestrian-cyclist interactions emerge. This section will first explore the types of infrastructure most likely to induce such interactions. In the second part, an overview of existing research on pedestrian-cyclist interactions will be presented, highlighting the approaches taken and identifying knowledge gaps that remain to be addressed.

2.1.1. Infrastructure

Since infrastructure is intended to guide movements and manage flows, it inherently influences where and how these interactions take place. Understanding these influences allows to identify the locations where interactions are most likely to occur. Several studies have examined the impact of different infrastructure layouts on movement patterns.

Duives et al. (2013) provides an overview of pedestrian flow dynamics, categorising them into uni-directional and multi-directional flows. The pedestrian movement is shaped due to guiding infrastructure consisting of solid walls, forcing flows in certain directions. Figure 2.1 illustrates various pedestrian flow scenarios at different intersection layouts. In the context of this research, the focus is on scenarios F, G, and H, where the infrastructure facilitates crossing flows. However, Duives et al. (2013) focusses solely on pedestrian movement, and does not consider cyclists as a distinct group. If cyclists are taken into account separately, several more conflicting categories could be distinguished.

W. Wang et al. (2024) considers both pedestrian and cyclist interactions, though not across all possible interaction scenarios. In this study, it is mentioned that the directions that pedestrians and cyclists follow are under influence of hard or soft separators. Hard separators, such as impenetrable walls, compel

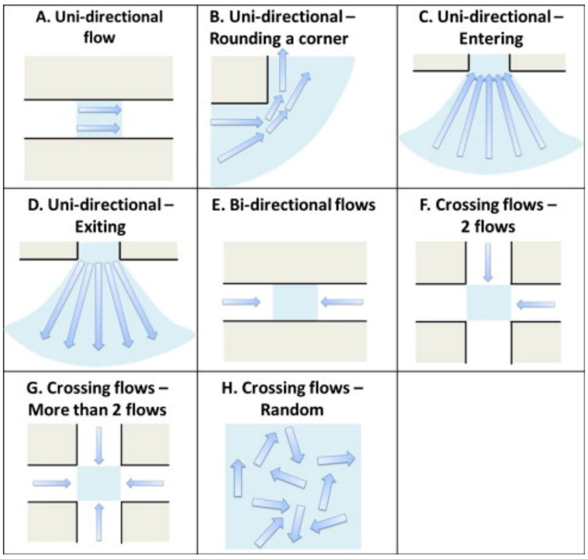


Figure 2.1: Pedestrian crowd flows from Duives et al. (2013).

cyclists and pedestrians to follow designated paths, following the flow scenarios described by Duives et al. (2013). In contrast, soft separators allow pedestrians to step off the infrastructure, though this research shows that they tend to stay near the separation edge, highlighting the role of infrastructure in guiding active modes of transport. The study compares the number of conflicts that occur on a straight road versus those at a T-intersection, while the flows of different modes are not separated. The conflicting numbers appear to be higher for the T-intersection, suggesting that interactions between cyclists and pedestrians tend to take place more often at intersections than on straight roads, as they facilitate encounters between active modes of transport.

The research of Wei et al. (2021) provides different possible intersection layouts for the purpose of autonomous driving strategies. Although this research focusses on motorised traffic, many of the intersection configurations it identifies can also be applied to active-mode interactions. Figure 2.2 presents an overview of different crossing layouts that could serve as a foundation for observing cyclist-pedestrian interactions, from Wei et al. (2021).

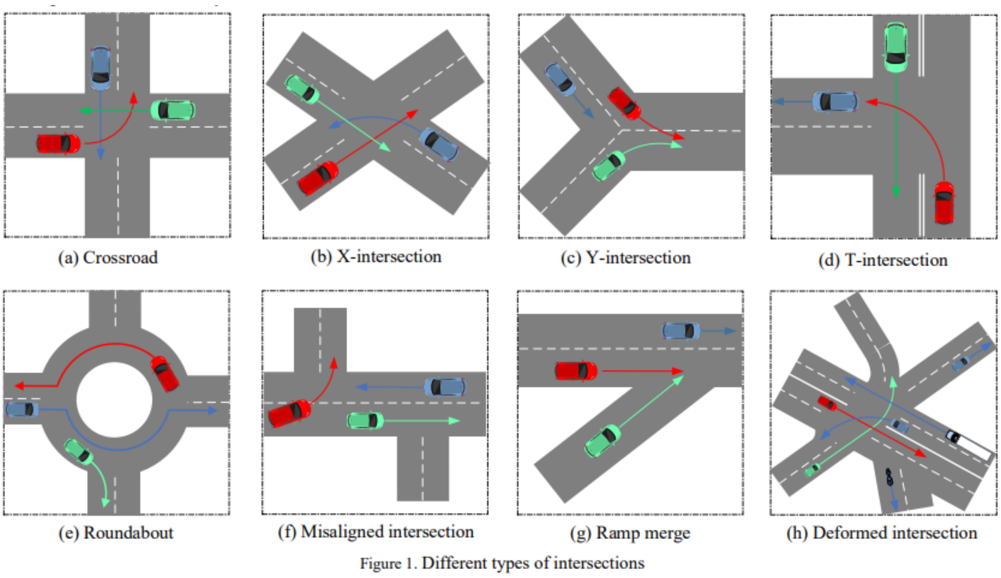


Figure 2.2: Intersection possibilities according to Wei et al. (2021).

Drawing on these studies, active-mode intersections are locations to observe pedestrian-cyclist interactions more frequently compared to straight roads. Although, to the writer's knowledge, there are no studies that distinguish all types of intersections suitable for various active modalities, the intersection types provided in aforementioned studies can be considered as a foundation for this research. In subsection 4.1, a new configuration of active mode intersections with separations of modes is presented. It is required for this research to make a selection of intersection locations to enhance the ability to capture sideways interactions and the resulting movement patterns.

2.1.2. Approach Cases in the Literature

A cyclist-pedestrian interaction can occur in different approaching situations. These can range from an approach that is head-on to one that is from behind and the approach can be in close range or from further away. In this section, active mode interactions that can be found in previous studies are highlighted, each focussing on different approach cases. In addition, the interactive scenarios that have not yet been investigated are addressed in this section.

In the paper of W. Wang et al. (2024) the approach cases that are analysed are between electric bicycles and pedestrians. The field study that has been conducted considers the approach cases to occur on a 'shared road', referring to a combined cycling and footpath, both in a straight road and T-intersection scenario. The straight road has a layout that enforces traffic participants into a head-on or rear-end conflict. Angled conflicts are considered in the sense that the time to collision is calculated, with which the severity of the conflict risk is determined, yet a specific analysis on what the angled conflict does to the movements of the pedestrian is missing.

A similar field study has been conducted by Dias et al. (2018) where, again, the focus was on a similar shared road of about 4.2 metres wide in which the flows of pedestrians, cyclists, and Segway riders were combined. In this case, site analysis was performed using an experimental setup. The movement scenarios that were analysed considered the pedestrian-cyclist interaction, more specifically the cyclist avoiding the pedestrian, and not vice versa.

Afghari et al. (2014) performed a field study of 158 minutes of video footage on pedestrian-cyclist interaction behaviour in Montreal. The case considered a zebra crossing from a side walk to a bus stop that crossed a bike path from the side, with a total number of 225 interactions, yet after filtering processes, only 17 interactions were left. The findings suggest that cyclists tend to maintain their speed and acceleration, whereas pedestrians make evasive actions by decelerating or accelerating or changing their movements with respect to the approaching cyclist. It is not specified in what way these movement changes occur. The bike path is furthermore under a slope, making a distinction between cyclist going up or down necessary. The authors acknowledge that the relatively short observation period and the limited number of close interactions restrict the generalisability of their conclusions and point to the need for further research.

In the study of Beitel et al. (2018), the risk of the cyclist-pedestrian interaction is assessed in shared spaces using video traffic data. Although this study does not provide information on the movements of cyclists or pedestrians, the conflict analysis could be a valuable asset to this research when studying the different types of conflict. Beitel et al. (2018) identifies three distinct types of conflicts between cyclists and pedestrians, classifying them into head-on, angled, and rear-end conflicts. Both head-on and rear-end conflicts consider one of the traffic participants approaching the other along the axis of direction of the other participant with a $\pm$ 30 degree angle. This leaves the angled approach to be a crossing that is perpendicular to the axis of direction with a $\pm$ 60 degree angle. In addition, the study differentiates these scenarios based on whether the pedestrian or the cyclist arrives at the conflict point first, leading to six different approaches. Figure 2.3 illustrates the different approach angles in a cyclist-pedestrian crossing scenario.

However, the classification of this angle of approach does not result in a description of behaviour for the cyclist or pedestrian. It is used to measure the safety risk of the type of conflicts that can occur. The risk assessment is done by determining the relative risk of each conflict approach and combining this with the risk of the value of the, so-called, post-encroachment time and the speed of the cyclist. The speed of the cyclist influences the risk of the conflict in terms of the impact that a higher speed has on the severity of a collision. The post-encroachment time is a measure to determine whether a conflict is nearby and is explained in more detail in subsection 2.1.3.

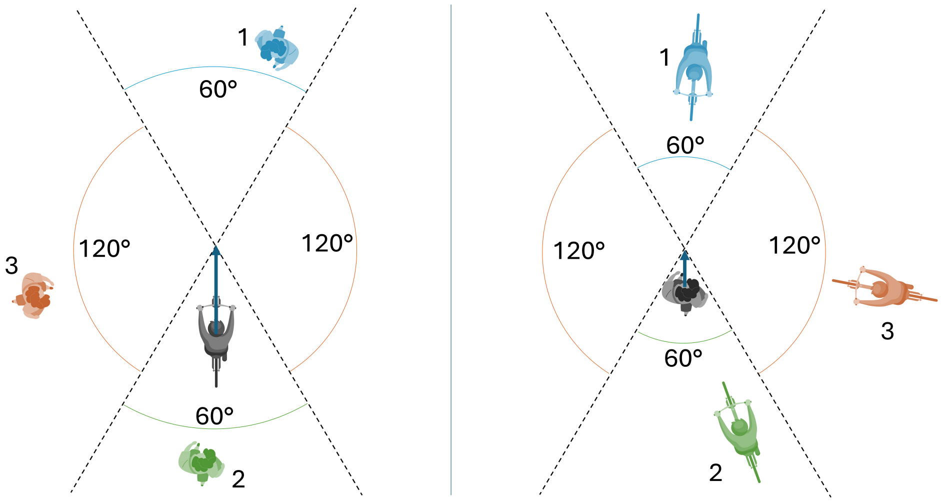


Figure 2.3: Categorisation of the possible crossings between cyclists and pedestrians with examples of a head-on (1), rear-end (2) and angled (3) conflict. Based on the paper of Beitel et al. (2018).

Summarising, the approach cases that have been studied in existing literature show a description of behaviour for head-on or rear-end conflicts, but often an analysis on sideways interactions is missing. The studies that do consider these sideways interactions either focus on the severity of conflicts, only inform on the movement changes that cyclists perform towards pedestrians, or in the case of Afghari et al. (2014), the study has too little data to generalise or specify the findings of the pedestrian behaviour.

2.1.3. post-encroachment time

To study and categorise cyclist-pedestrian movement patterns, several metrics have been used in the literature to assess the nature and intensity of their interactions. One of the adopted measures is the post-encroachment time (PET). Although PET is commonly used as a safety indicator to quantify the severity or proximity of conflict, it can also function as an indicator for identifying behavioural adaptations, such as stopping or slowing down. For this research, PET serves not only to assess risk, but also to classify behavioural responses based on the timing and spatial relationship between the two road users. These behavioural interpretations tied to PET are therefore relevant when examining the movement patterns that relate to pedestrian-cyclist interaction.

The PET is defined as the time it takes between the first traffic participant to leave the conflict area and the second traffic participant to reach this same area. It is a measure to determine the proximity of a collision with a small PET referring to a near conflict and vice versa. Beitel et al. (2018) prefers the measure of PET to the measure of time-to-collision (TTC) in which the time it takes for both modes to collide is calculated if their path continues. This preference is because PET considers more interactive scenarios than TTC and no motion predictions have to be made. However, for studying adaptations in the behaviour over time, these predictions could prove to be key for understanding the changes to movements that pedestrians and cyclists tend to make.

The risk assessment based on the value of the PET is done by classifying the conflicts into two categories of a PET between 0 and 2 s (conflicting) and between 2 and 5 s (interacting). Several studies use different threshold values for PET. For example, the paper of Tageldin and Sayed (2016) only considers the most dangerous conflicts, up to 1.5 s, for pedestrians interacting with several other modes when studying their evasive manoeuvres at a busy intersection. In the study of Afghari et al. (2014), 3 seconds is used as a PET threshold for safe interactions. A study of Zangenehpour et al. (2016) considers similar thresholds (1.5, 3 and 5 seconds) for cyclist-vehicle interaction. The debate around these different PET values explains that the circumstances of the crossing layout, intensities of the traffic, and the modalities that are studied could influence the threshold values for classifying certain interactions. However, the studies do agree that from a PET value above 5 s, it is not considered an interaction.

2.2. Social Forces

This section reviews the methodologies employed in previous studies that have applied the Social Forces Model (SFM), or a variation of this model, to simulate the behaviour of cyclists and pedestrians. It begins by outlining the core principles of the SFM, followed by an examination of how various studies have adapted or extended the model to capture cyclist-pedestrian interactions. By comparing their approaches, assumptions, and limitations, this section aims to identify both the required elements for model implementation and the behaviours that remain underexposed in current research.

The social forces model was first introduced by Helbing and Molnar (1995) in which three main force components are introduced: the driving, repulsive, and attractive forces. The driving force is the force that has a direction towards the goal of the agent, which results in an acceleration to a certain desired speed and direction. The repulsive forces are generated by objects or other agents that the agent in question needs to avoid in order to complete its trip. The attractive forces are the forces that the agent is drawn to during its trip, e.g. other persons or objects.

Helbing et al. (2002) refines the formulation of the social forces by introducing a term that monotonically decreases the exerted force to another agent depending on the approaching angle, which is the highest at the front of the pedestrian and zero at the back. Further enhancements of the social forces model for pedestrians have been developed by Xi et al. (2010), in which vision fields and grouping behaviour are included. The social forces model is being widely used when analysing pedestrian interactions on a microscopic scale. Although there is a discussion about the applicability of the social force model in crowd motion cases (Duives et al., 2013), the focus on individuals provides for mostly accurate one-on-one interactions between pedestrians. With the advancement of the SFM, other modes of transportation, including cyclists, have been attempted to be added to these models (M. Li et al., 2011, LIANG et al., 2012).

However, it remains a challenge for the social forces model to incorporate cyclists. The model would at least require the implementation of additional rules, as the social forces on an agent are induced by traffic participants in the vicinity (Y. Li et al., 2021). For realistic cyclist movement or pedestrian movement towards cyclists, this nearby range is not sufficient since the speed of cyclists force them and other participants to make decisions on movements further ahead in time.

Y. Li et al. (2021) tackles this issue for the movement of cyclist agents by introducing a modified social forces model. The main research case focusses on the lateral dispersion of cyclist groups when crossing an intersection following three (interactive) cases: freely moving, following, and overtaking. The additional elements of the SFM introduced in this study are the dynamic boundary model and a behaviour force that incorporates decision making. The model bases the behaviour force on the findings of a decision model, in which these findings are then modified to a formulation of a social force. This suggests that cyclists tend to follow a process of making their movements based on the choice between a set of options, rather than a movement based on gradual changes in the (near) environment as with the SFM.

W. Wang et al. (2024) uses a similar approach to model cyclist behaviour on e-bikes by introducing a modified version of the SFM. This study introduces the modification of the SFM for both cyclists and pedestrians and therefore also focusses on their interacting forces. For pedestrians, the study assumes most of the forces introduced by Helbing and Molnar (1995) with its subsequent modification (Helbing et al., 2002). An adjustment that is introduced is the force from soft separators which refers to a passable border stepping off the infrastructure. For pedestrians, the goal force and the force from other pedestrians remains similar. The force of cyclists is introduced as a force increased in size pointing from the cyclist to the pedestrian. This force is enlarged by a scale factor, because the higher speed and swaying characteristic of the cyclist make the pedestrian more reluctant towards cyclists.

Furthermore, W. Wang et al. (2024) proposes to significantly change the social forces exerted on cyclists, with the introduction of a following or overtaking force, a boundary force and a modified pedestrian-cyclist force. The pedestrian-cyclist force is based on a calculation of the cyclist predicted to coincide with the personal safety buffer of the pedestrian. If the cyclist is predicted to collide with this pedestrian or the space directly around it, the cyclist will adjust it's direction according to the closest escape direction (in front or behind the pedestrian). It is notable that this study accounts for an adjustment in the cyclist's movement toward the pedestrian based on predictive movement. In contrast, pedestrian movement towards cyclists is influenced solely by a force exerted from the current position of the cyclist to the pedestrian.

The study of Y. Wang et al. (2024) examines the safe space of pedestrians and e-bicyclists that they prefer to maintain during crossing. The results show that pedestrians have a semicircular safe space with uniform distance in all directions, while e-bicyclists require a semi-elliptical safety zone with greater distance in the forward direction. As speed increases, the safe areas for both pedestrians and e-bicyclists expand, with the forward travel distance of e-bikes being most significantly affected. This examination of safe spaces shows that people tend to remain at a certain distance from each other and that this distance is dependent on the speed of that particular person.

Other studies account for the speed and direction of movement as well when simulating pedestrian-cyclist interaction. The paper of Yuan et al. (2019) analyses the interaction between pedestrians and cyclists in a shared space and proposes a social force formulation that bases the movements of an agent on the anticipated movement of other traffic participants. The study highlights that accounting for the direction and speed of other traffic participants will improve the simulation of the interaction.

Dias et al. (2018) performs a similar approach in the interactions between cyclists, pedestrians and segway riders on a shared road. Here, the interactive force is based on the relative position of two agents and the relative velocity of these two. The final formulation considers a force exerted from one agent to the other with the size of the force dependent on an elliptical shape with the major axis in the movement of direction. This elliptical shape is combined with the formulation of Helbing et al. (2002) in which the force monotonically decreases with the angle of approach, so backward forces are minimised. However, this formulation is only tested by exerting the interactive force from the pedestrian towards the cyclist and not the other way around.

Summarising, several studies propose adaptations to the original social forces model to incorporate cyclists. The original SFM does not suffice to simulate cyclists in a proper way due to the difference in characteristics, so adaptations of the social forces or the implementation of additional rules are necessary. Many of the studies propose a highly adapted model for the movement of the cyclist, yet in some studies the changes this induces to pedestrian behaviour is not or barely considered. Other studies that do propose an adaptation of the forces exerted to both cyclists and pedestrians show that the velocity and the prediction of movement has an impact on the social force formulation. Table 2.1 provides an overview of the adaptations that the different studies applied, what mode their focus is on, and whether they make use of a prediction of movement.

Table 2.1: Summarising table of the studies focussing on a social forces and the adaptations thereof.

Study	Social Forces Adaptations	Focus on ped/cyc	Movement prediction (ped / cyc)
Helbing (1995)	Introducing three types of forces	Pedestrian	No / -
Helbing (2002)	Angular dependence	Pedestrian	No / -
Xi (2010)	Vision field, grouping behaviour	Pedestrian	No / -
Duives (2013)	-	Pedestrian	-
Li (2011)	Border force	Cyclist (& cars)	- / Yes
Liang (2012)	Trajectories choice model and Psychological-physical force model	Cyclist	No / -
Li (2021)	Dynamic boundary model, behaviour force based on decision making	Cyclist	- / Yes
W. Wang (2024)	Force of soft separators, following or overtaking force	Pedestrian & cyclist	No / Yes
Y. Wang (2024)	-	Pedestrian & e-bike	-
Yuan (2019)	Force formulation based on the next time step	Pedestrian & cyclist	Yes / Yes
Dias (2018)	Force based on elliptical shape in movement direction	Pedestrian & cyclist	No / Yes

2.3. Conclusion

The sub-research question corresponding to this literature review was: what types of movement patterns between cyclists and pedestrians have been identified in existing research, and how have these been implemented in social force-based models?

Some research has been performed in cyclist behaviour and much more in pedestrian behaviour, yet the interaction between cyclists and pedestrians has been studied limitedly, both in quantities of studies, as well as types of approaching cases. Approaches from behind and head-on have been studied in several shared environments. However, an environment where cyclists and pedestrians are separated by pavement and where they cross each other sideways is not considered extensively when observing behaviour. Many of the studies follow an approach to classify the interactions based on the proximity of the conflict. Post-encroachment time (PET) is a commonly used measure to create this classification, but the studies do not seem to agree upon the exact threshold value for this classification. This study seeks to address these missing links in the literature, which is further detailed in the Methodology.

Several adaptations of a cyclist model in social forces have been proposed to mimic certain behavioural aspects of the cyclist towards pedestrians better. Many studies use modification of the social forces or are even overruling the social forces suggesting that cycling behaviour cannot be modelled with social forces alone, because the formulation requires major changes. With this focus on modifications applied to the behaviour of cyclists, the behaviour of the pedestrian towards the cyclist is thereby often underexposed. Because cyclists commonly have higher speeds, the simulated behaviour is adapted in such a way that a prediction of movement is necessary for implementing cyclists and the behaviour of pedestrians towards these cyclists properly.

In short, the gaps in the literature motivates a focus on finding a PET threshold value based on pedestrian-cyclist interactions. It furthermore stresses that behavioural understanding is required for modelling practice, especially in the context of pedestrian responses towards cyclists.

3 Methodology

This chapter discusses the methodology framework that is used to answer the research question: *In what way do pedestrians change their movements when approaching and crossing bike paths with oncoming cyclists?*. The framework, provided in Figure 3.1, explains how it addresses the sub-research questions (SRQs) to describe movement changes from pedestrians towards cyclists on cycling paths.

From the Literature Review (Chapter 2), a literature gap was identified to answer the first sub-research question (SRQ1), which involved reviewing existing studies on interactions between cyclists and pedestrians and the modelling thereof.

Next, the framework focuses on selecting the data from specific locations and crossing scenarios based on a list of requirements. SRQ2: *What data is needed to describe the movements of pedestrians towards cyclists?* is thereby answered by determining the types of data needed to describe pedestrian movements towards cyclists.

Following this, SRQ3: Which data processing and enrichment methods are required to prepare pedestrian and cyclist data for a behavioural analysis? is addressed through a structured data preparation process. The selected data undergoes enrichment to add meaningful variables. The preprocessing step ensures that the dataset is sufficiently detailed to support a thorough behavioural analysis.

This analysis section involves classifying the data to organise it into meaningful categories. By finding threshold values based on previous research, benchmarks for this classification are developed together with findings in the data. This classification helps to understand different types of movements that pedestrians exhibit when approaching cyclists in sideways conflicts. With this insight, SRQ4 can be answered: *What movement changes does a pedestrian perform when approaching a cyclist in a sideways conflict?*. Statistical testing is applied to the classified data to identify possible relationships.

In the final step, the model's social force formulation is analysed and adapted on the basis of the findings of the behavioural analysis and the literature. Thereby, SRQ5 is addressed: *What adaptations can be done to the social forces model to mimic the behaviour of pedestrians towards approaching cyclists?*.

The framework as described is visualised in Figure 3.1, where the different steps with the related sub-research questions are shown in a flow chart.

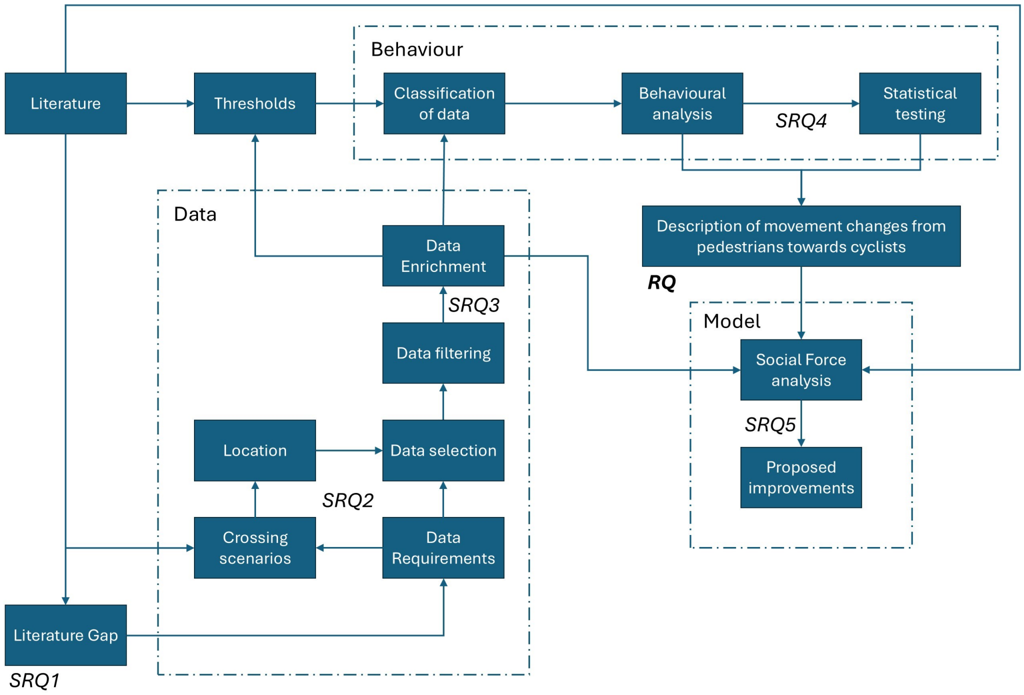


Figure 3.1: Framework of the Methodology

This chapter starts by explaining the concept of behavioural treats that is expected to occur in the interactions between pedestrians and cyclists. Next, it is clarified what the data requirements are to enable a focus on the sideways interaction. Subsequently, the preprocessing steps behind the data are explained. Hereafter, the method for deriving the behavioural insights from the data is presented. Finally, the method for setting up a test case in MassMotion is provided.

3.1. Conceptual Framework

To create an overview of the behavioural aspects that determine the pedestrian and cyclist interaction, a conceptual framework is constructed in which the different decision steps are highlighted that a pedestrian and cyclist are confronted with during their interaction. With this conceptual framework, it is clarified what specific types of variables need to be extracted from the data to be able to classify the interactions of pedestrians and cyclists and the changes of movements that follow from this interaction. The framework is highlighted in Figure 3.2. The process of interaction is schematised in a feedback loop in which the changes in behaviour result in new behaviour for the next time-step.

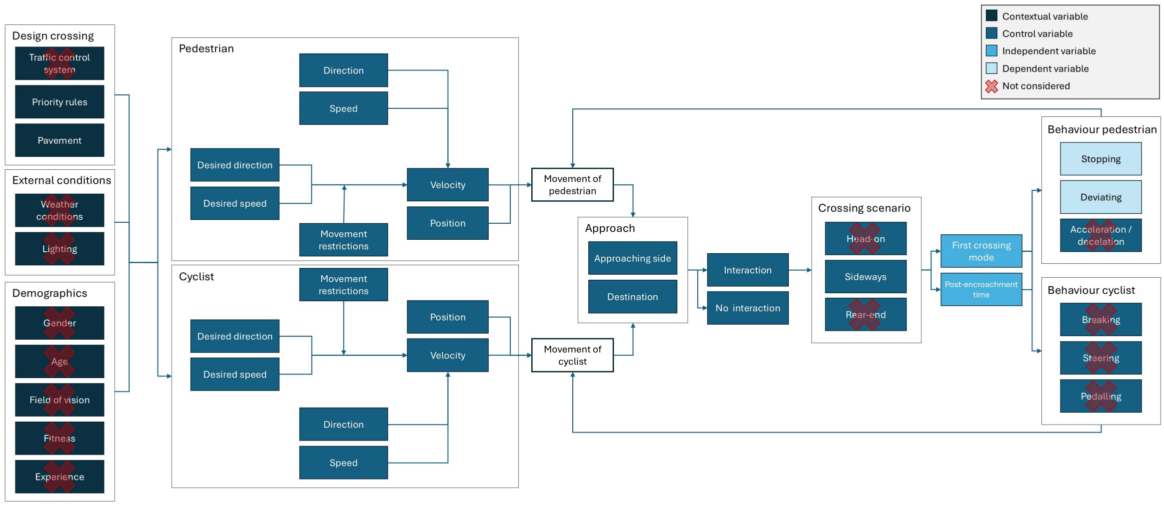


Figure 3.2: Conceptual framework of an interaction between a cyclist and a pedestrian.

The process consists of four types of variables: contextual, control, independent and dependent variables. The dependent variable is the focus of this research and follows from the outcomes of the research. This dependent variable is affected by the independent variables, and this research aims to explore the relationships between them. The independent variables are furthermore influenced by the control variables. To reduce the impact of these control variables on the relation, these variables are kept constant, or these will be used as variables for further categorisation. The interaction between pedestrians and cyclists are lastly influenced by contextual variables that determine the (external) circumstances traffic participants are exposed to.

The control variables can be detected by observing movements. Both pedestrians and cyclists will display initial movements that can be classified into speed and direction. Every pedestrian and cyclist has a certain speed and direction for each time-step, but also a desired speed and direction, which might deviate from the actual speed and direction due to the contextual variables listed in Figure 3.2. This research will mostly not consider these variables so a more general result is provided. The speed and direction combined determine the initial movement of both pedestrians and cyclists. This movement could then be influenced by the presence of other traffic participants, yet this research only considers the cyclist-pedestrian approaching scenario. This study distinguishes between the different approaching sides and the intended destination between a cyclist and a pedestrian, which is explained in more detail in 4.2. Once a pedestrian and cyclist approach one another, a change in the initial movement that was intended can occur. The approach can either result in a situation where no physical interaction takes place, therefore no significant movement changes have to be made, or an interaction does take place and behaviour will likely be adapted.

The interaction can be classified into three distinct categories, following the study of Beitel et al. (2018): a head-on, sideways and rear-end interaction are considered as the crossing scenarios between a cyclist and a pedestrian, of which the sideways crossing is studied. The severity of the sideways crossing can be expressed in the post-encroachment time, the measure to determine the time-wise crossing proximity of two traffic participants. A second measure to determine the severity of the crossing can be based on the mode that crosses first. The knowledge on the severity of the potential conflict helps to determine the movement changes that can be expected from a pedestrian when in conflict with a cyclist.

These movements can be categorised into several different behaviours, yet the actual physical result of the decisions that are made by pedestrians in this research are distinguished into three categories: stopping, deviating and accelerating/decelerating. For cyclists, the resultant behaviour can be divided in breaking, steering and pedalling. How the movement changes for the cyclist during an interaction is not considered, so these movements are assumed to be fixed or predetermined. Therefore, the movements of the cyclists are relevant to study with respect to that of the pedestrian to see how the pedestrian acts towards the presence and movements of this cyclist.

The stopping and deviating behaviour for the pedestrians are the dependent variables in this study. The occurrences of stopping and deviating are defined by threshold values. The relation towards the independent variables is used to detect the extent that stopping and deviating behaviour occurs.

3.2. Data Requirements

Following the conceptual framework, the movements of individual pedestrians and cyclists need to be captured accurately to be able to detail the behaviour towards each other. This should be done in a paved environment with dedicated bike and footpaths. The interaction process described by the conceptual framework should take place often enough so significant information is collected. This reasoning shapes the following list of requirements that serves as a guide for selecting proper data to analyse pedestrian-cyclist interactions:

Trajectories Pedestrian's and cyclist's movements need to be analysed following their position in time during interactions. Because in this study the focus is on individual interaction between a cyclist and a pedestrian, the trajectories need to be precise to be able to detect the slightest movements of both modes during their interaction. The data thus requires a high resolution (time and space-wise). As noted in the Literature Review (Subsection 2.1.3), almost all interactive behaviour occurs within a 5-second window before the other participant arrives. Therefore, pedestrian movements should in general be recorded for at least this duration.

Individual data The pedestrians and cyclists need to be distinguished from one another either visually or by registration of an ID.

Paved environment The environment that is analysed needs to have the proper infrastructure for both modes to be able to perform their regular movements.

Sufficient quantities of active modes The location needs to have a sufficiently large group of pedestrians and cyclists to be able to analyse common behaviour. The study of Afghari et al. (2014) with a quantity of 225 interactions was after a filtering process left with 17 interactions. This study therefore aims to find data of at least double the amount of 225 interactions for a certain scenario.

Separation of modalities The environment should have an indicated separation between cyclists and pedestrians by pavement, and it should induce the need for a pedestrian to cross the bike path, following the Literature Review conclusion (section 2.3) on the lack of knowledge on interactions at separated cycling paths.

No traffic control system The location should be free of any traffic control system that influences the behaviour of pedestrians and cyclists.

The list is used in the process of the data selection in Chapter 4. This chapter discusses a method for distinguishing intersection configurations and the possible crossing scenarios that could take place thereon. From there, the definitive site and data selection is presented that stems from the chosen crossing scenarios and the data requirements.

3.3. Data Enrichment

To categorise the data effectively, information must be extracted to distinguish between cyclist and pedestrian trajectories, identify interaction cases, and classify different behavioural patterns in these interaction cases. After the data and site selection, the data thus undergoes a process of enrichment. Below, a list of necessary information that is extracted from the data is provided:

Speed This information helps to differentiate cyclists from pedestrians. It also offers insights into the general behaviour, e.g. stopping, decelerating, accelerating, minimum and maximum speeds. The speed of person is determined for each trajectory point based on the locations and timing of the surrounding points.

Location The locations, x and y positions, are already provided in the data, yet with respect to the layout of the crossing, the location is still unknown. When this is clarified, it contributes to determining the type of mode based on their position on the cycling or footpath. It can also provide insight in adherence to pavement guidance. The location will be determined by fitting a blueprint of the crossing to a random sample set of trajectory points to match the pavement borders of the area.

Origin and Destination Following the locations of the flows of people, the origins and destinations can be determined, each referring to the branches of the crossing. These can specify the approach between a cyclist and a pedestrian. The origin and destination are derived from the location and direction of the first and last trajectory points.

Crossing Point The crossing point between the cyclists and pedestrians can be used to approximate the conflict area and thus where the pedestrian and cyclist are in relation to this point. The crossing point is estimated to be the point where the two segments of both trajectories cross, calculated by linear interpolation.

Crossing Angle The trajectories of a pedestrian and a cyclist that cross each other from the side are the trajectories that are being analysed. The crossing angle is based on the exact two segments of the trajectories that cross each other and need to be at least 30 degrees.

Post-encroachment Time (PET) The PET provides information on the severity of the conflict between a cyclist and a pedestrian with a lower PET resulting in a more dangerous conflict. The PET is calculated by taking the difference in time it takes for both modes to reach the crossing point, the time being calculated by linear interpolation.

First Crossing Mode The conflicts where pedestrians cross in front of cyclists, will result in different behaviour than the crossings where cyclists cross first. The first crosser is determined by comparing the times that both modes reach the crossing point (same time for the PET calculation), of which the lowest time corresponds to the first crosser.

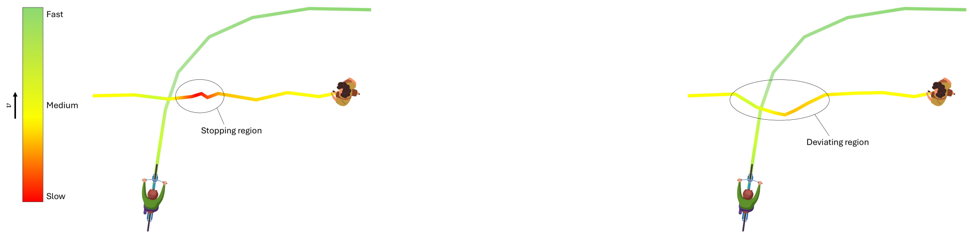
Predicted post-encroachment time The predicted PET can contribute to determine how the movement with respect to the approaching cyclist changes over time. The predicted PET is calculated by considering a prediction of the movement direction combined with the speed, and follows a similar approach as the actual PET.

The process of enriching the data with these values is explained in more detail in Chapter 5. The acceleration and rotations of trajectories are not determined, because these do not provide information on the gradual developments of a trajectory that are necessary to study the stopping and deviating behaviour of pedestrians.

3.4. Behavioural Analysis

This section discusses the method for deriving the behaviour of pedestrians towards cyclists through an analysis on the movement changes that pedestrians tend to make when being confronted with an approaching cyclist. This section aims to describe how pedestrians adapt their movements to avoid potential conflicts by specifically considering stopping and deviating behaviour. Also, a method is introduced to highlight more nuanced movement changes by considering a prediction of PET.

To create insight into the pedestrians that interact with a cyclist, trajectory plots of pedestrians crossing with cyclists are made that visualise the estimated speed at every measured time instant. The speed is visualised with a colour plot. Through such trajectory plots, sections can be highlighted that show stopping, slowing down, and deviating from the main direction. Two schematic examples are given in Figure 3.3 of trajectories of a pedestrian that stops 3.3a and deviates 3.3b with respect to an oncoming cyclist. The method for analysing the behaviour is further detailed in the subsections 3.4.1, 3.4.2 and 3.4.3



(a) Pedestrian stopping trajectory

(b) Pedestrian deviating trajectory

Figure 3.3: Two samples of a pedestrian crossing with a cyclist by performing stopping behaviour (a) and deviating behaviour (b).

3.4.1. Stopping Behaviour

The interactions that are detected through the visualisation require a categorisation that comprises the different severities of conflicts introduced in section 2.1.2. PET is used as a measure to distinguish the severity of conflicts, however multiple studies differ in opinion on the exact threshold(s) for PET. Therefore, this study performs a method for deriving this PET threshold using a data-based approach, thereby assuming that the severity of the conflict has an influence on the exhibited behaviour found in the data. In addition to PET, there is a distinction of conflicts where the pedestrian crosses before and after the cyclist, following the categorisation of Beitel et al. (2018).

For all PET scenarios and both the scenarios of the pedestrian crossing in front of or behind the cyclist, the stopping distances that pedestrians tend to keep are calculated. The stopping distance is in this case defined as the distance between the crossing point of the two trajectories to the trajectory point that gets below a threshold speed that is considered stopping. Because there is no clear reason to assume that pedestrians tend to stop whenever they cross before a cyclist, the stopping distances that are important for understanding pedestrians behaviour is in the case where the cyclist crosses first.

The pedestrians are assumed to maintain a certain amount of distance before they cross, with each individual having a slightly different preference on the distance they keep to maintain a safe feeling. It is for this reason that the stopping distance is assumed to follow a normal distribution. The stopping distances for each individual are plotted in a histogram to visually explain the normal distribution assumption and a statistical test is performed to determine its significance.

The statistical test follows a null hypothesis (H_0) that the distance that pedestrians tend to slow down from is normally distributed. The alternative hypothesis (H_1) is that this distance is not normally distributed. The performed test is a normality test which uses a combination of the D'Agostino and Pearson's test (Strangman, 2002).

If the resulting p-value is more than 0.05, the null hypothesis cannot be rejected, and there's no significant evidence that the data is not normally distributed.

The distances that pedestrians tend to keep from cyclists can be translated into MassMotion. However, it is unlikely that all pedestrians make the decision to stop when encountered with an approaching cyclist. Therefore, it is essential to estimate how many individuals actually make the decision to stop. This understanding helps clarify the range of behavioural options pedestrians consider in such situations.

3.4.2. predicted PET

Stopping is however only a part of how pedestrians could behave to eventually cross in front or behind the cyclist. This eventual outcome can be the result of several movements along the path. By studying a more gradual development of an approach between a cyclist and pedestrian, insights can be gained in the different movements that are performed besides stopping.

The predicted PET is a measure to study the development of an approach between a cyclist and a pedestrian. The actual PET is calculated for every crossing cyclist-pedestrian pair, and with some assumptions on the future movements of the pedestrian and cyclist, this can be done similarly for the predicted PET. A prediction of PET per time step provides the proximity of a conflict if two participants were to continue on their course. If this predicted PET changes, this means that the pedestrian or cyclist is adapting its movement either in speed or direction. The advantage of this method is that even the slightest movement changes by pedestrians are translated into this prediction of PET, which allows an analysis on more nuanced movement changes.

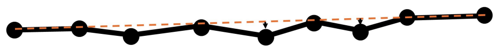
The assumptions for predicting PET are presented in Chapter 5 and the results of the development of this PET for the pedestrians are then explained in Chapter 6.

3.4.3. Deviating Behaviour

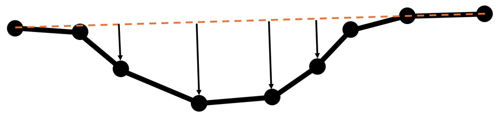
Apart from the stopping behaviour of pedestrians towards cyclists, the deviating behaviour is analysed. Deviation is assumed to occur in the cases a pedestrian crosses a cyclist from behind and is considered as an alternative to coming to a full stop. Instead of the pedestrian waiting for the cyclist to pass, the pedestrian will walk in the direction of the cyclist to be able to cross after the cyclist earlier. Another situation of deviation is assumed to occur where the pedestrian crosses in front of the cyclist, but it deviates away from the cyclist to avoid a collision.

Whether the pedestrian performs deviation will be studied only for the origin-destination pairs that go straight. For the south-west path at the Lorentzweg crossing, where the most optimal or convenient route is ambiguous, calculating meaningful deviation becomes unreliable.

The deviation of the path is determined as the lateral distance that all trajectory points have from the straight path. The straight path is determined as a straight line between the first and last point of the trajectory. Figure 3.4 provides an example of a non-deviating trajectory 3.4a and a deviating trajectory 3.4b with a visualisation of these lateral distances for each trajectory point. The value for deviation of a trajectory is calculated by using the method of the root mean squared deviation and by considering the maximum deviation, which is presented in the Results chapter.



(a) Non-deviating trajectory



(b) Deviating trajectory

Figure 3.4: Two samples of a trajectory without (a) and with (b) significant deviation.

An assumption is made on the deviation values of pedestrians interacting with cyclists compared to non-interacting pedestrians. Since, for the non-interacting pedestrians, there are less reasons to deviate from the intended path towards the destination, it is assumed that the distributions of deviating values are significantly different between interacting and non-interacting pedestrians. The null hypothesis thereby becomes (H_0) that there is no-significant difference in the distribution of deviation from a straight path towards the destination between an interacting and non-interacting pedestrian. The alternative hypothesis (H_1) states that there is a significant difference in the distribution between these two groups of pedestrians.

If the resulting p-value is more than 0.05, the null hypothesis cannot be rejected, and there's no significant evidence that both distributions are different.

3.5. MassMotion test case

The MassMotion environment is used to perform a test case on a cyclist-pedestrian interaction based on the main findings on stopping and deviating behaviour in the data. MassMotion consists of a user interface (UI) and a software development kit (SDK), in which it is possible to create and change the digital environment and the behaviour of the agents. This section discusses the set-up of the test case that is performed in MassMotion to detect potential issues in the current model.

The process of performing the test case consists of four phases:

1. A test sample is selected from the data that represents the behaviours that are the focus of this research.
2. The environment of the crossing is recreated in the UI to fit the layout of the actual crossing.
3. A test is performed that uses the current formulation of the model, scripted in the SDK, as input for the movement of the pedestrian and cyclist agent.
4. Adaptations are suggested in the SDK to improve the projected behaviours in the model.

For both the current and adapted formulation, the results are compared with the selected test set of the data. This is done by performing a visual and numerical comparison of the trajectories for both simulations with the actual trajectory.

4
Data selection

This Chapter provides the answer to sub-research question two: What data is needed to describe the movements of pedestrians towards cyclists? Thus, data needs to be collected that can capture the scenario described in the conceptual framework. It is essential to select suitable locations where the data collection will take place. The choice of crossings must be justified to ensure that they provide representative data on pedestrian-cyclist interactions. In addition, specific approach scenarios must be identified at these crossings to extract a variety of encounters, and thus behavioural patterns.

First, the optimal crossing configurations to detect sideways interactions are selected. Second, the multiple scenarios in which cyclists and pedestrians can cross each other on these intersections are presented and a few of these approaches are selected for studying in detail. Lastly, the method of collecting and collection site is detailed following from the configurations and the list of requirements presented in section 3.2.

4.1. Intersection Selection

As discussed in the Literature Review (subsection 2.1.1), infrastructure plays a guiding role in shaping the behaviour of both pedestrians and cyclists during their journeys. Active mode intersections are locations where the paths are more likely to intersect and because of these converging paths, these locations offer opportunities to observe a variety of behaviour. To better understand the nature of these interactions, this research focusses on selecting an appropriate intersection.

The intersections that are considered are based on those of the study of Wei et al. (2021) with the addition of the shared space, supported by the *Crossing flows - Random* category of Duives et al. (2013). Figure 4.1 visualises the different crossing types that will likely induce a crossing scenario between cyclists and pedestrians. On these intersections, the likelihood of a sideways interaction with at least a 30 degree angle increases when the legs of the intersection are positioned perpendicular to each other. Therefore, the merge, X-, and Y-intersection are less suitable crossings for observation.

For the multi-way intersections, the sideways interactions can be observed, though the complexity of these crossings increases with the number of paths connected to the intersection. When the complexity of a crossing increases, the pedestrians and cyclists might behave in a hesitant way when approaching the crossing, having to scan the traffic from all directions. Although this behaviour would be relevant for this study, it is beyond the scope of this research, where the focus should be on a one-on-one cyclist-pedestrian interaction without other disturbances. Shared spaces are being applied more often at squares to let the flows be regulated by interactive negotiations between traffic participants. Although the shared space provides many of these interactions, several studies have analysed different interactive scenarios in these environments and this study aims to analyse the pedestrian-cyclist interaction by analysing locations with a separation of modes.

The remaining three intersections are the crossroad, the T-intersection, and the misaligned intersection. Although the misaligned intersection most likely induces other approaching behaviour than the

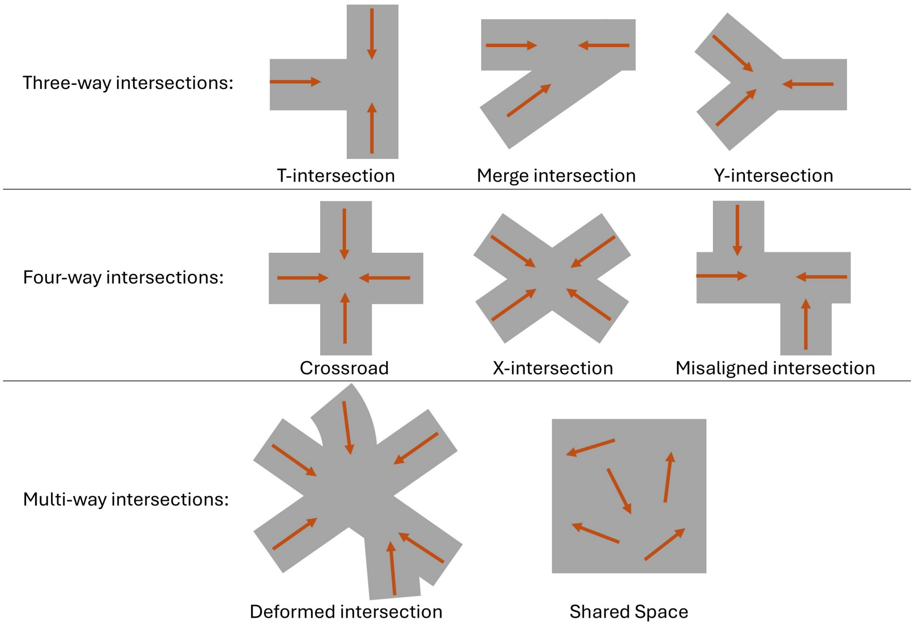


Figure 4.1: Intersection possibilities where flows of people tend to cross each other. Based on the paper of Wei et al. (2021) where the roundabout has been left out and a shared space is introduced.

T-intersection, the intersection is for the categorisation of crossing flows assumed to be a double T-intersection. This leaves two categories of intersections that are studied: the T-intersection and the crossroad intersection.

For the situation on the T-intersection, the topology creates a non-symmetric situation, with a major road going straight and a minor road that is perpendicular to it. Important to note here is that this study considers T-intersections where no specific hierarchy is given to one of the legs of the crossing. The priority rules that apply are the same for pedestrian-vehicle crossings where the right of way to pedestrians is only given in the scenario if the pedestrian approaches the cyclist head-on or from behind and the cyclist needs to cross the pedestrian by turning off the straight. In every other approach, the cyclist will have right of way because the space is specifically assigned to cyclists by a cycling path.

4.2. Crossing Scenarios

The focus of this study is on the behaviour of the pedestrian towards the cyclist. A categorisation of the crossing scenarios has been proposed to distinguish between different circumstances for the pedestrian. This categorisation will thus need to take into account the intended path of the pedestrian and how the cyclist approaches this pedestrian, influencing the behaviour of the pedestrian.

Therefore, a distinction for the crossing scenarios can be made based on two categories: the intended movement of the pedestrian and the respective approaching branch of the cyclist towards the crossing. The movement of the pedestrian is dependent on the origin and destination. The origin can be either one of the two branches on the major road going straight or the minor road that is the side branch of the T-intersection, the same applies for the destination. If the pedestrian were to interact with the cyclist, the possible crossing scenarios for the pedestrian need to be outlined. The pedestrian could either go straight on the major road, turn away from the major to the minor road, or make a turn from the minor road to the major road. For each of these categories, the pedestrian will need to cross the cycling path somewhere. This leaves six scenarios for the movement of the pedestrian, assuming that the mirrored image is a similar case. Figure 4.2 gives an overview of these six scenarios of pedestrians crossing the cycling path:

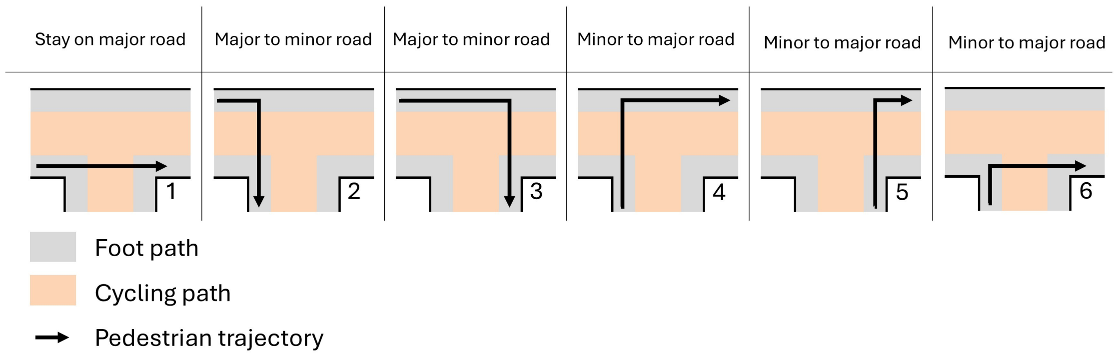


Figure 4.2: Possible ways a pedestrian can cross a T-intersection.

It has to be noted that both the origin and destination of the pedestrian matter for the type of crossing, because the actual trajectories are likely not a straight line, but rather a diagonal or curved path thereby either crossing already at the branch of the crossing or more in the centre of the crossing.

The cyclist that approaches the pedestrian can in this case come from the same branch, a sideways branch or the branch opposite of the pedestrian and it will need to cross the path of the pedestrian. Figure 4.3 below shows the possible crossing cyclist paths which origin is either from the same branch (from behind), from a side branch (sideways), or from the opposite branch (Head-on). Some cases have multiple crossing cyclist trajectories due to its different destinations or (sideways) origins, these are indicated with the letters a to d, starting from the north direction going clockwise. In the cases numbered 16 to 18, the crossings are not possible, because the cyclist cannot approach from an opposite branch.

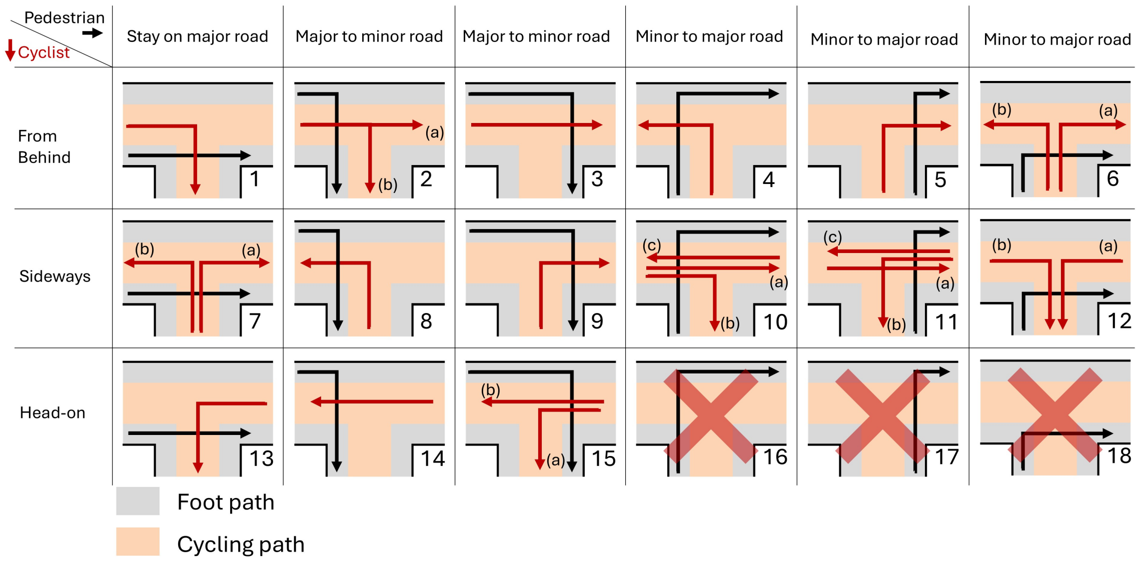


Figure 4.3: Possible approaching situations between a cyclist and a pedestrian on a T-intersection.

The crossroad scenario has many more crossing scenarios compared to the T-intersection due to the additional branch, yet many of the interactive situations are mirrored situations of others. Therefore, for the purpose of visualising, the starting branch for the pedestrian will be the same in all scenarios, only the destination will differ. The directional options the pedestrian has are shown in Figure 4.4 below with the pedestrian starting in the bottom right corner and crossing the intersection to the directions in a clockwise sequence:

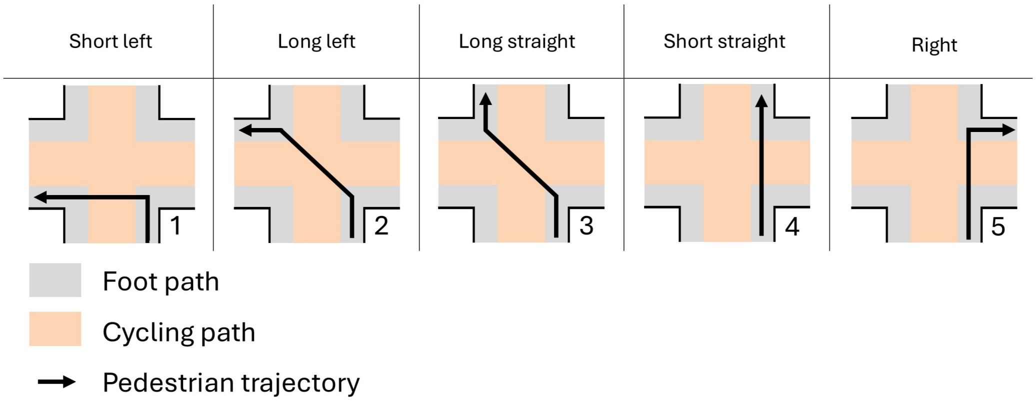


Figure 4.4: Possible ways a pedestrian can cross a crossroad intersection.

Again, the cyclist can approach the intersection from the branches categorised behind, sideways or the front. The cyclist can in multiple cases have several routes that lead to a crossing scenario, which are labelled from a to d in Figure 4.5.

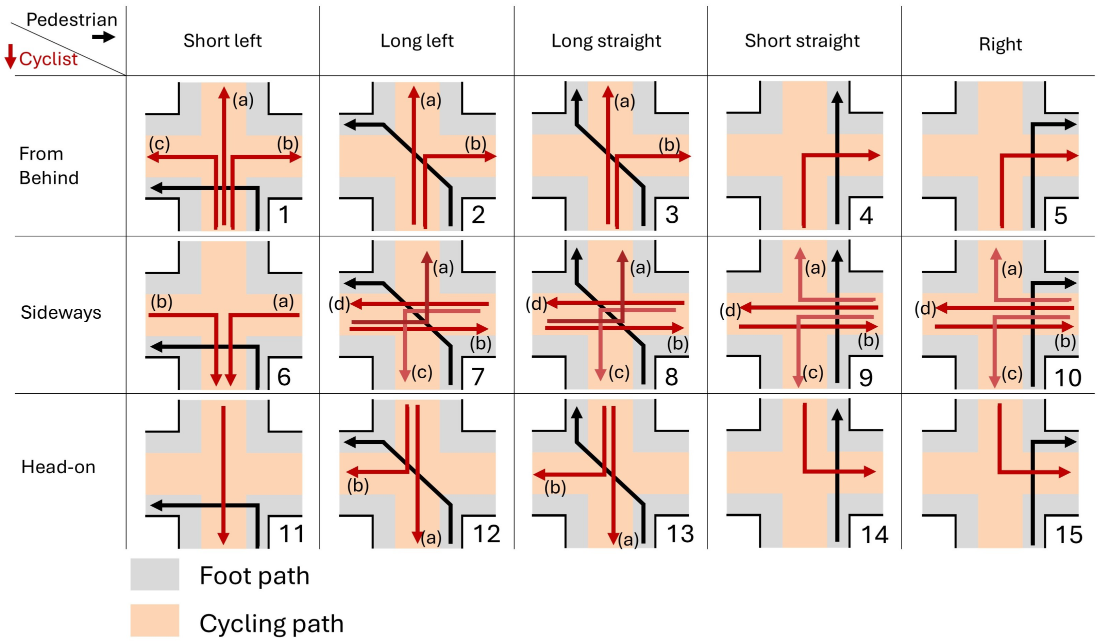


Figure 4.5: Possible approaching situations between a cyclist and a pedestrian on a crossroad intersection.

The analysis of pedestrian behaviour begins by considering the pedestrian paths at these two types of crossings that were studied. This pedestrian movement can vary depending on the presence of a cyclist. To capture this variation, several crossing scenarios have been selected from both a T-intersection and a crossroad. From the crossroad, a considered simple sideways approach with both modes going straight is chosen, since the behaviour should be described in the most generic cases of crossings. This setup is represented in scenario 9b and 9d of Figure 4.5. In addition, two interaction paths from the T-intersection are included. The first involves a pedestrian turning onto the major road, as illustrated by both scenarios (a and b) of situation 6 and 12 in Figure 4.3. The second involves a pedestrian continuing along the major road, corresponding to all the scenarios of 1, 7, and 13. The analysis of the crossing scenarios starts from the pedestrian movements and will from there on distinguish between the different cyclist approaches.

4.3. Data and Site Selection

The choice of data collection method stems directly from the previously defined data requirements, which emphasize the need to observe a large number of interactions between pedestrians and cyclists across diverse user groups. The TU Delft campus has an extensive amount of active mode infrastructure, with high daily volumes of people, primarily students (TU Delft, 2024), using this infrastructure. This quantity of data is preferred to be captured by an automated sensor set-up, instead of the request for an observer, limiting the amount of interactions that are captured.

Newer technologies such as as Wi-Fi detection systems could be used to trace cyclists and pedestrians, but the precision of these detection systems is most likely not sufficient to detect pedestrian and cyclist movements with a high resolution in time and space. For analysing more specific movements video data can be used that captures the movements of pedestrians and cyclists. However, the use of the imagery raises privacy concerns if large quantities of people are involuntarily being recorded.

The TU Delft has installed smart sensors across the campus that trace both pedestrians and cyclists without the collection of actual video footage. Therefore, the data collected by these sensors is used. These sensors track the heads of cyclists and pedestrians with a frame rate of approximately 1 frame per second (FPS), while recent instalments can be upgraded to a frame rate of 4 FPS. The computer attached to the sensor assigns a specific trace ID to each trajectory, but it is not specified whether the trajectory is from a cyclist or a pedestrian.

The campus consists of various parts with a mix of transportation modes, but the specific intersections which are solely accessible by bike and by foot are the intersections located at the Mekelpark in the middle of the campus. The crossings need to align with the topologies chosen in subsection 4.1 and include sideways interactions between cyclists and pedestrians. Two intersections along the Stieltjesweg match these criteria, one for each of the topologies. The following paragraphs provide the information on the layout of these two crossings.

The crossroad, shown in Figure 4.6, is chosen as the location to analyse the sideways interaction between cyclists and pedestrians. This interaction of pedestrians approaching the cyclist track from the side and both modes going straight aligns with the indicated situation in subsection 4.2. Most pedestrians are likely to cross the bike path from the north side, needing to go to the bus stop south of the bike path, but also the north side will attract pedestrians, mostly during the lunch break when there are food trucks along this path. The pavement clearly separates the two flows making distinctions of the type of mode possible. The crossing is close to the Mekelweg, where the bus lane is located, therefore, this crossing will from here on be referred to as the Mekelweg crossing.

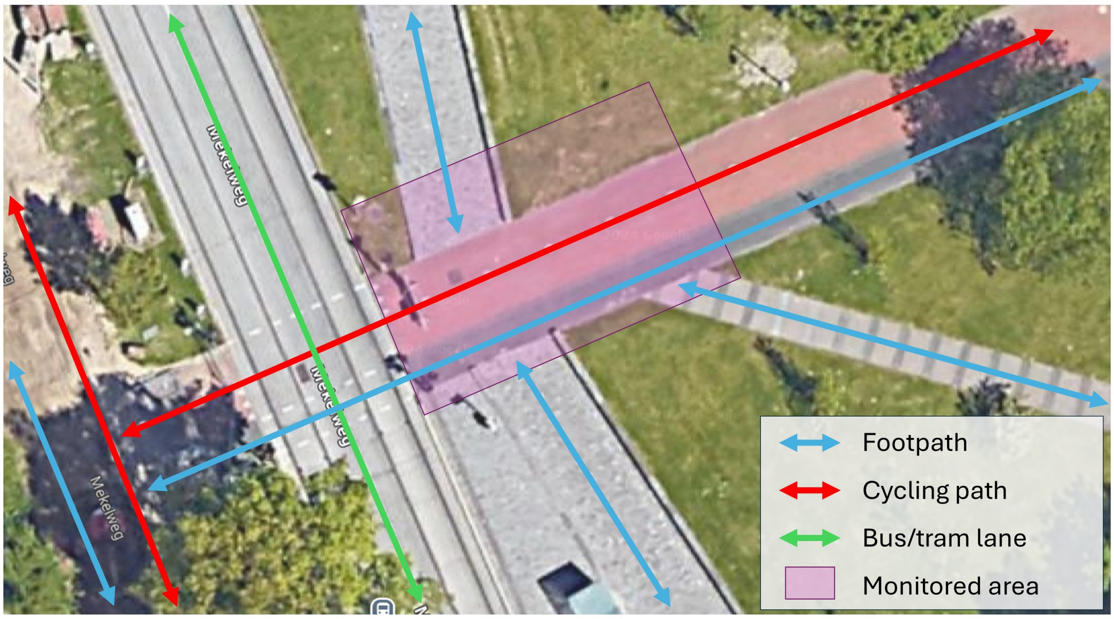


Figure 4.6: Blueprint of the Mekelweg crossing at the TU Delft campus where a cycling path crosses a footpath nearly perpendicular.

The crossing at current day (2025) is however under the influence of a construction project of a tramline, making the footpath inaccessible and transitioning the bike path into a steep ramp where cyclists are likely to brake. The behavioural changes induced by these construction works are not representative for this particular study; therefore, dates before the constructions started will be requested when accessing the data, which is before April 2023 (Wassink, 2023).

The other crossing that has been chosen to collect data from is shown in Figure 4.7. This crossing uses a combination of two sensors that have visuals on both T crossings with an overlapping area in between. An advantage of the data being collected at this intersection is that the frame rate has been set to 4 FPS, while the tram construction works are not of influence, so data of the current day can be collected. The Stieltjesweg crosses here with the Lorentzweg from the north, therefore this crossing will be referred to as the Lorentzweg crossing.

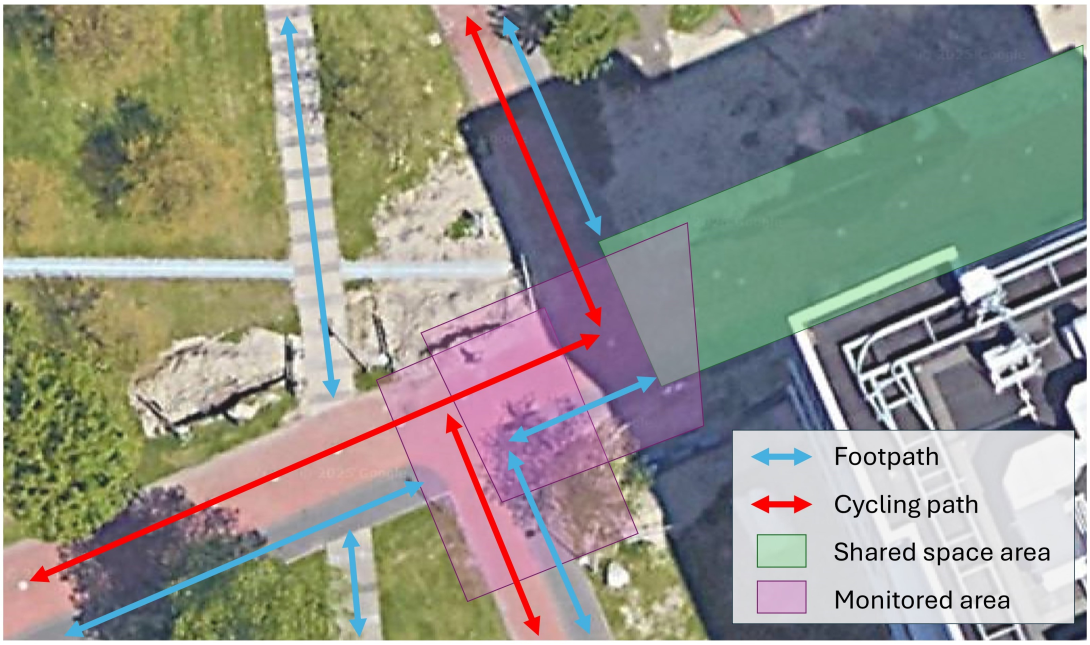


Figure 4.7: Blueprint of the Lorentzweg crossing with a larger detection area and more directions possible.

The specific interaction that is required from this double T-intersection is the pedestrian coming from the minor road in the south and turning onto the major road to the west. One drawback of the origin-destination combination for the pedestrian is that the pedestrian could have taken a shortcut just before this crossing, yet an observational study at this crossing has shown that many pedestrians still take this route. The other interaction that is analysed here is from the pedestrian coming from the east and needing to go to the west while crossing with a cyclist.

The four scenarios of pedestrian movements at the Mekelweg and Lorentzweg crossing are visualised in Figure 4.8 with a sample trajectory of a crossing cyclist. Chapter 6 will present the behavioural results from these crossing scenarios.

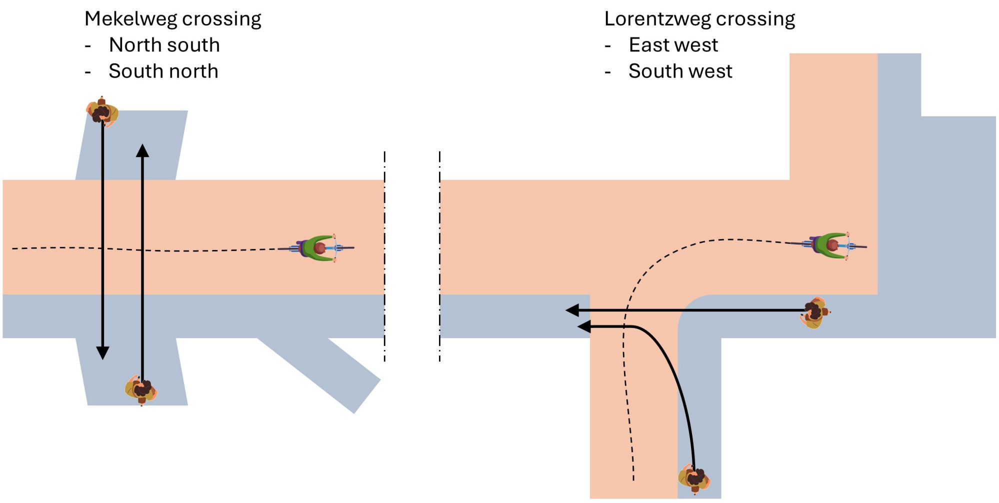


Figure 4.8: The four crossing scenarios for the pedestrian that are the focus point of this study with in both cases an example trajectory of a crossing cyclist.

4.4. Conclusion

The sub-research question related to the data selection was: What data is needed to describe the movements of pedestrians towards cyclists?

The data that is needed stems from the list of requirements that is proposed in this section. Trajectory data of individual pedestrians and cyclists with a separation of modes indicated by pavement is required to analyse pedestrian behaviour towards cyclists on bike paths. The environment should be a crossing without traffic control systems, so the interactions take place purely based on the negotiations between the two modes of traffic. The crossings that are analysed are a crossroad, which is the Mekelweg crossing, and T-crossing, which is the Lorentzweg crossing, at the TU Delft campus area, where multiple interactive scenarios are studied.

5 Data Preparation

Before the behaviour in the data can be classified and analysed, it requires steps of pre-processing, following Section 3.3 of the Methodology. Only after adding the values of speed, location, origins and destinations, and the different variables to distinguish types of conflicts, the stopping and deviating behaviour can be analysed and explained. This Chapter therefore aims to answer sub-research question 3: *Which data processing and enrichment methods are required to prepare pedestrian and cyclist data for a behavioural analysis?* by first providing the general filtering steps and quality assessment of the data. Secondly, the method for adding the different variables to the whole dataset is explained. Lastly, the circumstances and required adaptations for every crossing scenario are specified to be able to analyse the movement appropriately.

5.1. Data Preprocessing

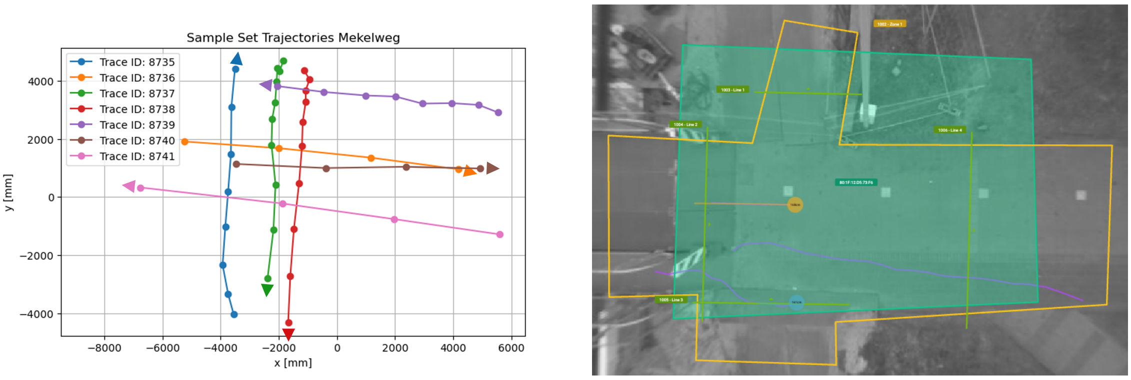
For this study, access has been granted by the TU Delft of trajectory data at the Mekelweg crossing from March 2023 to detect pedestrian-cyclist interactions. As mentioned in section 4.3, this month was chosen because of the current construction works. For the Lorentzweg crossing, more recent data from January 17 until March 3 2025 has been acquired. This section discusses the preprocessing steps required to take before this data can be enriched. This will be done by assessing the quality of this data and presenting steps to filter redundant data for the purpose of improving computational time.

5.1.1. Data Quality

The raw data is stored in separate json files for every minute of recording. The data that is registered provides x and y positions at a certain time belonging to a trace ID. The data that has been gathered at the Lorentzweg crossing not only provides insights into the trajectories, it also provides information on peoples height and whether there is a group formation or an individual person in the area. This study however, focusses on the interactions between individual cyclists and pedestrians. Therefore the group data is left out of the dataset. The heights of the participants are not taken into account as a contributing factor for the interactive behaviour.

This filtering step is processed in Python, in which the trace ID, date, time, x, and y position of every participant is stored in a single DataFrame per crossing. The Mekelweg crossing positions are given in millimetres, and the Lorentzweg crossing in meters, both with six decimal places, which suggests a very high precision. The sensor accuracy is however limited due to the nature of detection, which follows heads that are sized slightly more than 15 cm wide and 20 cm long on average (Reference, 2025). The eventual results are therefore rounded to one decimal place in meters for a more realistic interpretation.

The data was analysed through multiple videos of random sample trajectories. In the videos, points sharing a trace ID are linked in time order to form several trajectories. The trajectories are assessed based on the detection timing and accuracy at the beginning and end, the time intervals, and the accuracy of the positions. A representative set of trajectories is used to illustrate the overall quality of the data. This sample set is provided in Figure 5.1a, relating this to a blueprint of the Mekelweg crossing from the perspective of the sensor in Figure 5.1b. The direction of all trajectories is manually shown through an arrow-like triangle after the end point.



(a) Sample trajectories

(b) Blueprint

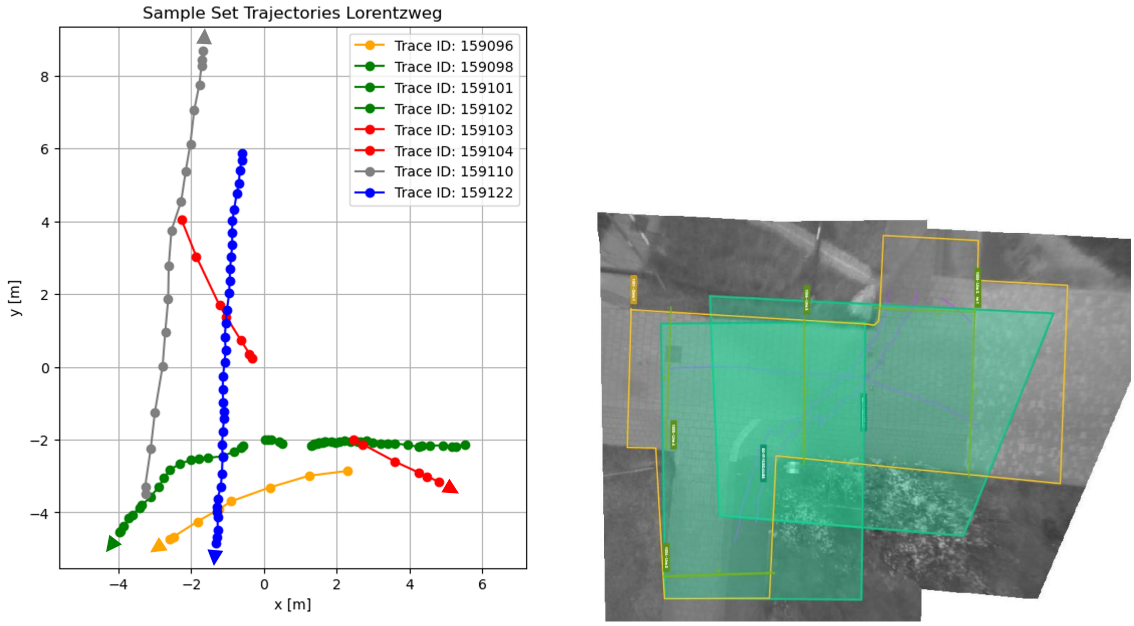
Figure 5.1: A sample set of trajectories at the Mekelweg crossing to enable a quality assessment of the data.

The beginning and end of the trajectories are in some cases not properly detected, especially at the edges of the detection area. For the Mekelweg crossing this is often the case for cyclists that enter the area with a high speed. Consequently, the first registered trajectory point can sometimes be already past another trajectory, which is the case for ID 8740 (brown) already having passed 8735 (blue), while it is very likely that these two paths would have crossed. Also in some cases, the trajectories seem to start late or end prematurely, which is likely due to the participant leaving the detection area, which happens for ID 8739 (purple).

The time interval of the trajectories at the Mekelweg was set to 1 frame per second (FPS), meaning that the time the location of a head is registered is every second. In the visualisation, it can be seen that some trajectories are therefore limited to just a few trajectory points, in this case trace ID's 8736, 8740, and 8741 moving in the x direction show only 4 points, which are likely derived from cyclists that reach higher speeds in general and are therefore leaving the detection area faster. In contrast, the movement of the pedestrians, likely indicated by the other trajectories 8735, 8737, 8738, and 8739 due to generally lower speeds, can be detected to distinguish stopping behaviour, which are visible in the parts of the trajectory that have dots closer together.

The axes of the trajectories do not match the blueprint of the crossing. The detection area indicated by the yellow line in Figure 5.1b shows that visible reach to the footpath to the north is higher than towards the south. This is opposite from the way most trajectories are projected, which extend more towards the south. The y values need to be flipped in order for the blueprint of the crossing to match the coordinate system. Further detailing of this fitting process from coordinates to blueprint is provided in subsection 5.2.1. The accuracy of the positions are determined by observing the course of the trajectories. The cyclist trajectories follow a mostly straight path, in line with the layout of the actual cycling track. The same applies for pedestrians, yet these trajectories seem to oscillate slightly. Especially when observing the trajectories where the pedestrian is likely standing still, which is for 8737 (green) and potentially 8738 (red), the position of the head does not remain in the exact same spot. Movements of the head are simultaneously realistic to occur, for example when the pedestrian aims to observe oncoming cyclists before crossing. Moreover, through every step the pedestrian takes, the head of the pedestrians show slight lateral swaying according to Parisi et al. (2016), which would align with the trajectories of these pedestrians showing similar movements. There is therefore no reason to assume that this oscillating arises due to a measuring error.

A similar plot of sample trajectories at the Lorentzweg is created in Figure 5.2a, next to the blueprint of this crossing in Figure 5.2b.



(a) Sample trajectories

(b) Blueprint

Figure 5.2: A sample set of trajectories at the Lorentzweg crossing to enable a quality assessment of the data.

A first remark to be made is that ID's 159098, 159101, and 159102 (green) are assumed to be of the same trace ID, as well as 159103, and 159104 (red). This assumption is based on a video analysis where the timing and position of the end of the first trajectory combined with its velocity almost perfectly matches with the start of the second trajectory. It is more likely these trajectories belong to the same participant, than that two participants are both detected only partially. This splitting of trajectories happened more often at this crossing and could be the cause of the merging of two sensor images into one, therefore sometimes missing a link between two trace ID's belonging to the same person. There are more consequences to the merging of these images, which are explained in section 5.3. Another issue with the beginning and end of the trajectories is that the speed seems to significantly reduce at both ends, while there is no clear reason for this reduction. This reduction happens for all participants, but can especially be detected for trajectories with a general higher speed, such as for 159096 (yellow), 159103 (red), 159110 (grey).

The time interval of the sensor at this crossing was set to 4 FPS. This is mainly beneficial for detecting more detailed movements, yet plotting and performing calculations increases the computational time per trajectory. It can be seen from trace ID's 159110 and 159122 that both the movements of the cyclist (grey) and pedestrian (blue) can be detected with this frame rate.

Similar to the Mekelweg crossing, the x and y axes of the trajectories, do not match with the orientation of the Lorentzweg crossing provided in Figure 5.2b. Trajectories plotted in the direction of the y-axis resemble the west-east direction, while the trajectories aligning with the x-axis direction are the actual north-south direction. The axes of the plots are therefore rotated 90 degrees to the right to match the blueprint of the crossing more accurately. The positions again seem to follow the expectancy of cyclist tracks, corresponding to 159096 (yellow), 159103 (red), and 159110 (grey), following a more stable course with little directional changes, compared to pedestrians (159098, 159122).

Concluding the findings of both crossings, the trajectory data covers the movements of cyclists and pedestrians with sufficient accuracy to enable a movement analysis. The trajectories are in some cases registered late, yet for studying the movements right before the moment of crossing, the consequences for this analysis are minimised, because these crossings often take place in the more central parts of the detected area. In some cases the late registration or the splitting of a participant's path in two or more trajectories might result in a missed crossing trajectory pair of a pedestrian and cyclist. This impact is considered small, since it slightly reduces the total number of registered crossings. For this reason, no steps are taken to for example connect the trace ID's that are assumed to belong to one participant, besides the consideration that this could be a time consuming and error sensitive process.

5.1.2. Data filtering

There are a significant number of traffic participants registered at both crossings that do not encounter other modes in the area. The trajectories displayed in this case are not relevant when investigating certain interactions between cyclists and pedestrians. Therefore, if the time of a traffic participant entering and leaving the area is not within a 5 seconds range of another participant, this trajectory is removed. This 5 seconds range is considered to be a wide margin for an influencing range, yet it only serves as a filter to remove redundant trajectories for interactional study. However, the filtered data will still be used for the purposes of deriving speed as reference material for information on unrestricted speeds.

The DataFrame is throughout the subsequent process supplemented with additional variables that help to explain the interactive behaviour between cyclists and pedestrians.

5.2. Data Enrichment

This section provides the methodology for deriving the variables that are required to categorise the interactions and presents these results.

The data is first enriched with the variables speed and location, which both are variables required to determine the type of mode. Determining the origin and destination to distinguish between the different directions is then discussed in the subsequent subsection. Furthermore, the crossing angle, post-encroachment time (PET), first crossing mode and crossing point are explained in subsection 5.2.3. Hereafter, a method for predicting PET is presented.

5.2.1. Speed and Location

Apart from providing insight into the movement of pedestrians and cyclists, the speed is considered a measure to distinguish pedestrians from cyclists. Together with the locations of trajectories, these variables are used to determine the type of mode. The probability of a trajectory belonging to a cyclist or pedestrian is calculated by using the average speed. Then the location is used to confirm this type of mode.

Since this study focusses on intersections where pedestrians and cyclists need to cross each other's paths, it is considered to assign the label cyclist and pedestrian to those following their respective bike and footpaths. The reason for consulting the location with respect to the crossing layout, is because the location could confirm that a certain traffic participant is indeed a cyclist or a pedestrian if the trajectory is fully on this respective path. In addition, a traffic participant having a high probability of being a certain mode based on the combination of the location and speed, could be a strong indicator for a trajectory being assigned the label cyclist or pedestrian.

For calculating the speed of the pedestrians and cyclists at the Mekelweg crossing, the forward difference derivative is proposed to calculate the instantaneous speed at a time t with the application of the following formula (Brorson (n.d.)):

$$v_t = \frac{\sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}}{\Delta t} \tag{5.1}$$

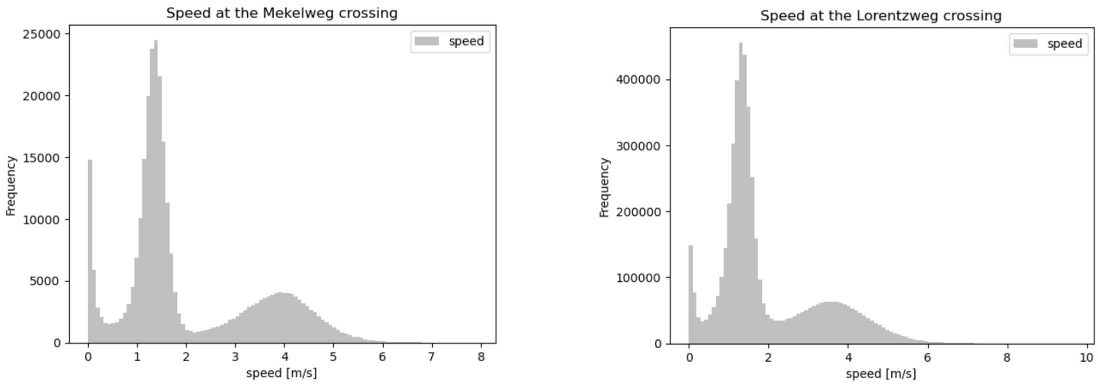
where $t+1$ denotes the next time step and Δt the difference between t and $t+1$. Especially for cyclists, the number of times a speed is calculated, is relatively low, since cyclists can pass the area within a few seconds.

To calculate the speed at the Lorentzweg crossing, the influence of measuring errors should be minimised. When performing the forward difference derivative method for this dataset, the speeds are more likely to approach unrealistic values, while the information of the trajectory can provide valuable interactions. Therefore, for the Lorentzweg crossing, a different method is chosen to determine the speed of the traffic participant: the central difference method. The formula for calculating the speed per time step in this case becomes:

$$v_t = \frac{\sqrt{(x_{t+1} - x_{t-1})^2 + (y_{t+1} - y_{t-1})^2}}{2\Delta t} \tag{5.2}$$

Where in this case the instantaneous speed is calculated by the previous and next trajectory point. The time difference will be about 0.5 seconds, which reduces the truncation error of the measuring device. Because information of the previous and next point is needed, the first and last trajectory point will not have a speed value, but since the frame rate of the cameras is 4 FPS instead of 1, there are more trajectory points to provide a general description of the speed and its development at this crossing.

Two speed histograms are created for the Mekelweg crossing and Lorentzweg crossing in Figure 5.3. Figure 5.3a shows the distribution of all instantaneous speeds registered at the Mekelweg crossing. Figure 5.3b shows this same distribution for the Lorentzweg crossing. The values of speeds found in the data align with previous studies on cyclist and pedestrian speeds by Chandra and Bharti (2013), Daamen and Hoogendoorn (2007), Yan et al. (2020) and Eriksson et al. (2019).



(a) Mekelweg crossing

(b) Lorentzweg crossing

Figure 5.3: Speed histograms of both the Mekelweg (a) and Lorentzweg (b) crossing.

The occurrences of speed show three clear peaks from left to right for people standing still, walking pedestrians and cycling cyclists. The left peak arises because people that have no movement through the environment, are spending a longer time in the area. This longer time spent here, makes for a bigger overall contribution of very low registered speeds, resulting in the peak at a speed of almost 0 m/s. The pedestrian speed data peaks at a lower speed value than the cyclist speed data, because of the common speed difference, confirmed in the studies of Chandra and Bharti (2013) and Daamen and Hoogendoorn (2007). The histograms support that based on only speed data, a first distinction of modes can be determined.

For the Mekelweg crossing, two distributions are derived which are visualised in Figure 5.4. The speed data of pedestrians and cyclists is assumed to follow normal distributions, which is also confirmed in the study of Chandra and Bharti (2013) and Nateghinia et al. (2024). The speeds of the traffic participants that stand still are excluded by setting the minimum speed at 0.77 m/s, based on the lower bound slow walking speed of Murtagh et al. (2021). After visual inspection of different minimum speeds, this value seems to match both normal distributions; see Appendix B. By applying a Gaussian Mixture Model, the speed data can be approximated by two curves, which sum is the combined distribution that follows the speed data.

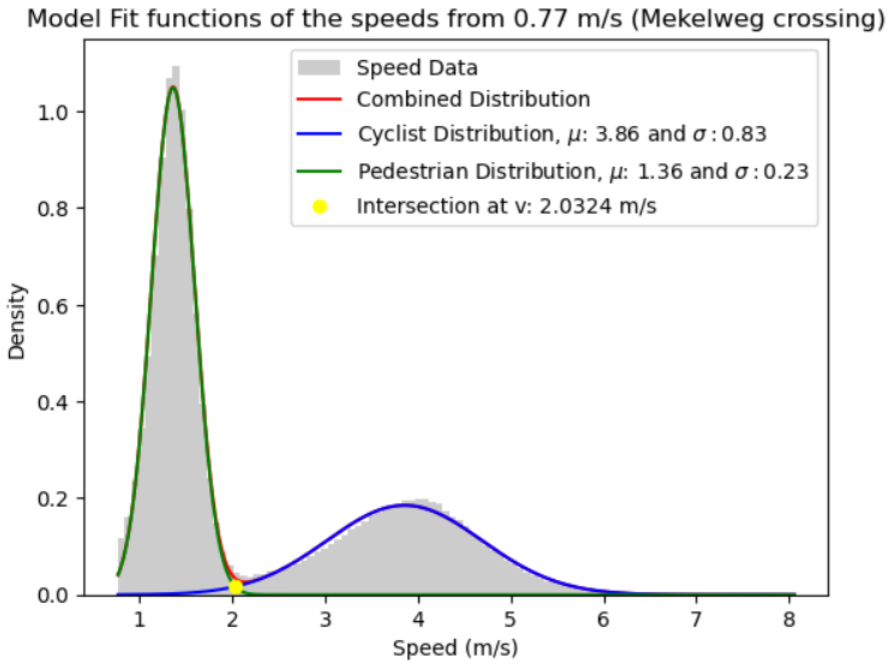


Figure 5.4: Relative frequency of the number of times a certain speed is registered at the Mekelweg crossing.

For the distinction between pedestrians and cyclists, the intersection point right of the pedestrian speed curve and left of the cyclist curve is calculated and the speed value attached to this intersection point is the threshold value. This threshold value theoretically means that from this speed, more cyclists reach this speed than pedestrians. For both crossings, the probability of the trajectory belonging to a cyclist is calculated by dividing the height of the cyclist distribution (blue) at the average speed of that cyclist divided by the total height of both distributions (red). In the same way, the probability of a trajectory belonging to a pedestrian (green) can be calculated.

The threshold for the speed at the Mekelweg crossing is calculated to be 2.03 m/s. To meet the requirement of a higher probability of being a cyclist, the participant should have an average speed of at least this value, or a particular part of the cyclist's track should have a speed greater than 2.5 m/s, because from this value, the chances of the ID belonging to a pedestrian are negligibly small (about 5 times the standard deviation from the mean).

The two curves for the Lorentzweg crossing are different from the Mekelweg crossing with the separation being less distinct. Figure 5.5 shows that the dip around 2 m/s is less deep, suggesting that more cyclists have lower speeds at this crossing. This could be explained by the different directions that cyclists have to go to, potentially having to decrease their speed to make a comfortable turn.

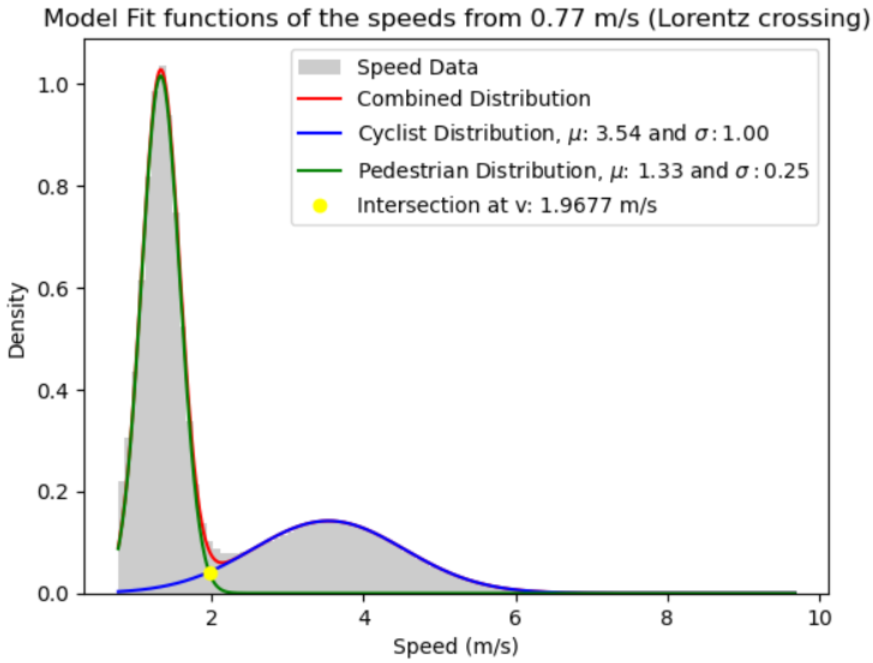


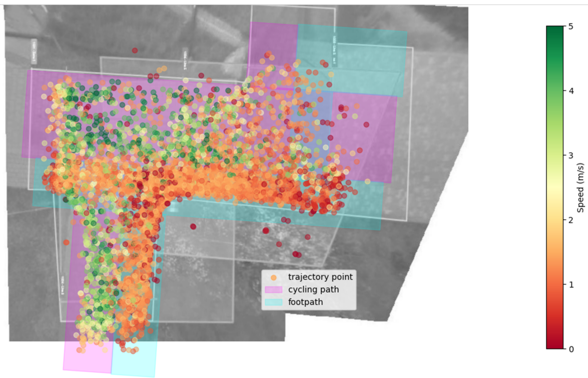
Figure 5.5: Relative frequency of the number of times a certain speed is registered at the Lorentzweg crossing.

As a consequence, the two speed curves intersect at a lower speed value. The minimum average speed threshold, which gives an equal probability of being a cyclist or pedestrian, is thus slightly lower at 1.97 m/s. For each trajectory, the probability of the average speed belonging to that of a cyclist or a pedestrian is calculated. This probability, together with the location are the variables that distinguish cyclists from pedestrians.

As mentioned in the Introduction (1.3), no distinction has been made in the different types of cyclist or pedestrian, for example pedelecs or runners. Since speed is the main indicator available that could contribute to this distinction, this cannot be assessed with sufficient certainty, because the speed distributions show large overlaps Schleinitz et al. (2017) and Smyth (2018). The number of appearances that these particular traffic participants have are furthermore very low.

Determining the locations with respect to the crossing layout will be done by extracting a random sample size of the data that indicates the density and speed of trajectory points of the monitored area. Based on the densities, the edges of the area, together with the locations of cycling and footpath, can be determined by scaling the crossing layout to a visually fitting size. The edges between the cycling and footpath are then furthermore determined through the colour distinction of the speed of the cyclist compared to the speed of the pedestrian. The locations of the foot and bike path are determined for the Mekelweg and Lorentzweg crossing, as shown in Figure 5.6a.

(a) Mekelweg crossing



(b) Lorentzweg crossing

Figure 5.6: Blueprints of the two crossing where the location of the bike and footpath are visually determined by a speed scatter plot of a sample of the data.

The trajectory dots indicate the edges of the area, as well as the edges of cycling to footpath by the separation of cyclists (green) and pedestrians (orange). Only when a trajectory is fully inside the area of the bike path (magenta) or footpath (cyan), the trace ID is assigned a label 'cyclist' or 'pedestrian' and the probability that this is the correct mode, initially based on the average speed, becomes 1. In all other scenarios, no definitive type of mode is assigned to the trace ID and instead the probability of speed remains the leading factor to determine the mode.

The result of this distinction in types of active modes is that in total 26090 trajectories are registered of pedestrians and 47910 of cyclists with full certainty at the Mekelweg crossing. The trajectories with less certainty of the modes are in total 5094 pedestrians and 2530 cyclists at the Mekelweg crossing. At the Lorentzweg crossing, the count of trajectories with full certainty of the mode are 42399 pedestrians and 94070 cyclists and with less certainty are 43311 pedestrians and 27812 cyclists.

These numbers suggest that the method of the location based modal split does not work as properly for the Lorentzweg crossing, since more than half of the pedestrians and about 20 % of cyclists are labelled without full certainty. From a field observation on site it was detected that the cyclists and pedestrians did not adhere to their respective paths at the Lorentzweg crossing, mainly due to shortcuts provided by the pavement of the other mode, see the samples of some trajectories in Figure 5.7.

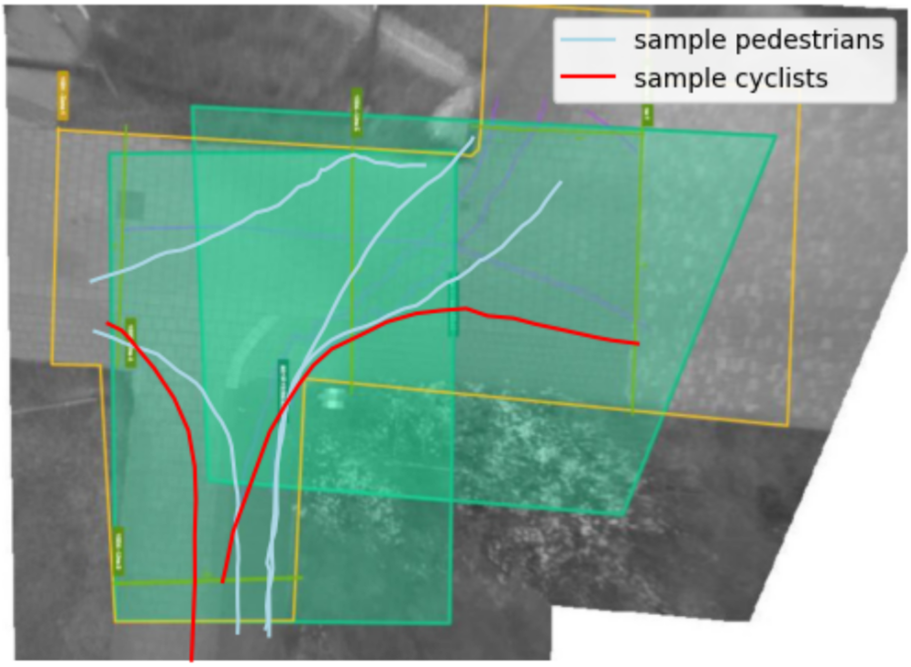


Figure 5.7: Samples at the Lorentzweg crossing with trajectories of cyclists and pedestrians making use of different pavement sections.

The Lorentzweg has this high uncertainty, because the trajectories of many pedestrians and cyclists do not match with the locations of the foot and cycling paths. It is therefore decided to lower the probability requirement for trajectories belonging to cyclists or pedestrians. This probability has been based on the speed distribution. Therefore, when lowering the probability threshold, the speed that is corresponding to this probability is changing as well. For example, a 98% probability considers a maximum speed of 1.48 m/s for pedestrians and a 2.25 m/s minimum speed for cyclists. When filtering out pedestrians up to an average speed of 1.48 m/s, a significant amount of trajectories is being removed. This is confirmed in the data and other studies (Chandra and Bharti, 2013 and Murtagh et al., 2021). Mainly fast walking pedestrians will be removed from the data, creating a possible bias in the behavioural patterns. The 95% probability corresponds with a maximum average speed of 1.65 m/s for pedestrians and 2.19 m/s minimum average speed for cyclists, aligning with the findings of Murtagh et al. (2021) on maximum pedestrian paces that are around 1.62 m/s. Therefore, this probability will be chosen as a threshold for distinguishing cyclists from pedestrians.

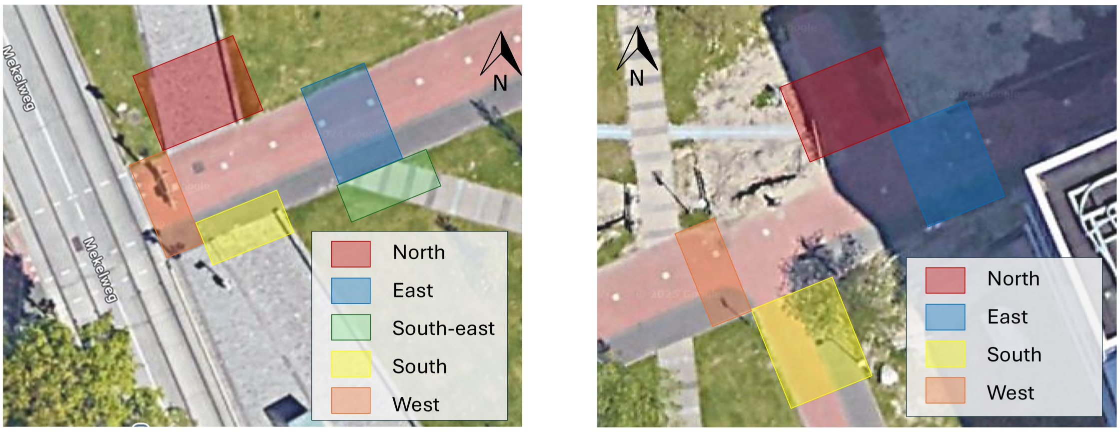
The count of trajectories within this probability margin then becomes 83517 pedestrians and 119160 cyclists and with less certainty, there are 2193 pedestrians and 2722 cyclists remaining.

5.2.2. Origins and Destinations

To distinguish the trajectory data between the four crossing scenarios of pedestrians as provided in section 4.3, the directions of the trajectories are determined.

The crossing at the Lorentzweg in total has 4 main directions that cyclists and pedestrians can come from and go to, referred to as the origin and destination respectively. The situation for each pair of origin and destination could have different behaviour as a consequence. For example, pedestrians coming from either the north, east or south and going to the west might have to cross the cycling path at the same location, but the approach case with respect to the cyclist and the bike path makes for a different situation, which are also highlighted in section 4. Therefore, further categorisation is applied where the origin and destination for each trajectory is determined. For each of the trajectories, the first and last points are assumed to be the origin and destination, respectively. For simplicity, the directions are referred to as the four cardinal directions, even though both crossings are slightly rotated and do not align perfectly with these directions.

The Mekelweg crossing is oriented in a similar rotation as the Lorentzweg crossing, but it has the addition of a south-east direction referring to the foot path providing a short cut for pedestrians (and unintentionally cyclists) coming from the west and going to the south of the TU Delft campus. All directions are visualised in Figure 5.8 with the trajectories origin and destination corresponding to the area that the first and last trajectory point is monitored respectively.



(a) Mekelweg crossing

(b) Lorentzweg crossing

Figure 5.8: Blueprints of the two crossing with an approximation of the area for determining the origin and destination.

The application of this method however, might miss certain origins or destinations due to the late registration that was mentioned in section 5.1. Therefore interactions for the directions of interest might not be considered, while they could prove to be valuable for this research, due to the increase in statistical significance. For this reason, an additional method to address these cases is introduced that takes into account the beginning and end direction of the trajectories that start or end in the centre of the detection area.

The first and last points of a trajectory are likely to point in the direction of the origin and destination. The exact directions can be calculated by applying the following formula:

$$\varphi = \arctan\left(\frac{\Delta y}{\Delta x}\right) \tag{5.3}$$

Where Δy and Δx are determined by: subtracting the x and y values of the second and second to last trajectory point from the first and last trajectory points respectively for the origin and destination. Because the Lorentzweg crossing has a higher frame rate, these first and last two trajectory points can deviate from the actual direction, most likely for pedestrians. For this reason, a larger time interval is used for calculating the direction, which means that not the second and second to last trajectory points are used, but the fourth and fourth to last for pedestrians, thereby requiring the trajectory to consist of at least 4 points.

The directions will then be compared to the four cardinal directions, with the same slight rotation. A direction is assigned to the trajectories that point to the cardinal direction with the slight rotation $\pm$22.5 degrees, dividing the 360 degrees into 8 segments of 45 degrees.

The result of all the origins and destinations is given in tables in Appendix C, where the cardinal directions and the directions that are left inconclusive are given for both pedestrians and cyclists for the Mekelweg and Lorentzweg crossing. The numbers of pedestrians and cyclists are slightly less than the numbers given in subsection 5.2.1, because participants starting or ending far outside the paved area were not considered.

5.2.3. Interacting Cyclists and Pedestrians

As referred to in the Introduction 1, the specific interactions that are analysed, are the sideways interactions of $\pm$ 60 degrees between individual cyclists and pedestrians. The paths of the pedestrian and cyclist need to cross within a reasonable amount of time for them to perform behaviour that aligns with an interaction. The crossings that occur within 5 seconds from each other are analysed, which is the interactive margin that is used in the studies of Beitel et al. (2018) and Zangenehpour et al. (2016) and a gap of 5 seconds or more at a road crossing of about 5 meters wide is very unlikely to be rejected by pedestrians according to Zhao et al. (2019), although this study focusses on pedestrian-vehicle interaction.

To determine whether a cyclist and pedestrian cross each others paths, it must be verified whether one of the segments of a pedestrian trajectory intersects with one of the segments of a cyclist trajectory within this 5 second range. To check whether the segments of two trajectories cross, the cross product method is applied (sqlpey, 2024), that uses the property of two points being on either side of a line segment in order for these to cross. Whenever two segments cross, the angle between the two segments needs to be at least 30 degrees in order for the interaction to be classified as a sideways interaction. The angle is calculated by applying the following formula (GeeksforGeeks (2024b)):

$$\varphi_t = \cos^{-1}\left(\frac{\vec{d_i} \cdot \vec{d_j}}{|\vec{d_i}| * |\vec{d_j}|}\right) \tag{5.4}$$

Where $\vec{d_i}$ and $\vec{d_j}$ are the directions of traffic participant i and j based on the position $\vec{p_i}$ and $\vec{p_j}$ right before and after crossing, referred to as time step $t - \frac{1}{2}$, and $t + \frac{1}{2}$. The directions therefore become:

$$\vec{d_i} = \vec{p_{i,t+\frac{1}{2}}} - \vec{p_{i,t-\frac{1}{2}}} \qquad \& \qquad \vec{d_j} = \vec{p_{j,t+\frac{1}{2}}} - \vec{p_{j,t-\frac{1}{2}}}$$

For each of the interactions, the post-encroachment time (PET) is calculated as the difference in time between the first exterior of the participant to leave the conflict area, and the second exterior of the participant to reach the area. However, the exact measures of the area are not fixed and unknown, since the person has no registered geometry. Therefore, the area of conflict is in this case simplified to the intersection point of the two trajectories, which gives an approximation of the centre of the conflict area. Based on this intersection point, the PET can be estimated by applying linear interpolation towards the intersection point for both trajectories. From this interpolation, a distinction is then made in the conflicts where the pedestrian crosses before and after the cyclist, because different behaviours are expected in both scenarios as explained in the Methodology.

Only the interactions that are within a range of 0 to 5 seconds, are studied further to detect interactive behaviour. This range is acquired from multiple studies (Beitel et al., 2018, Zangenehpour et al., 2016, and Tageldin and Sayed, 2016) that consider time ranges of 0 to 5 seconds as the interaction threshold for a cyclist and pedestrian. Because the studies do not agree on potential intermediate threshold values for PET on the severity of the conflict, this study will base these values on the outcome of stopping trajectories for PET's between 0 and 5 seconds, which is explained in section 6.1.1.

5.2.4. Predicted PET

As mentioned in Chapter 2, the post-encroachment time by definition does not require calculations for the predictions of movement. However, when studying movement changes of pedestrians, a measure of prediction could prove to be of value for the actual outcome of an interaction. A certain PET value cannot for example explain whether the pedestrian has stopped along it's path to provide the cyclist right of way or that it has continued on it's path with the same speed and direction. With the addition of a predicted PET, the development of this movement can be studied and patterns could be recognised that explain the outcomes of eventual conflicts. For this reason, this study will apply a method to predict the PET for every time step.

There is a risk with predicting the PET that the development of this value progresses unstable, because the predicted path can deviate from the actual path. Therefore, a few assumptions have to be made on what movements can be expected in next time steps:

1. The eventual predicted path of the pedestrian is directed towards its destination, which in this case is estimated to be the last trajectory point.

2. The speed of the pedestrian at the current time step is assumed to remain the same for the time predictions of the remaining straight direction towards the end goal.
3. In contrast to the predicted pedestrians path, the cyclist path is considered fixed in space.
4. Similar to the pedestrian, the current speed of the cyclist is used as a measure to predict the timing of the eventual path by maintaining the same speed across the whole trajectory.
5. If the cyclist is not (yet) present in the area when calculating the predicted PET for the pedestrian, the cyclist is assumed to follow the eventual trajectories speed and direction.

By excluding a prediction of the cyclists path, the prediction of PET therefore becomes less of an actual prediction, but the advantage of this method is that the course of the predicted PET is more stable compared to a calculation where the full movement is predicted. Besides, a predicted path in the way that it is predicted for a pedestrian is less likely to be the path that is predicted for the cyclist. In practice, pedestrians tend to make intuitive assumptions about the cyclist's path, and these assumptions are likely closer to the actual trajectory than a simplistic straight-line model.

The prediction of a trajectory can occur in many different ways. For example, a pedestrian approaching the crossing with a certain speed, could be predicted as the first mode to cross based on the calculations of its current speed and eventual goal, yet whenever the pedestrian decides to wait for the cyclist, the outcome of the prediction inverts and the pedestrian eventually crosses after the cyclist. A plot is made in Figure 5.9 that shows this development of predicted PET for three sample trajectories. On the y-axis, the predicted PET is shown with the positive predicted PET values referring to a pedestrian crossing before a cyclist and a negative predicted PET to that of a pedestrian crossing after the cyclist. This development is only calculated for the central area before the crossing point, excluding possible speed deviations at the beginning and end of the trajectory, which is why the time does not start at zero.

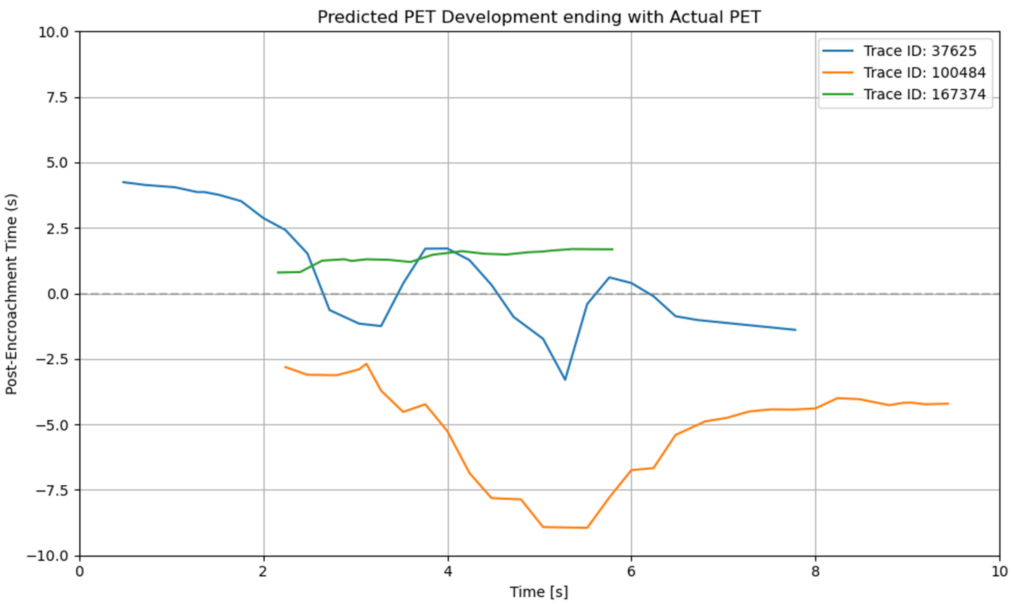


Figure 5.9: Samples at the Lorentzweg crossing of predicted PET development over time, ending with the actual PET.

A similar development profile is created for the speed in these parts of the trajectories for these three samples in Figure 5.10. This speed plot contributes to understanding the specific movements that the pedestrian makes when encountered with a cyclist.

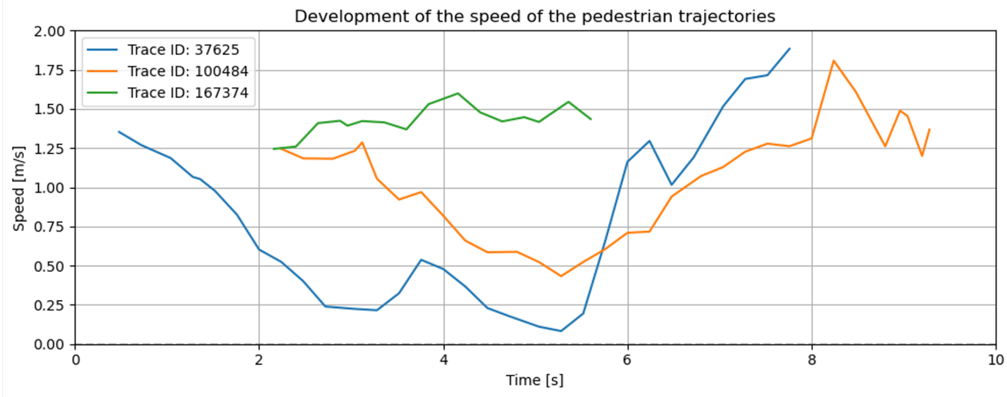


Figure 5.10: Samples at the Lorentzweg crossing of speed development over time.

A short interpretation of the predicted PET combined with the development of the speed of the three trajectories is listed below:

ID 167374 The pedestrian is predicted to cross before the cyclist and does eventually cross before it. A slight increase at the beginning in predicted PET can be detected, which in this case is due to a slight increase in speed, yet this could have also been caused by the cyclist slowing down.

ID 100484 The pedestrian is predicted to cross behind the cyclist and crosses behind the cyclist as well. It does so with quite a steep drop in the middle of the predicted PET indicated by the pedestrian slowing down significantly here, resulting in the temporary predicted PET to be much more negative than the eventual PET. The drop in speed has also likely caused the eventual PET to be lower compared to the predicted PET at the start.

ID 37625 The pedestrian was predicted to be first given the predicted PET at the beginning, yet eventually the pedestrian crosses after the cyclist. This is due to the pedestrian slowing down before the crossing, not wanting to take the risk of going before the cyclist.

The fluctuation of the predicted PET of the pedestrian with ID 37625 can be further clarified when studying the actual trajectory. Figure 5.11 shows the course of the trajectory of the pedestrian from east to west and the crossing cyclist going from south to north with the speed indicated by colour. The trajectory of the pedestrian consists of different phases, starting with the pedestrian approaching the crossing with a fairly regular walking speed, thereby predicting the pedestrian to cross before the cyclist. When the pedestrian comes near the edge of the footpath, it decides to come to a stop, decreasing its speed significantly, thereby predicting the PET to be in favour of the cyclist instead of the pedestrian. After this short stop, the pedestrian decides to increase it's speed again to cross behind the cyclist, yet in the way the PET is predicted, this increase in speed, together with the direction towards the eventual goal results again in a slight peak of the predicted PET, favouring the pedestrian for a short period of time. The pedestrian however aims to go around the cyclist from behind, eventually resulting in the PET value to be negative.

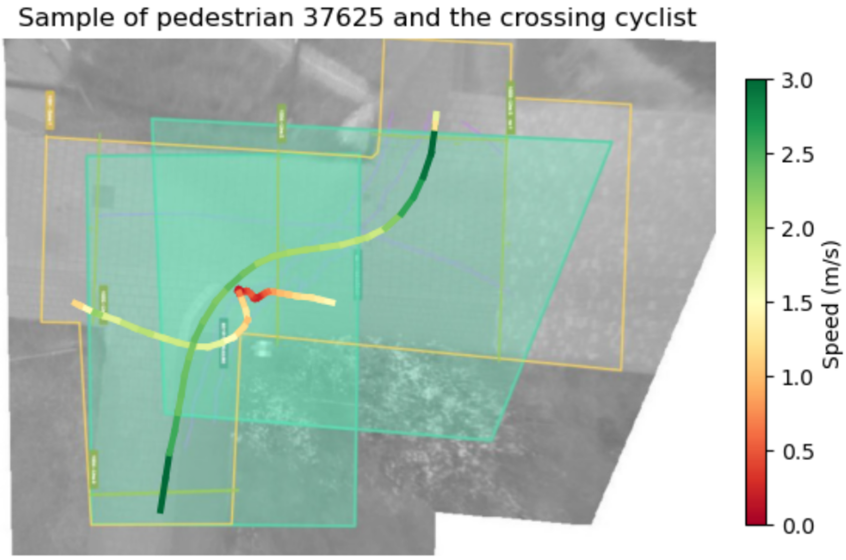


Figure 5.11: Pedestrian 37625 crossing with an approaching cyclist

5.3. Identification of Crossing Trajectories

All crossing trajectories that were registered within 5 seconds at the Mekelweg crossing between cyclists and pedestrians sum to a total of 5942 from the 81624 captured trajectories of cyclists and pedestrians going in all directions. For the Lorentzweg crossing there are 28085 crossing trajectories out of 207592 total trajectories. Tables 5.1 and 5.2 below show the number of total pedestrians and the number of pedestrians crossing with cyclists only for the directions of interest.

Table 5.1: Numbers of pedestrians crossing with a cyclist at the Mekelweg crossing.

Mekelweg crossing	South - North	North - South
Total trajectories	3200	3468
Crossing trajectories	1183	625
Percentage	37%	18%

Table 5.2: Numbers of pedestrians crossing with a cyclist at the Mekelweg crossing.

Lorentzweg crossing	South - West	East - West
Total trajectories	1688	17984
Crossing trajectories	448	5054
Percentage	27%	28%

The east-west direction of the Lorentzweg crossing show a significantly higher number of pedestrians (17984) and a higher number of pedestrians crossing with a cyclist (5054) compared to the other origin-destination scenarios. Thus, the interactions captured here provide a richer basis for identifying behavioural patterns. For this reason, the analysis that will be mostly shown is from this group, while the figures of the other crossing scenarios are included in the appendix to maintain clarity within this and the subsequent chapters. Comparisons between the different directions by means of descriptive statistics are provided in subsequent chapters.

Trajectory plots of pedestrians walking from east to west at the Lorentzweg which cross with a cyclist are shown in Figure 5.12. The colour indicates the speed of the pedestrians. A figure with higher opacity is shown in Appendix G to visualise the density of the trajectories. The head and tail of the trajectories indicate the sudden speed change, which is likely due to a measuring error as mentioned in section

5.1. Additionally, even though pedestrians may slow down or deviate in this area, it is improbable that a cyclist is the cause, given the distance to the eventual conflicting point exceeds 6 meters. For this reason, a frame is introduced indicated by the vertical blue lines in between which the stopping and deviating behaviours are considered. At the end of the trajectories, the left side of the figure, an evenly distributed strip of increased speed can be detected in the darker green area. This strip aligns with the border of the monitored area covered in the north-east of this crossing, therefore it is likely that the merging of the two images of trajectories produces a consistent measuring error for people going from east to west. However, when studying the stopping and deviating behaviour, that occurs mostly after crossing, so this does not have a big influence on the outcomes.

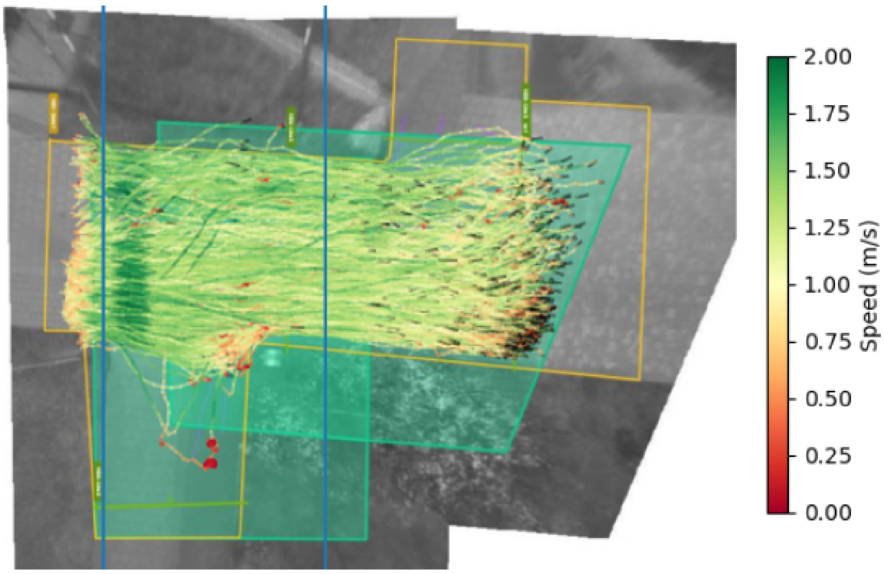


Figure 5.12: Trajectory plots of all pedestrians walking from east to west at the Lorentzweg crossing with the frame of behavioural consideration in between the blue lines

5.4. Conclusion

The pedestrian and cyclist data is prepared for a behavioural analysis by performing a quality assessment, filtering step, and by enriching the data. Thereby, the following sub-research question is answered: Which data processing and enrichment methods are required to prepare pedestrian and cyclist data for a behavioural analysis?

The data quality was assessed using video footage and trajectory plots at the Mekelweg and Lorentzweg crossings. The trajectories appear accurate: cyclist paths are smooth, while pedestrians show more lateral swaying, as expected. Moreover, the values of speeds align with existing literature. Some trajectories are registered late, possibly missing crossings. The Lorentzweg benefits from a higher frame rate, capturing more detail, but occasionally splits a single participant into multiple trace ID's, reducing interaction data.

In general, the trajectory data from the smart sensor is suitable to detect movement changes for pedestrians in their approach to cyclists on bike paths. The stopping and deviating movements that are considered can be captured with good accuracy, though for the Lorentzweg crossing the influence of the inaccurate tips of the trajectories should be minimised. This is done by narrowing the area to detect movement changes. In a later stage, crossing trajectories are compared with non-crossing ones to provide a reference point for interpreting movement patterns and to account for measurement inconsistencies.

Speed and location variables are added to the data to distinguish pedestrians from cyclists. Besides, speed is vital for understanding general movements. Origins and destinations stem from the location by considering the begin and end of each trajectory. These origins and destinations are important for selecting the intended interaction scenarios.

Pedestrian-cyclist crossings of at least a 30 degree angle are calculated and the crossing point and timing between both modes (PET) are derived from the crossing segments. PET provides the value to determine the proximity of the conflict and with a few assumptions a prediction of PET is possible. The predicted PET contributes to a more common understanding of the development of an interaction.

6 Results

This chapter presents the results of the data analysis to address the fourth sub-research question: What movement changes does a pedestrian perform when approaching a cyclist in a sideways conflict? Of the 7310 crossing pedestrians in four different crossing scenarios, 4780 crossed with a single cyclist.

First, the stopping behaviour of pedestrians is analysed by introducing a PET threshold value, the quantities of stopping and the stopping distance. Then, the gradual development of the pedestrian movements are analysed with respect to the cyclists by visualising the predicted PET with a numerical analysis. Lastly, the general deviating behaviour that could be detected is provided by again introducing threshold values and qualitatively assessing the directions where deviation occurs most.

6.1. Stopping Behaviour

In this section, the stopping behaviour of pedestrians is analysed. This analysis will be done based on the enriched data that is created as described in Chapter 5. The stopping analysis starts with a further classification of the trajectory data. All of the crossing trajectories are classified based on the mode that crosses first from the perspective of the pedestrian as described in section 3.4 of the Methodology. Following this first distinction, this section presents the method for deriving a PET threshold by studying the relation between the PET and the stopping behaviour. After this, the general stopping behaviour in terms of quantities and stopping distances is presented.

6.1.1. Threshold definition

This subsection compares the number of stopping pedestrians relative to the total number of crossing pedestrians to the values of PET. This analysis is solely done for the east west crossing pedestrians at the Lorentzweg crossing. The threshold used for a stopping trajectory is 0.77 m/s, which means that a trajectory reaching a speed below this value, is registered as a stopping trajectory. This threshold is based on the lower bound 95% interval of slow walking pedestrians by Murtagh et al. (2021). If a participant reaches a speed that is below this threshold value, it is labelled as a stopping trajectory.

The interaction cases are classified based on the PET values for pedestrians crossing first and second. Figure 6.1 shows the percentage of stopping trajectories of pedestrians when they cross before the cyclist (red) and after the cyclist (blue). The bars with less than 10 trajectories for comparison are removed due to a risk of insignificant values.

In the scenario where the pedestrian crosses before the cyclist, no clear trend in stopping trajectories can be detected, while the percentage of stopping trajectories is also significantly lower. These pedestrians most likely do not have a reason to stop, because they intend to be crossing before the cyclist. Thereby, a relation between PET and these rarely occurring stopping movements is very unlikely to arise. The remaining stopping behaviour is most likely the result of a few measuring errors inside the considered area or some regular stopping unrelated to the approaching cyclist. In the scenario where the pedestrian crosses after the cyclist, an overall decreasing trend can be detected.

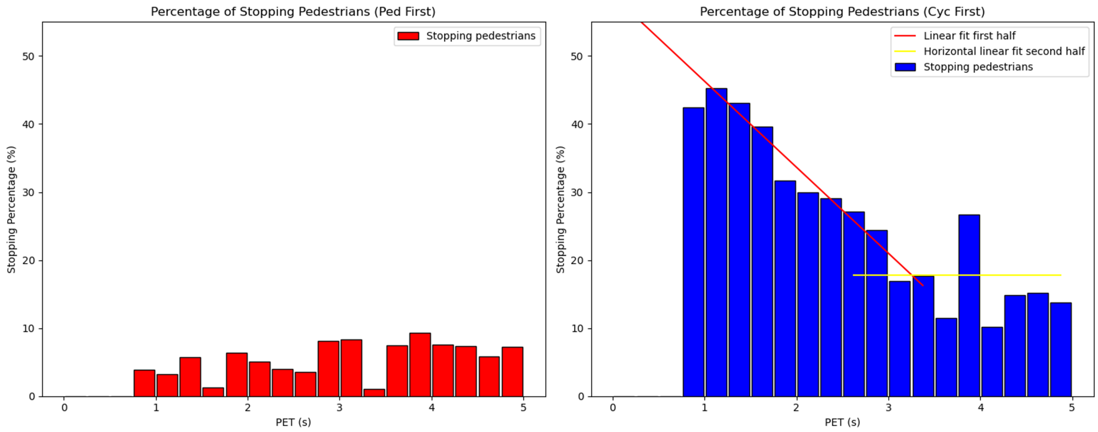


Figure 6.1: Percentages of stopping trajectories for different PET values when the pedestrian crosses first (red) and when the cyclist crosses first (blue).

An approximation of the course of the stopping percentages is used to create a threshold value for the PET. In the scenario where the pedestrian crosses after the cyclist, this pedestrian has likely done this by adapting its movement through slowing down. The closer the conflict is in terms of timing, the more likely it is that this pedestrian experiences the urge to come to a stop to cross behind a cyclist, because otherwise a collision might have happened. For this reason, there is likely a decreasing trend in the lower PET regions. According to Figure 6.1 halfway the PET values of 0 to 5 seconds, the trend seems to stabilise. The value for the average number of stopping trajectories is in this particular scenario 17.8%. The course of the trend after 5 seconds is unknown, while previous research tends to agree upon the 0 to 5 second range to be the range between which pedestrians and cyclists perform interaction, as presented in section 2.1.3, yet when observing the dataset of pedestrians that do not cross within 5 seconds with a cyclist for the east west crossing pedestrians, the average stopping percentage is 8.4% which is slightly lower than the 17.8%.

The trend is assumed to be linearly decreasing from the lowest PET values that were detected, because the higher a PET value becomes, the less likely it is that the pedestrian needs to stop. The stabilisation in the right part of the histogram can be approximated by a separate linear trend line. By approximating both of these sections with a trend line, the PET value separating these both sections can be determined at the intersection point of both lines. This intersection theoretically means that from that corresponding PET value, the amount of stopping trajectories, and thus stopping pedestrians, will change at a different rate once the PET increases. In Appendix D, the result of several fitting approaches are shown with the bin width, the PET end point of the first fitting line and the PET starting point of the second fitting line as variable inputs. Additionally, because it is unclear whether the amount of stopping stabilises halfway or whether it is decreasing, both trend types are considered as well. The results indicate that many of the intersection points show a changing trend for a PET value of around 3 seconds, which aligns with the findings of Tageldin and Sayed (2016) and Zangenehpour et al. (2016). Therefore, this PET value is considered a threshold value for the classification of different interactions.

6.1.2. Stopping results

The threshold value of PET is used to indicate the difference in approach cases of a conflicting situation (PET < 3.0 s) and an interactive situation (3.0 s < PET < 5.0 s). In this analysis, the consequence of the general categorisation is presented in visualisations. Then, the quantities of stopping are provided for all crossing scenarios. Hereafter, the stopping distances are presented.

Figure 6.2 shows the trajectories of the pedestrians going from east to west at the Lorentzweg crossing interacting with a cyclist within a PET of 0 - 3 seconds in the first column and a PET of 3 - 5 seconds in the second column. The first row indicates the approaching situations where the pedestrian leaves the conflict area first and is therefore crossing before the cyclist, while the second row shows the pedestrians that cross behind the cyclist. Similar plots of these trajectories are created for the other three crossing scenarios, which can be found in Appendix G.

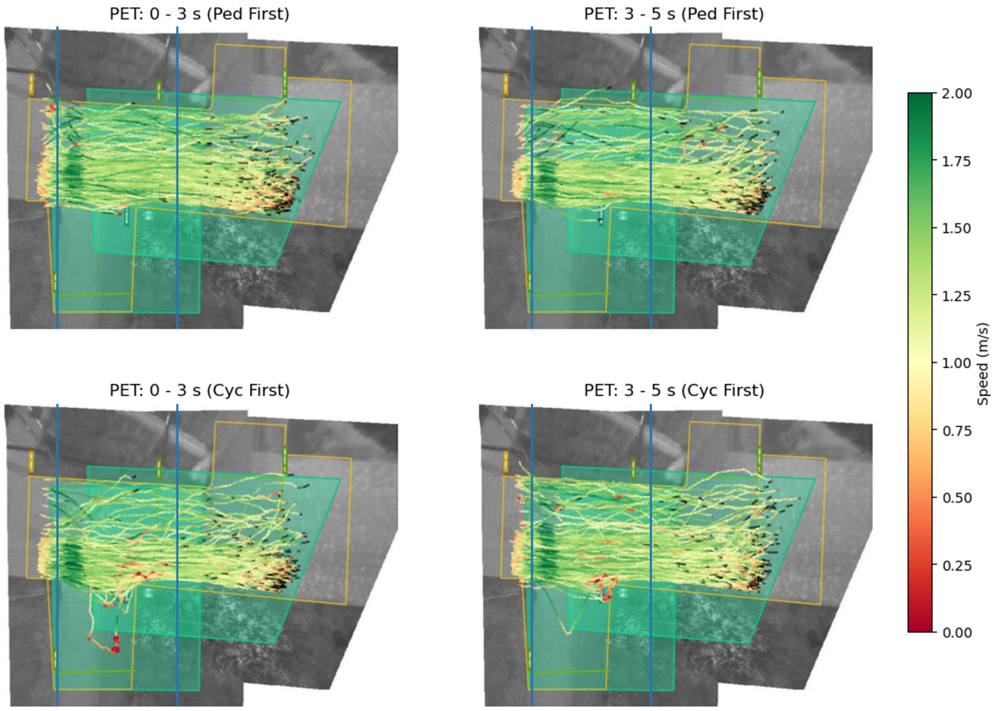


Figure 6.2: Pedestrians walking from east to west at the Lorentzweg crossing with a distinction based on PET value and whether the pedestrian crosses before or after the cyclist.

In many of the cases, the pedestrian seems to go straight towards its goal, yet in some cases the pedestrian either stops or sometimes seems to deviate from a straight trajectory. This deviation is addressed in section 6.3. Stopping is mostly observed right before the pedestrian enters the bike path, shown in the red parts of the trajectories. This significant stopping behaviour can only be observed in the cases where the pedestrian crosses after the cyclists, which is confirmed by the stopping percentages in Table 6.1 below.

Table 6.1: percentage of crossing trajectories that stop within each category of PET and first crossing mode

East - west (stopping trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	4.4%	6.5%
Cyclist crosses first	32.8%	15.6%

The category of pedestrians crossing after the cyclist in a conflicting situation tend to stop on average 32.8% of the time. This indicates that a significant number of pedestrians also do not stop, so they either slow down or even maintain their speed. If the pedestrian crosses before the cyclist, a slightly lower percentage of stopping is observed compared to non-crossing pedestrians, which was 8.4%. This possibly implies that some pedestrians confronted with an approaching cyclist feel slightly more urge to walk on than if this confrontation does not happen.

For the situations of pedestrians crossing after the cyclists, the relative number of stopping pedestrians is the highest. However, compared to the other crossing scenarios, this number is relatively low. Table 6.2 shows the stopping percentages for all different crossing scenarios. The percentages are based on the stopping and total trajectories given in Appendix G. Besides the stopping quantities of the crossing pedestrians, the quantity of stopping of the non-crossing pedestrians is provided as a reference.

Table 6.2: percentage of crossing trajectories that stop for all other crossings

Crossing	Origin - Destination	Stopping percentages		
		PET 0 - 3 s	PET 3 - 5 s	Reference non-crossing
Lorentzweg	East - west	32.8%	15.6%	8.4%
	South - west	51.9%	45.3%	34.4%
Mekelweg	South - north	32.2%	16.8%	9.4%
	North - south	47.6%	23.8%	13.2%

The percentage of the south-west crossing pedestrians at the Lorentzweg crossing has in this case the highest value for both the crossing and non-crossing pedestrians. The crossing pedestrians are however stopping relatively seldom because of an approaching cyclist, when considering the reference group that also frequently stops.

For the behaviour that pedestrians tend to perform, it is key to understand the stopping distances that pedestrians tend to keep. The distance that the pedestrian has towards the cyclist can differ significantly, depending on the approaching side of the cyclist. This is why it is proposed to measure the stopping distance of the pedestrian towards the crossing point. Besides, the stopping distance with respect to the cyclist can in some cases not be determined if the cyclist is not present in the area or not yet detected by the sensors. The distance from the crossing point to the pedestrian can be calculated with the following formula:

$$d_t = \sqrt{(x_{crosspoint} - x_t)^2 + (y_{crosspoint} - y_t)^2} \tag{6.1}$$

Where $x_{crosspoint}$ and $y_{crosspoint}$ denote the x and y value in the grid of the intersection where the pedestrian and cyclist in the future will cross each others paths. x_t and y_t denote the x and y position of the pedestrian in this same grid at the time instant t that the pedestrian slows down to a speed below 0.77 m/s.

For all stopping pedestrians in their respective category, histogram plots are created that visualise the distance pedestrians tend to keep with respect to the crossing point. Figure 6.3 shows these histograms for the east west direction. The other histograms are shown in Appendix H.

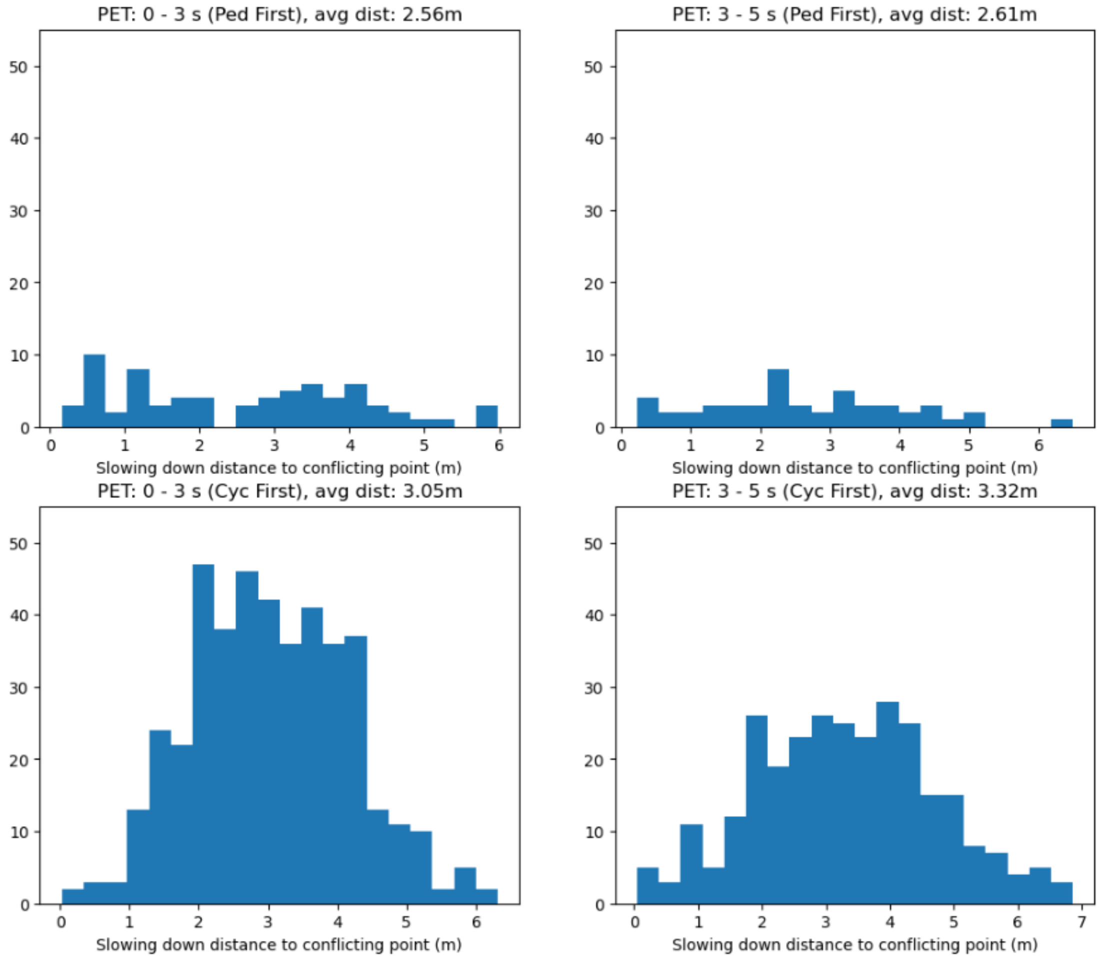


Figure 6.3: Frequency of stopping distances for pedestrians walking from east to west crossing with a cyclist at the Lorentzweg crossing.

From the histograms where the cyclists cross first, a distribution that resembles a normal distribution can be detected. Normality is assumed based on the idea that each pedestrian maintains a slightly different preferred distance from the crossing point, driven by personal comfort and perceived safety. These individual differences are expected to cluster symmetrically around a central tendency leading to a distribution that resembles a normal distribution. As mentioned in the Methodology, the statistical significance of the hypothesis is tested that the data potentially follows a normal distribution. A combination of D'Agostino and Pearson's test is performed that tests whether the distance with respect to the crossing point that pedestrians tend to slow down from is normally distributed (H_0). The alternative hypothesis (H_1) is that the distance does not follow a normal distribution. If the resulting p-value of the test is more than 0.05, the null hypothesis cannot be rejected, and there's no significant evidence that the data is not normally distributed. The results of all categories is given in Table 6.3.

Table 6.3: Statistical tests performed on normality of the stopping distance for pedestrians walking from east to west at the Lorentzweg crossing.

First crossing mode	PET category	Test statistic	P-value	p >0.05	Statistical significance
Ped first	0 - 3 s	8.135	0.017	No	Not normal
	3 - 5 s	3.986	0.136	Yes	Normal
Cyc first	0 - 3 s	0.785	0.676	Yes	Normal
	3 - 5 s	6.729	0.035	No	Not normal

The normality test for the pedestrians crossing behind the cyclist for a PET between 0 to 3 seconds results in a p-value above 0.05, therefore there is no significant reason to reject the null hypothesis, making it possible that the stopping distance up to the crossing point follows a normal distribution in these close conflicts. The characteristic values of the normal distribution for pedestrians crossing behind the cyclist for a PET between 0 and 3 seconds are listed in Table 6.4 below for all crossings.

Table 6.4: Mean and standard deviation of the stopping distance up to the crossing point for pedestrians crossing behind the cyclist within a PET of 0 to 3 seconds.

Crossing	Origin - Destination	Mean	Standard deviation
Lorentzweg	East - west	3.1 m	1.1 m
	South - west	2.5 m	1.2 m
Mekelweg	South - north	3.3 m	0.78 m
	North - south	3.5 m	0.83 m

The average stopping distance of the pedestrian towards the cyclist is on average lower at the Lorentzweg crossing compared to the stopping distances at the Mekelweg crossing in the close conflicting cases. The deviation of the stopping distance is however larger at the Lorentzweg crossing compared to the Mekelweg crossing, potentially due to the variety of directions that the cyclist can approach from and go to.

6.2. Gradual movement adaptation

The course of the trajectories provides insight into the eventual outcomes of the behaviour. This section provides the quantities of modes that cross first and explains this by reasoning from the perspective of more gradual movement adaptations.

Table 6.5 below shows the total number of trajectories for both PET categories for the pedestrians walking from east to west. For the other directions, Appendix G provides the total number of crossing trajectories.

Table 6.5: Total number of crossing trajectories within each category of PET and first crossing mode

East - west (total trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	1082	704
Cyclist crosses first	811	800

The category of pedestrians crossing first within 0 to 3 seconds from the cyclist has the highest number of pedestrians and this decreases for the range of 3 to 5 seconds. In the cases where the pedestrian crosses after the cyclist, the quantities are very similar, potentially due to pedestrians preferring to remain at a larger distance from the cyclist, thereby crossing after the cyclist in the 3 to 5 seconds range, which is also supported by the quantity of stopping trajectories which is generally more (15.6%) than the reference non-crossing pedestrians (8.4%).

The division of first and second crossing pedestrians does however differ when considering the different directions that the cyclist can approach from. Table 6.6 shows the quantities of pedestrians walking at the Lorentzweg from east to west while going first and second depending on the approaching direction of the cyclist.

Table 6.6: Division of pedestrians crossing before (first) and after (second) the approaching cyclist, with a separate category for the approach of the cyclist

Origin cyclist	Ped first	Ped second	Total	Ped first percentage	Ped second percentage
West	672	252	924	72.7%	27.3%
South	754	994	1748	43.1%	56.9%
East	46	64	110	41.8%	58.2%
North	174	169	343	50.7%	49.3%

The quantities and percentages in pedestrians crossing first/second show large differences. The pedestrians crossing with a cyclist from the south cross mostly after the cyclist, whereas the pedestrians crossing with the cyclist from the west more often cross first. This potentially arises due to the right of way that pedestrians have in the second scenario, though this is not reflected in the scenario where the cyclist approaches from the east. For the north direction no strong preference seems to be present for one of the two participants. A possible explanation for the divide of these last two directions is that the approach from the east and north cause a complicated traffic situation, where it is unclear for both modes which mode has priority. Therefore, the pedestrian might hold back more often due to the higher speed of the cyclist, even though it has priority.

Following the outcomes of first and second crossing pedestrians however does not explain the underlying approach of both modes on how the order of crossing was determined. The predicted PET can provide an insight into these approach cases to estimate the process behind the eventual outcome of a conflict. Figure 6.4 shows the predicted PET values for pedestrians crossing with a cyclist walking from east to west at the Lorentzweg crossing. This is the only scenario for which this calculation was conducted. Again, the trajectory points inside the considered frame up until the crossing point are considered which in this figure are evenly spread on a normalised scale of 0 to 1, to be able to visually compare all the developments of trajectories. The trajectories indicated in red cross before the cyclist, whereas the blue trajectories eventually cross behind the cyclist. An opacity has been applied to get an understanding of the density of predicted PET values.

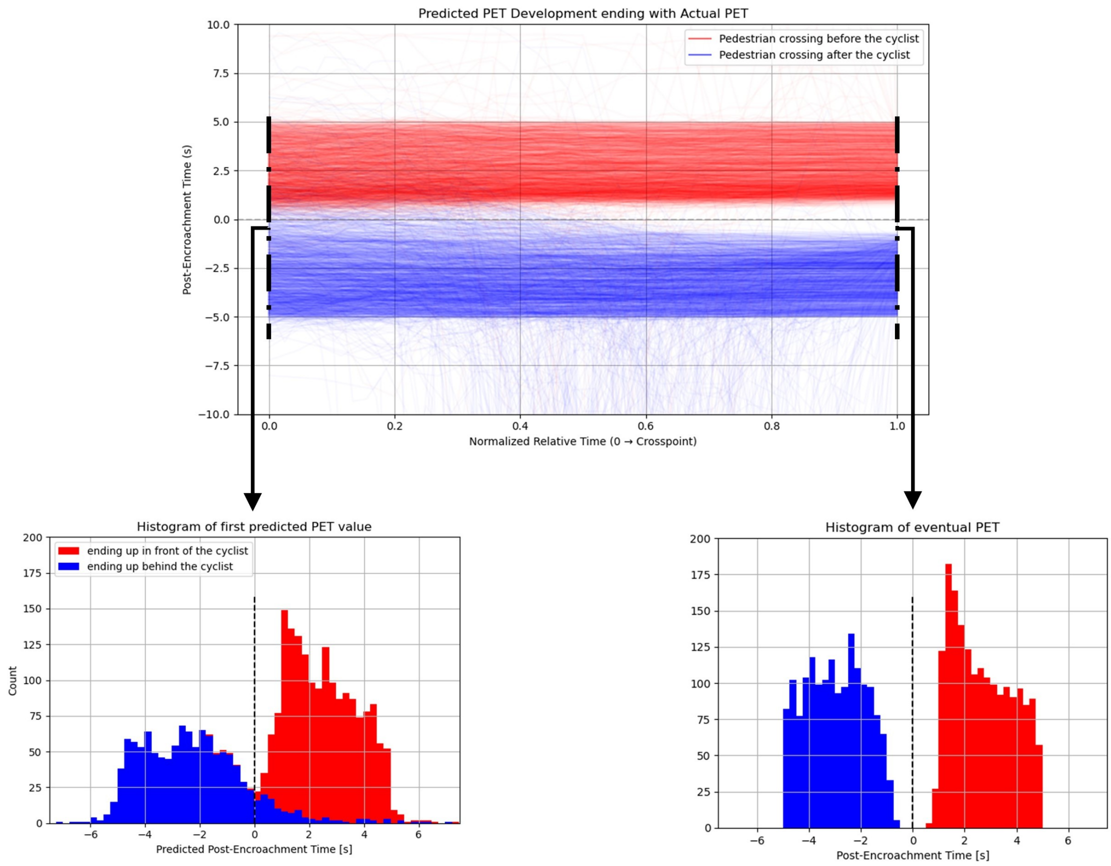


Figure 6.4: Development of the predicted PET for all trajectories of pedestrians walking from east to west at the lorentzweg crossing while crossing with a cyclist.

Following the general course of the predicted PET values, the trajectories seem to gradually divide themselves either to a positive PET value of more than 1 s or a negative PET value less than -1 s. This is likely because a PET value in between this range would indicate a collision since the exteriors of the two traffic participants are then in too close range from each other. Because the traffic participants want to prevent the collision from happening, they make decisions in their speed and direction in such a way that a collision does not occur. This figure indicates that often times the pedestrian crosses the cyclist from behind if a predicted collision would have occurred. This can be seen in the primarily blue coloured predicted PET values in between -1 and 1 s at x = 0. Two (stacked) histogram plots are shown in Figure 6.4 as well with the counts of these trajectories at the beginning (x = 0) and end (x = 1, or the actual PET), where it is confirmed that most of the predicted PET trajectories in between -1 and 1 end up behind the cyclist (again indicated by the blue colour).

The actual PET values (right histogram) show two peaks on both sides of the zero PET value, albeit that the blue peak is less distinct. The peak in red suggests that unlike pedestrians, cyclists tend to cross pedestrians sooner after the pedestrian has passed the conflicting point. The red peak corresponds with the higher quantity of first crossing pedestrians between 0 to 3 seconds from Table 6.5. In the predicted scenario (left histogram), this peak is also already noticeable. This would suggest that before the pedestrian reaches the point to change its movements, the cyclist has already made the decision to yield to the pedestrian and has behaved accordingly.

To evaluate the consistency of the predicted PET values across the trajectories, the predicted first crossing mode is quantitively compared with the actual first mode. The predicted first mode is derived based on the sequence of predicted PET values along the trajectory by evaluating the sign of predicted PET across multiple time steps. Table 6.7 shows how often the pedestrian was temporarily predicted as crossing first and second compared to the eventual outcome. These tables are provided below.

Table 6.7: Count of pedestrians predicted first and second compared to the actual outcome considering several consecutive trajectory points.

Number of consecutive trajectory points	Predicted	Actual first	Actual second
1	First	1727	155
	Second	59	1456
2	First	1747	134
	Second	39	1477
3	First	1754	119
	Second	32	1492
4	First	1761	99
	Second	25	1512

The values in each section of the table on the anti-diagonal show that in all cases, it occurs more often that during a certain part of the trajectory the pedestrian is predicted to cross first, but ends up second than the other way around, again suggesting that the pedestrian is more likely to yield to the cyclist, than the other way around.

6.3. Deviating Behaviour

In this section, an approach is presented to distinguish deviating trajectories from non-deviating ones. This is done by comparing the deviation of crossing and non-crossing trajectories to find a threshold value. Thereafter, the approach of the cyclist is used as a variable to distinguish the scenarios where deviating behaviour occurs most often.

6.3.1. Deviating threshold

The trajectory plots of Figure 6.2 show a number of trajectories that deviate from a straight path towards the end goal. Particularly, the paths that align with the foot path seem to have a region with many trajectories that deviate southwards after the footpath makes a turn in this direction. For calculating the deviation of the pedestrians walking from east to west at the Lorentzweg crossing, only the paths that start and end at the height of the foot path in between the specified frame in Figure 5.12 are considered. For the scenarios of the Mekelweg crossing, the full trajectory is considered.

A measure for the deviation is the root mean squared deviation (RMSD). The RMSD can be calculated with the lateral distance of each trajectory point from a centre line. The centre line is in the case of pedestrians walking from east to west defined as the line in between the first and last trajectory point in between the specified blue frame. A visualisation of this centre line is made of a sample trajectory in Figure 6.5.

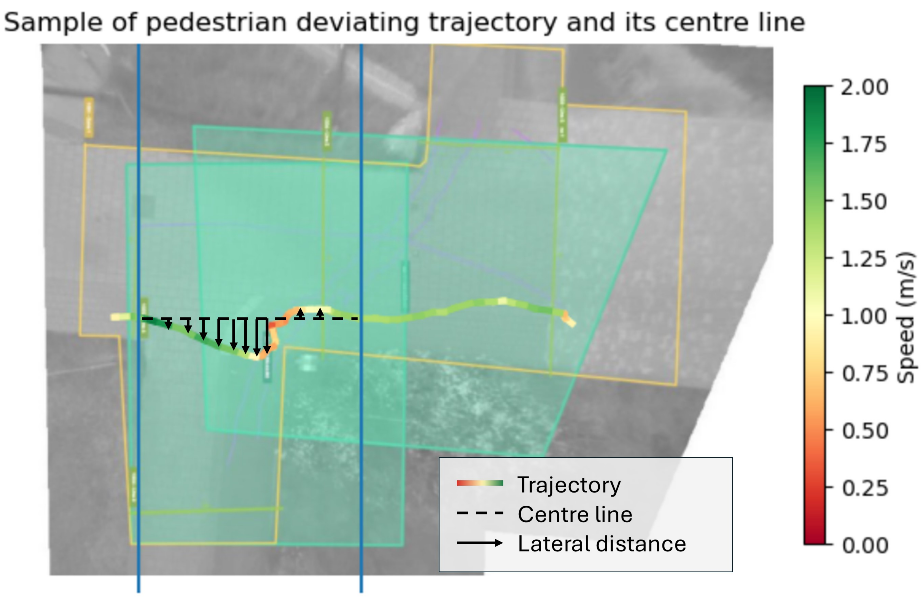


Figure 6.5: Sample trajectory of a pedestrian walking from east to west with the highlighted centre line and lateral distances of the trajectory points.

The lateral distance from each trajectory point to this line can be determined by applying the distance formula from a point to a line (GeeksforGeeks, 2024a):

$$d = \frac{|ax_p - y_p + b|}{\sqrt{a^2+1}} \tag{6.2}$$

Where x_p and y_p are the coordinates of that trajectory point and a and b are the slope and y-intercept of the straight line respectively.

To compare the distances for a full trajectory with respect to the centre line, the formula for the root mean square deviation is applied that provides the overall magnitude of deviation from the central line (Ather, 2022):

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n-2}} \tag{6.3}$$

Where n is the number of trajectory points within the considered frame. The first and last trajectory point are not considered in the total deviation calculation ($n-2$), because the distance to the central line has to be zero if this line is based on these two points.

In some cases, the overall deviation of a trajectory might not have a high value, while the trajectory might deviate significantly at a specific location, e.g. right before entering the bike path. For this reason, not only RMSD, so overall deviation, is considered, but also the maximum deviation of a trajectory, measured with the same lateral distances calculated with formula 6.2.

For both the RMSD and the maximum deviation, it needs to be verified what a suitable threshold could be for appointing deviating trajectories. Determining the threshold values is done for the pedestrians walking from east to west at the Lorentzweg crossing. This is because it has a larger number of pedestrians, making the findings statistically more relevant. These thresholds are then applied to all three straight crossing scenarios.

First, insight in the deviation values for crossing pedestrian trajectories is needed and this is compared to the general deviating values of non-crossing pedestrians. Histogram plots are made of the RMSD and maximum deviation of all trajectories of pedestrians walking from east to west. Figures 6.6 and 6.7 show these plots respectively, with a log-normal fitting function used to approximate the course of these deviations. Deviations often exhibit right-skewed behaviour, so small deviations are common, but large deviations are rare and occur with decreasing frequency. The log-normal distribution models this kind of asymmetry.

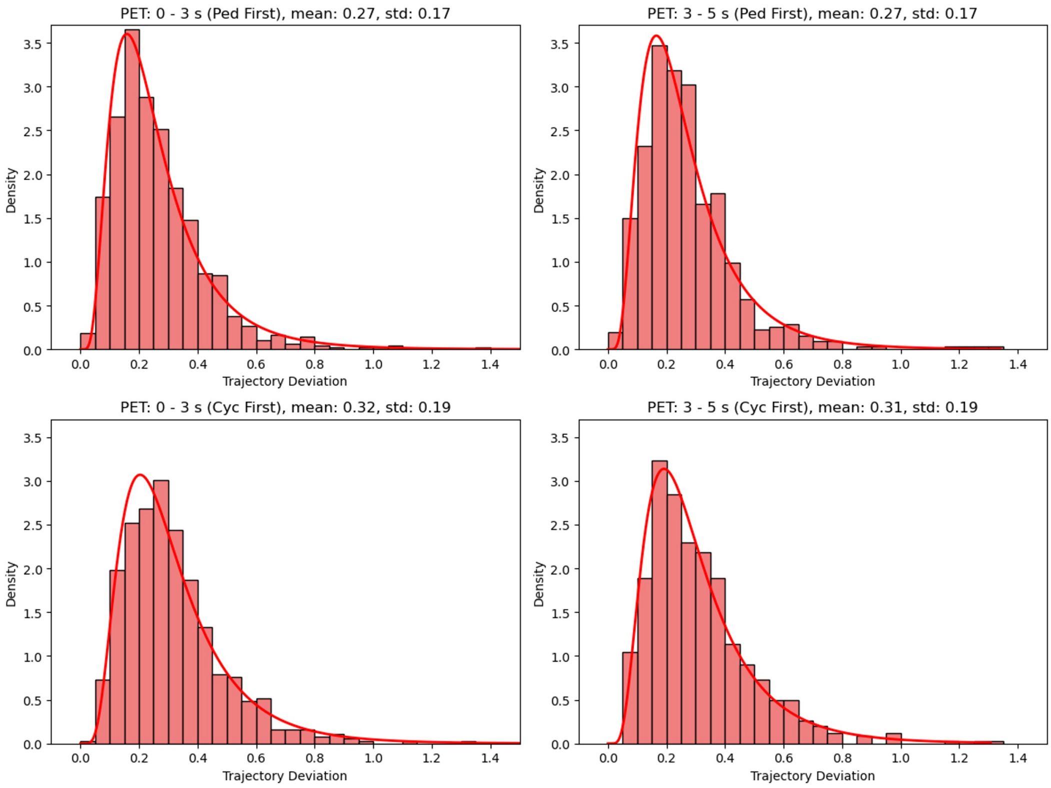


Figure 6.6: Histograms of the root mean squared deviation of crossing trajectories

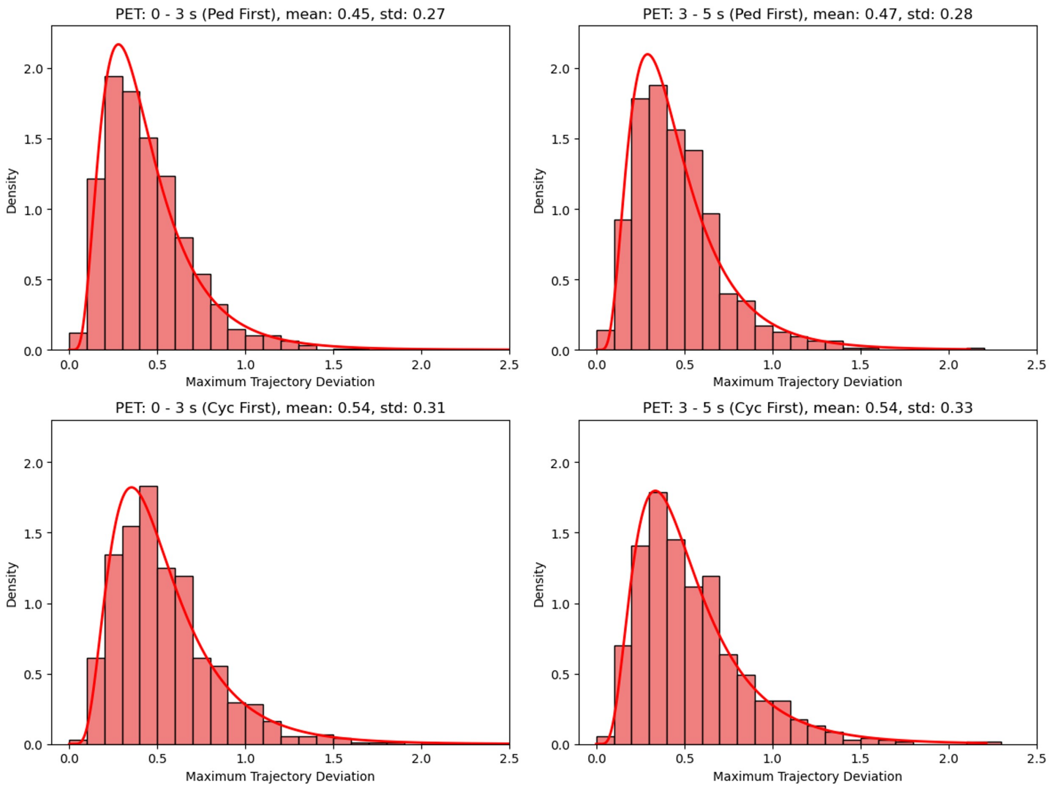
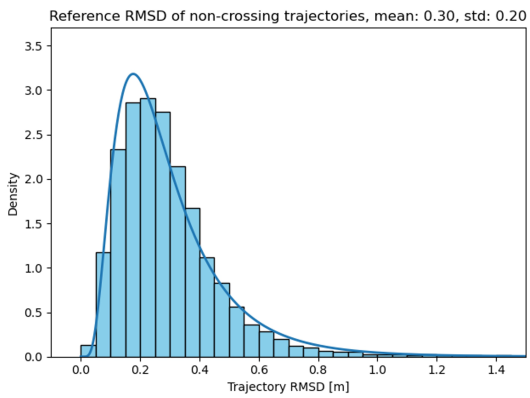
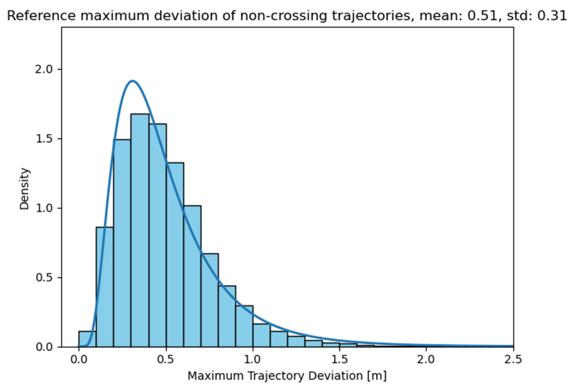


Figure 6.7: Histograms of the maximum deviation of crossing trajectories

The RMSD and maximum deviation both have a higher mean and standard deviation in the scenarios where the cyclist crosses first, suggesting that pedestrians tend to deviate from their path more frequently and that this deviation is larger if the cyclist crosses first than if the pedestrian crosses first. It is however key to understand how these numbers compare to the pedestrians that do not cross with a cyclist, to detect whether movement changes occur. A similar histogram can thus be made analysing the same framed area of pedestrians going east to west that do not cross with cyclists from the side for both the RMSD and the maximum deviation. Figures 6.8a and 6.8b show these histograms for the RMSD and the maximum deviation respectively.



(a) RMSD



(b) Maximum deviation

Figure 6.8: Histograms of the RMSD (a) and maximum deviation (b) of non-crossing pedestrians.

Although the figures seem to be similar in shape, the minor differences in means and standard deviation together with the quantity of the data make for a different distribution for all scenarios, so the deviation of non-crossing pedestrians, pedestrians crossing first and pedestrians crossing second. To test whether the distributions are significantly different, a Kolmogorov-Smirnov-test (KS-test) is performed on the dataset.

The distribution of deviations among non-crossing pedestrians serves as the reference group. The null hypothesis (H_0) for each comparison is that the distribution of trajectory deviations (both RMSD and maximum deviation) for crossing pedestrians does not significantly differ from that of non-crossing pedestrians. The alternative hypotheses (H_1) state that crossing pedestrians exhibit significantly different deviation distributions compared to non-crossing ones.

The test results of each of the hypotheses are shown in Tables 6.8 and 6.9 highlighting that all distributions are significantly different (p-value < 0.05).

Table 6.8: Significance test for the trajectory RMSD for crossing pedestrians compared to non-crossing pedestrians.

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.072	8.02e-5	Yes	Significantly different
	3 - 5 s	0.083	3.32e-4	Yes	Significantly different
Cyc first	0 - 3 s	0.104	1.89e-7	Yes	Significantly different
	3 - 5 s	0.060	1.30e-2	Yes	Significantly different

Table 6.9: Significance test for the trajectory maximum deviation for crossing pedestrians compared to non-crossing pedestrians.

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.067	3.71e-4	Yes	Significantly different
	3 - 5 s	0.066	8.48e-3	Yes	Significantly different
Cyc first	0 - 3 s	0.103	3.30e-7	Yes	Significantly different
	3 - 5 s	0.077	5.38e-4	Yes	Significantly different

This significant difference is not always present for the Mekelweg crossing scenarios, of which the histogram plots and the resulting statistical tests can be found in Appendix F. It can be noted here that a significant difference between two distributions is more likely to arise for higher amounts of trajectories, so the values of Tables 6.8 and 6.9 are assumed to provide the most accurate results.

Because the crossing and non-crossing pedestrians are significantly different in their deviating behaviour, this difference can be used to determine a threshold for deviating trajectories. The threshold that is determined is based on the difference between the non-crossing pedestrians and the pedestrians that cross after the cyclists. The pedestrians that cross before the cyclists also show a significant difference in deviating behaviour, but this is because these pedestrians actually tend to walk in a more straight path than non-crossing pedestrians do. This is different from the assumption made in section 3.4.3, where the pedestrian was assumed to deviate slightly away from the approaching cyclist if it decides to cross before the cyclist. Possibly, this is due to these pedestrians preventing a collision by walking straight to their goal, instead of a gentle walk where more lateral deviation could occur.

The contrast between pedestrians who cross behind a cyclist and those who do not cross can be characterized by comparing their fitted log-normal distributions for trajectory RMSD and maximum deviation. Appendix E provides a visual overview of the steps used to establish a threshold based on these distributions.

To better understand where deviating behaviour when crossing diverges most significantly from the reference behaviour, a graph is created that maps this divergence across the full range of RMSD and maximum deviation values. Rather than relying on absolute differences a relative comparison is made, because the absolute difference does not necessarily emphasise the most distinct cases. The graph highlights where crossing pedestrians exhibit disproportionately frequent deviations compared to non-crossing individuals. The point of greatest relative difference is then selected as the threshold, as it represents the region where crossing behaviour deviates the most from the reference behaviour. Figures 6.9 and 6.10 show the graphs of the relative density difference with the threshold value highlighted as the maximum for each of the scenarios.

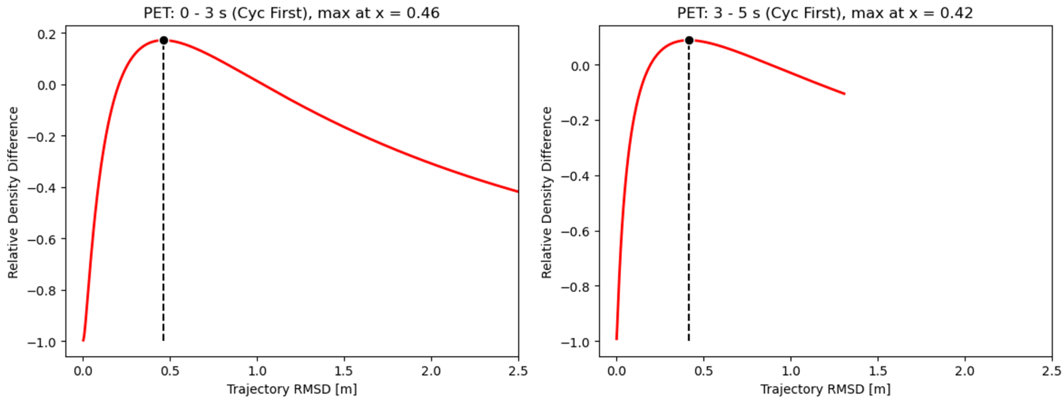


Figure 6.9: The density difference of the RMSD distributions between crossing pedestrians and non-crossing pedestrians relative to non-crossing pedestrians.

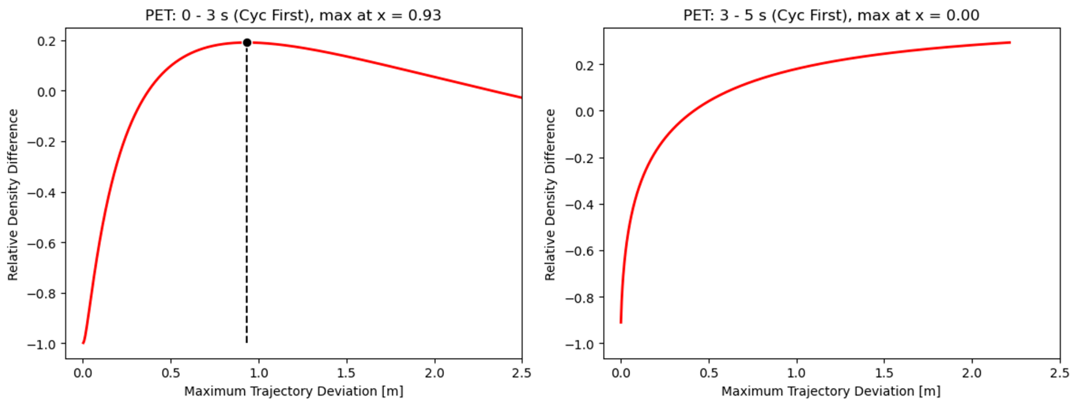


Figure 6.10: The density difference of the maximum deviation distributions between crossing pedestrians and non-crossing pedestrians relative to non-crossing pedestrians.

The threshold value that is chosen for RMSD is chosen in between the two founded values, which in this case is set at 0.45 m. For maximum deviation, no clear threshold value was found in the scenario of PET 3 - 5 seconds, which is why the threshold value in this case is set at 0.93 m, based on the PET 0 - 3 second scenario.

6.3.2. Deviating directions

The threshold values are used as a basis for extracting trajectories that deviate substantially. The deviation not only happens to certain extents, the directions to which pedestrians tend to deviate is also of importance when creating an understanding of this type of behaviour. In the Methodology, an assumption was made on the direction of the pedestrian, depending on the origin of the cyclist. The direction of the cyclist could be an indicator for the pedestrian to deviate towards the cyclist, in the case that the pedestrian crosses behind the cyclist. Therefore, subplots of the deviating trajectories are created that distinguish the four cardinal directions that the cyclist approaches from. These subplots are shown in Figure 6.11.

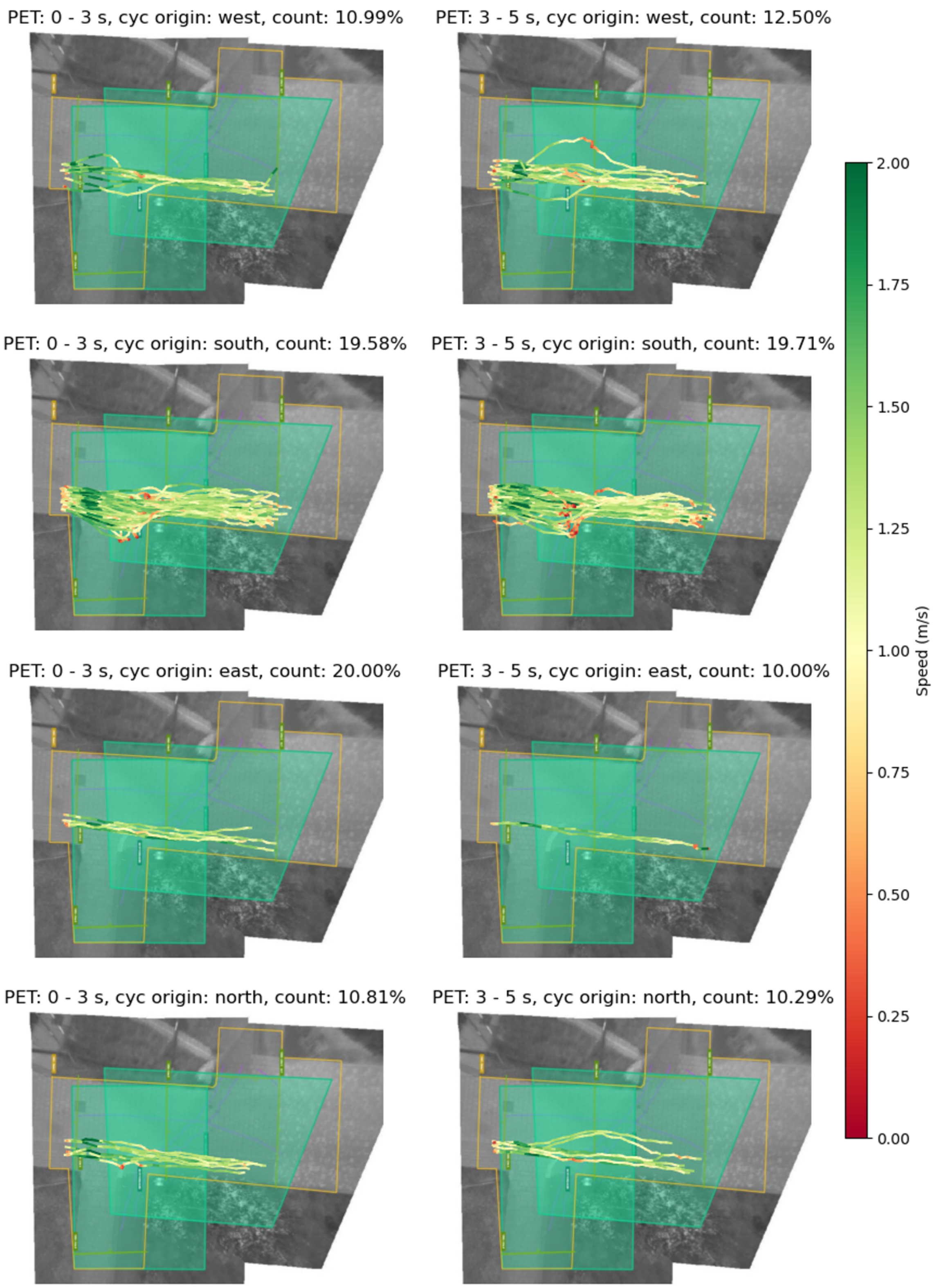


Figure 6.11: Subplots of the pedestrians walking from east to west crossing behind a cyclist that approaches from one of the four cardinal directions.

The percentages of deviating trajectories compared to the total number of pedestrian east-west crossings with a cyclist from one of the cardinal direction is given as the 'count'. From these percentages it can be seen that in the case that a cyclist approaches from the south, the pedestrian is deviating relatively often compared to the other directions. This deviation seems to be mostly southwards, so in the direction of where the cyclist is coming from. This behaviour is likely due to a pedestrian wanting to cross the cyclist earlier in time than it would if it would come to a complete stop. The percentage for the east direction is also 20%, yet a visual inspection of this category does not indicate clear deviations.

For the other crossing scenarios, no clear pattern could be detected in the direction of deviation based on the same threshold values, except for the pedestrians crossing from south to north at the Mekelweg. In the specific approaching case where the cyclist approaches from the east at this crossing, the pedestrians again seem to deviate slightly in the direction of the approaching cyclist.

Commonly, the percentages of deviating trajectories were higher, yet no real directional preference could be detected for all other crossing scenarios. The trajectories of pedestrians crossing behind the cyclist at the Mekelweg that deviate substantially are provided in Appendix I with again the origin of the cyclist used to classify the trajectories.

Even though deviation can be detected in some approaching cases, in general it occurs less often compared to pedestrians that stop in their approach to cycling paths.

6.4. Conclusion

In this section, an answer is provided to the sub-research question: what movement changes does a pedestrian perform when approaching a cyclist in a sideways conflict?

A pedestrian can perform multiple changes to its movement when it approaches a cyclist. The results show that stopping behaviour occurs approximately 30 to 50% of the time in the case that the pedestrian crosses after the cyclist for a crossing happening within a PET of 0 to 3 seconds from each other. Stopping happens slightly less often in the case that the pedestrian crosses before the cyclist compared to pedestrians that do not cross with a cyclist.

The stopping distance that pedestrians tend to keep towards the eventual conflicting point fluctuates around a mean of about 3 meters, but is dependent on the circumstances of the crossing scenario. Figure 6.12 summarises the different stopping percentages for pedestrians crossing after a cyclist within 0 to 3 seconds, together with the average stopping distances that are kept for all scenarios.

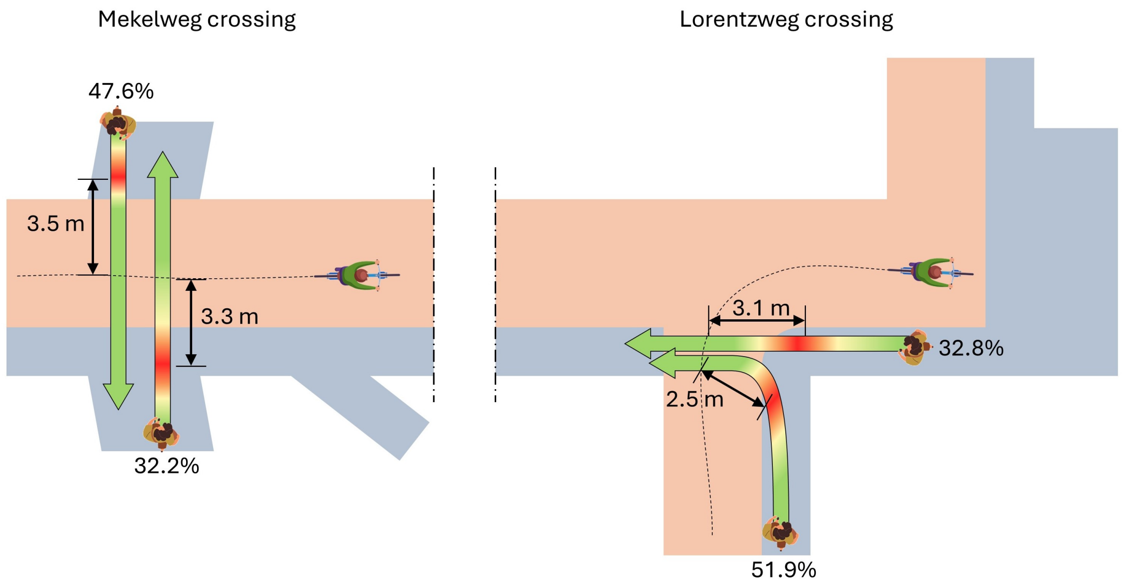


Figure 6.12: All stopping percentages and mean stopping distances of the different analysed crossing scenarios.

The data suggests that the pedestrians more often yield to cyclists in the cases that a close conflict is predicted (predicted PET in between -1 and 1 second). This is confirmed by observing and quantifying the general course of the number of predicted first and second pedestrians.

Deviation happens less often compared to stopping, yet there is a significant difference in deviating behaviour for crossing pedestrians compared to non-crossing pedestrians. The pedestrian tends to walk a straighter path when it decides to cross before the cyclist than when it does not cross at all. Deviation mostly happens if the pedestrian crosses after the cyclist while the cyclist approached from a branch sideways to the approach of the pedestrian. The deviation is then mostly towards the direction of the approaching cyclist.

7
Modelling Test Case

In this chapter, a test case is presented that was set-up in MassMotion. In this test case, the aim is not to propose a generalised adaptation of the Social Force Model, but rather to explore how targeted changes can improve the behaviour in a controlled scenario. A trajectory pair of an interacting pedestrian and cyclist was chosen that acts as a representative example. While this setup does not result in a broadly applicable model on its own, it serves as an illustrative example for how such adjustments could be incorporated into a more generic model in future work.

Following this case, sub research question 5 is attempted to be answered: *What adaptations can be done to the social forces model to mimic the behaviour of pedestrians towards approaching cyclists?*. The environment of the Lorentzweg crossing is digitally recreated. The pedestrian and cyclist movement of a single sideways crossing is attempted to align with the findings in the data. The steps that are performed are presented in Chapter 3.5 and follow the structure of the selection of the trajectories, the set-up of the environment, the initial formulation test and the adapted form of this test.

7.1. Sample from the data

To choose a representative interaction between a pedestrian and cyclist from the data, the intersection and crossing scenario needs to be determined. The main focus on this research has been on the pedestrians walking from east to west at the Lorentzweg crossing, because this scenario provided the richest amount of interactions between cyclist and pedestrians. Furhtermore considering that the trajectory data has a higher frame rate, covers a larger area and is therefore able to visualise the movement of the cyclist properly, the Lorentzweg crossing is chosen as the intersection of interest with the specific focus of the pedestrian walking from east to west.

The cyclist can approach the pedestrian from the 4 cardinal directions, yet when observing the deviating behaviour and pedestrians crossing behind the cyclist depending on the approaching direction of the cyclist (Table 6.6), the cyclist approaching from the south induces the most movement changes for the pedestrian in terms of stopping and deviating. This is the behaviour that is attempted to be tested in this case. Thus, the cyclist approaching from the south is selected as the to be analysed interaction.

The interactive cyclist-pedestrian pair that has been chosen is shown in Figure 7.1 with the pedestrian walking from east to west and the cyclist going from south to west. This example was chosen because the pedestrian got to a speed below 0.77 m/s with a stopping distance to the crossing point that is within the expected range of the normal distribution. Besides, a slight deviation southward could be detected at the instant the cyclist was approximately in front of the pedestrian. Furthermore, a video inspection of that time instant proved that there was no interference of other participants, neither by crossing nor by approaching in close range. The pedestrian crosses after the cyclist with a PET value of 2.9 seconds.

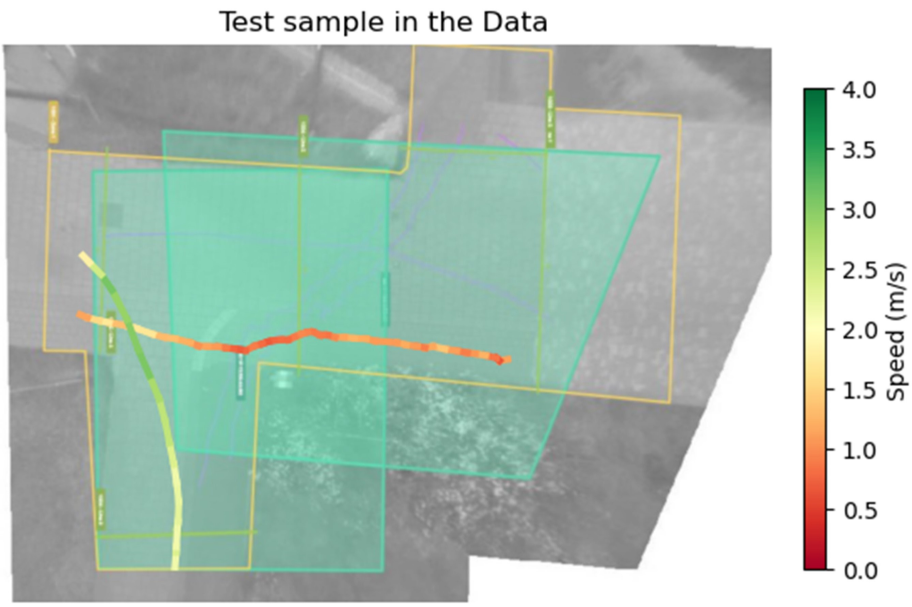


Figure 7.1: Chosen trajectory pair of a cyclist going from south to west crossing with a pedestrian going east to west.

7.2. Environmental set-up

The environment of the crossing is recreated in MassMotion by projecting a scaled map of the intersection as a blueprint while matching the position and orientation from the map to the axis of the data collected by the smart sensor.

Agents are placed within defined areas that represent the cycling and footpaths that constrain the agents to remain in that area. To allow movement between the pedestrian and cycling paths, transitions are only enabled at the specific crossing points. Since the focus of this study is on pedestrian behaviour in response to cyclists, the cyclist's path is fixed and follows the trajectory recorded in the data. As a consequence for the behaviour of the pedestrian, it is reactive instead of interactive, because the cyclist will not adapt its movement based on the presence of the pedestrian. Therefore, every time step, the pedestrian will change its movement to the cyclist and its current direction. Figure 7.2 shows the set-up of the environment.

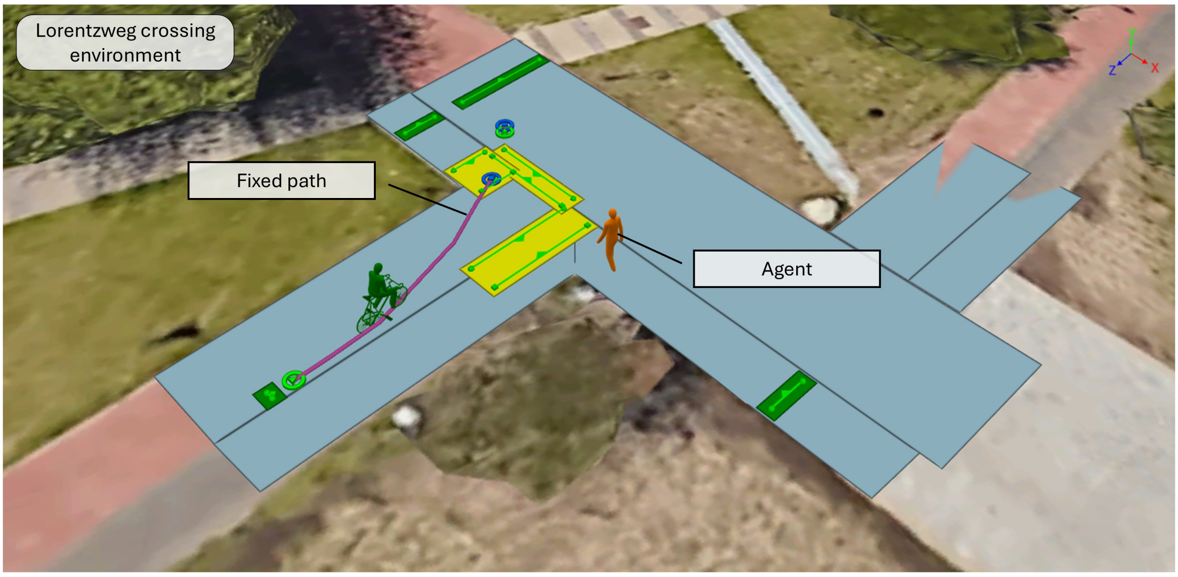


Figure 7.2: Set-up of the environment in MassMotion

After the crossing point, the cyclist did not perfectly adhere to the fixed path that was created for the cyclist, because it had to be split into two sections. This is why the trajectory of the cyclist does not perfectly align with the trajectory from the data. The speed of the cyclist can furthermore not be adapted during its trip, therefore, the speed at the beginning of the trajectory is the speed that the cyclist mostly contains during the whole trajectory.

The appearance of agents can be timed with a precision of one second. Creating an agent from a starting point with this precision might result in timing issues for the conflict. Therefore, an estimation of the location of the portals has to be made from where the pedestrian and cyclist agent will start, to follow the correct location and timing of the trajectories in the data.

7.3. Original formulation test

The original formulation test that is performed considers the script that was created in the SDK to make a first prediction of movements for cyclists and the responding pedestrians, created by Oasys. For every time step (0.2 seconds) in the simulation, the speed and direction of the pedestrian are determined based on the presence of the cyclist, together with all other already existing forces in the model. The speed of the cyclist varies with a mean of 4 m/s, a standard deviation of 0.108 m/s and has a minimum and maximum value of 2 and 7 m/s respectively. This standard deviation is likely to be higher when considering the findings of the speed distributions of cyclists at both crossings in the data (Chapter 5.2.1). For the purpose of this test however, this deviation is set to an even lower value, to match the speed of the cyclist in the sample.

The influence from the cyclist to the pedestrian is in the SDK formulated as an overruling statement on the initial model, so no force is formulated. The direction that the pedestrian tends to go to is driven by a combination of frontal, sideways and rear distances that the pedestrian agent tries to maintain from the cyclist, which are 6, 0.5 and 2 meters respectively. If the pedestrian approaches the cyclist within these regions, it will react by moving with the same speed towards the nearest escape direction. This escape direction is perpendicular to the direction of the cyclist and always points away from this axis of movement. The script also calculates the collision courses and sets the TTC value of 3 seconds as a threshold below which a similar fleeing reaction is performed with an additional increase in speed of 50%. Lastly, if the pedestrian gets in the range of 1 meter around the cyclist, the pedestrian will again move in the escape direction with a speed increase of 50%.

An issue with the fleeing reaction based on the sideways approach arises in the simulation where unexpected behaviour of the pedestrian is detected. When performing the simulation, the trajectories show the movement patterns as illustrated in Figure 7.3.

Figure 7.3: Trajectory pair of simulated cyclist agent crossing with the pedestrian agent for the old situation in MassMotion.

The trajectory of the pedestrian shows fleeing behaviour. This can be seen in the red part of the trajectory overlapped by an orange part, indicating that the pedestrian turned around twice to not get near the cyclist. This behaviour occurs while the cyclist is not in the near regions of the pedestrian. The simulation in this case does not properly reflect the behaviour that was detected in the data, where the behaviour of the pedestrian was more in stopping and a slight deviation southwards.

The velocity vector of the cyclist is rotated 90 degrees to the left and normalised. Between this vector and the vector of relative position (from the cyclist to the pedestrian agent) the dot product is calculated to consider whether the pedestrian is below the 0.5 meter range from the cyclist. Two visualisations of this calculation are shown in Figure 7.4.

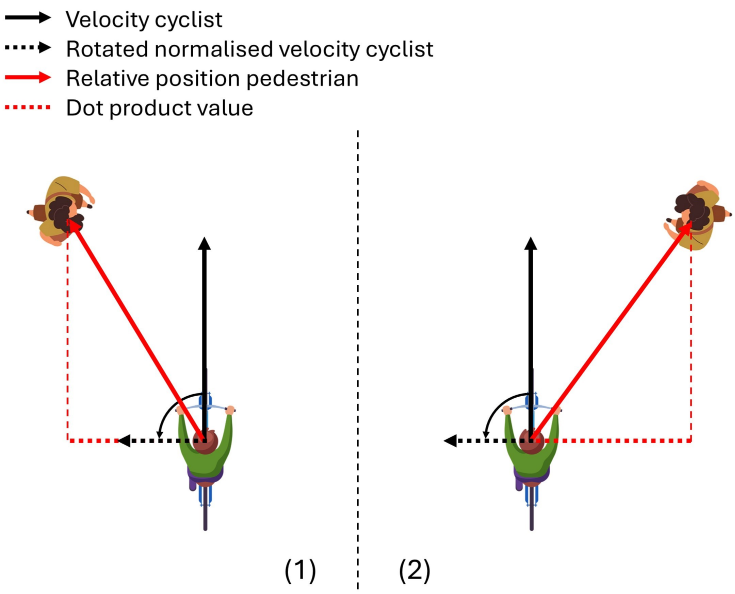


Figure 7.4: Two situations where the dot product is calculated as a projection of the cyclist-pedestrian vector on the normalised rotated cyclist direction vector. The value can be both positive (1) as negative (2).

This method only partially works for an interaction between a cyclist and a pedestrian. The dot product is a measure to determine the common direction two vectors have with each other. When the two vectors are pointing in opposite directions (situation 2), the value of the dot product takes on negative values, thereby being below 0.5, inducing pedestrians to flee in the perpendicular direction from the cyclist velocity direction. An approaching pedestrian from the right from the perspective of the cyclist reacts therefore in an unusual way, by fleeing in the opposite direction.

To correct this behaviour, the dot product can be made positive by taking the absolute value bars. Although this adaptation should contribute to more suiting behaviour of the pedestrian, the fleeing behaviour in general has almost not occurred in the data. The formulation in general therefore requires several adaptations that potentially improve the movements of the pedestrian.

7.4. Adapted formulation test

The new formulation of the simulation takes into account the findings of the data and attempts to translate these into a new formulation. This formulation is performed in the SDK. Two main topics are addressed in this new simulation: the speed and the direction of movement for the pedestrian.

The pedestrian is likely to decrease its speed when encountered with an approaching cyclist. It is therefore important to distinguish in which situations it is likely that the pedestrian is about to slow down. In the previous formulation, the pedestrian either changed its movement due to its distance with respect to the cyclist or because of a collision course expressed in TTC. This formulation will make use of predicted PET instead of TTC, because predicted PET also distinguishes potential near collisions that pedestrians might want to avoid. The predicted PET that is calculated in this scenario, considers only the speed and movement direction of the current time step to calculate the timing in between the pedestrian and cyclist agent to cross each others paths. The crossing point between these two modes is calculated for every time step as well. This can then translate into the distance from which the pedestrian tends to slow down or stop.

For this particular situation, the distance where the pedestrian reaches a speed below 0.77 m/s of the pedestrian is 3.86 meters, as established from the trajectory analysis. The adapted formulation introduces conditions based on predicted PET and distance to the conflicting point to trigger a speed change in the pedestrian movement. In this scenario, the speed is reduced by 37.5% if the pedestrian PET is predicted to be less than 3 seconds and the distance to the collision point is less than 3.5 meters.

The 37.5% reduction is based on an estimate of the difference between the pedestrian's speed before and after the deceleration phase in the trajectory, which ranges from 1.2 m/s to 0.75 m/s respectively, see Figure 7.5. The PET threshold of 3 seconds reflects observed fluctuations in the predicted PET around that value during the encounter, see Figure 7.6, while the distance threshold of 3.5 meters is slightly lower than the observed 3.86 meters to compensate for the more immediate reaction of pedestrians changing speed in the simulation, which tend to be less gradual than the case from the data.

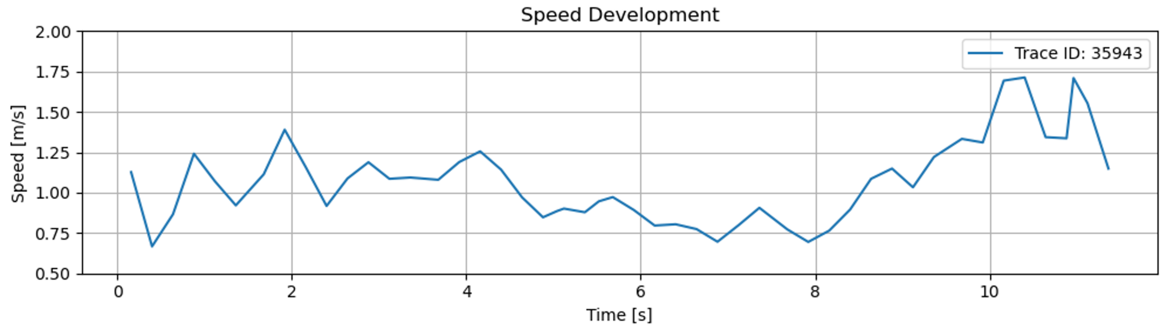


Figure 7.5: Development of the speed from the sample of the data.

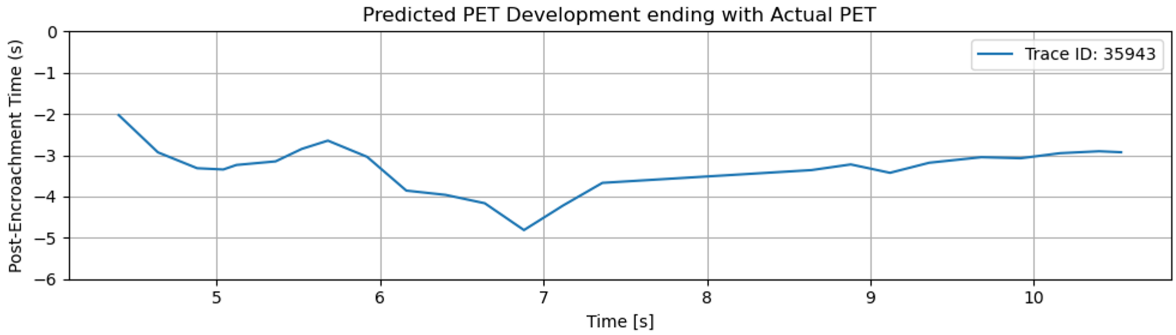


Figure 7.6: Development of the predicted PET from the sample of the data.

Besides the speed, the direction of the pedestrian needs to be determined as well. The regular formulation of social forces directs a force from the cyclist to the pedestrian, creating a movement of the pedestrian parallel to the path of the cyclist, see Figure 7.7, situation 1. The previous method explained in section 7.3 prevented this situation by using a perpendicular force from the direction of the cyclist, yet this would possibly not result in the deviating behaviour that could be observed in the data. For this reason, an additional point is considered in front of the cyclist that influences the preferred direction of the pedestrian. Instead of a force directed solely from the cyclist, this method uses a combination of a force exerted from the cyclist and from the additional point in front of the cyclist. This is shown in the second situation in Figure 7.7.

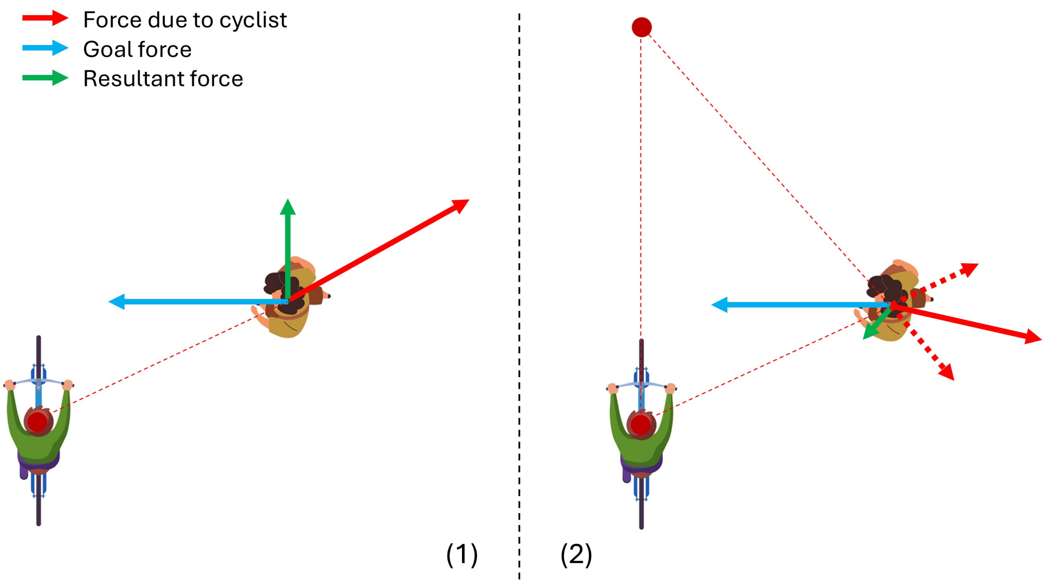


Figure 7.7: Two situations where the pedestrian is tempted to walk along the direction of the cyclist (1) and where the pedestrian is tempted to go behind the cyclist (2).

The position of the additional point is based on the current speed and direction of the cyclist. Specifically, it is placed at a distance of five times the cyclist's velocity vector at that time step, meaning that the cyclist's position is predicted five seconds into the future. This multiplier can be interpreted as a predictor: it approximates the future position the cyclist would reach if it maintained constant speed and direction. The resulting influence on the pedestrian is calculated from vectors pointing from both the cyclist's current and projected positions toward the pedestrian, which together define a preferred 'escape' direction. The resulting direction of the pedestrian is then compiled of its previous and escape direction. The choice of a five-second projection is mostly based on visual evaluation of the trajectories, but can also be interpreted as modeling an anticipatory behaviour in which the pedestrian responds to the space the cyclist is expected to occupy in the near future. As the two influence vectors average out, the resulting directional influence is approximately aligned with the cyclist's position 2.5 seconds ahead, reflecting an assumed anticipatory window in the pedestrian's decision-making.

The final direction is then calculated by multiplying both the initial and escape directions with weights based on the predicted PET. The weights for the escape and previous directions are calculated by the following formula:

$$weight_{escape} = \frac{1}{|PET| + 2} \quad \& \quad weight_{previous} = 1 - weight_{escape} \tag{7.1}$$

The weight of the escape direction cannot be higher than the weight of the current direction, therefore, the turning around behaviour of the pedestrian is prevented.

The eventual outcome of the simulated trajectories are shown in Figure 7.8. The trajectory of the pedestrian shows slowing down behaviour and a slight bit of deviation southwards, which is similar to the observed behaviour.

Figure 7.8: Trajectory pair of simulated cyclist agent crossing with the pedestrian agent for the new situation in MassMotion.

Visually, the trajectory of the adapted formulation aligns with the sample from the data. Furthermore, a comparison of the speeds for all three cases, shows that the adapted formulation and test sample from the data follow a similar speed path. Figure 7.9 provides a visual comparison of all three speed profiles.

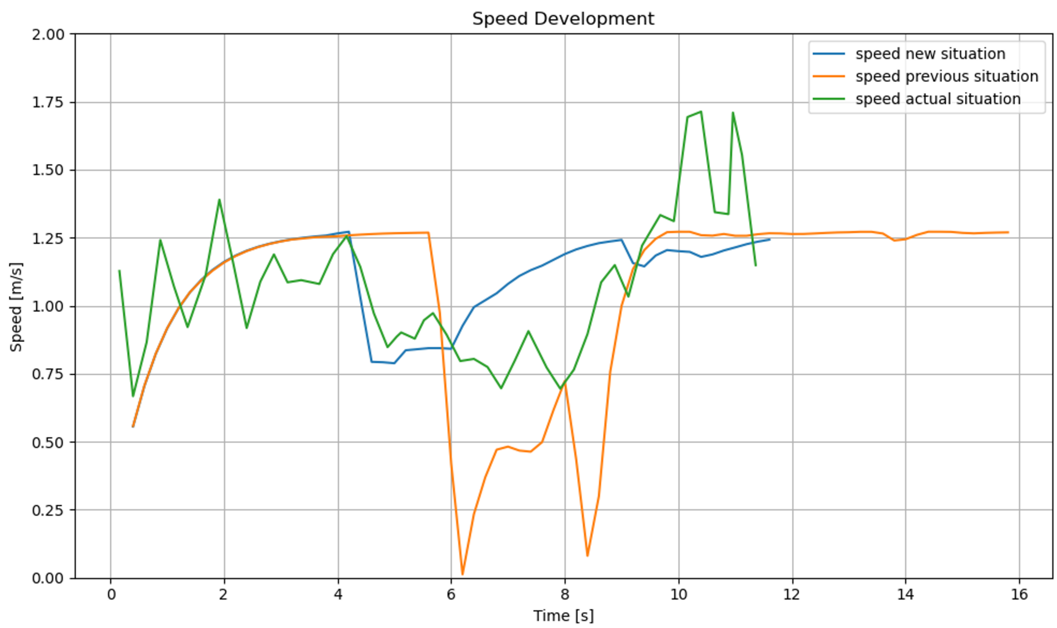


Figure 7.9: Comparison of the speeds of the three situations.

The speed profiles of the previous and new situation in Massmotion start exactly the same until the interaction with the cyclist takes place, from here on, the new situation resembles the speed profile of the actual situation better compared to the old situation. After 6 seconds, the pedestrian agent in the new situation slowly recovers to its original speed, whereas in the data the pedestrian seems to keep a low speed for a longer time period. As a result, the eventual post-encroachment time of the simulated situation is slightly lower compared to the actual situation. The total time spent in the area of both the pedestrian agent in the new situation and the pedestrian from the sample of the data are the same, meaning that both pedestrian's trip takes the same time, while also the distance that is covered by the pedestrians is similar. Table 7.1 below provides an overview of the pedestrian's PET value, total time spent in the area, and the covered distance for all three situations.

Table 7.1: Numerical comparison of the PET, total time spent and total distance covered by the pedestrian (agent) in the detected area.

	Old situation MassMotion	New situation MassMotion	Actual data
PET	4.56 s	2.36 s	2.93 s
Total time spent	15.8 s	11.6 s	11.6 s
Distance covered	16.6 m	12.6 m	12.4 m

7.5. Conclusion

The adaptations that are proposed in this Chapter align with the findings of the data that the pedestrian is tempted to slow down or stop when confronted with an approaching cyclist. Also, a method is introduced that induces slight deviations in the direction of the approaching cyclist. Thereby, an answer is provided to the sub-research question: what adaptations can be done to the social forces model to mimic the behaviour of pedestrians towards approaching cyclists?

Figure 7.10 shows the three situations next to each other, with the new situation to the right.

Figure 7.10: Trajectory comparison of the cases (from left to right): actual trajectory, previous set-up, new set-up.

However, the method that is proposed in this thesis requires significantly more test cases to properly calibrate the parameters and validate the results of the simulation. The parameters that should be properly calibrated are the slowing down percentages, the multiplication factor of the cyclists velocity for determining the additional point in front of the cyclist, and the weight of the escape vector based on the PET value. It should be considered what the effect is that the new formulation has on the movements of pedestrians right in front of the cyclist and whether it should be adjusted for these cases.

Also, the methods introduced in this chapter have been attempted to align with a single sample pedestrian-cyclist pair, whereas the results of all the interactions show a variety in behaviour. Thus, a proper model should consider this stochasticity of for example threshold values for preferred stopping distances and for predicted PET.

8 Discussion

In this Chapter, the findings of this study are discussed and related to the existing knowledge on pedestrian-cyclist interactions. A critical overview is provided by comparing the findings to existing literature and highlighting the limitations of this study with respect to the collected data and applied methods.

8.1. Thesis in light of the existing literature

The studies that have been examined show an understanding of the movement of cyclists in many different approaching cases with a pedestrian, yet this thesis can supplement the general knowledge on pedestrian-cyclist interactions by focussing on the behaviour of the pedestrian. This section considers how the findings relate to earlier work on pedestrian behaviour.

For stopping behaviour, a threshold value was set and the quantities of stopping pedestrians were compared in different crossing scenarios. Then the stopping distances that pedestrians tend to keep towards the crossing point and distributions were created to detect whether the type of conflict has an influence on the stopping distance. Stopping behaviour specifically for pedestrians approaching cyclists has to the writer's knowledge not yet been studied, yet methods of mimicking the behaviour of pedestrians towards cyclists in a social forces model have been proposed.

The findings of the PET thresholds are in accordance with some of the PET values that have been determined in previous studies. The studies of Zangenehpour et al. (2016) and Tageldin and Sayed (2016) agree that the conflicts within a PET of 0 and 3 seconds are dangerous conflicts. Therefore the findings of the data shown in subsection 6.1.1 align with this threshold value that from a PET value of 3 seconds the amount of stopping trajectories seem to drop less significantly, suggesting that the severity of the conflict influences the stopping behaviour. These same studies however also consider conflicts with PET below 1.5 seconds. In relation to the stopping percentage, no significant change in stopping behaviour was found in the data for this PET value. Also, the study of Beitel et al. (2018) presented a threshold value of 2 seconds, for which again no significant change in stopping behaviour was found. All studies seem to agree on 5 seconds to be the maximum value of PET that is considered to be an interaction. However, in this research there seemed to be a difference between the amount of stopping for non-crossing trajectories and for the 5 second PET interactions, namely 8.4% against 17.8% respectively in the most significant approaching scenario. It is unknown whether from a PET higher than 5 seconds a significant drop in stopping trajectories can be detected.

Furthermore, it is assumed that the PET classification of 0 to 3 and 3 to 5 seconds can also be applied for deviating behaviour. It is however unclear whether this assumption is justified and possibly further research could explicitly show a relation between PET and deviation.

Deviation of pedestrians paths have been detected in the study of Afghari et al. (2014), though it was not specified to what extent and in which direction this deviation takes place. The deviating direction of the pedestrian with respect to the cyclist is regularly considered in the studies on social forces. Some of these studies (W. Wang et al., 2024, LIANG et al., 2012) assume that the force exerted from the cyclist to the pedestrian acts along this direction. The resulting movement of pedestrians in a sideways conflict would be that the pedestrian moves along in the movement direction of the cyclist. However, the deviating behaviour that was detected in this study suggests that the pedestrian moves in the direction that the cyclist comes from, thereby contradicting the outcomes of these studies. Other pedestrian studies (Dias et al., 2018, Yuan et al., 2019) show that the deviating behaviour can indeed be towards the direction the crossing cyclist is coming from, because a future position is considered.

This study has been conducted at cyclist-pedestrian intersections with a separation of both modes by indicated coloured pavement and a ribbed strip of pavement in between. From other literature it is unclear how this specific type of infrastructure influences the choices of pedestrians and cyclists compared to other (separation) methods, such as kerbs (vertical or sloped), solid barriers or no separation. A separate study, focussing on these types of pavement designs could clarify the specifics of the expected behaviour at each of these types.

8.2. Limitations of the Data

The TU Delft campus has a bias considering the population distribution, with relatively many young adults making use of the infrastructure. According to statistics of the TU Delft (TU Delft, 2024), the student population is significantly larger than the personnel population, which on average are younger (KorteAntwoorden, 2016 and Strategic Development, 2015) and the people are predominantly male (Data-Insights, 2025). The effect that these relatively young men have on the overall data is most noticeable in speeds and agile movements. Potentially older people tend to take less risks, as a result having pedestrians tend to stop more often, potentially decreasing the amount of deviation as an alternative for stopping.

Also, the trajectories at the Lorentzweg crossing (Figure 5.12) show a fixed measuring error when the trajectory transitions from one camera to the other and another error at the end of the trajectory, as explained in Chapter 5.1. Although it was mentioned that this would not affect the movement changes that the pedestrian makes before the crossing, it has an influence on the spread of the speed distributions. Because the standard deviation of the speed becomes larger, the overlapping region of speeds between pedestrians and cyclists becomes larger, therefore cyclist and pedestrian trajectories could be unintentionally filtered out or incorrectly labelled a cyclist or pedestrian.

Data on interaction might not always be fully captured in the smart sensor. Trajectories do not give any information on the negotiations that pedestrians and cyclists tend to perform when in a real life situation. The pedestrian might make a choice based on visual cues, e.g. eye contact, or even verbal interaction. Although the result of the behaviour is projected in the data, the methods that were applied to arrive to that physical behaviour is not always clear.

Furthermore, interaction is also an iterative process that changes due to the behaviour of the other mode. It is therefore a big assumption in this research that the trajectories of cyclists are treated as somewhat fixed or predetermined movements, whereas these are normally also determined based on the movement of the oncoming traffic participants. In section 6.2, it was suggested that cyclists perhaps made their decisions before entering the detection area, which could mean that cyclists tend to yield more than is suggested based on the findings in this research.

The data limitations could be solved by applying another collection methodology, such as conducting the experiment at another sight, with other equipment or with observers, or a different set-up of the smart sensors. A comparative study at different crossings outside of the campus could reveal what movements are performed by a representative sample of the population. Otherwise, the demographics of the traffic participants could be taken into account as an influencing variable by performing a field study with observers. An additional advantage of this collection method is that it can provide more nuanced insights into the way the interactions unfold and thus provide more qualitative insight as well. Another way to circumvent the limitations is to reinstall the smart sensors in such a way that the images that are created are properly connected, through calibration and validation of the trajectories. This connection of multiple frames can then furthermore be used to detect cyclist movements further down their path, if installed with sufficient distance coverage.

8.3. Limitations of the Method

During this investigation, some limitations were discovered in the method, which are outlined in this section.

A first limitation in the method is in the use of fitted distributions for classification, especially when the fit does not align well with the empirical data. The distribution fitted to the speed data may not exactly reflect the observed speeds. In the case of the speed distributions of all trajectory points, the fitted distribution of the cyclist's speed appears to underestimate the average speed compared to what a visual inspection of the histogram would suggest (Figure 5.4). As a result, the threshold used to distinguish between pedestrians and cyclists may be slightly lower than ideal, potentially leading to some misclassification.

A similar fitting problem seems to arise for the deviating functions, which are expressed by a log-normal distribution. The distribution function mostly does not perfectly align with the reference data (Figures 6.8a and 6.8b), overestimating the deviation at the higher end and therefore underestimating the relative difference that is used to determine the threshold. This could have been prevented by creating more visually fitting functions, rather than relying on the fitting functions that are generated based on the dataset with pre-written functions in Python.

Pedestrians and cyclists are now distinguished mostly based on their speed, where the 95 percent interval as a consequence has that the speeds that are considered are only from pedestrians that do not walk too fast and cyclists that do not go too slow. The group of pedestrians going fast could however be pivotal when studying the different behaviours of pedestrians. A fast walking group might have had an influence on the number of pedestrians crossing before the cyclist, because the expectancy is that this group is tempted to cross before the cyclist more often than to decrease their speed and go behind it. A consideration that can be made is to decrease the certainty interval determining the type of mode based on speed. To compensate for this, other methods could be considered to distinguish cyclists from pedestrians, for example by introducing additional factors such as acceleration or rotation. Another method to distinguish cyclists from pedestrians is by application of machine learning, which has been done in the study of W. Wang et al. (2024), that could automatically learn motion patterns and distinguish between the two based on several more features. However, such an approach requires an existing dataset in which each trajectory is reliably labelled as either pedestrian or cyclist. Since this is not available in this study, such ground truth would need to be manually established, for example by visually inspecting and labelling each trajectory. This would enable the application of supervised learning, though it comes with other challenges, such as subjectivity on the distinction between a cyclist and pedestrian and it is likely labour-intensive.

Only a distinction is made on individual interactions taking place within 5 seconds, but say a pedestrian encounters two consecutive cyclists in a row and it is not accepting the gap in between, the first cyclist that passes might have left the conflicting point more than 5 seconds after the pedestrian has passed that point, wrongly stating that the pedestrian only crossed one trajectory. This could have been prevented by calculating the interactions within a longer range, yet this would increase the computational time.

Also, crossings should happen with at least a 30 degree angle, which is very locally determined, but if a cyclist crosses with a pedestrian under a smaller angle, this interaction is not registered, thereby sometimes filtering out some specific movements that are common for these specific interactions. These issues could be prevented by performing iterative steps in which the founded crossing trajectories are reconsidered in the context of the full dataset, to detect whether no other traffic participant could have been of influence on the particular crossing scenario. Another way to approach the interaction patterns is by only considering the origins and destinations and the angles that these crossing paths likely make, while not fixating on a very local angle.

9
Conclusion

This study investigated how pedestrians adjust their movements when crossing bike paths with oncoming cyclists. The analysis focused on three types of behavioural responses: stopping, deviating, and a more subtle form of gradual yielding, where pedestrians adjust their speed or direction. The aim was to answer the main research question:

In what way do pedestrians change their movements when approaching and crossing bike paths with oncoming cyclists?

A total of 4780 individual pedestrian-cyclist crossings have been analysed spread across two active mode intersections at the TU Delft campus. This study derived movement indicators to quantify these behavioural changes. The key findings are presented below, structured around the three main types of behaviour observed. These findings serve to summarise the outcomes presented throughout the report and highlight the most important insights in relation to the research question. The chapter concludes by answering the main research question and then provides recommendations for future research and practice.

9.1. Stopping Behaviour

Pedestrians perform different behaviours depending on the type of conflict with the cyclist. For the conflicts between a single pedestrian and cyclist, only a part of the pedestrians come to a stop. If the conflict is happening within 0 to 3 seconds, the average stopping percentage of pedestrians crossing behind a cyclist is 32.8% for the crossing scenario with the most interactions, where the pedestrians cross the minor road of a T-intersection, referred to as the east-west connection at the Lorentzweg crossing. For the other crossing scenarios this percentage is 51.9%, 32.2%, and 47.6% at the Lorentzweg crossing for pedestrians going from south to west, and at the Mekelweg crossing going from south to north and north to south respectively. Especially, the stopping percentage difference between the south-west crossing pedestrians and the straight crossing scenarios is large. The stopping percentage is likely higher because the pedestrians are making a sharp turn, therefore being more often registered as a stopping trajectory. This is confirmed when comparing the stopping percentages with the reference group of pedestrians that do not cross, which is 34.4% and is much higher compared to the other scenarios which are around 10%.

The stopping percentage for crossing pedestrians at the Lorentzweg crossing going from east to west shows a gradual decline within a PET range of 0 to 3 seconds, where the closest interactions within 1 second of each other have a stopping percentage of around 45% and the interactions that have a timing difference of around 3 seconds have a stopping percentage of approximately 20%.

If the pedestrian makes the decision to stop and wait for a passing cyclist, the stopping distance that the pedestrian tends to keep in a conflict between 0 to 3 seconds towards the crossing point fluctuates around 3 meters for the pedestrians at the east-west connection of the Lorentzweg crossing. The standard deviation of this distance is a bit more than 1 meter.

Depending on the type of origin-destination pair, the stopping distances for the pedestrians in the other crossing scenarios differ slightly from the east-west connection. The south-west crossing pedestrians at the Lorentzweg intersection show a slightly lower mean stopping distance to the conflict point compared to the east-west crossing pedestrians. This difference is likely explained by the geometry of the intersection. Pedestrians crossing from the east to the west approach the conflict point along a path that is perpendicular to the bike path, so their stopping distance is measured directly along that axis. In contrast, pedestrians crossing from the south to the west approach the same conflict point at an angle, meaning their distance to the point of conflict is measured diagonally. As a result, this straight line distance is smaller.

For the crossing scenarios at the Mekelweg (south-north and north-south) the mean of the stopping distance is slightly higher, while the standard deviation tends to be lower. This is possibly due to the width of the bike path and the different destination options at the Lorentzweg crossing in comparison to the Mekelweg crossing. The width of the bike path that is crossed at the Mekelweg is larger than the Lorentzweg crossing (approximately 4 meters and 2.5 meters respectively), so cyclists approaching from the right, from the perspective of the pedestrian, have a broader strip of bike path in between the pedestrian and cyclist. The pedestrian is likely to wait somewhere at the edge of the footpath for the cyclist to cross, as was seen in the visualisations of the speed-trajectory plots (Figure 5.12 and Appendix G). Therefore, the distance that the pedestrian keeps towards this crossing cyclist is on average larger. At the Lorentzweg, the width of the path is smaller, yet the multiple directions that cyclists have as origin or destination possibly create the larger spread of stopping distances that pedestrians tend to keep, with cyclists possibly turning away or towards the pedestrian.

Figure 9.1 summarises the percentages of stopping trajectories and distances that pedestrians on average tend to keep towards the conflict point for the crossing scenario where the pedestrian crosses after the cyclist within a PET of 0 to 3 seconds.

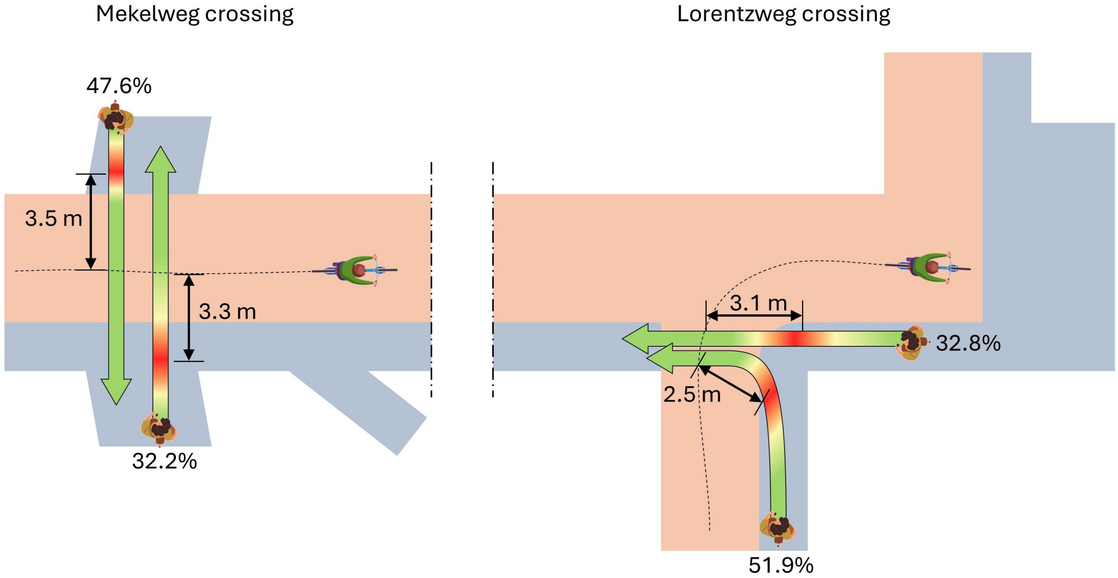


Figure 9.1: All stopping percentages and mean stopping distances of the different analysed crossing scenarios.

9.2. Gradual yielding

Other ways that pedestrians and cyclists tend to cross with each other are in the form of a more gradual self-organising process. Since stopping does usually not happen more than half of the times a pedestrian crosses behind the cyclist, it suggests that it happens sufficiently frequently that pedestrians either slow down slightly or that they maintain their speed on a constant level to eventually cross behind the cyclist.

This is supported when analysing the development of the predicted post-encroachment time (PET) for the Lorentzweg crossing for pedestrians crossing from east to west where the development of this value for many trajectories remains constant throughout time, suggesting that pedestrians maintain their speed in the cases where no collision or near collision course is predicted.

The predicted PET furthermore provides the insight that in the cases that a pedestrian and cyclist are predicted to get in very close range of each other (PET between 0 and 1 seconds), the pedestrian is more likely to yield to the cyclist, thereby crossing the cyclist from behind. This either happens by stopping, thereby causing a severe drop in predicted PET, or by the pedestrian slowly changing its movement direction, speed or both to end up behind the cyclist, thereby observing a more gradual decline of the predicted PET.

9.3. Deviating Behaviour

Pedestrians can deviate from their intended straight path by deviating in a certain direction to prevent a collision with a cyclist. Deviation happens only significantly often in the cases that the pedestrian crosses after the cyclist. The data even suggests that pedestrians that cross before the cyclist maintain a straighter path than pedestrians that do not cross with cyclists. At the locations with fewer total interactions, the Mekelweg crossing from south to north and north to south, the distinction between crossing and non-crossing trajectories was less pronounced.

The cases that deviation happens most often is when the cyclist approaches the pedestrian from a sideways branch of the crossing. If the cyclist approaches from a branch behind or in front of the pedestrian, deviation occurs less often and with less magnitude.

Deviation to a significant amount happens in about 20% of the cases where the cyclist approaches the pedestrian from a sideways branch and the cyclist crosses first. The pedestrian in this case often tends to deviate in the direction that the cyclist is coming from to be able to cross earlier.

9.4. Answer to the Main Research Question

Previously listed findings ultimately lead to the answer to the main research question: *In what way do pedestrians change their movements when approaching and crossing bike paths with oncoming cyclists?*

Based on the findings, pedestrians adapt their movements in several ways when encountering oncoming cyclists. The pedestrian either comes to a stop, slows down, deviates from its intended path or uses a combination of these movements in the case that the pedestrian decides to yield to the cyclist. This movement is often projected when the cyclist and pedestrian are predicted to get in very close range of each other and it especially occurs when the cyclist approaches from a sideways branch with respect to the approach of the pedestrian. While stopping is more common when the post-encroachment time becomes small, deviation is most often used in a situation where a slight adaptations are necessary to prevent a collision and is only done when more space is available. In the cases where the pedestrian has a reasonably safe opportunity to cross before the cyclist, it does so with determination by keeping a straight path and almost no stopping.

The findings of this study highlight the anticipatory nature of pedestrian behaviour in encounters with cyclists, particularly in environments where no clear right-of-way is established. This has implications for how active-mode infrastructure is designed and regulated. Despite the legal equality of both pedestrians and cyclists as vulnerable road users, the behaviour observed suggests that pedestrians are often cautious towards cyclists, acknowledging the speed advantages of cyclists. This shows that there may be a gap between how the law treats pedestrians and cyclists equally, and how they actually behave in practice. For urban designers and policymakers, understanding the patterns of behaviour is essential in creating crossings that accommodate safe interaction. Infrastructure should not only separate modes where needed but also provide clear visual cues that induce anticipation, which is especially valuable in an urban environment where a modal shift to active mobility is emerging.

9.5. Recommendations

A proposal in MassMotion has been created that could potentially translate the stopping and deviating behaviour that has been found in the data. However, the current implementation has only been partially calibrated for a single scenario and should be viewed as an illustrative example rather than a validated model. The parameters that govern the speed reduction and directional changes, which regularly occurs in the analysed data, need to be tuned to fit multiple scenarios. It is therefore proposed to perform further research into the calibration and validation of the proposed changes to the formulation to actually capture the effect that it has on the simulation performance. Calibration involves adjusting the model parameters to reflect the observed behavioural patterns. Validation tests whether the model can accurately reproduce these behaviours in different or unseen scenarios. The calibration can be based on the statistics of the data used in this research, which are the stopping percentages, preferred stopping distances and descriptive statistics of deviations and should be compared with these statistics to make an overall comparison. A stochastic approach is needed to capture the natural variation in pedestrian behaviour, such as whether an individual chooses to stop, slow down, or deviate, based on probabilities derived from the observed data. This stochasticity can furthermore be applied using the statistics of the normally distributed stopping distances that pedestrians prefer. Validation could be performed by applying the calibrated model to a separate dataset and comparing the simulated outcomes, such as stopping quantities, deviating patterns, or overall flow, to those observed in reality.

Furthermore, from the data it was observed that there is a large group of people that actually tend to cross in front of the cyclist, which show less stopping and deviation compared to non-crossing pedestrians. For further testing of the MassMotion model, the behaviour of pedestrians crossing after the cyclist should potentially be distinguished from the pedestrians crossing in front of the cyclist. This may require the development and testing of a separate behavioural formulation which is most likely based on the value of the predicted PET. When the predicted PET is sufficiently high, in the sense that the pedestrian is predicted to cross first, the model should distinguish between the pedestrians that are tempted to cross before the cyclist and the ones that still decide to wait and cross after the cyclist.

Further research should not only be done based on the findings of this study, it could explore other interaction types as well. This study specifically analysed individual pedestrian-cyclist interactions, whereas Oasys could also encourage studies on mutual behaviours of groups of cyclists, or the interaction these groups perform towards pedestrians. Behaviour in groups can be significantly different from individual behaviours, because of the collective expectations the people in these groups have towards each other. Massmotion would benefit from these different behavioural studies for creating a more inclusive model.

Besides research in the modelling, there are also other perspectives possible when analysing the behaviour of a pedestrian-cyclist interaction. While this study focussed on pedestrian responses, future research could explore the behaviour of cyclists during crossing scenarios by analysing their movements on the bike path further before reaching the crossing, because the findings of the data suggested that cyclists might make their decisions further away from the conflicting point. As mentioned in the Discussion, this could be done by using multiple smart sensors, which are properly calibrated to each other, across a cycling path to detect the potential anticipatory movements of cyclists towards pedestrians.

Future research could further explore shared decision-making processes between the two modes. Particularly, interactions could be studied more in depth by focussing on verbal and non-verbal negotiations that pedestrians and cyclists perform to each other to grant or take their right of way. This is more likely to be properly detected by observers, that have the ability to classify the interactive behaviours based on nuanced communication between cyclists and pedestrians.

Another follow-up study could use the findings of stopping and deviating behaviour to do a design-oriented research of bike path crossings, not only considering the layout of the crossing itself, but also taking into account the effect the placement of a crossing has on surrounding areas. This study could explore how different crossing configurations influence pedestrian comfort and safety.

Oasys could develop their software in such a way that the behavioural findings of this study could find an application for the users. First of all, the cyclist agent can be implemented within the software, by providing a realised cyclist agent, with corresponding speeds and movement restrictions, that does not require steps of adapting from a pedestrian agent. Secondly, if the cyclist were to be implemented correctly, a modification of the social forces is likely to be necessary by implementing cyclist specific behavioural rules for both cyclists themselves and pedestrian responses towards that cyclist. Lastly, options could be added to define probabilities or distributions for agent decisions under certain conditions (e.g. a PET less than 3 seconds leading to a 30% chance of stopping).

On another note, Oasys could refine their software by enabling the use of trajectory data. This could be done by implementing trajectory data, similar to the data used in this study, to enable calibration of different parameters in the model when existing interactions need to be improved or new interactions need to be tested.

Urban planners, designers and policymakers should be aware of the interactive behaviours observed at active mode crossings, as highlighted in this study. Municipalities could use this behavioural analysis to better understand pedestrian-cyclist interaction and inform the design of active mode crossings. The results show that pedestrians more often yield, which implies the need for designs that could equalise both modes, such as surface markings that prioritise pedestrian awareness or signage clarifying expectations. on the contrary, if this precautionary behaviour is actually desired in certain circumstances, crossings should be designed to provide good sight lines and predictable cyclist paths to enable these behaviours safely. On a larger scale, practitioners should be aware of the impact the integration of a cycling path has on the pedestrian responses and on the potential accessibility issues it causes. Urban environments where pedestrians and cyclists coexist should be designed with careful consideration of the behavioural dynamics to ensure both safety and inclusiveness.

References

Afghari, A. P., Ismail, K., Saunier, N., Sharma, A., & Miranda-Moreno, L. (2014). Pedestrian-cyclist interactions at bus stops along segregated bike paths: A case study of montreal. *Transportation Research Board (USA) Annual Meeting*, 1–1. https://eprints.qut.edu.au/79841/

ARAG. (2025). *Wet beschermt voetganger en fietser bij ongeval*. Retrieved July 4, 2025, from https://www.arag.nl/reizen-en-verkeer/voertuigschade/ongeval-automobilist-met-fietser-of-voetganger/#:~:text=Wettelijk%20is%20de%20kwetsbare%20voetganger%20en%20fietser%20beschermd,aan%20hebt%20gedaan%20om%20de%20aanrijding%20te%20voorkomen.

Ather, S. H. (2022). *How to calculate rmsd*. Retrieved March 24, 2022, from https://www.sciencing.com/calculate-rmsd-5146965/

Beitel, D., Stipancic, J., Manaugh, K., & Miranda-Moreno, L. (2018). Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. *Transportation Research Part D: Transport and Environment*, *65*, 710–724. https://doi.org/https://doi.org/10.1016/j.trd.2018.10.001

Brorson, S. (n.d.). *Forward euler method*. https://math.libretexts.org/Bookshelves/Differential_Equations/Numerically_Solving_Ordinary_Differential_Equations_%28Brorson%29/01%3A_Chapters/1.02%3A_Forward_Euler_method

Chandra, S., & Bharti, A. K. (2013). Speed distribution curves for pedestrians during walking and crossing [2nd Conference of Transportation Research Group of India (2nd CTRG)]. *Procedia - Social and Behavioral Sciences*, *104*, 660–667. https://doi.org/https://doi.org/10.1016/j.sbspro.2013.11.160

Daamen, W., & Hoogendoorn, S. P. (2007). Free speed distributions — based on empirical data in different traffic conditions (Peter, K. Hermann, S. M. W. Nathalie, & Gattermann, Eds.). *Pedestrian and Evacuation Dynamics 2005*, 13–25.

Data-Insights. (2025). *Gender diversity dashboard*. Retrieved June 1, 2025, from https://www.tudelft.nl/en/about-tu-delft/strategy/diversity-inclusion/facts-and-figures-1

Dias, C., Nishiuchi, H., Hyoudo, S., & Todoroki, T. (2018). Simulating interactions between pedestrians, segway riders and cyclists in shared spaces using social force model [International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)Emerging Transport Technologies for Next Generation Mobility]. *Transportation Research Procedia*, *34*, 91–98. https://doi.org/https://doi.org/10.1016/j.trpro.2018.11.018

Donaire-Gonzalez, D., de Nazelle, A., Cole-Hunter, T., Curto, A., Rodriguez, D. A., Mendez, M. A., Garcia-Aymerich, J., Basagaña, X., Ambros, A., Jerrett, M., & Nieuwenhuijsen, M. J. (2015). The added benefit of bicycle commuting on the regular amount of physical activity performed. *American Journal of Preventive Medicine*, *49*, 842–849. https://doi.org/https://doi.org/10.1016/j.amepre.2015.03.036

Duives, D. C., Daamen, W., & Hoogendoorn, S. P. (2013). State-of-the-art crowd motion simulation models. *Transportation Research Part C: Emerging Technologies*, *37*, 193–209. https://doi.org/https://doi.org/10.1016/j.trc.2013.02.005

Eriksson, J., Forsman, Å., Niska, A., Gustafsson, S., & Sörensen, G. (2019). An analysis of cyclists' speed at combined pedestrian and cycle paths. *Traffic Injury Prevention*, *20*, 56–61. https://doi.org/10.1080/15389588.2019.1658083

GeeksforGeeks. (2024a). *Distance of a point from a line*. Retrieved July 19, 2024, from https://www.geeksforgeeks.org/distance-of-a-point-from-a-line/

GeeksforGeeks. (2024b). *Sangle between two vectors formula*. Retrieved April 10, 2024, from https://www.geeksforgeeks.org/angle-between-two-vectors-formula/

Helbing, D., Farkas, I., Molnar, P., & Vicsek, T. (2002, January). Simulation of pedestrian crowds in normal and evacuation situations. Springer.

Helbing, D., & Molnar, P. (1995, January). Social force model for pedestrian dynamics. Physical Review E.
Huyghebaert, P. (2021). *Houden fietsers genoeg rekening met voetgangers?* Retrieved July 1, 2025, from https://www.vrt.be/vrtnws/nl/2021/09/21/broos-en-fietersbond-over-column/
Klüpfel, H. (2014). Large scale multi-modal simulation of pedestrian traffic [The Conference on Pedestrian and Evacuation Dynamics 2014 (PED 2014), 22-24 October 2014, Delft, The Netherlands]. *Transportation Research Procedia*, *2*, 446–451. https://doi.org/https://doi.org/10.1016/j.trpro.2014.09.058
KorteAntwoorden. (2016). *Hoe oud zijn studenten in nederland?* Retrieved August 24, 2022, from https://korteantwoorden.com/hoe-oud-zijn-studenten-in-nederland/
Letsel Hulp Service. (n.d.). *Aanrijding tussen fietser en voetganger: Wie is aansprakelijk?* Retrieved July 4, 2025, from https://letselhulpservice.nl/aangereden-als-fietser/veelgestelde-vragen/aanrijding-fietser-voetganger-wie-aansprakelijk
Letsel & Schade. (2025). *Voetganger aangereden op fietspad*. Retrieved July 1, 2025, from https://letsel.info/voetganger-aangereden-op-fietspad/
Li, M., Shi, F., & Chen, D. (2011). Analyze bicycle-car mixed flow by social force model for collision risk evaluation. *3rd International Conference on Road Safety and Simulation*, 1–22.
Li, Y., Ni, Y., & Sun, J. (2021). A modified social force model for high-density through bicycle flow at mixed-traffic intersections. *Simulation Modelling Practice and Theory*, *108*, 102265. https://doi.org/https://doi.org/10.1016/j.simpat.2020.102265
LIANG, X., MAO, B., & XU, Q. (2012). Psychological-physical force model for bicycle dynamics. *Journal of Transportation Systems Engineering and Information Technology*, *12*, 91–97. https://doi.org/https://doi.org/10.1016/S1570-6672(11)60197-9
Murtagh, E., Mair, J., Aguiar, E., Tudor-Locke, C., & Murphy, M. (2021). Outdoor walking speeds of apparently healthy adults: A systematic review and meta□analysis. *Sports Medicine*, *51*, 1–31. https://doi.org/10.1007/s40279-020-01351-3
Nateghinia, E., Beitel, D., Lesani, A., & Miranda-Moreno, L. F. (2024). A lidar-based methodology for monitoring and collecting microscopic bicycle flow parameters on bicycle facilities. *Transportation*, *51*, 129–153. https://doi.org/10.1007/s11116-022-10322-8
Ng, K. M., Yuen, C. W., Onn, C. C., & Ibrahim, N. I. (2024). Urban mobility mode shift to active transport: Sociodemographic dependency and potential greenhouse gas emission reduction. *Sage Open*, *14*, 21582440241228644. https://doi.org/10.1177/21582440241228644
Parisi, D. R., Negri, P. A., & Bruno, L. (2016). Experimental characterization of collision avoidance in pedestrian dynamics. *Phys. Rev. E*, *94*, 022318. https://doi.org/10.1103/PhysRevE.94.022318
Pisoni, E., Christidis, P., & Cawood, E. N. (2022). Active mobility versus motorized transport? user choices and benefits for the society. *Science of The Total Environment*, *806*, 150627. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.150627
Reference. (2025). *What is the average size of the human head?* Retrieved May 20, 2025, from https://www.reference.com/science-technology/average-size-human-head-62364d028e431bf3
Schleinitz, K., Petzoldt, T., Franke-Bartholdt, L., Krems, J., & Gehlert, T. (2017). The german naturalistic cycling study – comparing cycling speed of riders of different e-bikes and conventional bicycles. *Safety Science*, *92*, 290–297. https://doi.org/https://doi.org/10.1016/j.ssci.2015.07.027
Shahhoseini, Z., & Sarvi, M. (2019). Pedestrian crowd flows in shared spaces: Investigating the impact of geometry based on micro and macro scale measures. *Transportation Research Part B: Methodological*, *122*, 57–87. https://doi.org/https://doi.org/10.1016/j.trb.2019.01.019
Smyth, B. (2018). Fast starters and slow finishers: A large-scale data analysis of pacing at the beginning and end of the marathon for recreational runners. *Journal of Sports Analytics*, *4*, 229–242. https://doi.org/10.3233/JSA-170205
sqlpey. (2024). *Solved: Top 12 methods to check if two segments intersect*. Retrieved November 6, 2024, from https://sqlpey.com/python/solved-top-12-methods-to-check-if-two-segments-intersect/
Strangman, G. (2002). *Normaltest*. Retrieved April 16, 2025, from https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html
Strategic Development. (2015). *Statistieken 2015* (tech. rep.). TU Delft.

Tageldin, A., & Sayed, T. (2016). Developing evasive action-based indicators for identifying pedestrian conflicts in less organized traffic environments. *Journal of Advanced Transportation*, *50*, 1193–1208. https://doi.org/https://doi.org/10.1002/atr.1397

TU Delft. (2024). *Facts and figures*. Retrieved June 1, 2025, from https://www.tudelft.nl/en/about-tu-delft/organisation/facts-and-figures

voetgangersvereniging Nederland. (2021). *Dilemma*. Retrieved July 4, 2025, from https://voetgangersverenigingnederland.nl/dilemma/

Wang, W., Zhou, H., Lo, J. T. Y., Lo, S. M., & Wang, Y. (2024). A modified social force model for pedestrian-bicycle mixed flows and its application on evaluating the conflict risk in shared roads. *Physica A: Statistical Mechanics and its Applications*, *643*, 129788. https://doi.org/https://doi.org/10.1016/j.physa.2024.129788

Wang, Y., Jia, Y., Chen, W., Wang, T., & Zhang, A. (2024). Examining safe spaces for pedestrians and e-bicyclists at urban crosswalks: An analysis based on drone-captured video. *Accident Analysis & Prevention*, *194*, 107365. https://doi.org/https://doi.org/10.1016/j.aap.2023.107365

Wassink, J. (2023). *Tram line 19 is finally coming… but first the rails will be removed*. Retrieved April 21, 2023, from https://delta.tudelft.nl/en/article/tram-line-19-finally-coming-first-rails-will-be-removed#:~:text=Twenty%20years%20after%20the%20decision%20to%20extend%20tram,demolition%20took%20place%20earlier.%20%28Photo%3A%20M.%20van%20Bekkum%29

Wei, L., Li, Z., Gong, J., Gong, C., & Li, J. (2021). Autonomous driving strategies at intersections: Scenarios, state-of-the-art, and future outlooks. https://doi.org/10.1109/ITSC48978.2021.9564518

Xi, H., Son, Y.-J., & Lee, S. (2010). An integrated pedestrian behavior model based on extended decision field theory and social force model. *Proceedings of the 2010 Winter Simulation Conference*, 824–836. https://doi.org/10.1109/WSC.2010.5679108

Yan, X., Chen, J., Bai, H., Wang, T., & Yang, Z. (2020). Influence factor analysis of bicycle free-flow speed for determining the design speeds of separated bicycle lanes. *Information*, *11*. https://doi.org/10.3390/info11100459

Yuan, Y., Bernat, G.-R., van Oijen, T. P., Daamen, W., Seer, S., & Hoogendoorn, S. P. (2019). Social force model describing pedestrian and cyclist behaviour in shared spaces (S. H. Hamdar, Ed.). *Traffic and Granular Flow ’17*, 477–486.

Zangenehpour, S., Strauss, J., Miranda-Moreno, L. F., & Saunier, N. (2016). Are signalized intersections with cycle tracks safer? a case–control study based on automated surrogate safety analysis using video data. *Accident Analysis & Prevention*, *86*, 161–172. https://doi.org/https://doi.org/10.1016/j.aap.2015.10.025

Zhao, J., Malenje, J. O., Tang, Y., & Han, Y. (2019). Gap acceptance probability model for pedestrians at unsignalized mid-block crosswalks based on logistic regression. *Accident Analysis I& Prevention*, *129*, 76–83. https://doi.org/https://doi.org/10.1016/j.aap.2019.05.012

A

Source Code

This appendix contains the most important lines of code for enriching the data.

```
"""
Calculating the time in seconds from the start.
"""
import pandas as pd

# Ensure the datetime column is in datetime format
df['datetime'] = pd.to_datetime(df['datetime'])

# Define the start time
start_time = df['datetime'].min()

# Compute the relative time in seconds
df['relative_time'] = (df['datetime'] - start_time).dt.total_seconds()
```

For speed, only the central difference method is shown. The forward difference method is very similar.

```
"""
Calculation for deriving the speed at the Lorentzweg crossing, using the central difference method
"""
import pandas as pd
import numpy as np
from tqdm import tqdm

def calculate_speed(p1, p2, t1, t2):
    # Calculates the speed between two points and their respective time
    p1 = np.array(p1)
    p2 = np.array(p2)
    time_difference = t2 - t1
    distance = np.linalg.norm(p1 - p2)
    return distance / time_difference if time_difference > 0 else np.nan

def determine_speed(df):
    # Group the dataframe by trace_id
    groups = df.groupby('trace_id')

    # Create a new column for speed
    df['speed'] = np.nan

    # Initialise a progress bar
    with tqdm(total=len(groups), desc="Processing␣Groups") as pbar:
        # Iterate through each group
        for trace_id, group in groups:
            # Sort the group by relative_time to ensure proper calculation order
            group = group.sort_values(by='relative_time')

            # Calculate the speed for each row (except the last row)
```

```
            for j in range(1, len(group) - 1):
                p1 = (group.iloc[j - 1]['x'], group.iloc[j - 1]['y'])
                p2 = (group.iloc[j + 1]['x'], group.iloc[j + 1]['y'])
                t1 = group.iloc[j - 1]['relative_time']
                t2 = group.iloc[j + 1]['relative_time']

                # Calculate speed and assign to the dataframe
                speed = calculate_speed(p1, p2, t1, t2)
                df.at[group.index[j], 'speed'] = speed

            # Update progress bar
            pbar.update(1)

    return df

# Example usage
df_speed = determine_speed(df)
```

```
"""
Determining the probability of a trajectory belongs to a cyclist or a pedestrian based on the average and maximum speed. Also calculates the probability of a trajectory belonging to a certain mode based on the pdf functions of both modes.
"""
import pandas as pd
import numpy as np
from tqdm import tqdm
from sklearn.mixture import GaussianMixture

intersection_speed = 1.9677 # [m/s]
minimal_maximum_speed_cyclist = 2.5 # [m/s]

# Fit a Gaussian Mixture Model with 2 components
gmm = GaussianMixture(n_components=2, random_state=1)

def is_it_cyclist_or_pedestrian(df, intersection_speed, minimal_maximum_speed_cyclist):
    # Create a list to store the mode type classification for each trace_id
    mode_classifications = []
    probability_of_mode = []

    # Group the dataframe by trace_id
    groups = list(df.groupby('trace_id'))

    # Initialise a progress bar
    with tqdm(total=len(groups), desc="Processing groups") as pbar:
        # Iterate through each group
        for trace_id, group in groups:
            group = group.reset_index()
            average_speed = group['speed'].mean()
            maximum_speed = group['speed'].max()

            # Determine if the trace_id corresponds to a cyclist or pedestrian
            if average_speed > intersection_speed or maximum_speed > minimal_maximum_speed_cyclist:
                mode_type = 'cyclist'
                pdf = (gmm.weights_[1] * (1 / (stds[1] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[1]) / stds[1])**2))) / (
                        gmm.weights_[1] * (1 / (stds[1] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[1]) / stds[1])**2)) + (
                        gmm.weights_[0] * (1 / (stds[0] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[0]) / stds[0])**2))
                        ))
            elif average_speed <= intersection_speed and average_speed > 0.77:
                mode_type = 'pedestrian'
                pdf = (gmm.weights_[0] * (1 / (stds[0] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[0]) / stds[0])**2))) / (
                        gmm.weights_[1] * (1 / (stds[1] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[1]) / stds[1])**2)) + (
                        gmm.weights_[0] * (1 / (stds[0] * np.sqrt(2 * np.pi)) * np.exp(-0.5 * ((average_speed - means[0]) / stds[0])**2))
                        ))
            elif average_speed <= 0.77 and average_speed > 0:
```

```python
                mode_type = 'pedestrian'
                pdf = 1.0
            else:
                mode_type = float('NaN')
                pdf = 0


            # Append the result for this trace_id
            mode_classifications.append((trace_id, mode_type))
            probability_of_mode.append((trace_id, pdf))

            # Update progress bar
            pbar.update(1)

    # Convert the classifications to a DataFrame
    classification_df = pd.DataFrame(mode_classifications, columns=['trace_id', 'type_of_active_mode'])
    probability_df = pd.DataFrame(probability_of_mode, columns=['trace_id', 'probability_of_mode'])

    # Merge the classifications back into the original dataframe
    df = df.merge(classification_df, on='trace_id')
    df = df.merge(probability_df, on='trace_id')

    return df

df_type_active_mode = is_it_cyclist_or_pedestrian(df_speed, intersection_speed, minimal_maximum_speed_cyclist)
```

The filter for no encounters is applied to pedestrians and cyclists that enter and exit the area with a time margin of 5 seconds of no other traffic participants found in this area.

```python
"""
Filtering out the trajectories of pedestrians and cyclists that do not encounter any other traffic participant 5 s before they entered or 5 s after they have left the area.
"""
import pandas as pd

time_range = 5 # seconds
range_limit = 100

def filter_time_range_limited(df, time_range, range_limit=100):
    # Reset index for consistent indexing
    df = df.sort_values(by='relative_time').reset_index(drop=True)

    # Set to store trace_ids of trajectories to keep
    keep_trace_ids = set()

    with tqdm(total=len(df), desc="Processing trajectories") as pbar:

        for index, row in df.iterrows():
            # Get rows within the index range limit
            lower_bound = max(index - range_limit, 0)
            upper_bound = min(index + range_limit + 1, len(df))

            # Subset the dataframe to rows within the range limit
            nearby_rows = df.iloc[lower_bound:upper_bound]

            # Exclude rows with the same trace_id
            nearby_rows = nearby_rows[nearby_rows['trace_id'] != row['trace_id']]

            # Check if any nearby row meets the time condition
            if ((nearby_rows['relative_time'] - row['relative_time']).abs() <= time_range).any():
                keep_trace_ids.add(row['trace_id'])

            # Update progress bar
            pbar.update(1)

    # Filter the DataFrame to include only rows with matching trace_ids
    filtered_df = df[df['trace_id'].isin(keep_trace_ids)].reset_index(drop=True)
```

```

    return filtered_df

df_filter_1 = filter_time_range_limited(df_type_active_mode, time_range, range_limit)
```

Determining the location happens in a few steps: first the cycling and footpath need to be determined visually, then the functions are created that can test whether a point is inside the rectangle, lastly it is tested for all trajectories to what extent the trajectory is on the foot or cycling path.

```
"""
Code for determining the locations of foot and cycling paths via a plot.
"""

import pandas as pd
import numpy as np
import matplotlib.image as mpimg

# show only one of every 100 points
df_filter_2 = df_filter_1.iloc[::100, :]

fig, ax = plt.subplots(figsize=(16, 9))

# import image
img = mpimg.imread('Location_new_crossing.jpg')
ax.imshow(img, extent=[-7, 14.5, -6.5, 10], alpha = 0.8)
ax.scatter(df_filter_2['y'], -df_filter_2['x'], label = 'trajectory points', alpha=0.8)

# Introduce all corner points of the different rectangles
point_1 = (-6, 2.08)
width_1 = 17
height_1 = 4

point_2 = (-4.1, -7.5)
width_2 = 2.21
height_2 = 9.5

point_3 = (4.3, 5.3)
width_3 = 2.21
height_3 = 3

point_4 = (-5.5, 0.045)
width_4 = 16
height_4 = 2

point_5 = (-1.9, -7.7)
width_5 = 2
height_5 = 7.5

point_6 = (6.5, 5.2)
width_6 = 5
height_6 = 3

point_7 = (6.2, 1.2)
width_7 = 1.8
height_7 = 4

# Rectangles are slightly rotated.
angle_degrees = 4
angle_radians = angle_degrees * np.pi / 180

# Adding rectangles for the cycling path
rect1 = mpatches.Rectangle(point_1, width_1, height_1, angle=-angle_degrees, rotation_point='xy', label= 'cycling path', edgecolor='red', facecolor='red', alpha=0.2)
ax.add_patch(rect1)
rect2 = mpatches.Rectangle(point_2, width_2, height_2, angle=-angle_degrees, rotation_point='xy', edgecolor='red', facecolor='red', alpha=0.2)
ax.add_patch(rect2)
rect3 = mpatches.Rectangle(point_3, width_3, height_3, angle=-angle_degrees, rotation_point='xy', edgecolor='red', facecolor='red', alpha=0.2)
ax.add_patch(rect3)
```

```python

# Adding rectangles for the footpath
rect4 = mpatches.Rectangle(point_4, width_4, height_4, angle=-angle_degrees, rotation_point='xy', label= 'footpath', edgecolor='yellow', facecolor='yellow', alpha=0.2)
ax.add_patch(rect4)
rect5 = mpatches.Rectangle(point_5, width_5, height_5, angle=-angle_degrees, rotation_point='xy', edgecolor='yellow', facecolor='yellow', alpha=0.2)
ax.add_patch(rect5)
rect6 = mpatches.Rectangle(point_6, width_6, height_6, angle=-angle_degrees, rotation_point='xy', edgecolor='yellow', facecolor='yellow', alpha=0.2)
ax.add_patch(rect6)
rect7 = mpatches.Rectangle(point_7, width_7, height_7, angle=-angle_degrees, rotation_point='xy', edgecolor='yellow', facecolor='yellow', alpha=0.2)
ax.add_patch(rect7)

plt.legend(loc='best', bbox_to_anchor=(0.55, 0.2, 0.5, 0.5))
plt.axis('scaled')
plt.show()
```

```python
"""
Defining different vectors that provide the borders of the rectangles. Also creating functions that test whether a point is inside the foot or cycling path.
"""
import numpy as np

#direction vectors for determining the exact area
rotation_matrix = np.array([[0, -1],
                            [1, 0]]) # Left turn
first_vector = [1, np.tan(- angle_radians)]
second_vector = np.dot(rotation_matrix, first_vector)
third_vector = np.dot(rotation_matrix, second_vector)
fourth_vector = np.dot(rotation_matrix, third_vector)

cycling_paths_first_point = (point_1, point_2, point_3)
foot_paths_first_point = (point_4, point_5, point_6, point_7)
cycling_paths_second_point_x = (point_1[0] + width_1 * np.cos(angle_radians) + height_1 * np.sin(angle_radians),
                                point_2[0] + width_2 * np.cos(angle_radians) + height_2 * np.sin(angle_radians),
                                point_3[0] + width_3 * np.cos(angle_radians) + height_3 * np.sin(angle_radians))
foot_paths_second_point_x = (point_4[0] + width_4 * np.cos(angle_radians) + height_4 * np.sin(angle_radians),
                             point_5[0] + width_5 * np.cos(angle_radians) + height_5 * np.sin(angle_radians),
                             point_6[0] + width_6 * np.cos(angle_radians) + height_6 * np.sin(angle_radians),
                             point_7[0] + width_7 * np.cos(angle_radians) + height_7 * np.sin(angle_radians))
cycling_paths_second_point_y = (point_1[1] - width_1 * np.sin(angle_radians) + height_1 * np.cos(angle_radians),
                                point_2[1] - width_2 * np.sin(angle_radians) + height_2 * np.cos(angle_radians),
                                point_3[1] - width_3 * np.sin(angle_radians) + height_3 * np.cos(angle_radians))
foot_paths_second_point_y = (point_4[1] - width_4 * np.sin(angle_radians) + height_4 * np.cos(angle_radians),
                             point_5[1] - width_5 * np.sin(angle_radians) + height_5 * np.cos(angle_radians),
                             point_6[1] - width_6 * np.sin(angle_radians) + height_6 * np.cos(angle_radians),
                             point_7[1] - width_7 * np.sin(angle_radians) + height_7 * np.cos(angle_radians))

# Check whether the full trajectory is on the cycling path
def trajectory_on_cycling_path(point):
    trajectory_on_cycling_path = False
    point = np.array(point)
    for i, check_point in enumerate(cycling_paths_first_point):
        location_wrt_first_point = point - np.array(check_point)
        location_wrt_second_point = point - np.array([cycling_paths_second_point_x[i],
```

```
            cycling_paths_second_point_y[i]])
        if np.dot(location_wrt_first_point, first_vector) > 0 and np.dot(
            location_wrt_first_point, second_vector) > 0 and np.dot(location_wrt_second_point
            , third_vector) > 0 and np.dot(location_wrt_second_point, fourth_vector) > 0:
            trajectory_on_cycling_path = True
    return trajectory_on_cycling_path

# Check whether the full trajectory is on the footpath
def trajectory_on_foot_path(point):
    trajectory_on_foot_path = False
    point = np.array(point)
    for i, check_point in enumerate(foot_paths_first_point):
        location_wrt_first_point = point - np.array(check_point)
        location_wrt_second_point = point - np.array([foot_paths_second_point_x[i],
            foot_paths_second_point_y[i]])
        if np.dot(location_wrt_first_point, first_vector) > 0 and np.dot(
            location_wrt_first_point, second_vector) > 0 and np.dot(location_wrt_second_point
            , third_vector) > 0 and np.dot(location_wrt_second_point, fourth_vector) > 0:
            trajectory_on_foot_path = True
    return trajectory_on_foot_path
```

```
"""
Confirming the mode based on the location together with the speed.
"""
import pandas as pd
import numpy as np
from tqdm import tqdm

def cyclist_on_cycle_path(df):
    # Group the dataframe by trace_id
    groups = list(df.groupby('trace_id'))

    # Create new columns for mode classification and probability
    df['mode_location_based'] = 'unclear'
    df['probability_location'] = 0.0  # New column for probability based on location

    # Initialise the progress bar
    with tqdm(total=len(groups), desc="Processing␣Groups") as pbar:
        # Iterate through groups and check for the location of each point of trajectory
        for i, (current_trace_id, current_group) in enumerate(groups):
            current_group = current_group.reset_index(drop=True)
            predicted_mode = current_group.iloc[0]['type_of_active_mode']
            cyclist_list = []
            pedestrian_list = []

            for k in range(len(current_group)):
                p1 = (current_group.iloc[k]['y'], -current_group.iloc[k]['x'])
                cyclist_list.append(trajectory_on_cycling_path(p1))
                pedestrian_list.append(trajectory_on_foot_path(p1))

            total_count = len(cyclist_list)
            true_count_cyclist = sum(cyclist_list)
            true_count_pedestrian = sum(pedestrian_list)

            probability_cyclist = true_count_cyclist / total_count if total_count > 0 else 0
            probability_pedestrian = true_count_pedestrian / total_count if total_count > 0
                else 0

            # Confirming the mode based on location
            if all(cyclist_list) and predicted_mode == 'cyclist':  # All points indicate
                cyclist
                df.loc[df['trace_id'] == current_trace_id, 'mode_location_based'] = 'cyclist'
                df.loc[df['trace_id'] == current_trace_id, 'probability_location'] =
                    probability_cyclist
                df.loc[df['trace_id'] == current_trace_id, 'probability_of_mode'] = 1.0
            elif all(pedestrian_list) and predicted_mode == 'pedestrian':  # All points
                indicate pedestrian
                df.loc[df['trace_id'] == current_trace_id, 'mode_location_based'] = '
                    pedestrian'
                df.loc[df['trace_id'] == current_trace_id, 'probability_location'] =
                    probability_pedestrian
```

```
                df.loc[df['trace_id'] == current_trace_id, 'probability_of_mode'] = 1.0
            else:  # Mixed points
                df.loc[df['trace_id'] == current_trace_id, 'probability_location'] = probability_cyclist if predicted_mode == 'cyclist' else probability_pedestrian

            # Update the progress bar
            pbar.update(1)

    return df.reset_index(drop=True)

# Usage
df_filter_location = cyclist_on_cycle_path(df_filter_1)
```

Multiple functions are required for determining whether two segments of trajectories are crossing. After defining these functions, it can be checked for each segment whether they cross with a segment of another mode within a certain time range.

```
"""
Multiple functions to define whether two segments cross based on the cross product method
"""
import numpy as np

# Calculate the cross product
def cross_product(o, a, b):
    return (a[1] - o[1]) * (b[0] - o[0]) - (a[0] - o[0]) * (b[1] - o[1])

# Apply the cross product and return a true if the first and last trajectory point actually cross (for computational efficiency)
def end_segments_cross(p1, p2, p3, p4):
    if (cross_product(p1, p2, p3) * cross_product(p1, p2, p4) < 0 and
            cross_product(p3, p4, p1) * cross_product(p3, p4, p2) < 0):
        return True
    else:
        return False

# Apply the cross product and return a true if the segments actually cross under an angle of at least 30 degrees (1/6 pi)
def segments_cross(p1, p2, p3, p4):
    if (cross_product(p1, p2, p3) * cross_product(p1, p2, p4) < 0 and
            cross_product(p3, p4, p1) * cross_product(p3, p4, p2) < 0):
        # intersection angle
        dotproduct = abs((p1[0] - p2[0]) * (p3[0] - p4[0]) + (p1[1] - p2[1]) * (p3[1] - p4[1]))
        p1 = np.array(p1)
        p2 = np.array(p2)
        p3 = np.array(p3)
        p4 = np.array(p4)
        combined_lengths = np.linalg.norm(p1 - p2) * np.linalg.norm(p3 - p4)
        angle = np.arccos(dotproduct / combined_lengths)
        return angle > 1/6 * np.pi
    else:
        return False

# Returns the intersection point based on interpolation
def intersection_point(p1, p2, p3, p4):
    a1 = (p1[1] - p2[1]) / (p1[0] - p2[0])
    b1 = p1[1] - a1 * p1[0]
    a3 = (p3[1] - p4[1]) / (p3[0] - p4[0])
    b3 = p3[1] - a3 * p3[0]
    x = (b3 - b1) / (a1 - a3)
    y = a1 * x + b1
    return x, y

# Calculates the post-encroachment time based on interpolation
def what_is_the_post_encroachment_time(t1, t2, t3, t4, p1, p2, p3, p4, intersection_x, intersection_y, current_mode, next_mode):
    #interpolate the time for both modes to reach the intersection point
    time_current_mode = t1 + (t2 - t1) * np.linalg.norm([intersection_x, intersection_y] - np.array(p1)) / np.linalg.norm(
```

```python
                                                        np.array(p2) - np.array(p1))
    time_next_mode = t3 + (t4 - t3) * np.linalg.norm([intersection_x, intersection_y] - np.array(p3)) / np.linalg.norm(
                                                        np.array(p4) - np.array(p3))
    # Calculate the post-encroachment time
    if time_current_mode < time_next_mode:
        first = current_mode
        post_encroachment_time = time_next_mode - time_current_mode
    else:
        first = next_mode
        post_encroachment_time = time_current_mode - time_next_mode
    return first, post_encroachment_time, time_current_mode, time_next_mode

# Calculates whether there is even a possibility for two trajectories to overlap by looking at the borders of the 'boxes' that the trajectories are in
def boxes_overlap(bbox1, bbox2):
    x_min1, y_min1, x_max1, y_max1 = bbox1
    x_min2, y_min2, x_max2, y_max2 = bbox2

    # Check if one box is completely to the left of the other
    if x_max1 < x_min2 or x_max2 < x_min1:
        return False

    # Check if one box is completely above the other
    if y_max1 < y_min2 or y_max2 < y_min1:
        return False

    return True
```

```python
"""
Applies all previous functions to determine whether two trajectories are crossing
"""
import pandas as pd
import numpy as np
from tqdm import tqdm

max_groups_to_check = 10
max_time_diff= 5
probability_threshold = 0.95

def keep_intersecting_ped_cyc(df, max_groups_to_check=10, max_time_diff=5, probability_threshold=0.95):
    groups = list(df.groupby('trace_id'))
    processed_trace_ids = set()

    # Initialize storage for crossing data
    crossing_data = {
        'crossing_trace_ids': {},
        'post_encroachment_time': {},
        'first_mode': {},
        'intersection_x': {},
        'intersection_y': {},
        'time_at_crosspoint': {}
    }

    for trace_id in df['trace_id'].unique():
        for key in crossing_data:
            crossing_data[key][trace_id] = []

    with tqdm(total=len(groups), desc="Processing Groups") as pbar:
        for i, (current_trace_id, current_group) in enumerate(groups):
            current_group = current_group.reset_index(drop=True)
            current_mode = current_group['type_of_active_mode'].iloc[0]
            if current_group['probability_of_mode'].iloc[0] < probability_threshold:
                continue

            for j in range(i + 1, min(i + 1 + max_groups_to_check, len(groups))):
                next_trace_id, next_group = groups[j]
                next_group = next_group.reset_index(drop=True)
                next_mode = next_group['type_of_active_mode'].iloc[0]

```

```
                if {'cyclist', 'pedestrian'} != {current_mode, next_mode}:
                    continue
                if next_group['probability_of_mode'].iloc[0] < probability_threshold:
                    continue

                current_bbox = (current_group['x'].min(), current_group['y'].min(),
                                current_group['x'].max(), current_group['y'].max())
                next_bbox = (next_group['x'].min(), next_group['y'].min(),
                             next_group['x'].max(), next_group['y'].max())

                if not boxes_overlap(current_bbox, next_bbox):
                    continue

                beginpoint_1 = (current_group.iloc[0]['x'], current_group.iloc[0]['y'])
                endpoint_1 = (current_group.iloc[-1]['x'], current_group.iloc[-1]['y'])
                beginpoint_2 = (next_group.iloc[0]['x'], next_group.iloc[0]['y'])
                endpoint_2 = (next_group.iloc[-1]['x'], next_group.iloc[-1]['y'])

                if not end_segments_cross(beginpoint_1, endpoint_1, beginpoint_2, endpoint_2):
                    continue

                for k in range(len(current_group) - 1):
                    p1, p2 = (current_group.iloc[k]['x'], current_group.iloc[k]['y']), (current_group.iloc[k + 1]['x'], current_group.iloc[k + 1]['y'])
                    t1, t2 = current_group.iloc[k]['relative_time'], current_group.iloc[k + 1]['relative_time']

                    for l in range(len(next_group) - 1):
                        p3, p4 = (next_group.iloc[l]['x'], next_group.iloc[l]['y']), (next_group.iloc[l + 1]['x'], next_group.iloc[l + 1]['y'])
                        t3, t4 = next_group.iloc[l]['relative_time'], next_group.iloc[l + 1]['relative_time']

                        if abs(t1 - t3) > max_time_diff:
                            continue

                        if segments_cross(p1, p2, p3, p4):
                            intersection_x, intersection_y = intersection_point(p1, p2, p3, p4)
                            first, pet, time_curr, time_next = what_is_the_post_encroachment_time(
                                t1, t2, t3, t4, p1, p2, p3, p4, intersection_x, intersection_y, current_mode, next_mode)

                            for tid, opp_tid, time_at_cross in [
                                (current_trace_id, next_trace_id, time_curr),
                                (next_trace_id, current_trace_id, time_next)
                            ]:
                                crossing_data['crossing_trace_ids'][tid].append(opp_tid)
                                crossing_data['post_encroachment_time'][tid].append(pet)
                                crossing_data['first_mode'][tid].append(first)
                                crossing_data['intersection_x'][tid].append(intersection_x)
                                crossing_data['intersection_y'][tid].append(intersection_y)
                                crossing_data['time_at_crosspoint'][tid].append(time_at_cross)

                            processed_trace_ids.update([current_trace_id, next_trace_id])
            pbar.update(1)

    # Fill empty lists with NaN
    for key in crossing_data:
        for tid, val in crossing_data[key].items():
            if not val:
                crossing_data[key][tid] = np.nan

    # Attach to dataframe
    for key in crossing_data:
        df[key] = df['trace_id'].map(crossing_data[key])

    return df.reset_index(drop=True)
```

```

# Usage
df_filter_crossings = keep_intersecting_ped_cyc(df_filter_location, max_groups_to_check, max_time_diff, probability_threshold)
```

In a similar way as for the location, the origin and destinations are determined based on the location of the first and last trajectory points, but also a second method is used that determines this based on the direction that the trajectory moves from or goes to.

```
"""
Defining rectangles inside which the origins or destinations are set. Also, a function is presented that provides an alternative for determining the origin or destination if the first method does not give a result.
"""
import numpy as np

#direction vectors for determining the exact area
rotation_matrix = np.array([[0, -1],
                            [1, 0]])
first_vector_south_west = [1, np.tan(-5.5 * np.pi / 180.0)]
second_vector_south_west = np.dot(rotation_matrix, first_vector_south_west)
third_vector_south_west = np.dot(rotation_matrix, second_vector_south_west)
fourth_vector_south_west = np.dot(rotation_matrix, third_vector_south_west)

first_vector_north_east = [1, np.tan(-10 * np.pi / 180.0)]
second_vector_north_east = np.dot(rotation_matrix, first_vector_north_east)
third_vector_north_east = np.dot(rotation_matrix, second_vector_north_east)
fourth_vector_north_east = np.dot(rotation_matrix, third_vector_north_east)

first_point_west = np.array([point_1])
second_point_west = np.array([point_1[0] + width_1 * np.cos(angle_125) + height_1 * np.sin(angle_125),
                              point_1[1] - width_1 * np.sin(angle_125) + height_1 * np.cos(angle_125)])

first_point_south = np.array([point_2])
second_point_south = np.array([point_2[0] + width_2 * np.cos(angle_125) + height_2 * np.sin(angle_125),
                               point_2[1] - width_2 * np.sin(angle_125) + height_2 * np.cos(angle_125)])

first_point_east = np.array([point_3])
second_point_east = np.array([point_3[0] + width_3 * np.cos(angle_34) + height_3 * np.sin(angle_34),
                              point_3[1] - width_3 * np.sin(angle_34) + height_3 * np.cos(angle_34)])

first_point_north = np.array([point_4])
second_point_north = np.array([point_4[0] + width_4 * np.cos(angle_34) + height_4 * np.sin(angle_34),
                              point_4[1] - width_4 * np.sin(angle_34) + height_4 * np.cos(angle_34)])

first_point_centre = np.array([point_5])
second_point_centre = np.array([point_5[0] + width_5 * np.cos(angle_125) + height_5 * np.sin(angle_125),
                              point_5[1] - width_5 * np.sin(angle_125) + height_5 * np.cos(angle_125)])

direction_list = [(first_point_west, second_point_west, first_vector_south_west, second_vector_south_west, third_vector_south_west, fourth_vector_south_west),
                  (first_point_south, second_point_south, first_vector_south_west, second_vector_south_west, third_vector_south_west, fourth_vector_south_west),
                  (first_point_east, second_point_east, first_vector_north_east, second_vector_north_east, third_vector_north_east, fourth_vector_north_east),
                  (first_point_north, second_point_north, first_vector_north_east, second_vector_north_east, third_vector_north_east, fourth_vector_north_east),
```

```python
                  (first_point_centre, second_point_centre, first_vector_south_west,
                      second_vector_south_west, third_vector_south_west,
                      fourth_vector_south_west)]

# Check the directions of the full trajectory
def point_direction(point):
    trajectory_direction = False
    point = np.array(point)
    for i, (lower_left_point, upper_right_point, first, second, third, fourth) in enumerate(
        direction_list):
        location_wrt_first_point = point - lower_left_point
        location_wrt_second_point = point - upper_right_point
        # print('respective location:', location_wrt_second_point)
        # print('inproduct:', np.dot(location_wrt_second_point, third))
        if np.dot(location_wrt_first_point, first) > 0 and np.dot(location_wrt_first_point,
            second) > 0 and np.dot(location_wrt_second_point, third) > 0 and np.dot(
            location_wrt_second_point, fourth) > 0:
            trajectory_direction = i + 1
            break
    return int(trajectory_direction)

quadrants = [
        (-5/8 * np.pi - 5.5 * np.pi / 180, -3/8 * np.pi - 5.5 * np.pi / 180),
        (-1/8 * np.pi - 5.5 * np.pi / 180, 1/8 * np.pi - 5.5 * np.pi / 180),
        (3/8 * np.pi - 5.5 * np.pi / 180, 5/8 * np.pi - 5.5 * np.pi / 180)
        ]

def alternative_direction(point1, point2):
    point1 = np.array(point1)
    point2 = np.array(point2)
    direction_vector = point1 - point2
    x_direction = direction_vector[0]
    y_direction = direction_vector[1]
    if x_direction == 0:
        return 4 if y_direction >= 0 else 2  # Assign direction values to special cases

    direction_angle = np.arctan(y_direction / x_direction)
    if x_direction < 0:
        direction_angle += np.pi if y_direction >= 0 else -np.pi

    if direction_angle >= 7/8 * np.pi - 5.5 * np.pi / 180 or direction_angle < -7/8 * np.pi -
         5.5 * np.pi / 180:
        return 1

    # Check which quadrant the direction_angle falls into
    for i, (start, end) in enumerate(quadrants):
        if start <= direction_angle < end:
            return i + 2  # Quadrants are 1-indexed
    return 5
```

```python
"""
Function that assigns the actual directions to the trajectories numbered as
1 = west
2 = south
3 = east
4 = north
5 = inconclusive --> potentially changed by the alternative method to become one of the 4
    cardinal directions.
"""
import pandas as pd
import numpy as np

def coming_from_going_to(df):
    # Group the dataframe by trace_id
    groups = list(df.groupby('trace_id'))

    # Initialize the acceleration column
    df['coming_from'] = np.nan
    df['going_to'] = np.nan

    # Initialize the progress bar
```

```
    with tqdm(total=len(groups), desc="Processing␣Groups") as pbar:
        # Iterate through groups
        for current_trace_id, current_group in groups:
            beginpoint_1 = (current_group.iloc[0]['y'], - current_group.iloc[0]['x'])
            endpoint_1 = (current_group.iloc[-1]['y'], - current_group.iloc[-1]['x'])

            # Determine mode
            mode = current_group.iloc[0]['type_of_active_mode']
            min_length = 4 if mode == 'pedestrian' else 2
            begin_iloc = 3 if mode == 'pedestrian' else 1
            end_iloc = -4 if mode == 'pedestrian' else -2

            # Calculate directions
            coming_from_value = point_direction(beginpoint_1)
            going_to_value = point_direction(endpoint_1)

            # Adjust if direction is 5 (inconclusive)
            if coming_from_value == 5 and len(current_group) >= min_length:
                beginpoint_2 = (current_group.iloc[begin_iloc]['y'], - current_group.iloc[
                    begin_iloc]['x'])
                coming_from_value = alternative_direction(beginpoint_1, beginpoint_2)

            if going_to_value == 5 and len(current_group) >= min_length:
                endpoint_2 = (current_group.iloc[end_iloc]['y'], - current_group.iloc[
                    end_iloc]['x'])
                going_to_value = alternative_direction(endpoint_1, endpoint_2)

            # Update only the relevant rows in df
            df.loc[df['trace_id'] == current_trace_id, ['coming_from', 'going_to']] = int(
                coming_from_value), int(going_to_value)

            # Update the progress bar
            pbar.update(1)

    return df.reset_index(drop=True)

df_with_begin_and_end = coming_from_going_to(df_filter_crossings)
```

Lastly, the code for calculating the predicted PET, which uses a slightly different method than the actual PET.

```
"""
Additional functions to calculate the predicted PET
"""
import pandas as pd
import numpy as np

# Function to calculate predicted PET based on the variables speed and direction of the
    pedestrian and speed for the cyclist
def predicted_PET(t1, t3, t4, p1, p2, p3, p4, intersection_x, intersection_y, v1, t_cyclist,
    distance_cyclist, v_cyclist):
    distance_pedestrian = np.linalg.norm([intersection_x, intersection_y] - np.array(p1))
    #interpolate the time for both modes to reach the intersection point
    time_pedestrian = t1 + distance_pedestrian / v1

    if np.isnan(v_cyclist):
        distance_cyclist = np.linalg.norm([intersection_x, intersection_y] - np.array(p3))
        time_cyclist = t3 + (t4 - t3) * np.linalg.norm([intersection_x, intersection_y] - np.
            array(p3)) / np.linalg.norm(
                                                            np.array(p4) - np.array(p3))
    else:
        time_cyclist = t_cyclist + distance_cyclist / v_cyclist

    post_encroachment_time = time_cyclist - time_pedestrian
    return post_encroachment_time, time_pedestrian, time_cyclist

# Function to determine what the PET is (including negative values for pedestrians crossing
    after the cyclist)
def what_is_the_post_encroachment_time(t1, t2, t3, t4, p1, p2, p3, p4, intersection_x,
    intersection_y, current_mode, next_mode):
    #interpolate the time for both modes to reach the intersection point
```

```python
    time_current_mode = t1 + (t2 - t1) * np.linalg.norm([intersection_x, intersection_y] - np.array(p1)) / np.linalg.norm(
                                                            np.array(p2) - np.array(p1))
    time_next_mode = t3 + (t4 - t3) * np.linalg.norm([intersection_x, intersection_y] - np.array(p3)) / np.linalg.norm(
                                                            np.array(p4) - np.array(p3))
    # Calculate the post-encroachment time
    if time_current_mode < time_next_mode:
        first = current_mode
        post_encroachment_time = time_next_mode - time_current_mode
    else:
        first = next_mode
        post_encroachment_time = time_current_mode - time_next_mode
    return first, post_encroachment_time, time_current_mode, time_next_mode

def calculate_predicted_PET(pedestrian_df, full_df):
    print(len(pedestrian_df))
    mask = pedestrian_df['post_encroachment_time'].apply(lambda x: is_single_pet_in_range(x, 0, 5))
    filtered_df = pedestrian_df[mask].copy()
    print(len(filtered_df))
    # filtered_df['post_encroachment_time'] = filtered_df['post_encroachment_time'].apply(extract_single_float_from_list)

    list_columns = ['crossing_trace_ids', 'post_encroachment_time', 'intersection_x', 'intersection_y', 'time_at_crosspoint']
    for col in list_columns:
        filtered_df[col] = filtered_df[col].apply(safe_eval_list)

    pedestrian_df_updated = filtered_df.copy()
    groups = list(pedestrian_df_updated.groupby('trace_id'))

    for current_trace_id, current_pedestrian in tqdm(groups, desc="Calculating␣PET"):
        current_pedestrian = current_pedestrian.reset_index(drop=True)
        first_row = current_pedestrian.iloc[0]

        crossing_cyclist_trace_id = first_row['crossing_trace_ids'][0]
        crossing_cyclist_group = full_df[full_df['trace_id'] == crossing_cyclist_trace_id].reset_index(drop=True)

        p2 = (current_pedestrian.iloc[-1]['x'], current_pedestrian.iloc[-1]['y'])
        time_at_crosspoint = first_row['time_at_crosspoint'][0]

        for k in range(len(current_pedestrian) - 1):
            row_k = current_pedestrian.iloc[k]
            p1 = (row_k['x'], row_k['y'])
            t1 = row_k['relative_time']
            v1 = row_k['speed']

            if t1 > time_at_crosspoint:
                break

            for l in range(len(crossing_cyclist_group) - 1):
                row_l, row_l1 = crossing_cyclist_group.iloc[l], crossing_cyclist_group.iloc[l + 1]
                p3 = (row_l['x'], row_l['y'])
                p4 = (row_l1['x'], row_l1['y'])
                t3, t4 = row_l['relative_time'], row_l1['relative_time']

                if segments_cross(p1, p2, p3, p4):

                    intersection_x, intersection_y = intersection_point(p1, p2, p3, p4)
                    pet, time_ped, time_cyc = predicted_PET(t1, t3, t4, p1, p2, p3, p4, intersection_x, intersection_y, v1)

                    idx = pedestrian_df_updated[
                        (pedestrian_df_updated['trace_id'] == current_trace_id) &
                        (pedestrian_df_updated['relative_time'] == t1)].index
                    if not idx.empty:
                        pedestrian_df_updated.at[idx[0], 'predicted_post_encroachment_time'] = pet
```

```

    return pedestrian_df_updated

predicted_interacting_pedestrian_east_west = calculate_predicted_PET(
    interacting_pedestrian_east_west, df_with_begin_and_end)
# Only taking the interacting pedestrians crossing from east to west
```

B

Fitting a Distribution Function

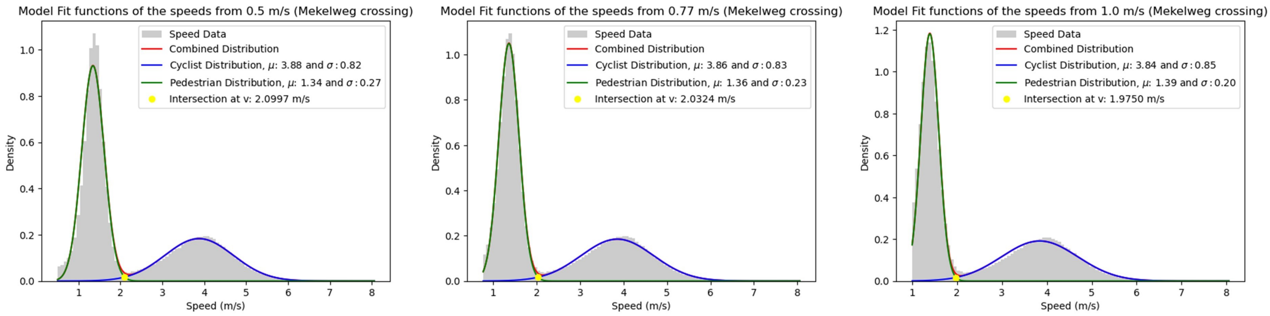


Figure B.1: Comparison for three different threshold values of the speed where the distribution functions are plotted on. The threshold values of 0.5 m/s and 1.0 m/s deliver distributions that are not accurate enough to describe the curves visualised by the histogram of the data

C

Origin-Destination Matrices for pedestrians and cyclists at the Mekelweg and Lorentzweg crossing

Table C.1: Origin-Destination Matrix of pedestrians at the Mekelweg crossing

↓ Origin \Destination →	West	South	South-east	East	North	Inconclusive
West	63	353	298	1184	202	59
South	469	1138	671	3004	2849	347
South-east	340	872	55	40	14	11
East	1764	3237	42	100	46	45
North	349	1564	110	95	438	747
Inconclusive	210	328	17	36	798	121

Table C.2: Origin-Destination Matrix of cyclists at the Mekelweg crossing

↓ Origin \Destination →	West	South	South-east	East	North	Inconclusive
West	0	38	1148	18984	1	560
South	63	39	145	215	17	126
South-east	41	6	3	2	0	2
East	14569	19	41	453	5	32
North	3	16	4	25	1	6
Inconclusive	3186	162	24	25	14	110

Table C.3: Origin-Destination Matrix of pedestrians at the Lorentzweg crossing

↓ Origin \Destination →	West	South	East	North	Inconclusive
West	372	715	19841	991	1401
South	1688	1403	4318	5631	4251
East	17984	3718	5465	173	339
North	1602	4522	90	1098	267
Inconclusive	1263	3851	447	359	1257

Table C.4: Origin-Destination Matrix of cyclists at the Lorentzweg crossing

↓ Origin \Destination →	West	South	East	North	Inconclusive
West	566	14957	29810	4450	1607
South	18822	2069	1526	4123	1426
East	21631	1027	478	42	72
North	4678	3528	94	88	294
Inconclusive	2455	3498	279	633	758

D

Fitting lines on PET vs stopping percentages

See next page.

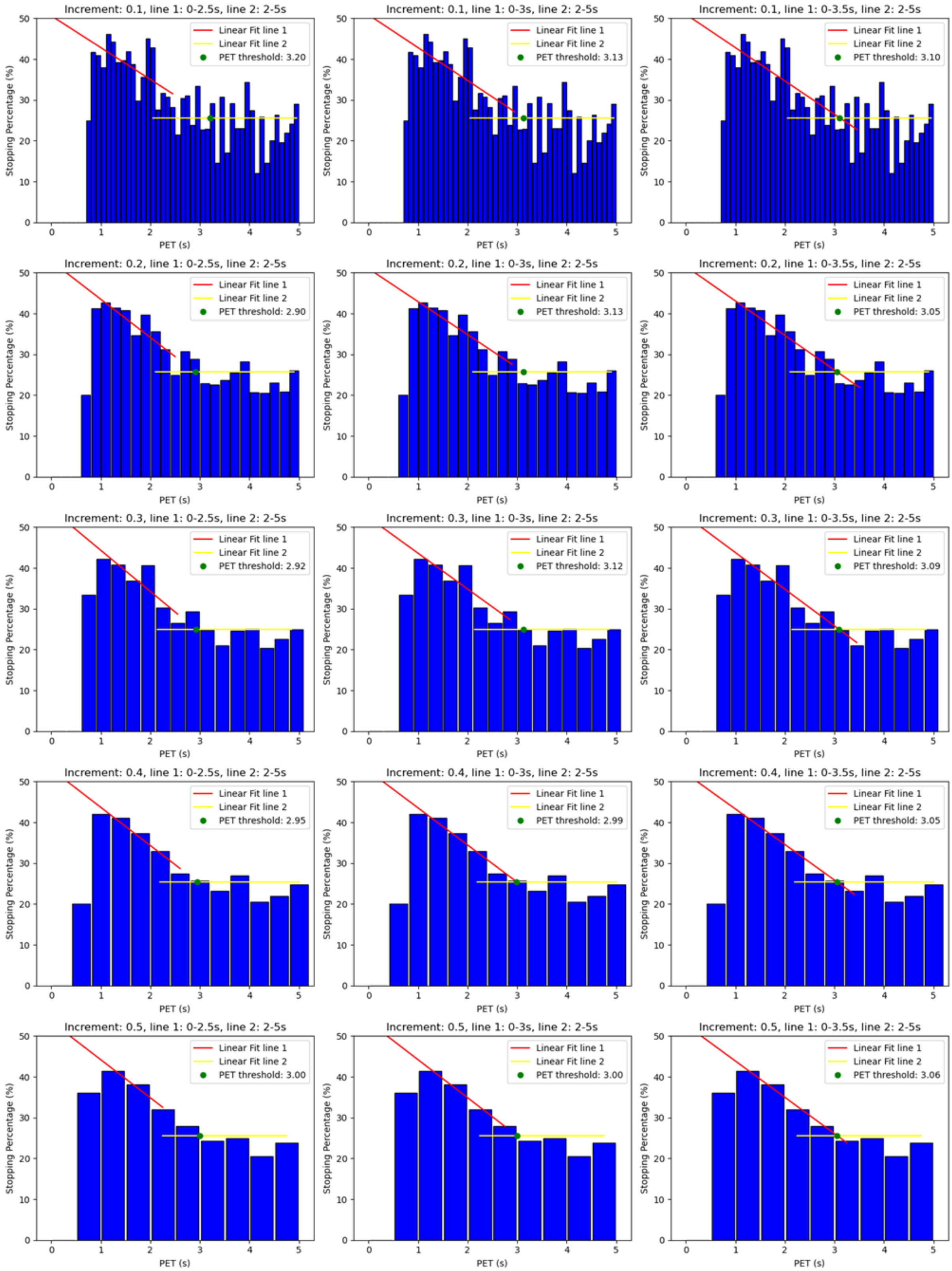


Figure D.1: Several fitting lines on PET vs stopping percentages, with the bar width (0.1 to 0.5) and the end of the first fitting line (2.5 to 3.5 s) as variables. The second (horizontal) fitting line starts at 2 seconds.

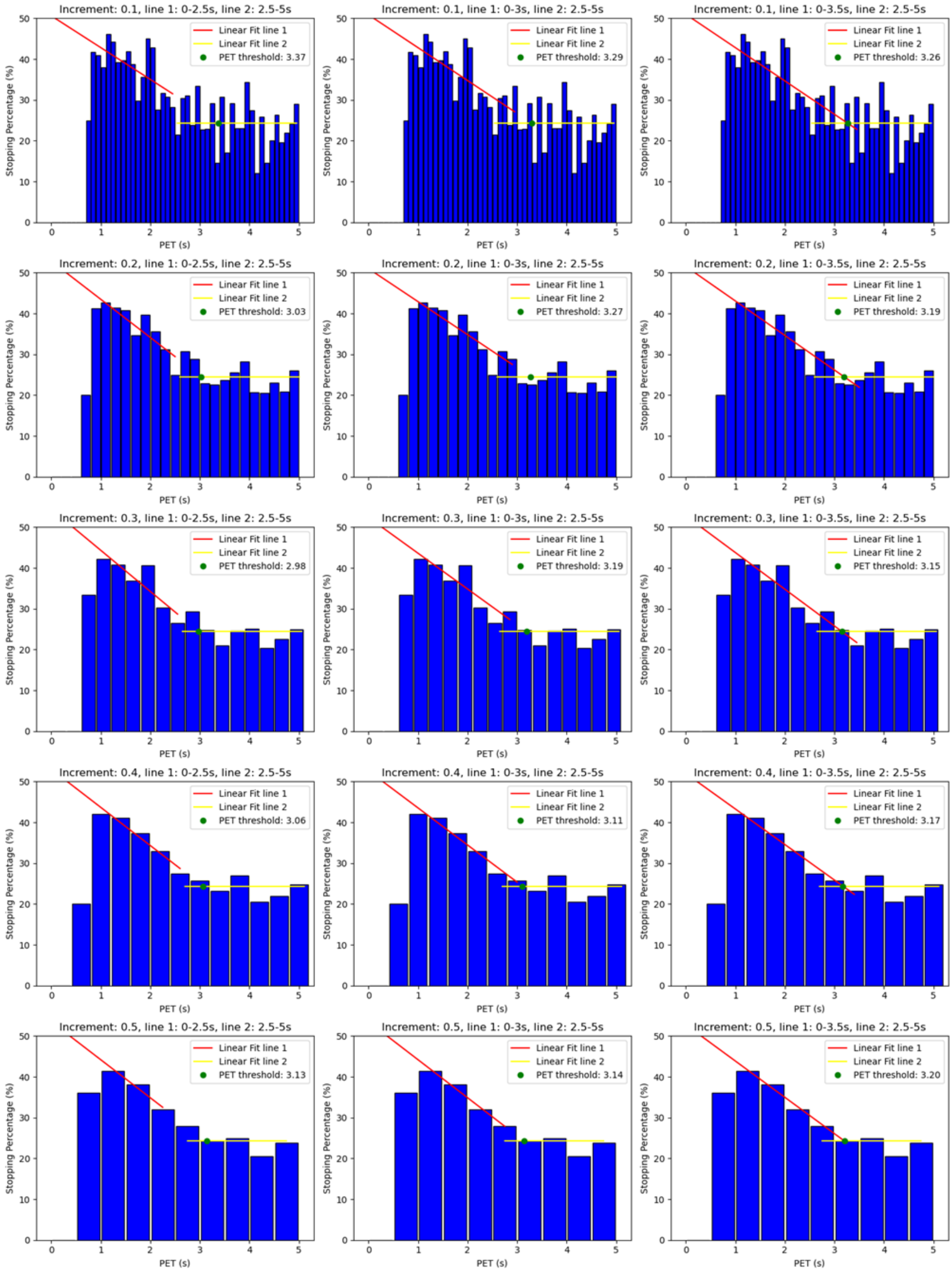


Figure D.2: Several fitting lines on PET vs stopping percentages, with the bar width (0.1 to 0.5) and the end of the first fitting line (2.5 to 3.5 s) as variables. The second (horizontal) fitting line starts at 2.5 seconds.

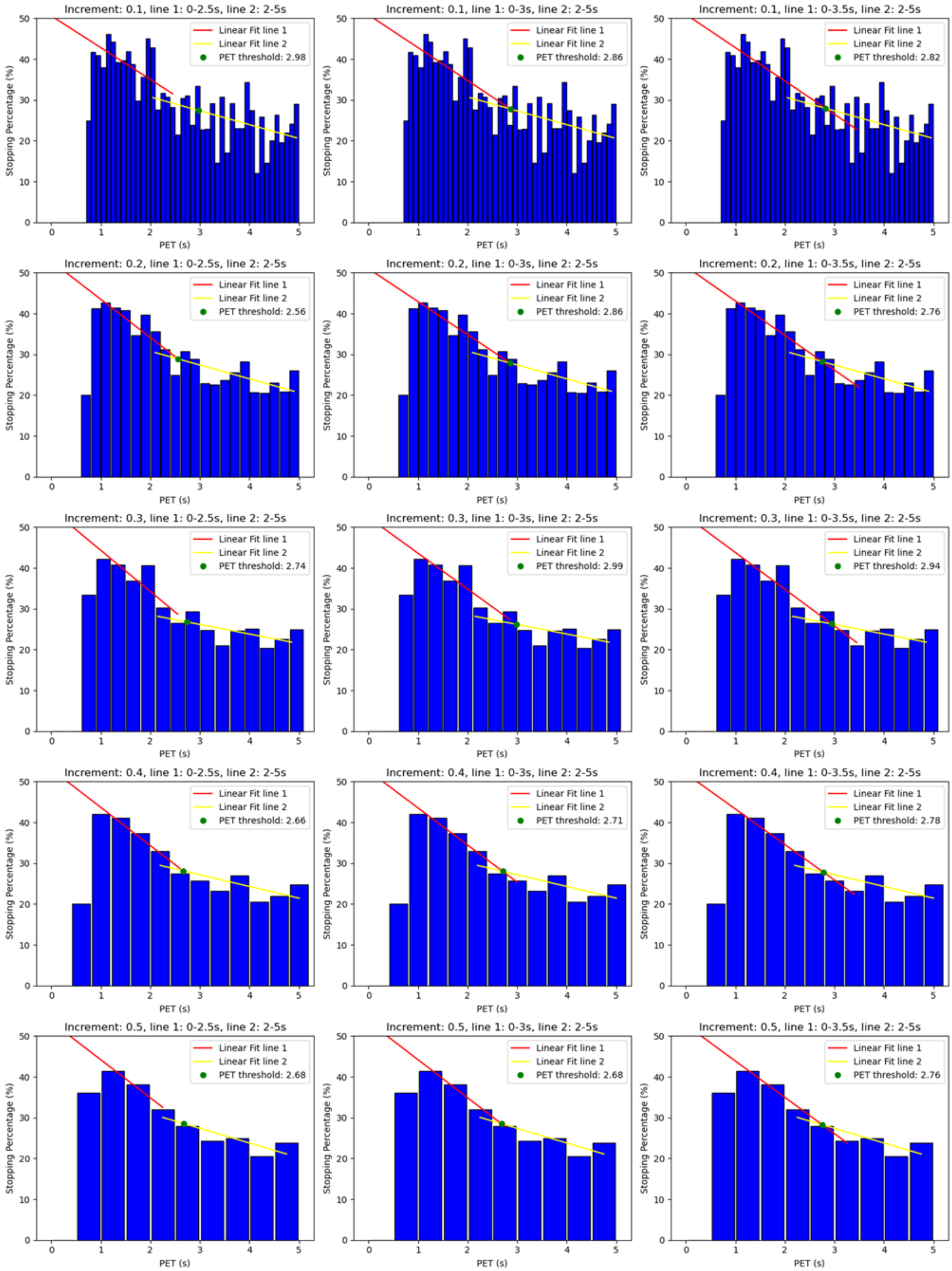


Figure D.3: Several fitting lines on PET vs stopping percentages, with the bar width (0.1 to 0.5) and the end of the first fitting line (2.5 to 3.5 s) as variables. The second fitting line starts at 2 seconds.

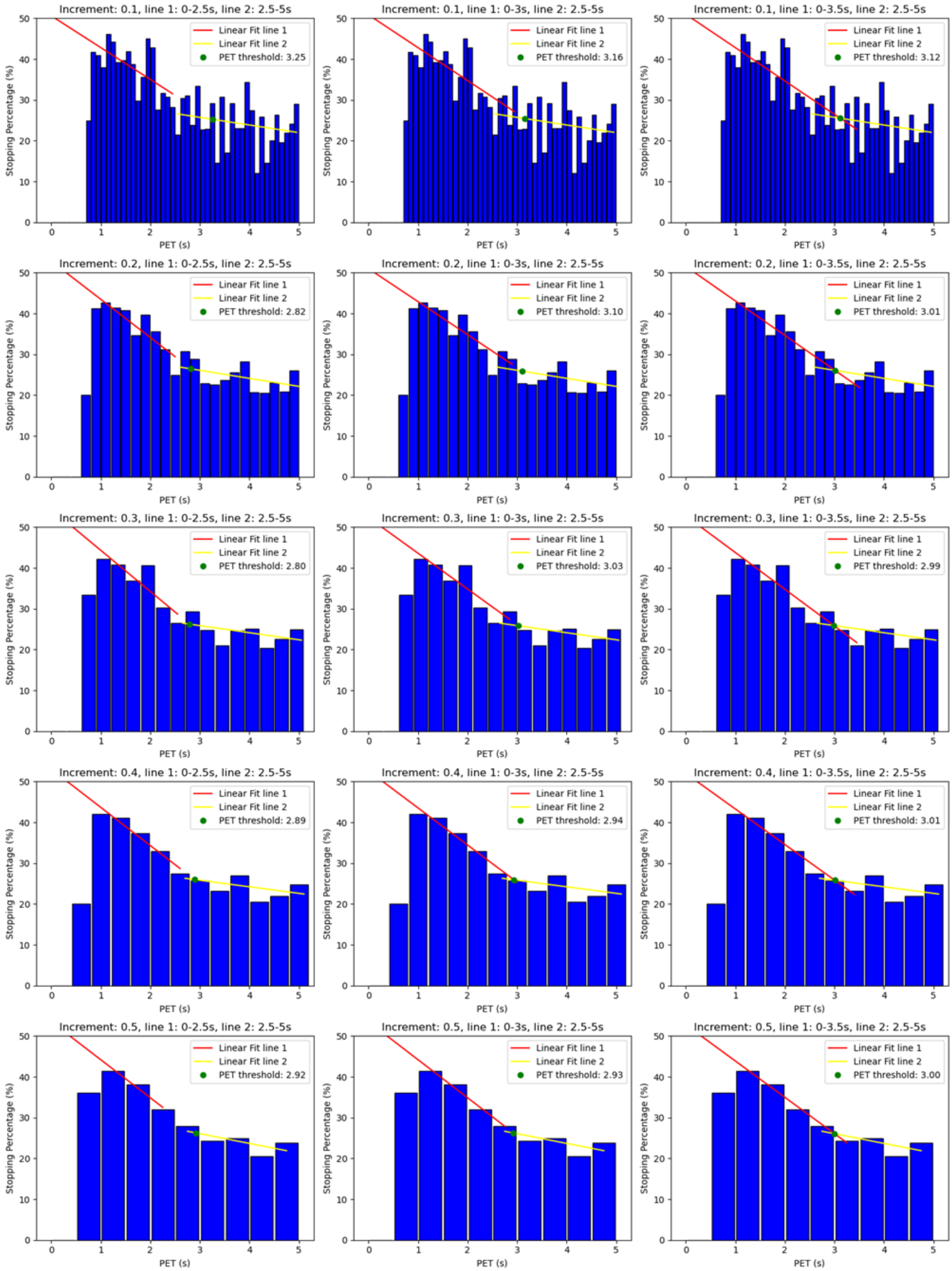


Figure D.4: Several fitting lines on PET vs stopping percentages, with the bar width (0.1 to 0.5) and the end of the first fitting line (2.5 to 3.5 s) as variables. The second fitting line starts at 2.5 seconds.

E

Determining threshold RMSD and maximum deviation

The example that is shown in this appendix is from the RMSD of pedestrians walking from east to west at the Lorentzweg crossing. The same method is applied for the maximum deviation. The trajectory RMSD for these pedestrians is given for the four scenarios:

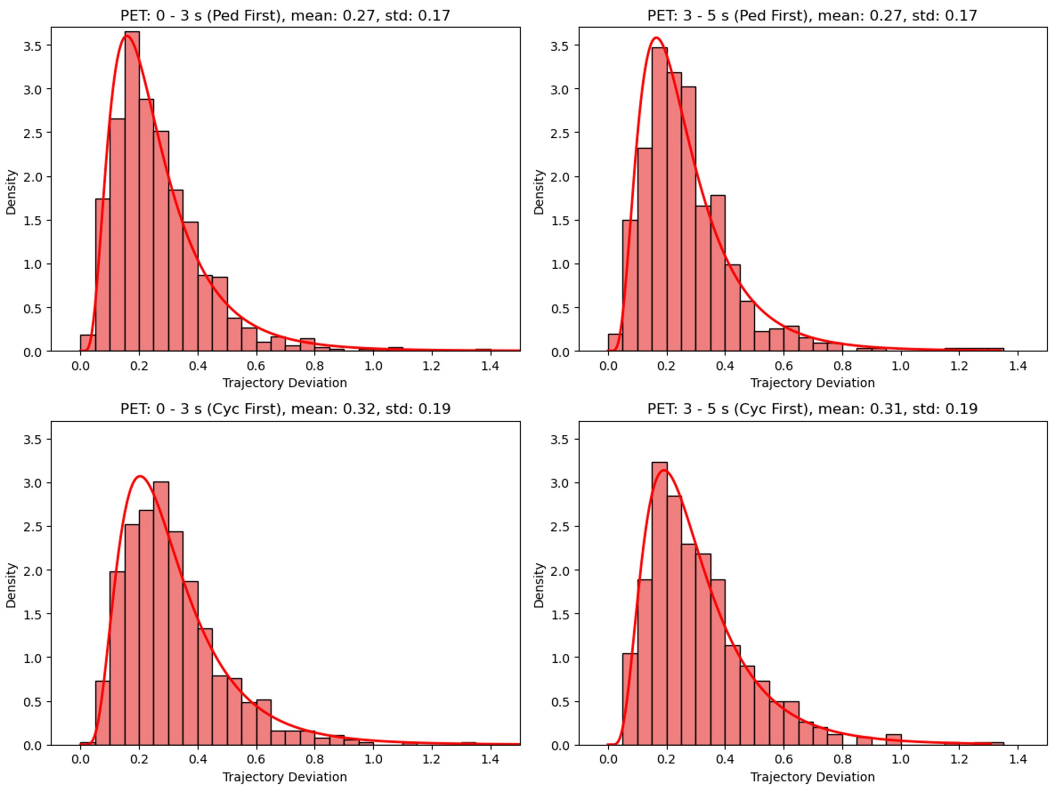


Figure E.1: Histograms of the root mean squared deviation of crossing trajectories

This is compared with the RMSD of all other trajectories of pedestrians walking from east to west at the Lorentzweg crossing, so non-crossing pedestrians. The comparison is made by overlaying the crossing pedestrian lognormal distributions with the non-crossing:

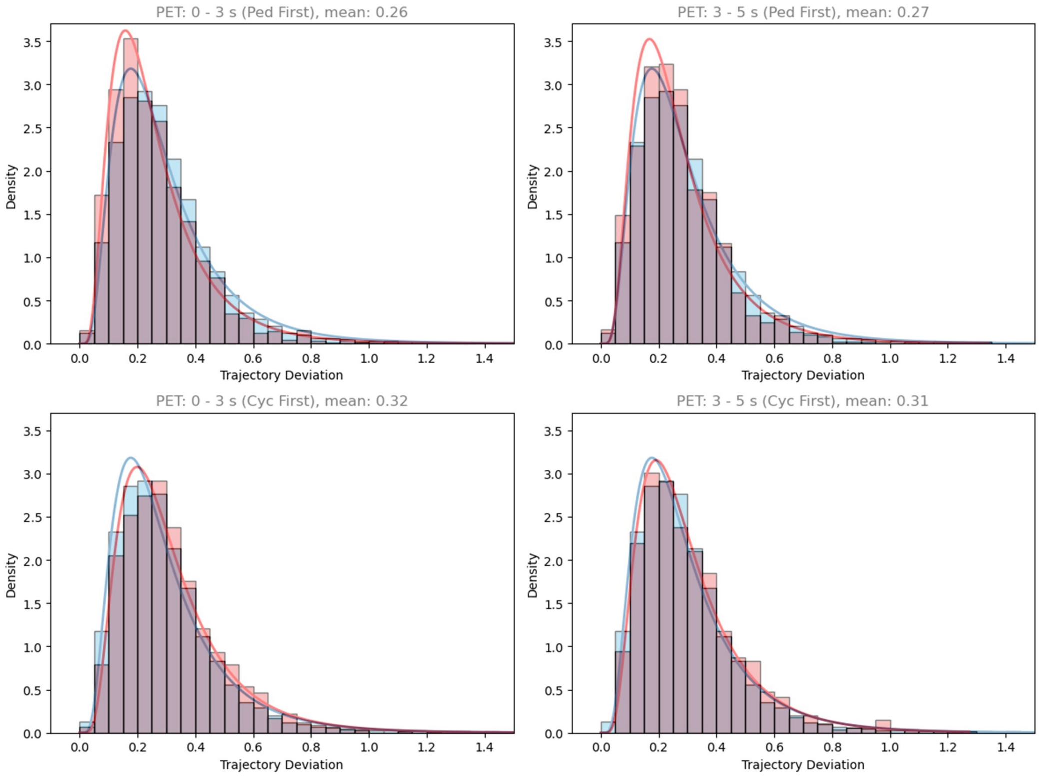


Figure E.2: Histograms of the root mean squared deviation of crossing trajectories vs non-crossing trajectories

This figure visualises the difference between the crossing and non-crossing trajectories and based on the distribution function, this difference can be calculated for the RMSD values. A new graph is made that shows this expected absolute difference between crossing and non-crossing trajectories:

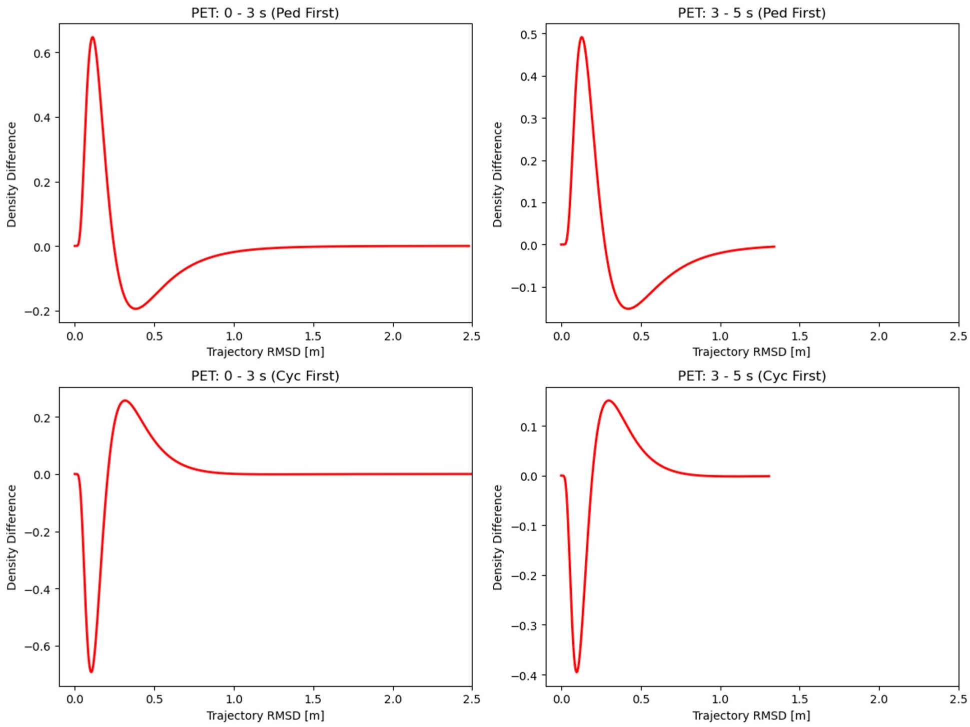


Figure E.3: Density difference between crossing pedestrians and non-crossing pedestrians

However, the absolute difference between the crossing and non-crossing pedestrians does not explain where the deviation is most different from the normal (non-crossing) situation. This is why this difference is measured relative to the non-crossing situation, to see at what amount of RMSD the trajectories of crossing pedestrians are occurring relatively more often compared to the non-crossing pedestrians

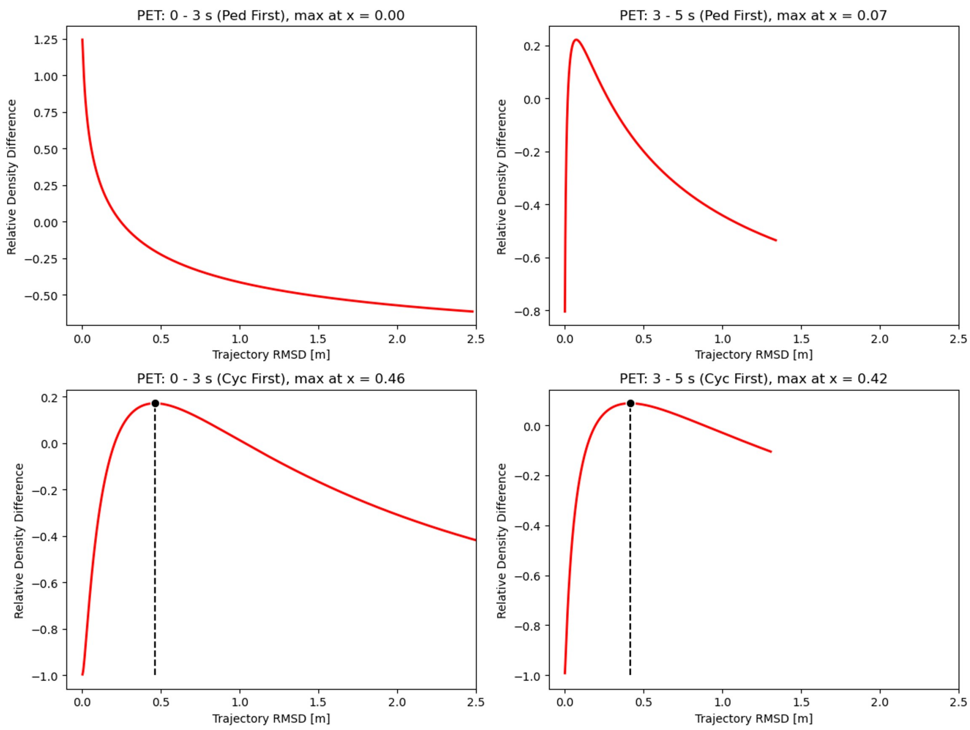


Figure E.4: Relative density difference between crossing pedestrians and non-crossing pedestrians

The peak in this figure is highlighted, which can be used as the threshold value, because in this situation, the most deviation takes place compared to a non-crossing situation.

F

Deviation Mekelweg crossing

The RMSD and maximum deviation for pedestrians crossing (south-north or north-south) at the Mekelweg crossing has lower significance compared to the east-west crossing pedestrians at the Lorentzweg crossing.

F.1. RMSD south-north

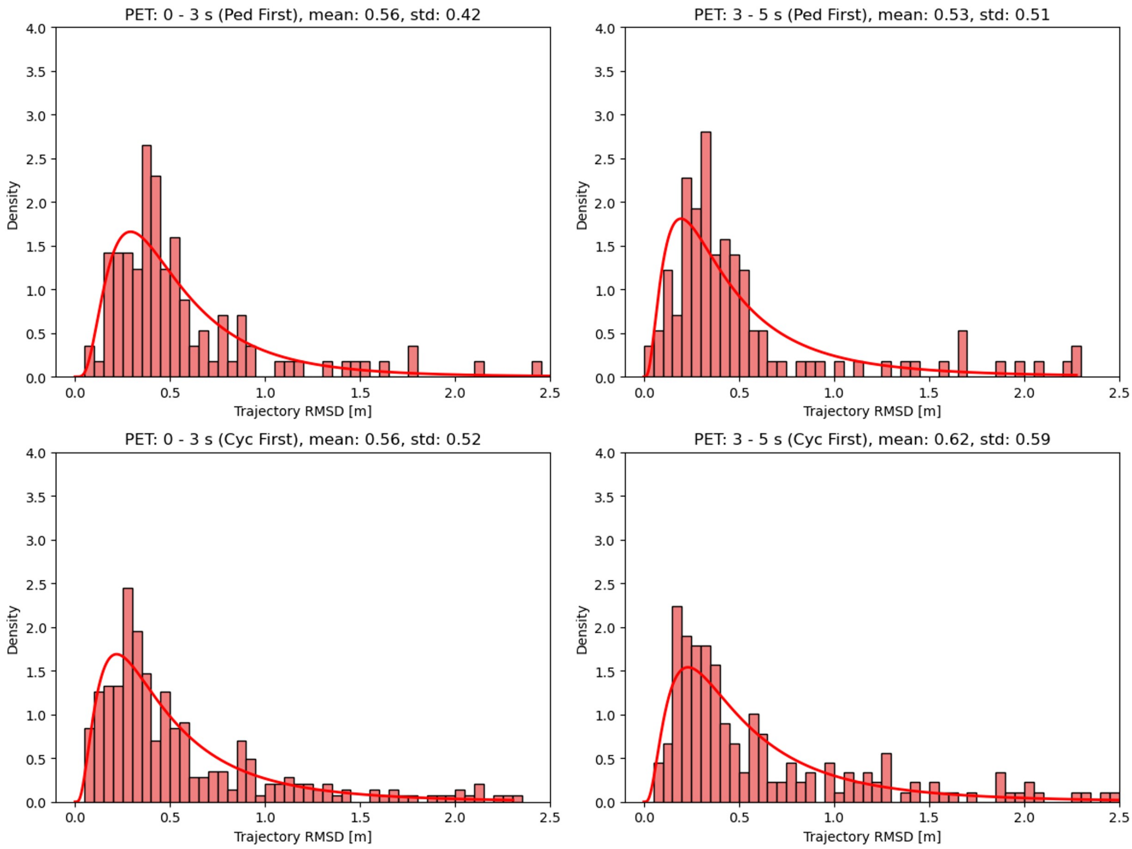


Figure F.1: RMSD of crossing pedestrians going south to north at the Mekelweg crossing

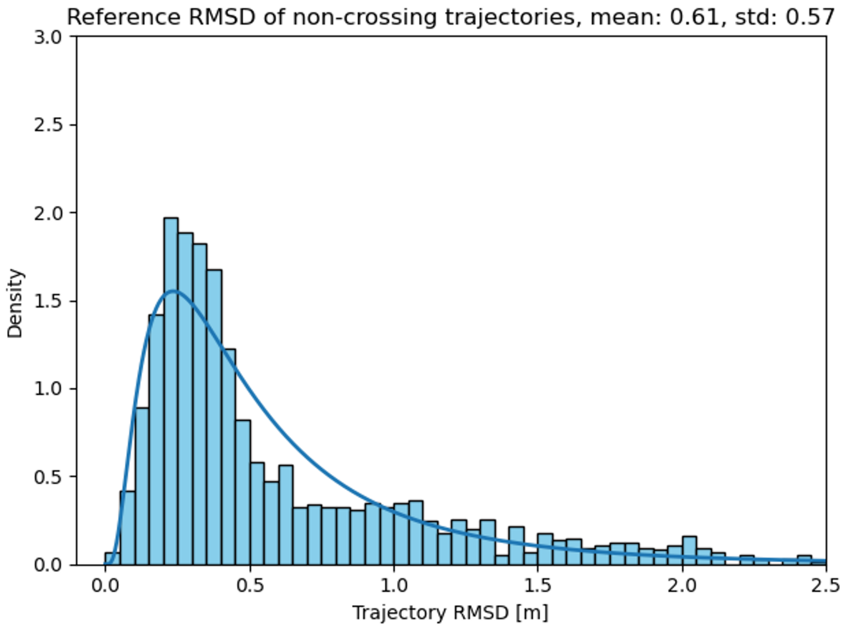


Figure F.2: RMSD of non-crossing pedestrians going south to north at the Mekelweg crossing

Table F.1: Significance test for the trajectory RMSD for crossing pedestrians compared to non-crossing pedestrians at the Mekelweg crossing

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.098	0.211	No	Not significantly different
	3 - 5 s	0.171	2.23e-3	Yes	Significantly different
Cyc first	0 - 3 s	0.094	0.011	Yes	Significantly different
	3 - 5 s	0.071	0.306	No	Not significantly different

F.2. Maximum Deviation south-north

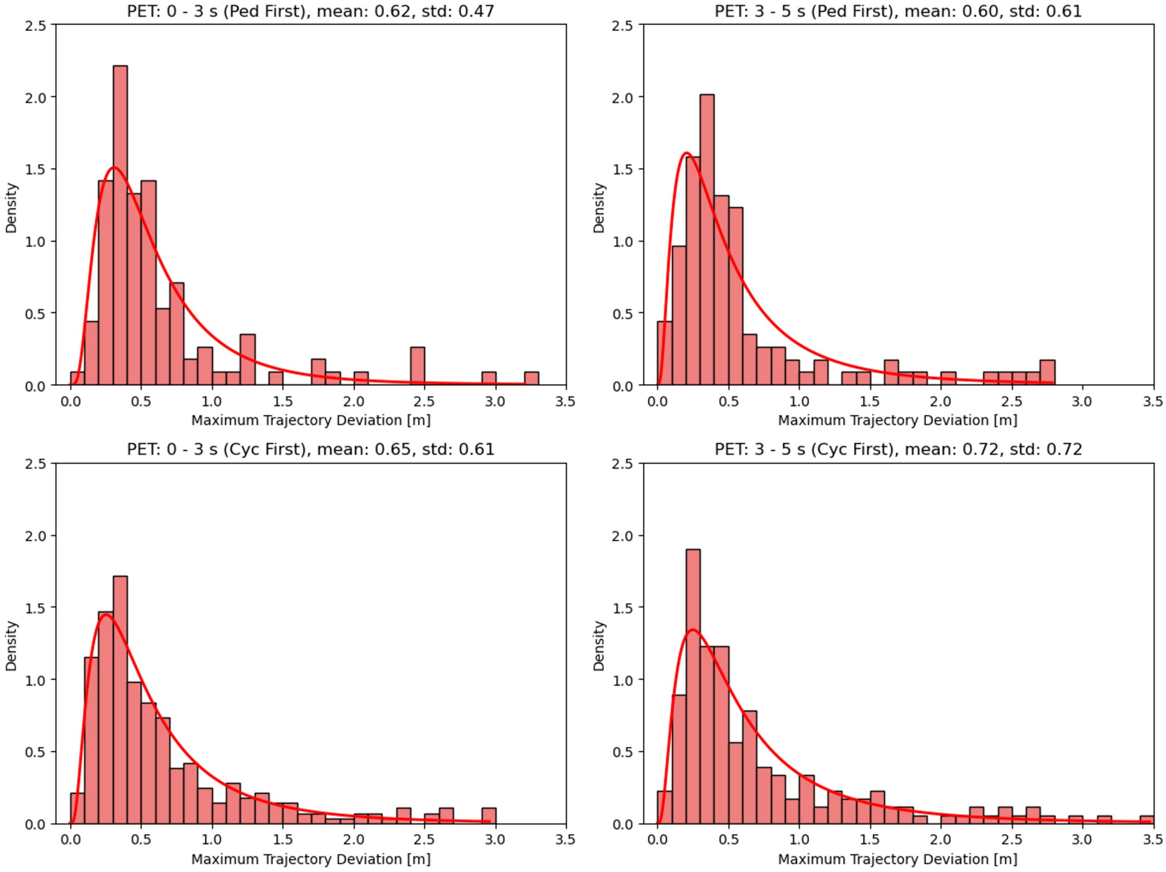


Figure F.3: Maximum deviation of crossing pedestrians going south to north at the Mekelweg crossing

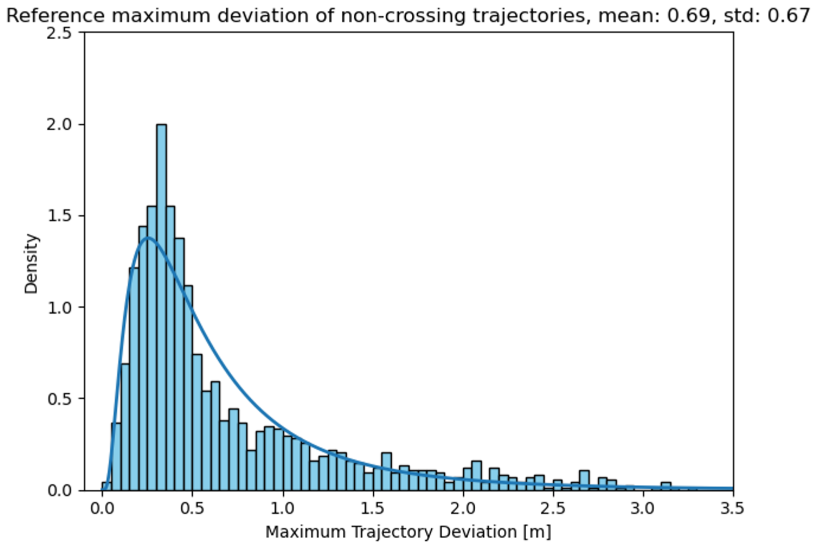


Figure F.4: Maximum deviation of non-crossing pedestrians going south to north at the Mekelweg crossing

Table F.2: Significance test for the trajectory maximum deviation for crossing pedestrians compared to non-crossing pedestrians at the Mekelweg crossing

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.118	0.079	No	Not significantly different
	3 - 5 s	0.166	3.16e-3	Yes	Significantly different
Cyc first	0 - 3 s	0.066	0.154	No	Not significantly different
	3 - 5 s	0.060	0.525	No	Not significantly different

F.3. RMSD north-south

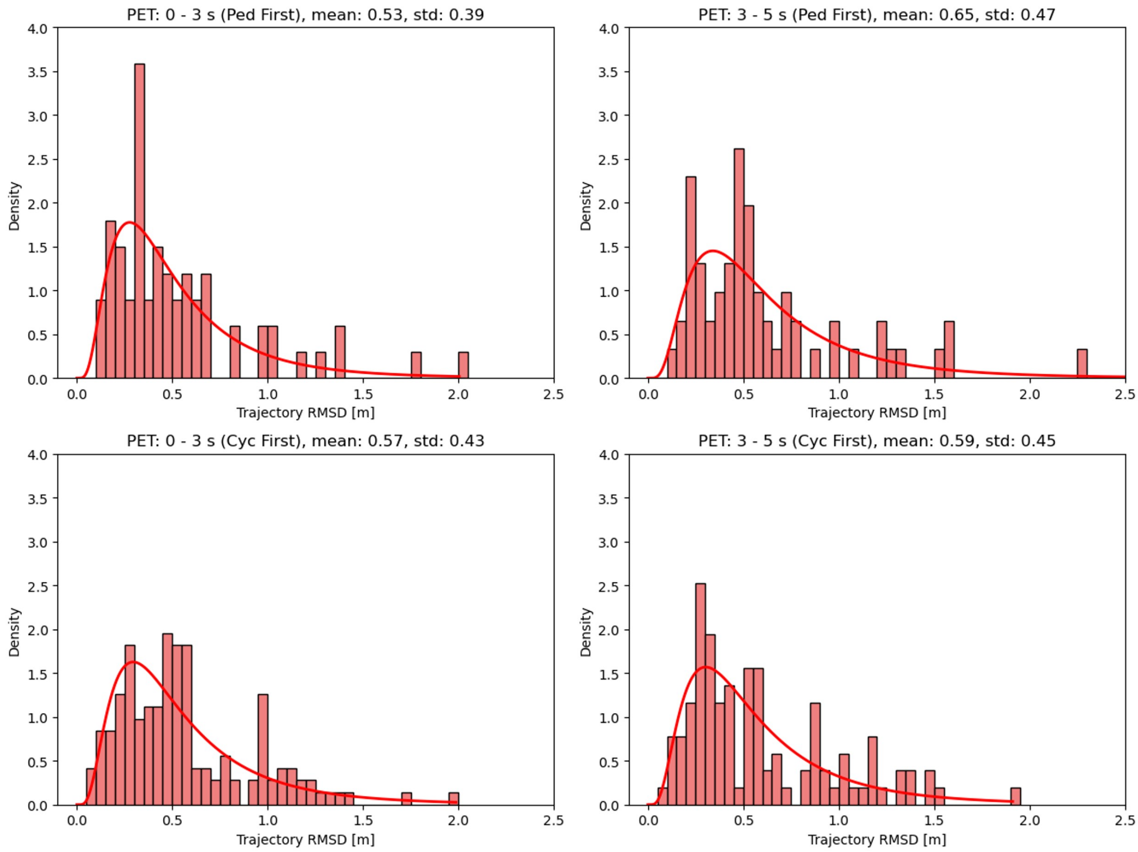


Figure F.5: RMSD of crossing pedestrians going north to south at the Mekelweg crossing

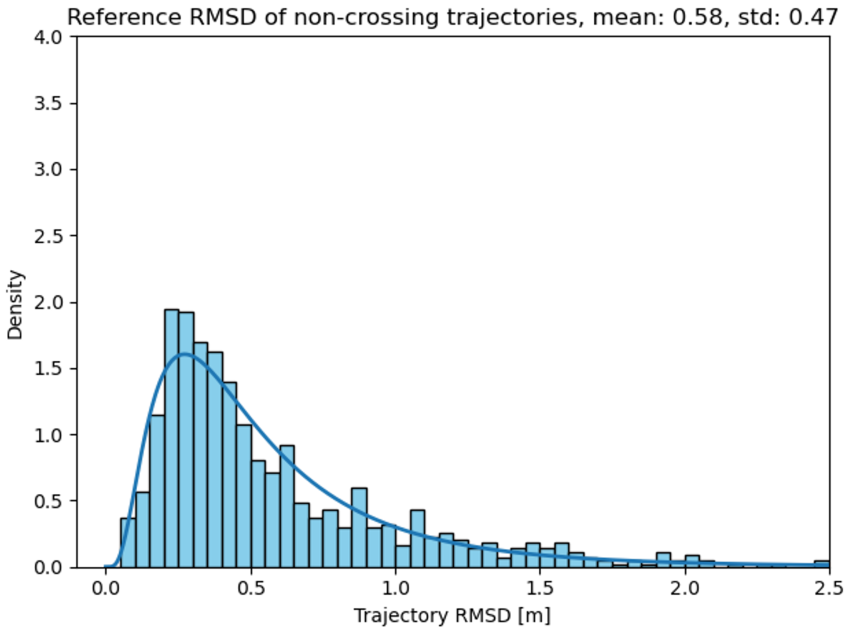


Figure F.6: RMSD of non-crossing pedestrians going north to south at the Mekelweg crossing

Table F.3: Significance test for the trajectory RMSD for crossing pedestrians compared to non-crossing pedestrians at the Mekelweg crossing

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.096	0.532	No	Not significantly different
	3 - 5 s	0.138	0.178	No	Not significantly different
Cyc first	0 - 3 s	0.091	0.175	No	Not significantly different
	3 - 5 s	0.085	0.428	No	Not significantly different

F.4. Maximum Deviation north-south

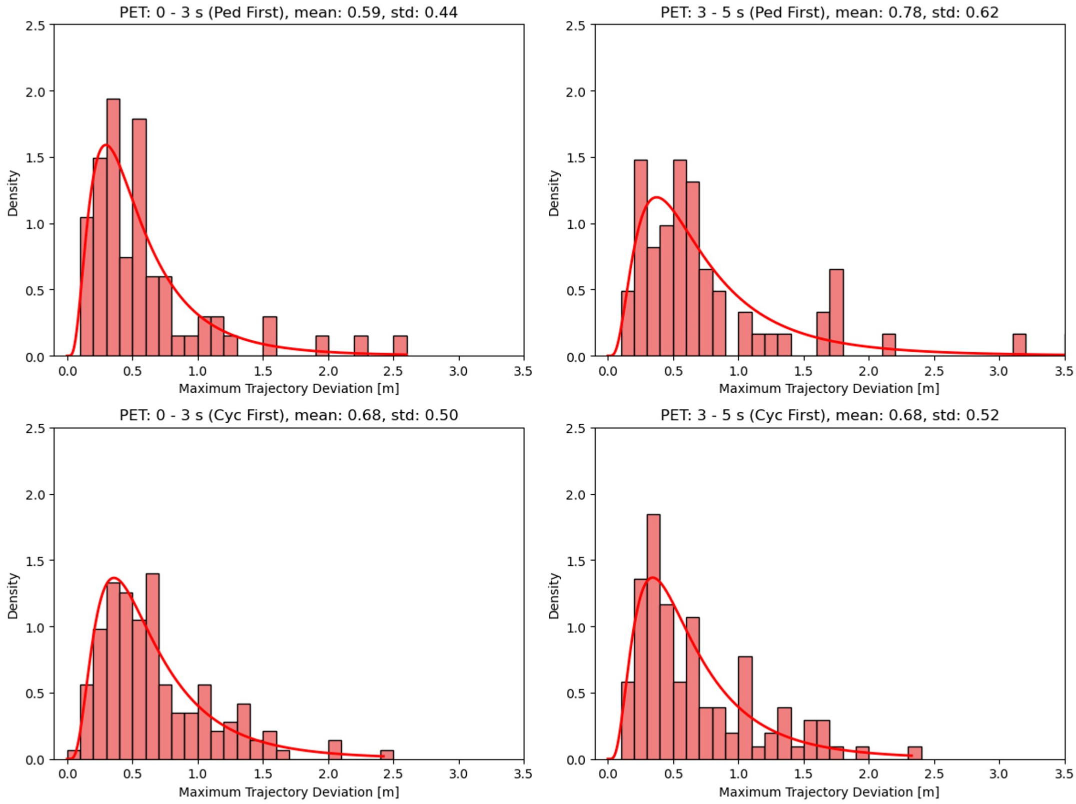


Figure F.7: Maximum deviation of crossing pedestrians going north to south at the Mekelweg crossing

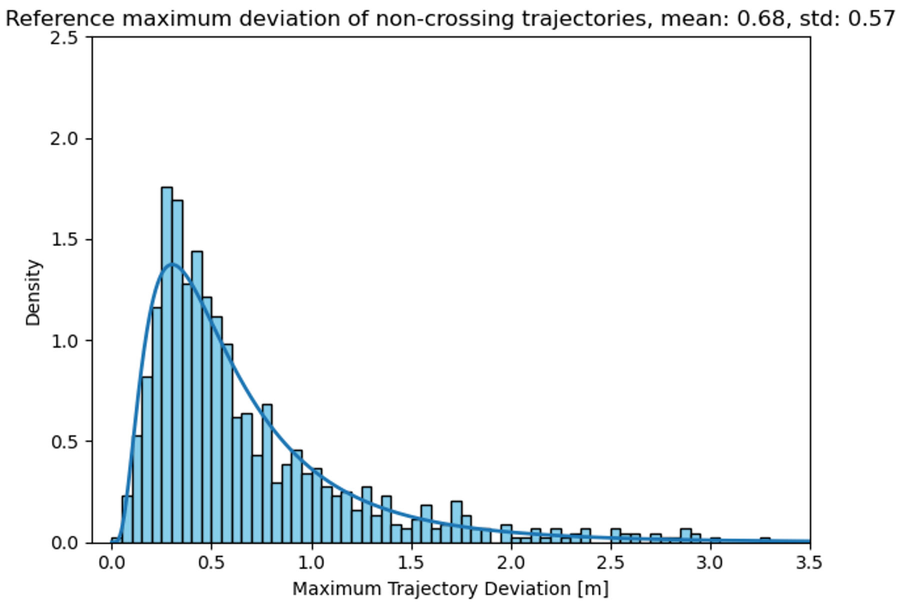


Figure F.8: Maximum deviation of non-crossing pedestrians going north to south at the Mekelweg crossing

Table F.4: Significance test for the trajectory maximum deviation for crossing pedestrians compared to non-crossing pedestrians at the Mekelweg crossing

First crossing mode	PET category	KS statistic	P-value	p <0.05	Statistical significance
Ped first	0 - 3 s	0.142	0.122	No	Not significantly different
	3 - 5 s	0.127	0.253	No	Not significantly different
Cyc first	0 - 3 s	0.096	0.133	No	Not significantly different
	3 - 5 s	0.066	0.742	No	Not significantly different

G

Trajectory plots crossing pedestrians Lorentzweg and Mekelweg crossing

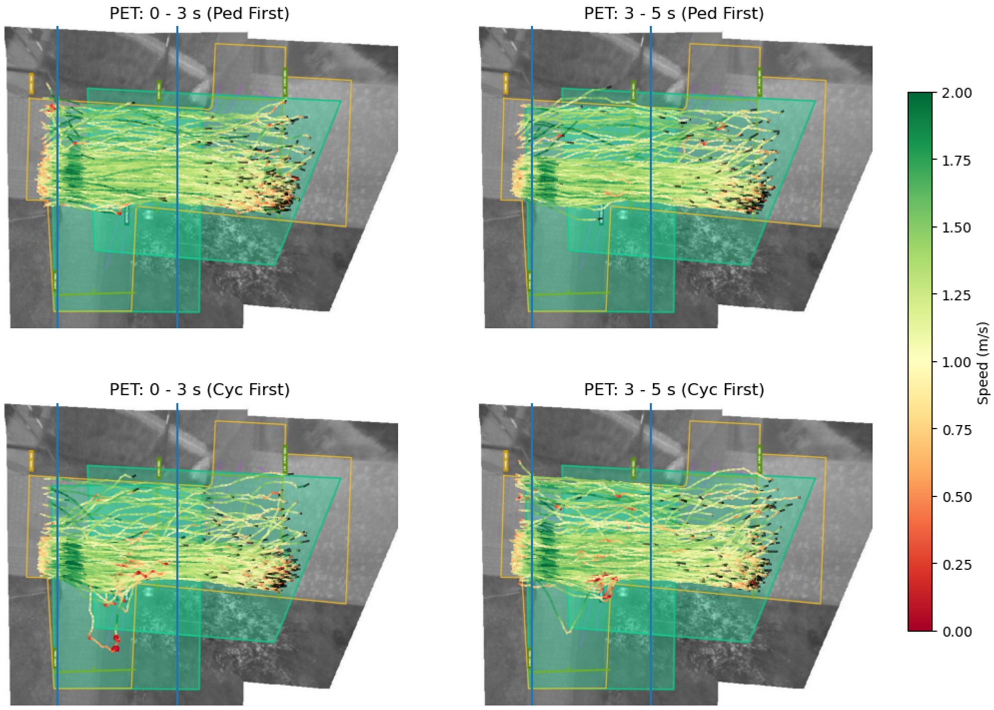


Figure G.1: All crossing pedestrians walking from east to west at the Lorentzweg crossing.

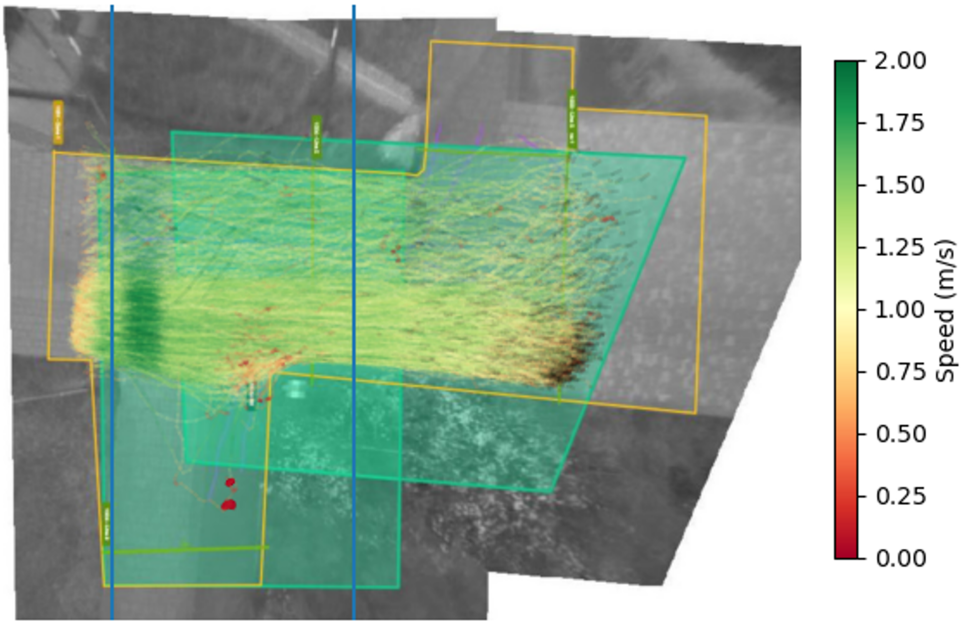


Figure G.2: Trajectory plots of all pedestrians walking from east to west at the Lorentzweg crossing with higher opacity

Table G.1: Total number of crossing trajectories within each category of PET and first crossing mode (pedestrians go from south to west) at the Lorentzweg crossing

East - west (total trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	1082	704
Cyclist crosses first	811	800

Table G.2: number of crossing trajectories that stop within each category of PET and first crossing mode (pedestrians go from south to west) at the Lorentzweg crossing

East - west (stopping trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	48	46
Cyclist crosses first	266	125

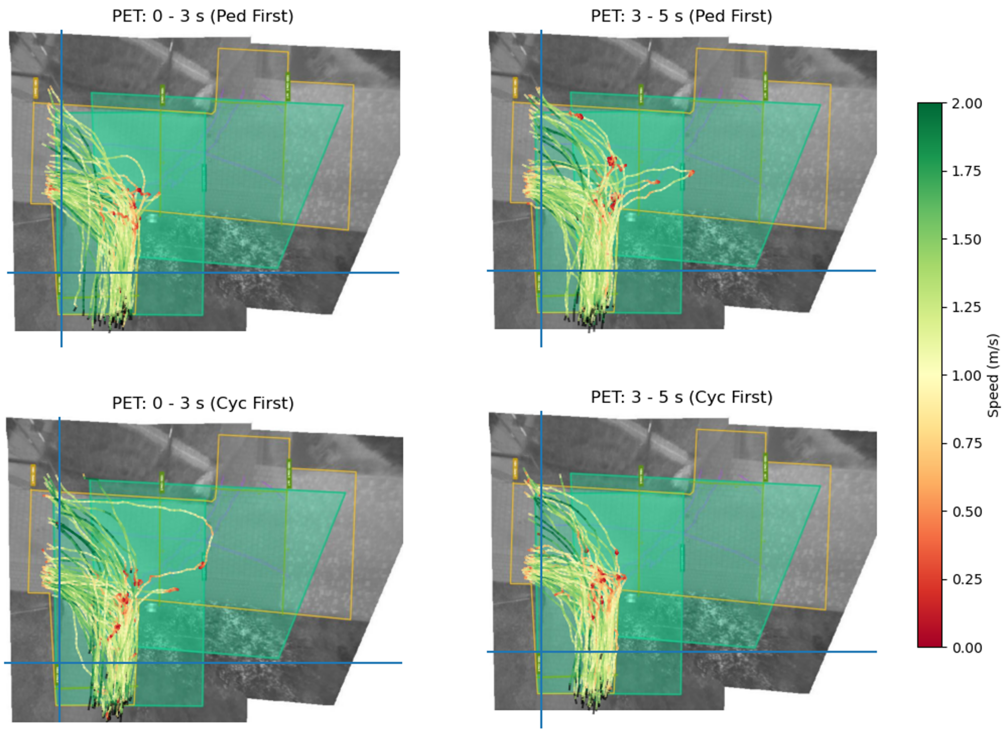


Figure G.3: All crossing pedestrians walking from south to west at the Lorentzweg crossing.

Table G.3: Total number of crossing trajectories within each category of PET and first crossing mode (pedestrians go from south to west) at the Lorentzweg crossing

South - west (total trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	70	74
Cyclist crosses first	77	64

Table G.4: number of crossing trajectories that stop within each category of PET and first crossing mode (pedestrians go from south to west) at the Lorentzweg crossing

South - west (stopping trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	28	32
Cyclist crosses first	40	29

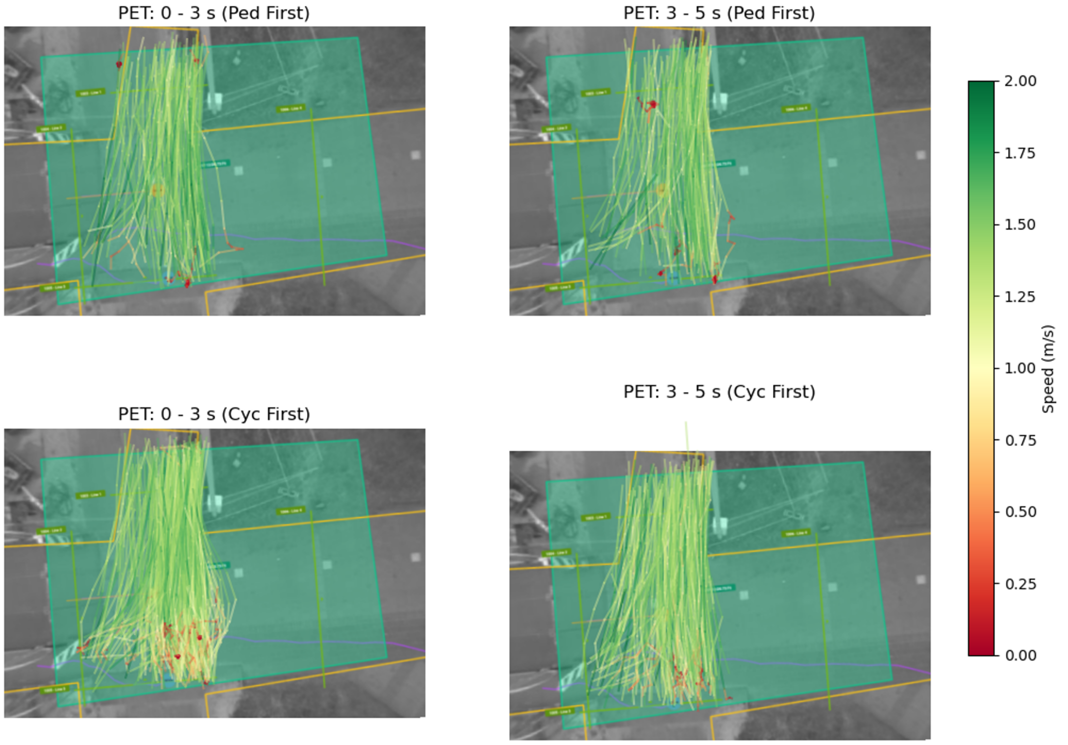


Figure G.4: All crossing pedestrians walking from south to north at the Mekelweg crossing.

Table G.5: Total number of crossing trajectories within each category of PET and first crossing mode (pedestrians go from south to north) at the Mekelweg crossing

South - north (total trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	116	117
Cyclist crosses first	292	190

Table G.6: number of crossing trajectories that stop within each category of PET and first crossing mode (pedestrians go from south to north) at the Mekelweg crossing

South - north (stopping trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	10	13
Cyclist crosses first	94	32

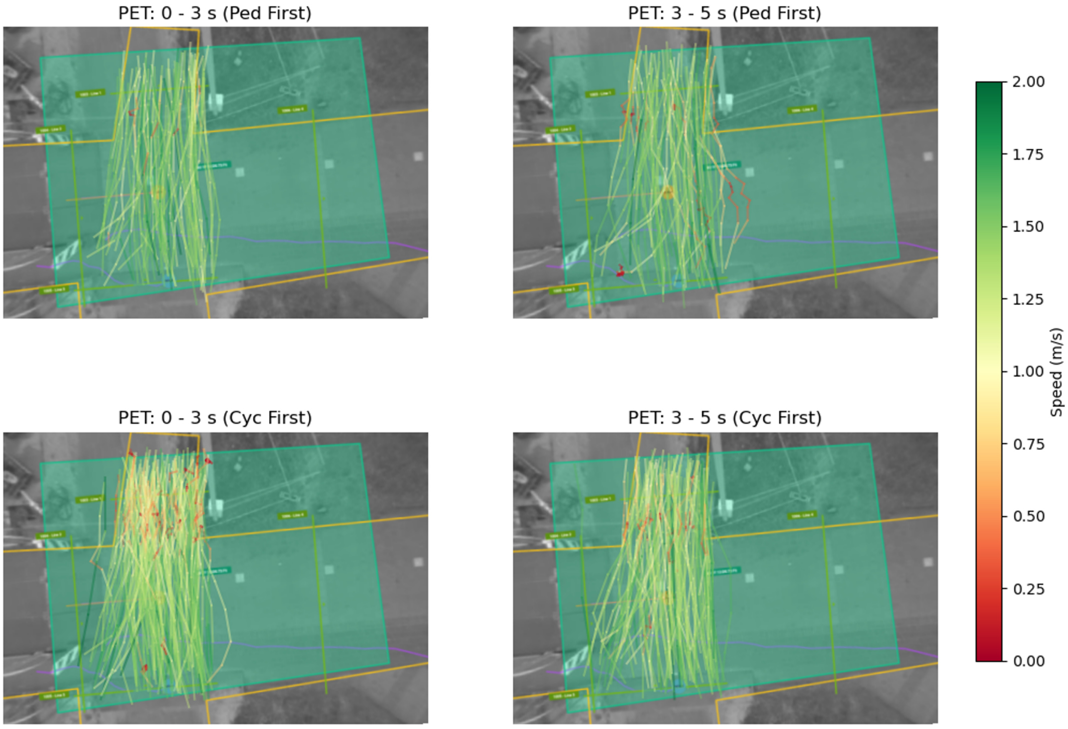


Figure G.5: All crossing pedestrians walking from north to south at the Mekelweg crossing.

Table G.7: Total number of crossing trajectories within each category of PET and first crossing mode (pedestrians go from north to south) at the Mekelweg crossing

North - south (total trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	71	64
Cyclist crosses first	143	105

Table G.8: number of crossing trajectories that stop within each category of PET and first crossing mode (pedestrians go from north to south) at the Mekelweg crossing

North - south (stopping trajectories)	PET 0 - 3s	PET 3 - 5s
Pedestrian crosses first	6	14
Cyclist crosses first	68	25

H

Histograms stopping distance pedestrians Lorentzweg and Mekelweg crossing

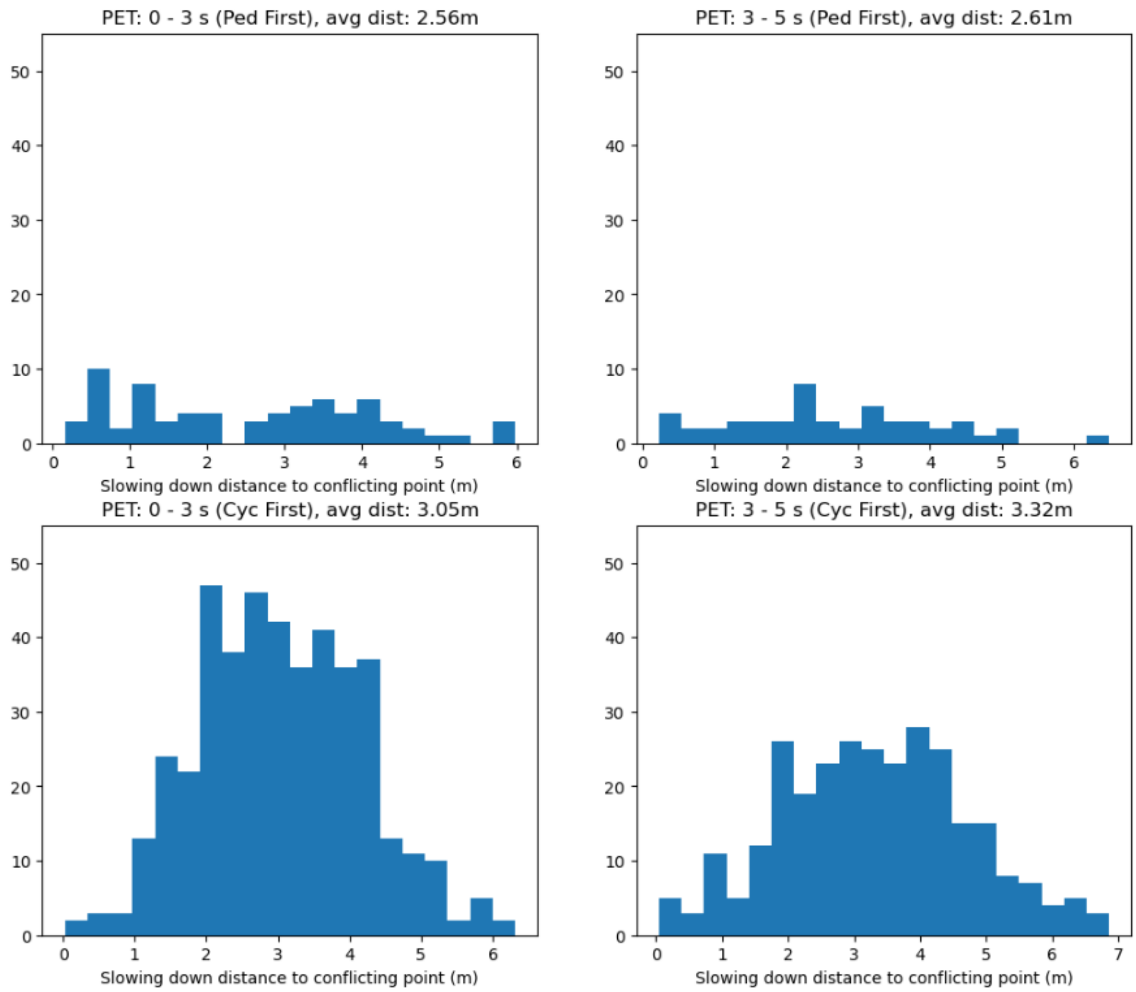


Figure H.1: Frequency of stopping distances for pedestrians walking from east to west crossing with a cyclist at the Lorentzweg crossing.

Table H.1: Mean and standard deviation of the stopping distance up to the crossing point for pedestrians walking from east to west at the Lorentzweg crossing when crossing behind the cyclist.

PET	Mean	Standard deviation
0 - 3 s	3.105	1.119

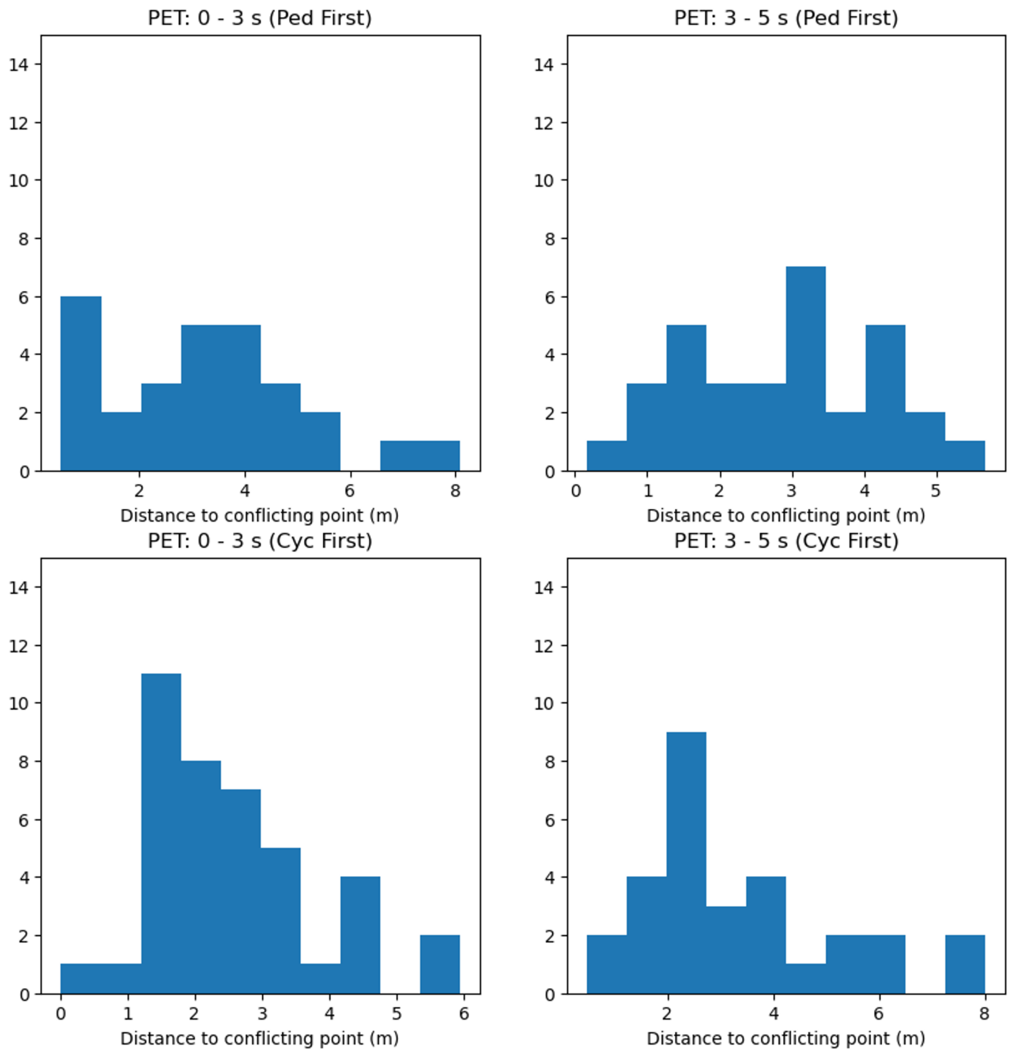


Figure H.2: Frequency of stopping distances for pedestrians walking from south to west crossing with a cyclist at the Lorentzweg crossing.

Table H.2: Mean and standard deviation of the stopping distance up to the crossing point for pedestrians walking from south to west at the Lorentzweg crossing when crossing behind the cyclist.

PET	Mean	Standard deviation
0 - 3 s	2.538	1.230

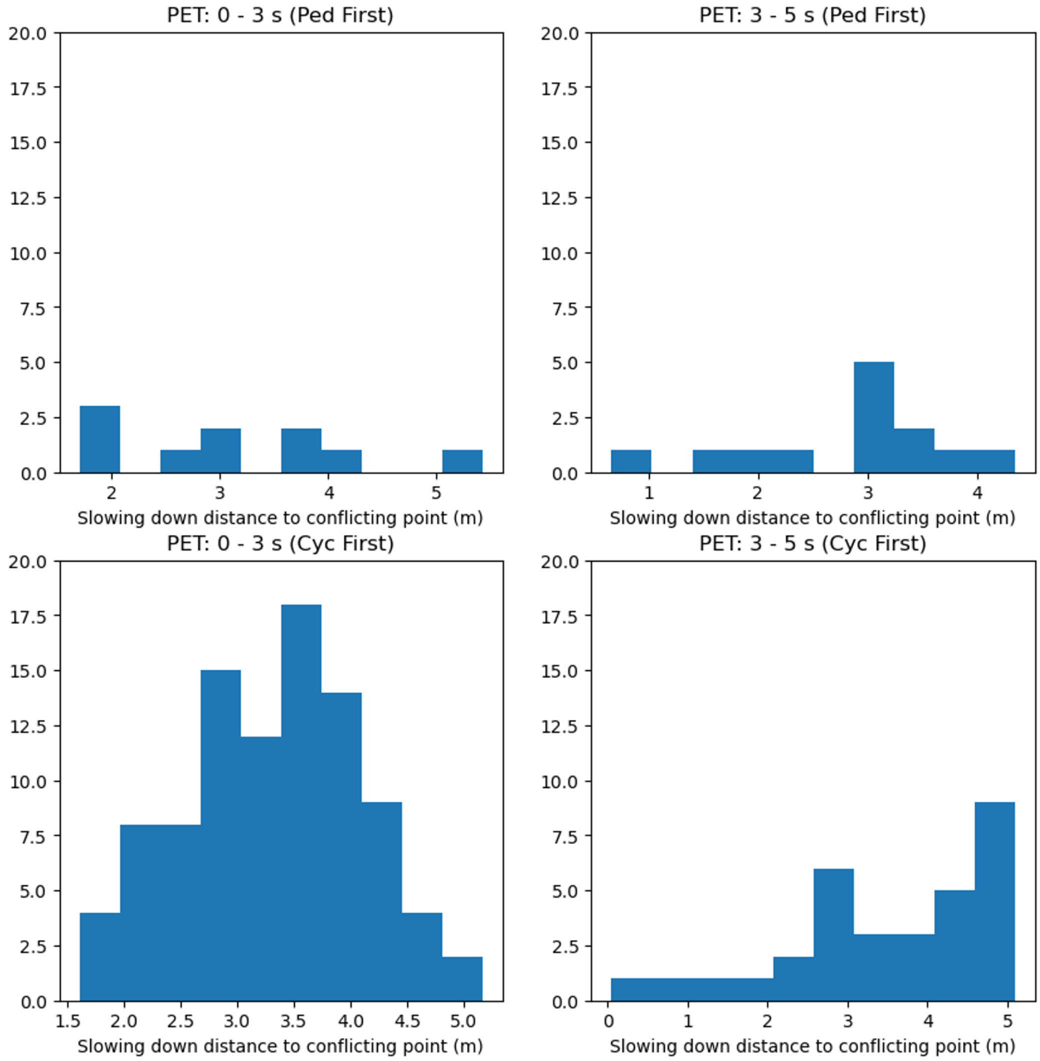


Figure H.3: Frequency of stopping distances for pedestrians walking from south to north crossing with a cyclist at the Mekelweg crossing.

Table H.3: Mean and standard deviation of the stopping distance up to the crossing point for pedestrians walking from south to north at the Mekelweg crossing when crossing behind the cyclist.

PET	Mean	Standard deviation
0 - 3 s	3.299	0.776

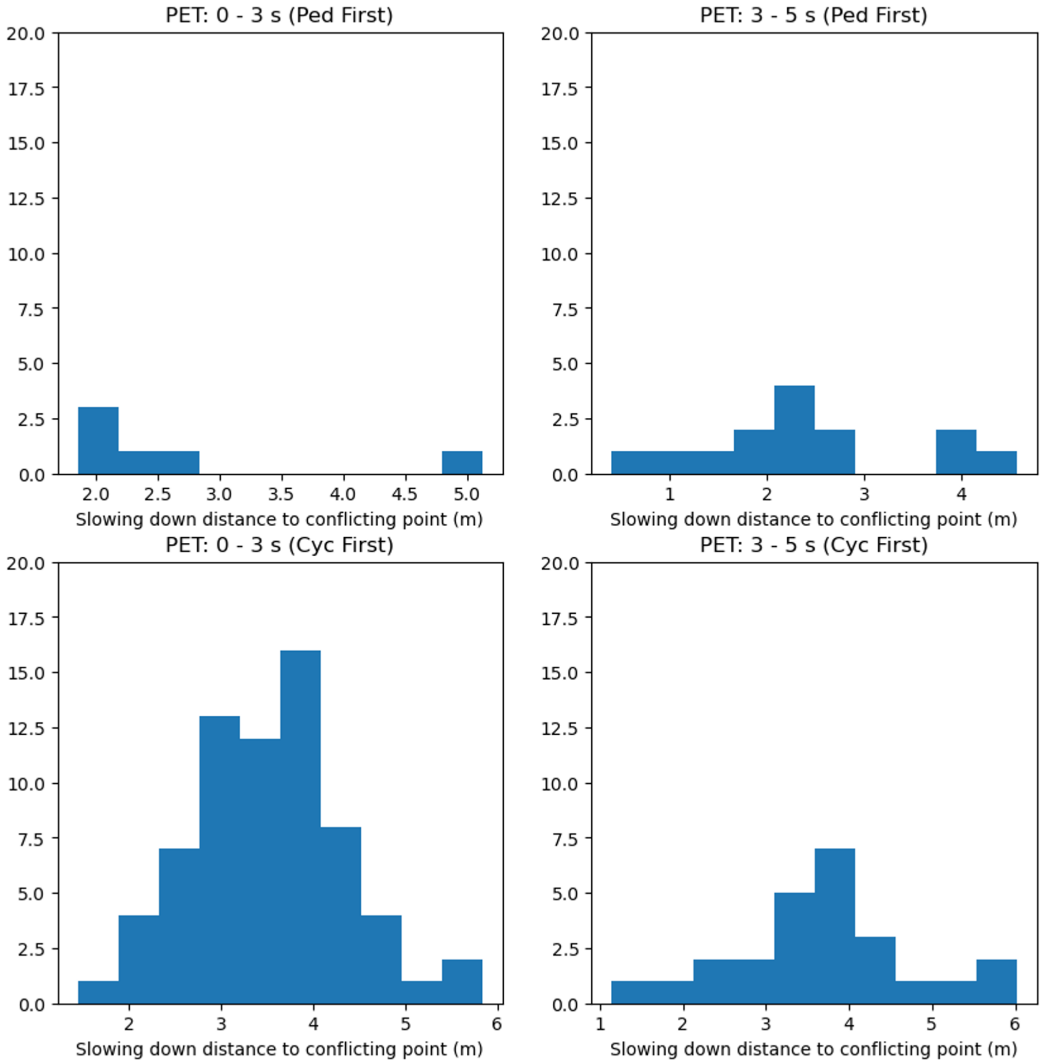


Figure H.4: Frequency of stopping distances for pedestrians walking from north to south crossing with a cyclist at the Mekelweg crossing.

Table H.4: Mean and standard deviation of the stopping distance up to the crossing point for pedestrians walking from north to south at the Mekelweg crossing when crossing behind the cyclist.

PET	Mean	Standard deviation
0 - 3 s	3.511	0.832

I

Deviating paths

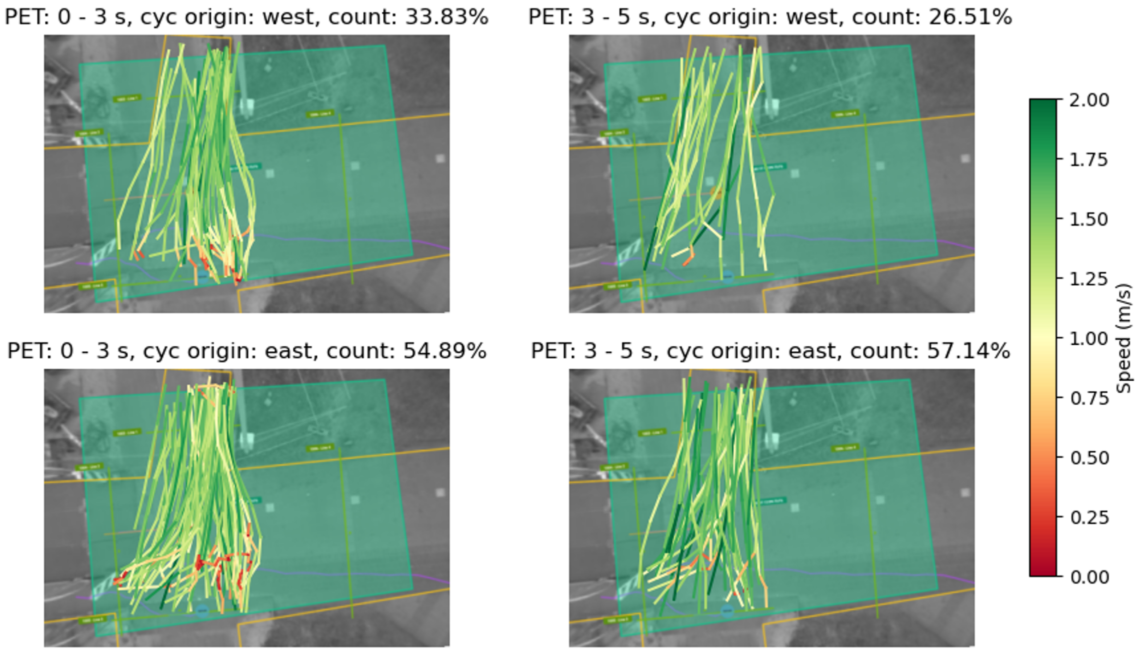


Figure I.1: Deviating trajectories going south to north at the Mekelweg crossing with the direction of the cyclist

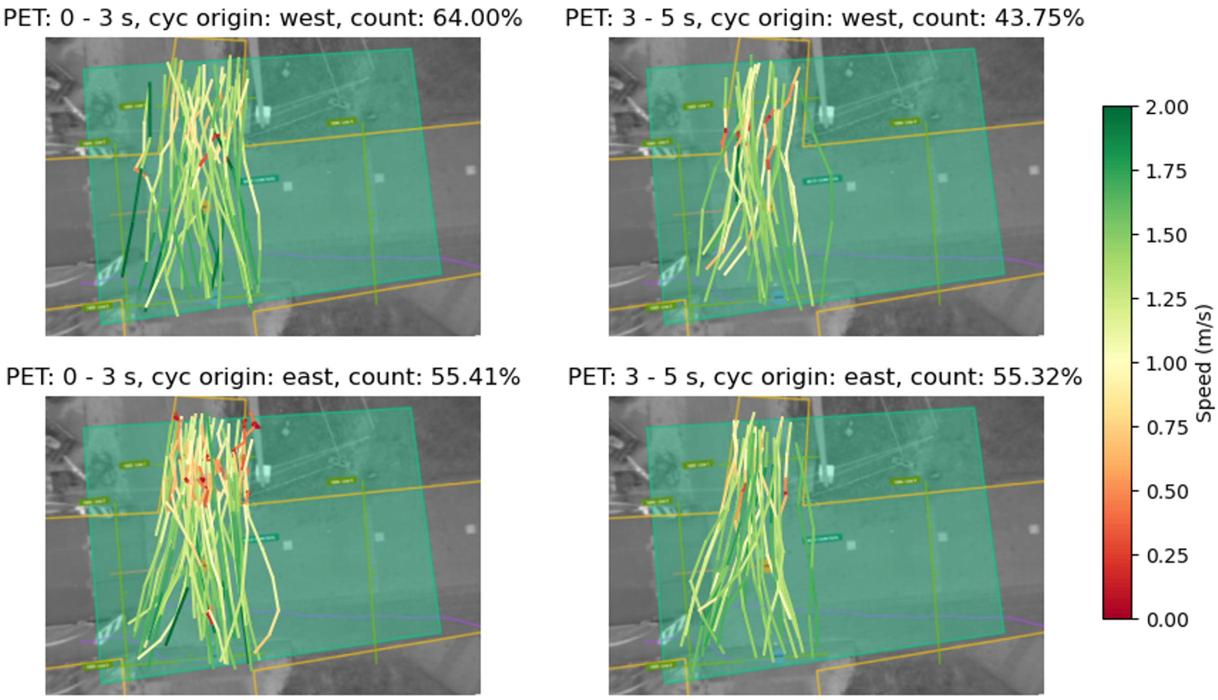


Figure I.2: Deviating trajectories going north to south at the Mekelweg crossing with the direction of the cyclist