TU Delft

Delft University of Technology

Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour

Gavriilidou, Alexandra; Daamen, Winnie; Yuan, Yufei; Hoogendoorn, Serge

DOI
10.1016/j.trc.2019.06.012

Publication date
2019

Document Version
Accepted author manuscript

Published in
Transportation Research Part C: Emerging Technologies

Citation (APA)
Gavriilidou, A., Daamen, W., Yuan, Y., & Hoogendoorn, S. (2019). Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. *Transportation Research Part C: Emerging Technologies*, *105*, 468-484. https://doi.org/10.1016/j.trc.2019.06.012

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Highlights

Modelling Cyclist Queue Formation using a Two-layer Framework for Operational Cycling Behaviour

A. Gavriilidou [a], W. Daamen [a], Y. Yuan [a] and S. P. Hoogendoorn [a]

- Defining and formulating operational behavioural level for cyclist traffic
- Developing a two-layer framework to capture tasks in the mental and physical layers
- Applying framework to model bicycle queue formation at a signalised intersection
- Estimating and validating discrete choice models per layer in a case study in Amsterdam

© 2019 Manuscript version made available under CC-BY-NC-ND 4.0 license
https://creativecommons.org/licenses/by-nc-nd/4.0/

Modelling Cyclist Queue Formation using a Two-layer Framework for Operational Cycling Behaviour

A. Gavriilidou [a*], W. Daamen [a], Y. Yuan [a] and S. P. Hoogendoorn [a]

[a] *Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands*

Abstract

Operational cycling behaviour is greatly understudied, even lacking a definition of what this behavioural level actually entails in terms of decision making. In this paper, we define the cyclist operational level and argue that it consists of two intertwined processes, a mental and a physical process. The mental process refers to path choices made within a route and the physical process refers to the bicycle control dynamics through pedalling and steering. We propose a novel two-layer framework, where each layer captures the tasks of one of the processes within the operational level. Discrete choice theory is proposed to model each layer. The plausibility of the framework is demonstrated through an application focusing on the queue formation process upstream of a red traffic light, including selecting a queuing position and cycling towards it. Models are estimated for the two layers using cyclist trajectory data collected at a signalised intersection in Amsterdam, the Netherlands. The models reveal the attributes that influence the decisions made in each layer and are face validated using simulation. The proposed framework and the (behavioural) findings of its application are the main

*Email: a.gavriilidou@tudelft.nl

scientific contributions of this paper, which pave the way for future research.

Keywords: operational cycling behaviour, two-layer framework, discrete choice theory, cyclist queue formation process, cyclist queue position choice, cyclist trajectories

---

1. Introduction

Though the interest in cycling in cities increases, research on bicycle traffic behaviour is still in its infancy. Insights into this behaviour, and understanding how cyclists interact with each other and make use of cycling infrastructure are crucial if cities are to be designed to accommodate large amounts of cyclists and ensure their safety. Since models can be used to evaluate different designs under varying traffic situations, this need for insights is linked to the need to create reliable and accurate models that can, for example, assess the capacity of intersections or predict the number of encounters on bi-directional cycle paths as a surrogate safety measure.

Research on how cyclists make use of the infrastructure is, however, limited. Among the few examples is Jiang et al. (2013), who studied the gap acceptance of cyclists against right-turning vehicular traffic at signalised intersections. They found that cyclists started decelerating when they are within 30m from the stop line and that their acceptance of a gap depends on the speed of the cyclist and of the motorised vehicle, as well as the size of the available gap. Kucharski et al. (2017) observed the formation of multiple channels in queues at signalised intersections and found that the number of channels formed correlated with the length of the queue. Since they only looked at a single intersection, it is possible that the effect of other factors,

such as the width of the cycle path, has not been identified. In line with this remark, the authors stressed the need for a bigger sample before a model could be formulated to describe the queue formation process.

More research effort has been put on modelling the bicycle control dynamics while riding and interacting with other road users, as several microscopic behavioural models have been developed. Early microscopic cyclist models made use of modelling paradigms developed for cars, such as Cellular Automata models, while adjusting their parameters to reflect the lower speeds of bicycles and their smaller size (Mallikarjuna and Rao 2009; Yao et al. 2009; Vasic and Ruskin 2012). However, the rules governing the movement between cells have not been adjusted to represent cycling behaviour. Another example is the car-following model that was derived for bicycle traffic by Andresen et al. (2013). Even though it was calibrated using empirical cyclist data, the model described single file bicycle flow which is generally not representative of cyclist movements on cycling infrastructure. In addition to these, models stemming from research on pedestrian dynamics were developed to model the microscopic cycling behaviour, such as social force models that determine the movement of cyclists based on attractive forces towards the desired destination and repulsive forces from obstacles and other traffic users, including other cyclists (Li et al. 2011; Liang et al. 2012; Huang et al. 2017). Utility-based models constitute another approach to describe pedestrian dynamics, but their application to cycling is so far scarce. Only one game theoretical approach has been applied and was deemed plausible (Gavriilidou et al. 2019b), but it should be extended to improve its behavioural realism, which is quite cumbersome, due to its complex mathematical derivation.

We claim that the decisions and actions taken by cyclists while riding and interacting with other traffic participants and with the infrastructure belong to the same behavioural level and should be, therefore, modelled together. We refer to this as operational cycling behaviour level and, since a proper definition of what it entails is still missing, we define it in this paper. At the same time we put forward a novel two-layer modelling framework that can be used to capture the mental and physical processes of operational cycling behaviour. Moreover, this paper proposes for the first time the use of discrete choice theory to identify and predict microscopic bicycle traffic flow operations. The third contribution is the application of the proposed approach to model the behaviour of cyclists when they approach a red traffic light. Discrete choice models are estimated for both layers based on trajectory data collected in Amsterdam, and describe the queue formation process, which includes selecting a queuing position and cycling towards it. The estimated models are face validated and reveal the factors that play a role in this process.

The paper is structured as follows. Section 2 defines the cyclist behavioural levels and describes the proposed modelling framework. In section 3 the proposed mathematical model is explained, followed by its application on a dataset described in section 4. The model estimation approach for each layer is discussed in section 5. In section 6 the results of the best performing estimated model for each layer are presented, along with simulation results for face validation. Finally, in section 7 conclusions are drawn and recommendations for future research are made.

2. Conceptual modelling framework

The focus of this paper is on modelling operational cycling behaviour. Since literature describing the behaviour of different modes is not aligned with respect to what operational behaviour entails, a definition needs to first be provided. Figure 1 shows the distinction of the behavioural levels used for car and pedestrian traffic by Michon (1985) and Hoogendoorn and Bovy (2004), respectively, and the one we propose for cycling traffic.

According to Rasmussen (1983), riding a bicycle is a combination of tasks executed based on rules to perform manoeuvres and automatic actions for split-second control of the bicycle. We, therefore, believe that they belong to the same level, the operational level. We adopt the definitions used for pedestrians with respect to the strategic and tactical level (Hoogendoorn and Bovy 2004), whose explanation goes beyond the scope of this paper, and focus on the operational level. The input to this level is the route from one origin to a destination. Within this level, two layers are distinguished, following the concept of a plan-action decision structure proposed by Choudhury et al. (2010) and applied to model pedestrian walking behaviour by Fukuda et al. (2014). In the upper layer, cyclists need to choose intermediate destinations and build up their path within the route while interacting with other traffic users and with the infrastructure. We call this the 'operational mental' layer. Path choices refer, among other things, to yielding, accepting a gap to merge or cross, stopping for a red traffic signal, turning, and overtaking. For the execution of each of these path choices, bicycle control dynamics in the form of pedalling and steering are necessary. This is the lower operational layer which we name 'operational physical' layer.

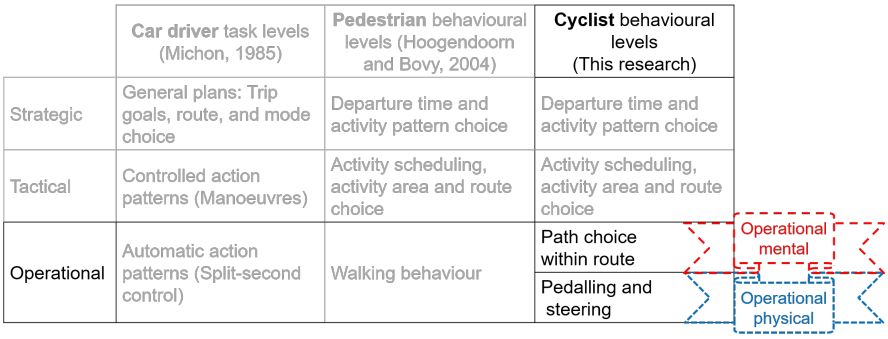


Figure 1: Distinction of behavioural levels for car (left) and pedestrian traffic (middle) found in literature and for cyclist traffic (right) proposed in this paper.

Given this definition, we build upon the conceptual model of Gavriilidou et al. (2019a), shown in Figure 2, which describes cycling behaviour at the operational level, and we fit the two proposed layers within the individual behaviour, as visualised in Figure 3.

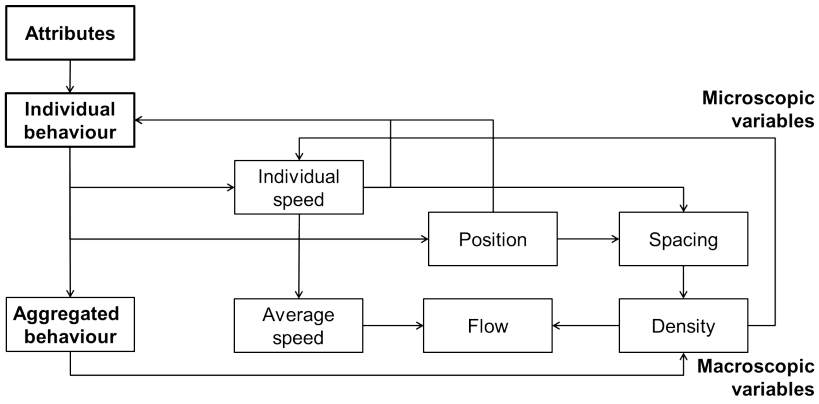


Figure 2: Conceptual model of operational cycling behaviour. Attributes are linked to individual behaviour. Collectively, they lead to aggregated behaviour. These behaviours can be observed via microscopic and macroscopic variables (Gavriilidou et al. 2019a).

Five types of path choices have been identified in the operational mental

layer and each of them is captured by a separate model. The choices correspond to situations when (i) cyclists decide to overtake, but also when they are approaching an unsignalised intersection intending to cross or merge and need to decide (ii) whether to accept a gap in the conflicting traffic stream. Another choice at unsignalised intersections which is at the discretion of the cyclists is to (iii) yield to oncoming traffic. Moreover, situations at a red traffic light are covered, where cyclists decide (iv) whether they stop and (v) where to position themselves in the queue.

We hypothesise that these choices depend on a set of attributes that need to be taken into account by the models. The attributes displayed in Figure 3 for the decisions to overtake, yield, and stop at a red traffic light are the outcome of a stated preference survey we conducted in the Netherlands, discussed in (Gavriilidou et al. 2019a). They still need to be validated with field data, but give good insights into the behavioural attributes. The gap acceptance attributes are taken from (Jiang et al. 2013), even though they studied interactions between bicycles and motorised traffic. For bicycle-to-bicycle interactions on designated cycling infrastructure this list needs to be further investigated. In the application of the framework in this paper, the attributes describing the queue position choice have been investigated and the findings are added to the figure.

The operational physical layer consists of the controls that each individual exerts once a path choice has been made. These controls are steering and pedalling to determine the cycling direction and speed, respectively. This layer is described by a single dedicated model covering steering and pedalling jointly. By applying these controls the state of each individual cyclist

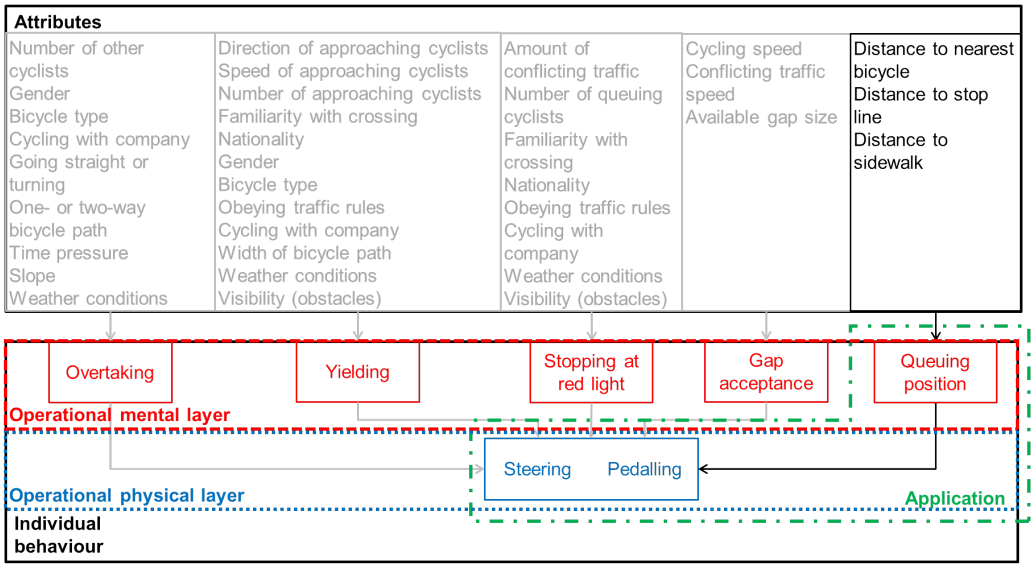


Figure 3: List of influential attributes per choice within the individual mental layer (coloured in red). The choice in the mental layer determines the choices within the operational physical layer (coloured in blue). The elements that go beyond the scope of this paper are coloured in gray, while the scope of the model application in this paper is framed within the green box (built upon (Gavriilidou et al. 2019a)).

(i.e., speed, position and headway) is affected. On an aggregated scale (see Figure 2) they have an effect on density and other macroscopic characteristics which can, then, result in changes in the choices made by each individual cyclist, thereby substantiating an interaction between the two layers. This interaction works in two directions: (i) the choice made in the mental layer is communicated into the physical layer, and; (ii) the new state of the system after applying the decision taken in the physical layer influences the new choice to be made in the mental layer.

3. Mathematical modelling

In order to select the mathematical model that best fits our framework, we should first identify our behavioural assumptions. Those are discussed in subsection 3.1, followed by the description of the models for each operational behaviour layer (subsection 3.2 for the operational mental layer and subsection 3.3 for the operational physical layer).

3.1. Behavioural assumptions

The following assumptions are made regarding the cycling behaviour at the operational level:

1. Cyclists are effort minimisers, motivated by the general principle of least effort. This holds for both layers, though the definition of effort might differ per layer.
2. Decisions in the two layers are made sequentially.
3. The decision made in one layer is input into the other layer.
4. The updating frequency of the mental layer is smaller (i.e., has a longer horizon) than that of the physical layer.
5. When making a decision, cyclists evaluate a set of alternatives using specific attributes.

In line with the framework presented in section 2, we propose a two-layer mathematical model and use discrete choice theory and utility maximisation to model each layer. This allows the identification of the key attributes of the decision making process in each layer. It should be noted that discrete choice models have been used to model the movement of pedestrians (Antonini et al.

2006) and motorcycles (Shiomi et al. 2012; Lee et al. 2009). In motorcycle research, a discrete choice model has also been estimated in the context of queue formation at an intersection (Lee and Wong 2016). One of the main reasons for the lack of such a cycling model is that discrete choice models require empirical data to be estimated and there has been a lack of cyclist trajectory data which we overcome in this paper.

3.2. Modelling the operational mental layer

The use of discrete choice theory to model the operational mental layer is demonstrated through its application on one of the path choices, namely the queue position (green box in Figure 3). The description in this subsection is qualitative, while the quantitative estimations are introduced in subsection 5.1.

In the operational mental layer of this application, cyclists need to decide where to stop in the queue formed upstream of a red traffic light. In our approach, the two-dimensional space is discretised in diamond-shaped cells, since we argue that they represent better the space a bicycle occupies than a rectangular grid. These cells compose the choice set. Each cell is assigned a (dis)utility based on cell attributes and characteristics of the decision maker (cyclist). Using discrete choice theory and the utility maximisation principle, a model can be estimated from cyclist trajectory data, revealing the significant attributes and their relative contribution to the overall cell utility. Availability conditions are also taken into account, since cells that are occupied cannot be re-assigned.

The diamond-shaped grid is visualised in Figure 4, where bicycles already present in the queue (coloured in black) are standing still and a cyclist

(coloured in green) approaches and needs to make a decision. Given this situation, the green cyclist will select a cell that is not yet occupied and gives the highest utility. In this case the red cell is selected and assigned as the intended queuing position of the green cyclist. This is the output of the operational mental layer that is used by the operational physical layer.

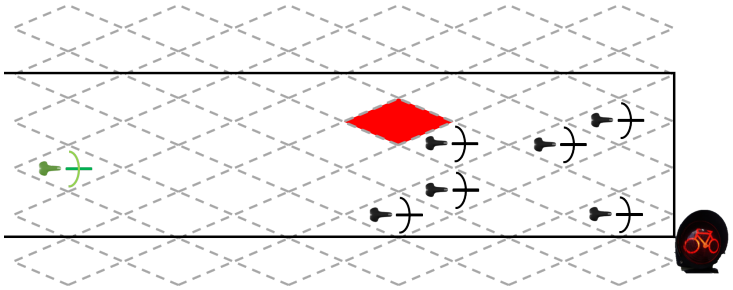


Figure 4: Schematic of the operational mental layer during the queuing process at a (red) traffic light. The green cyclist approaching the traffic light decides in this layer the intended queuing position (red cell) based on the characteristics of the cells, the availability conditions and utility maximisation.

3.3. Modelling the operational physical layer

As already mentioned, the intended queuing position is fed as input to the discrete choice model of the operational physical layer, where the cyclist decides upon the controls to reach this position. The controls are a combination of pedalling and steering, which are expressed as changes in speed and direction relative to the speed and direction, respectively, at the moment the decision is made. The justification of the choice of speed and direction difference as controls over the choice of their corresponding absolute values or the choice of a new position in the two-dimensional space in the next time step relates to the assumption that cyclists are effort minimisers. This means that

they choose the relative effort they are willing to exert in each time step and that goes through changes in pedalling and steering rather than anticipation of their future position.

The choice alternatives are visualised in the fan-shaped individual-specific grid in Figure 5. The fan shape is selected because it reflects the angular movements that characterise cyclist motion. The angular sections capture the radial directions accessible with appropriate changes in steering. The number of angular sections and arched zones is only illustrative and should be determined dependent on the application. In this example, the middle angular section corresponds to no change in the direction, two sections to the right correspond to a small and a bigger steering movement towards the right, and sections to the left steering to the left. The arched zones represent possible relative speed regimes that can be reached through pedalling or braking. The arch closest to the bicycle corresponds to speed reduction (deceleration), the arch furthest away corresponds to an increase in speed (acceleration) and the middle arch corresponds to a choice of no change in speed.

Two more aspects are demonstrated in the figure. One is the sequence of decisions in time within this layer (a lighter shade of grey is given to the grid for decisions to be made in each future time step). The sequence of positions resulting from these choices leads the cyclist to the final position and together forms the cyclist trajectory. The other aspect demonstrated is that the grid is always aligned with the cycling direction of the cyclist at the moment the decision is made. This is shown by the rotation of the grid in each time step, such that the 'no change in direction' alternative is

a continuation of the change in direction that was chosen in the previous time step. The centre of the grid in each time step corresponds to the new location of the cyclist, which depends on the time step, the cycling speed and the choice of change in speed made in the previous step. As the figure is illustrative and no numerical values for speed and time are assigned, the distance separating the different grids is only qualitative.

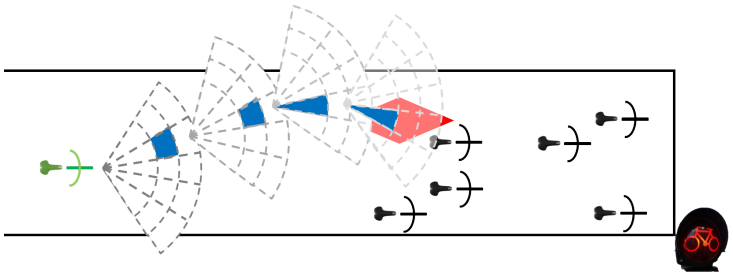


Figure 5: Schematic of the operational physical layer during the queuing process at a traffic light given the intended queuing position (red cell) provided by the operational mental layer. A sequence of decisions (blue cells) is made that corresponds to the combination of angle and speed difference with the highest utility at each time step (grey-scale fans). The sequence of positions resulting from these choices forms the cyclist trajectory.

At each time step, the cell in the fan with the highest utility is selected and the position of the cyclist is updated for one time step, when a new decision is required. It is possible that in between these decision moments there is an interaction with the operational mental layer if a situation occurs that was originally not anticipated by the cyclist and necessitates the estimation of a new intended queuing position (e.g., if another cyclist occupies the originally intended position).

In any case, the result of this interaction between the two layers and the decisions made over time is the cyclist trajectory to reach the intended queu-

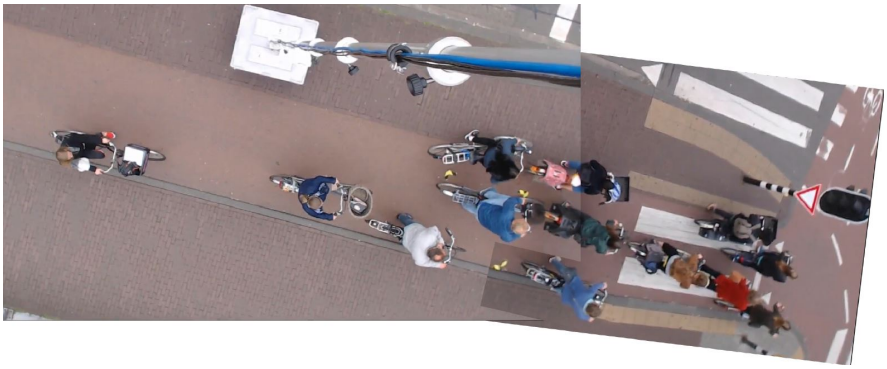
ing position. This trajectory together with the final queuing position fully describe the operational cycling behaviour. A quantitative application of the proposed framework is presented in section 5, where specific models have been estimated using field data from Amsterdam, the Netherlands. Prior to that, the data available for the model identification is described in section 4.

4. Data on queue formation process

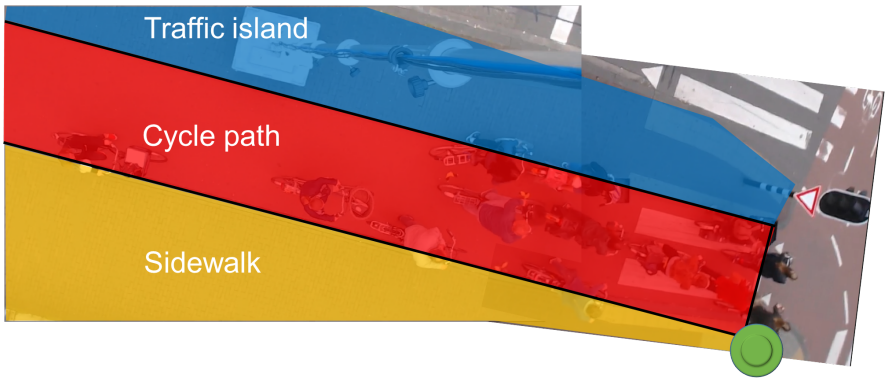
This section presents the dataset used for the model estimation and validation. First, both the site and dataset are introduced (subsection 4.1), followed by a description of the data processing to prepare the dataset to estimate the models (subsection 4.2).

4.1. Site and dataset description

The dataset used for the model estimation contains cyclist trajectories that have been extracted from video camera footage at a signalised intersection in Amsterdam, the Netherlands (Figure 6(a)). The site consists of a 2m-wide unidirectional cycle path that is separated on the left from motorised traffic (even though the cycle path itself can be used by scooters) via a traffic island that has a different surface type but no height difference and can be used by pedestrians and cyclists who queue. The cycle path is also separated from the sidewalk by a curb on the right. These areas are illustrated in Figure 6(b). With two cameras elevated over the cycle path, top-view video images were recorded during an afternoon. The combined view of both cameras covers a length of 20m upstream of the stop line at the traffic light.



(a) Combined view from front and back camera at the intersection. The yellow tape on the cycle path marks the overlapping area between the two cameras.



(b) Study areas of the traffic island, the cycle path and the sidewalk. The black lines denote the edges of the cycle path and a green button is used to show the location of the 'request-green' button.

Figure 6: Top view and areas of interest at the site.

As the aim of our application is to use the two-layer framework to model the queue formation process (i.e., choosing a location to queue and cycling towards it), trajectories of cyclists approaching the traffic light during the red-light phase are collected. We focus on pure bicycle-to-bicycle interactions and therefore, only bicycle trajectories are extracted. Red-light phases where

more than one scooter was present are omitted, as well as phases when there is interaction with crossing pedestrians. The phases with one scooter are kept, assuming that one scooter does not have an effect on the results when considering pure bicycle-bicycle interactions. Only the position it occupies is tracked, so that it is not available to cyclists that arrive later. This way the sample of tracked bicycles increases, and the trajectory of the scooter is ignored.

The last criterion for a red-light phase to be removed from the dataset is the presence of disturbances, such as pedestrians crossing and creating conflicts, bicycles joining the queue from another side or even the sidewalk where they were parked, and cyclists that decided to run the red light even though they had originally queued, thereby initiating movements within the queue.

The final dataset consists of 46 red-light phases with 454 cyclists and 18 scooters in total queuing up. It should be noted that cyclists arriving after the traffic light turned green were not included since their intended queuing position, if any, was not observed. The transition from video files to microscopic cyclist trajectories comprises six steps, which were performed as follows:

1. Decomposition of videos into frames with an average frame rate of 6 frames per second (fps).
2. Manual tracking per frame of the head of each cyclist who approaches the intersection during a red-light phase until standstill.
3. Height transformation to project the trajectories at the head positions to the ground.

4. Orthorectification to correct for the distortion due to the fact that the cameras were placed at an angle and did not point vertically downwards to the cycle path, as well as to compensate for the lens distortion.
5. Time conversion from frame number to seconds.
6. Trajectory merging of the two cameras for each cyclist.

For more information about the data collection and the extraction steps, the reader is directed to the paper by Goñi Ros et al. (2018).

4.2. Data processing

Figure 7 shows the cyclist trajectories, speeds and steering angles of one red-light phase. The extracted raw trajectories contain noise ('*' in Fig. 7), which necessitates a smoothing process to be applied on the data. The smoothing is done by means of a moving average with a fixed-length sliding window across the trajectory data vectors (x: vector of positions in the horizontal direction, y: vector of positions in the vertical direction, t: vector of time instances corresponding to each position). The calculation of the mean is performed for each element of the original vector while centring the window around the corresponding element. Different sliding window lengths were compared (see appendix Appendix A). The results only have limited difference and favour the smoothing of the trajectories over a duration of 6 frames. The smoothed data points are marked by '+' in Figure 7.

Although the average frame rate was 6fps, it was not constant over time, resulting in data points that are not separated by the same time gap. For our application, it is crucial to have a consistent time discretisation throughout all cyclist trajectories as each point corresponds to a moment at which a

decision was made. These points are derived by taking the instant right before these homogenised timestamps and projecting the trajectory in (x,y) using the smoothed speed at that instant. Different time steps were compared (see appendix Appendix A) and a time step of 1 second was found to be best at muting the noise. This means that each cyclist is assumed to make a new decision regarding the steering angle and speed difference every second. The final data points are marked by 'o' in Figure 7.

The trajectories show the path each cyclist followed from the moment they were detected by the back camera up to their final queuing position. The values in both axes have been adjusted for the visualisation, such that the (0,0) point coincides with the location of the 'request-green' button, while in the actual dataset they have positive values that increase in the direction of cyclist movement. The speed and steering angle are computed between consecutive data points and are visualised relative to the horizontal distance that the cyclist has traversed. The horizontal axis of these two graphs has been offset such that all trajectories end at the same point, which facilitates the comparison between the original, the smoothed and the final dataset. A number of observations can be made:

- The trajectories show a good match between the original, the smoothed and the final dataset.
- The smoothing helps reducing the noise in the speed and steering angle.
- The speed is decreasing throughout the observed trajectory which is in line with the findings of Jiang et al. (2013) that deceleration occurs within the 30m upstream of the intersection.

- The steering adjustments are small at the beginning of the trajectory.
- When the speed is low, the steering adjustments increase and they are maximum at the end of the trajectories where the speed is the lowest and the head is swaying more.

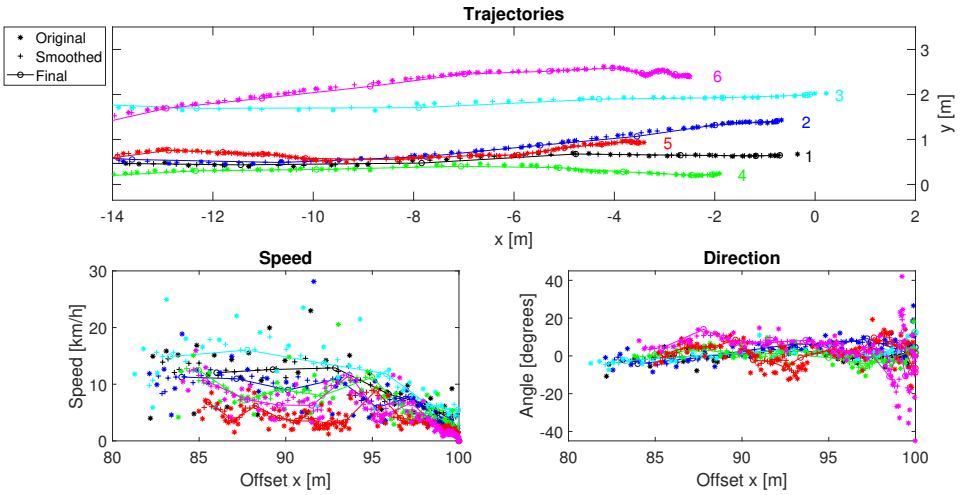


Figure 7: Cyclist trajectories numbered based on their order of arrival (top), speed (bottom left) and steering angle (bottom right) when approaching a red traffic light in the original dataset, when smoothed with a sliding window length of 6 frames and finally when homogenised with time step of 1s. In the top figure the point (0,0) is the location where the stop line meets the curb of the sidewalk, while at the bottom figures the positions in x have been offset such that they end at the same arbitrary point for all cyclists.

These final trajectory points are used for the model estimation. The operational mental layer requires only the last point which corresponds to the queuing position of each cyclist. The operational physical layer takes into account every point as they have been assumed to correspond to a decision point. It should be noted that the reason why the operational mental layer

does not make use of the original dataset is to guarantee the consistency between the two layers.

5. Model estimation approach

In this section the estimation approach for each layer (subsection 5.1 for the operational mental layer and subsection 5.2 for the operational physical layer) is discussed. It includes a justification of the grid choice and selection of attributes to explain the corresponding behaviour, as well as assumptions specific to the model estimation.

5.1. Operational mental layer estimation approach

This layer aims to capture the decision making when joining a queue, where cyclists have to choose their queuing position, as introduced in Figure 4. The observed queuing positions (i.e., last trajectory point of the processed dataset when cyclists are at standstill) are visualised in Figure 8, where the point (0,0) is the location of the 'request-green' button and the red lines indicate the boundaries of the cycle path. As expected, positions next to the 'request-green' button are the most frequently selected. Other positions at the stop line are also favourable, as well as positions on the traffic island, especially for cyclists who want to make a left turn at the intersection. As the queue increases in length, there is a preference for a position next to the curb of the sidewalk rather than a position in the middle of the cycle path.

These choices can be analysed to identify which attributes have an influence, and to what extent, on the queuing position choice by estimating a choice model. The estimation requires the definition of the choice set (5.1.1),

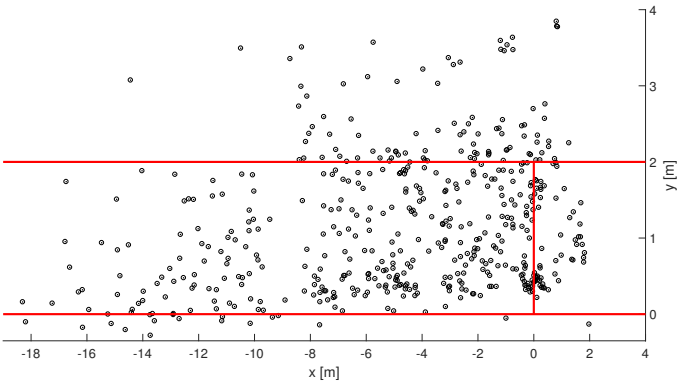


Figure 8: Observed queuing positions. The point (0,0) is the location where the stop line meets the curb of the sidewalk (also the location of the 'request-green' button). The red lines indicate the boundaries of the cycle path.

the specification of the utility functions (5.1.2) and the demarcation of the estimation assumptions (5.1.3).

5.1.1. Choice set definition

The cycle path and surrounding areas (sidewalk and traffic island next to the cycle path) are discretised using the aforementioned diamond-shaped grid. The resulting cells correspond to the choice alternatives of a cyclist. As previously mentioned, this grid better captures the shape of the bicycle (compared to a rectangular grid) and allows for a more realistic representation of queuing.

Each cell is scaled such that it can fit one cyclist based on the standard dimensions of 2m length and 70cm handlebar width (CROW 2016). These dimensions also show a good match with the average observed spacing in the longitudinal and lateral direction between stopped cyclists in our dataset.

The grid is generated in such a way that there is a cell right next to the

sidewalk whose middle crosses the stop line. This is because the 'request-green' button is located right next to the stop line so cyclists stop on top of the line rather than behind it.

The real queuing position is then projected to this grid and assigned to the cell whose centroid is closest to it. This projection results in the top plot of Figure 10, which will be explained in the subsection 6.1. The choice set for each cyclist comprises the cells that are not already occupied by others.

5.1.2. Utility specification

The attributes (cell characteristics) that are hypothesised to capture the attractiveness (utility) of a cell are the distance to the stop line, the distance to the edges of the cycle path and the presence of other cyclists in the queue. Moreover, the behaviour of the first cyclist is hypothesised to be different from the behaviour of the rest, since the first arriving cyclist needs to stop next to the 'request-green' button to be able to press the button to request green. For this reason, the attributes related to the distance to the stop line and the distance to the edges of the cycle path are separately estimated for the first cyclist and for the rest. Given these hypotheses, a general description of the specific attributes is first provided, followed by a full list of the detailed attribute notation and definition.

Regarding the distance to the stop line, cells whose centroid is downstream the stop line are differentiated from those that are upstream. This way the former, i.e. stopping after having crossed the stop line, can be penalised and avoided as a queuing position.

With respect to the distance to the edges of the cycle path, it is hypothesised that the effect on utility is not symmetrical as the distance increases

within and outside of the two edges, because being on the cycle path is desirable, while being on the sidewalk is less comfortable due to the presence of pedestrians and being on the traffic island increases the proximity to motorised traffic. Based on this hypothesis, the area covered in the choice set is subdivided into four sublanes, namely the sidewalk, the right lane of the cycle path, the left lane and the traffic island. The right edge of the cycle path is taken as reference for the first two sublanes and the left edge as reference for the last two.

The presence of other cyclists can be represented in several ways. Therefore, more than one attribute is defined. One way is the distance to the nearest cyclist in the queue. Another considers the number of cyclists within each of the aforementioned sublanes, as the more cyclists stopped within a sublane, the less attractive the sublane becomes, because cyclists cannot manoeuvre to overtake and would need to join the end of the queue. This end of the queue may also be seen as an offset of the stop line within each sublane, i.e. cyclists have a higher utility in stopping closer to the end of the queue. An attribute is therefore added that considers the distance to the cyclist at the back of the queue of the sublane.

The full list of attributes for this layer is given below:

- X_{F} [-]: dummy indicating whether the cyclist is the first one arriving.
- X_{button} [-]: dummy to denote if a cell is the cell next to the 'request-green' button ($X_{\mathrm{button}} = 1$).
- X_{d2stop} [m]: longitudinal distance between the location of the stop line and the centroid of the considered cell.

- X_{up} [-]: dummy indicating whether the considered cell is upstream or on the stop line.
- X_{onside} [-]: dummy to denote if a cell is on the sidewalk ($X_{\text{onside}} = 1$).
- X_{onisland} [-]: dummy to denote if a cell is on the traffic island ($X_{\text{onisland}} = 1$).
- X_{rightln} [-]: dummy to denote if a cell is on the right lane of the cycle path ($X_{\text{rightln}} = 1$).
- X_{leftln} [-]: dummy to denote if a cell is on the left lane of the cycle path ($X_{\text{leftln}} = 1$).
- X_{d2Redge} [m]: absolute lateral distance between the location of the right edge of the cycle path and the centroid of the considered cell.
- X_{d2Ledge} [m]: absolute lateral distance between the location of the left edge of the cycle path and the centroid of the considered cell.
- $X_{\text{d2nearEucl}}$ [m]: Euclidean distance between the centroid of the considered cell and the one closest to it that is occupied by cyclists already standing in the queue.
- X_{d2nearX} [m]: minimum absolute longitudinal distance between the location of the centroid of the considered cell and those already occupied by cyclists standing in the queue.
- X_{d2nearY} [m]: minimum absolute lateral distance between the location of the centroid of the considered cell and those already occupied by cyclists standing in the queue.

- X_{total} [cyclists]: the total number of cyclists within the sublane of the considered cell.

- X_{d2lastX} [m]: longitudinal distance between the last queuing cyclist in the sublane where the considered cell belongs and the centroid of the considered cell, if the considered cell is upstream.

In the construction of the systematic part of the utility functions (V), interaction terms among these attributes are used in a linear weighted summation. The weights (coefficients) are denoted by β and are generic for all alternatives as there is no straightforward way to classify them in nests that would acquire alternative specific weights. An example utility function of a cell c (V_{c}) in a model where only the interaction term between the dummy X_{F}, the dummy X_{up}, and the variable X_{d2stop} is considered, is shown in Equation 1.

$$\begin{aligned} V_{\text{c}} = & \\ & \beta_{\text{upF}} * X_{\text{F}} * X_{\text{d2stop}_{\text{c}}} * X_{\text{up}_{\text{c}}} + \beta_{\text{upR}} * (1 - X_{\text{F}}) * X_{\text{d2stop}_{\text{c}}} * X_{\text{up}_{\text{c}}} \\ & + \beta_{\text{downF}} * X_{\text{F}} * X_{\text{d2stop}_{\text{c}}} * (1 - X_{\text{up}_{\text{c}}}) \\ & + \beta_{\text{downR}} * (1 - X_{\text{F}}) * X_{\text{d2stop}_{\text{c}}} * (1 - X_{\text{up}_{\text{c}}}) \end{aligned} \tag{1}$$

5.1.3. Estimation assumptions

When estimating a model within this layer, the following assumptions are made to simplify the estimation process:

1. There is no correlation between the alternatives (Independence of Irrelevant Alternatives, IIA property) and therefore, a multinomial logit model can be used.

2. Queuing spots are assigned upon the entrance of a cyclist in the camera vision field, which means that this decision precedes any decisions on the operational physical layer and that the assignment of spots follows the order of arrival of cyclists.
3. The assigned queuing spots are not updated over the course of the cyclist trajectory. This means that the interaction between the two layers is in this application one-way.
4. Cyclists are assumed to be aware of the spot selected by their predecessors. This is imposed through availability conditions in the logit model, which remove cells that are already assigned to a cyclist from the choice set of oncoming cyclists.

These assumptions can later be relaxed, e.g., by considering spatial correlation between the diamond-shaped cells and by allowing the updating of the decision for the intended queuing position. The spatial correlation should be considered in future research as cyclists do not see the diamonds on the cycle path but might apply different discretisation of space in areas combining several cells. This would require the definition of those areas and the estimation of a mixed logit model. The decision updating becomes relevant when unanticipated changes take place, such as a cyclist entering the cycle path from another direction and occupying the originally desired position. Another reason to consider updating is when speed differences are large and overtaking might place. In this case, the first come first serve rule might need to be replaced by a rule based on cycling speed. Since the dataset does not contain disturbances of sudden appearing cyclists and there is no information on desired queuing positions other than the revealed one, these effects are

righteously ignored.

5.2. Operational physical layer estimation approach

Within this layer, the intended queuing position is known and the cyclist decides in every time step the changes in pedalling and steering until the next time step. In order to reach the intended queuing position, a sequence of time steps, and corresponding decisions, is needed and results in the cyclist trajectory.

The estimation requires the definition of the choice set (5.2.1), the specification of the utility functions (5.2.2) and the demarcation of the estimation assumptions (5.2.3).

5.2.1. Choice set definition

In the physical layer, cyclists decide whether they will change their speed and direction at the current time step. By looking at the combinations that occur in the processed dataset the discretisation of the fan-shaped grid of Figure 5 can be motivated.

The observed choices are visualised by the blue dots in Figure 9. It shows that most observations are concentrated around zero in the angle difference and more specifically in the boundary of [-15,15] degrees. Larger changes in the steering angle only take place when there is no, or very small, change in the speed (ΔSpeed between [-2,2] km/h). From further inspection of the dataset, it is noted that this coincides with very low speeds. Moreover, regarding the speed changes, most observations are negative, which is in line with the findings of section 4 that cyclists are already decelerating when entering the observed area.

Following these insights, for our application the fan-shaped choice grid is defined to range from speed changes of -12 km/h to +8 km/h with a step of 2 km/h, and the steering angle changes included are {-45,-30,-15,-10,-5,0,5,10,15,30,45} degrees. The observed choices are then assigned to their closest grid point and choices that would result in a negative cycling speed are made unavailable.

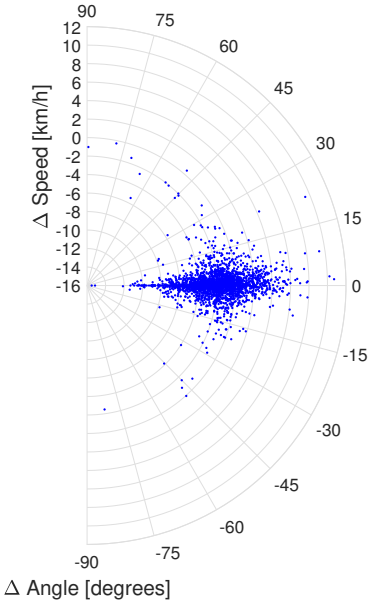


Figure 9: Observed choices of changes in steering angle and speed.

5.2.2. Utility specification

Based on the observation that the distance to other cyclists and the curb play a role in the decision for the queuing position, we hypothesise that they also affect the path that is chosen to reach that position. In this decision layer, the position of the stop line is less relevant but what might have an effect is the distance to the intended queuing position. When considering

this distance, one can differentiate between choices that result in passing the intended queuing position and choices that do not. This way the former can be penalised and avoided towards the end of the trajectory. Additionally, it is hypothesised that the behaviour towards moving cyclists differs from the behaviour against stopped cyclists or obstacles. This discrepancy is captured by the distance to the nearest bicycle and the difference in cycling speed, which is calculated separately for moving and for stationary bicycles. The full list of attributes is given below:

- X_{d2dest} [m]: Euclidean distance between the destination (i.e., the intended queuing position) and the location to be reached within a time step if the considered change in speed and angle is chosen.

- X_{pass} [-]: dummy indicating whether in the next step, given the considered change in speed and angle, the cyclist will have traversed a longer longitudinal distance than needed to reach the intended queuing position.

- X_{d2Mov} [m]: minimum Euclidean distance between the anticipated positions of the cyclists in front who are moving, and the location to be reached within a time step if the considered change in speed and angle is chosen.

- X_{d2Stop} [m]: minimum Euclidean distance between the anticipated positions of the cyclists in front who are stopped, and the location to be reached within a time step if the considered change in speed and angle is chosen.

- X_{spdMov} [m/s]: maximum speed difference between the considered speed and that of cyclists in front who are moving.

- X_{spdStop} [m/s]: speed to be reached if the considered change in speed is chosen and if there are cyclists in front who are stopped. Since the stopped cyclists have no speed, this attribute reflects the disutility of having the considered speed when others have stopped.

- X_{step} [-]: dummy indicating whether in the next step the cyclist will need to get on or off the curb of the sidewalk.

- X_{offpath} [-]: dummy indicating whether in the next step the cyclist will need to get on or off a the traffic island.

These attributes are used in a linear weighted summation to construct the systematic part of the utility functions (V), with the exception of the distance to the destination, which is covered by an interaction term between the dummy X_{pass} and the X_{d2dest}. An example utility function of alternative a (V_{a}) in a model where only this interaction term is considered, is shown in Equation 2.

$$V_{\text{a}} = \beta_{\text{over}} * X_{\text{d2dest}_{\text{a}}} * X_{\text{pass}_{\text{a}}} + \beta_{\text{under}} * X_{\text{d2dest}_{\text{a}}} * (1 - X_{\text{pass}_{\text{a}}}) \qquad (2)$$

5.2.3. Estimation assumptions

When estimating a model within this layer, we make the following assumptions:

- Changes in cycling speed and direction are decided simultaneously.

- Cyclists do not move backwards and so no negative speeds are allowed.

- Decisions made by the same person at different time steps are independent. As no serial correlation is assumed, a multinomial logit model is to be estimated.
- When other cyclists are present, there is full knowledge of their current speed and direction.
- Zero acceleration is assumed for the other cyclists and so their anticipated position within one time step can be estimated based on their current cycling speeds.
- There is no memory from previous time steps. Only cyclists within the vision field at the current position affect the decision to be made.
- The vision field contains everything that is in front of or at least at the same longitudinal position as the cyclist making a decision.

The pitfall of ignoring the serial correlation is that bias of an individual towards a certain type of behaviour is overlooked, and the risk of inconsistent behaviour between time steps is introduced. Even though it decreases the model realism, we argue that it is an acceptable simplification to get first insights into the operational physical layer. If this assumption is relaxed and panel data are considered, a mixed logit model should be estimated. In terms of anticipation, the assumption of zero acceleration is reasonable as it cannot be expected that the intentions of others are known. The full knowledge of the speed and position assumption could be relaxed by considering some noise rather than the exact measurements. Last but not least, future research should introduce a memory function to improve the anticipatory skills of cyclists and increase the model realism.

6. Results and discussion

This section provides and discusses the model estimation results for each layer. Models have been estimated using Python Biogeme (Bierlaire 2016). The best performing model is found based on goodness of fit measures, i.e. $\bar{\rho}^2$, AIC and BIC criteria, and is presented in subsection 6.1 for the operational mental layer and in subsection 6.2 for the operational physical layer. Both models are face validated by means of simulation with Biogeme. The simulation results are discussed in subsection 6.3.

6.1. Operational mental layer model

The estimated values of the coefficients of the best performing model are shown in Table 1, along with their robust statistics. The model consists of 12 parameters, one third of which captures the behaviour of the first arriving cyclist. These four attributes are differently weighed from the corresponding ones for the rest of the cyclists, which confirms the hypothesis made in subsection 5.1.2 that the behaviour of the first cyclist is different.

The cell next to the 'request-green' button has a positive coefficient ($\beta_{\text{buttonF}} = 1.24$) for the first cyclist, while it does not affect the utility for the rest of the cyclists. Moreover, the utility decreases the further the queuing position is from the stop line. This disutility is greater in the case of crossing the stop line and stopping further downstream ($\beta_{\text{downF}} = -2.13$) compared to stopping upstream of the stop line ($\beta_{\text{upF}} = -1.18$). Regarding the distance to the edges, as most cyclists stop within the cycle path, the only attribute with sufficient observations to estimate a coefficient is the distance from the curb of the sidewalk to a position within the right lane of

the cycle path. This coefficient has a positive value ($\beta_{\text{rightlnF}} = 4.91$), which shows that first arriving cyclists prefer to be close to the middle of the cycle path (i.e. at the end of the right lane). This is reasonable since it serves the purpose of stopping next to the 'request-green' button.

For the rest of the cyclists, three coefficients are estimated concerning the distance to the edges. There is an increase in utility by being on the right lane ($\beta_{\text{rightlnR}} = 1.21$), and a decrease by being on the sidewalk ($\beta_{\text{onsideR}} = -6.46$) and on the traffic island ($\beta_{\text{onislandR}} = -1.85$). The difference in magnitude of the inflicted disutility can be explained by the fact that the sidewalk is primarily intended for use by pedestrians, while the traffic island can be used by cyclists, especially if they want to turn left at the intersection. Also for these cyclists, there is a disutility the further downstream the queuing position is from the stop line ($\beta_{\text{downR}} = -1.29$). The coefficient of the attribute describing the distance to the stop line for cells upstream the stop line is positive ($\beta_{\text{upR}} = 0.30$), which might seem counter-intuitive, but can be explained by its negative correlation with the distance to the nearest bicycle, as well as with the distance to the last stopped cyclist within a sublane. These two have negative coefficients $\beta_{\text{d2nearX}} = -0.53$ and $\beta_{\text{d2lastX}} = -0.22$, respectively, which indicates that cyclists want to stay close to each other in the queue in the longitudinal direction. Since the last stopped cyclist within a sublane is considered as an offset of the stop line, it is reasonable that the coefficient is negative and the arriving cyclist wants to stay as close as possible to the adjusted stop line. The last attribute of the model captures the effect of the number of queuing cyclists within a sublane and expected has a negative coefficient ($\beta_{\text{total}} = -0.39$). This shows that the more cyclists stopped within

a sublane, the lower the utility of that sublane and therefore, it is more likely that the arriving cyclist will choose to stop in another sublane.

Table 1: Estimated model parameters for the operational mental layer.

Coefficient name	Coefficient value	Robust standard error	Robust t-test	Robust p-value
β_{buttonF}	1.24	0.37	3.31	0.00
β_{upF}	-1.18	0.26	-4.58	0.00
β_{downF}	-2.13	0.55	-3.90	0.00
β_{rightlnF}	4.91	0.75	6.57	0.00
β_{upR}	0.30	0.03	9.70	0.00
β_{downR}	-1.29	0.31	-4.14	0.00
β_{rightlnR}	1.21	0.21	5.89	0.00
$\beta_{\text{onislandR}}$	-1.85	0.23	-8.05	0.00
β_{onsideR}	-6.46	1.04	-6.19	0.00
β_{d2nearX}	-0.53	0.05	-10.90	0.00
β_{total}	-0.39	0.06	-6.43	0.00
β_{d2lastX}	-0.22	0.04	-6.30	0.00

6.2. Operational physical layer model

The estimated values of the coefficients of the best performing model are provided in Table 2, along with their robust statistics. All values are statistically significant, which confirms that the hypothesised attributes influence the choices made with respect to changes in speed and cycling direction.

Moreover, the coefficient values prove that the behaviour towards stopped

and moving cyclists is indeed different, especially when considering the distance. There is a higher disutility when getting closer to a stopped cyclist than to a moving one. This can be explained by the fact that the moving cyclist continues to change position, while stopped cyclists form a (static) obstacle when the intended queuing position is not adjacent to them.

With respect to the distance to the intended queuing position, there is a penalty for changes in speed and angle that increase this distance. The penalty is bigger when the position is passed, which is reasonable. Cyclists should not be willing to cycle further than their intended queuing position.

Another valuable insight is the disutility of having to go on and off the cycle path at consecutive time steps. As expected, the disutility is much higher on the side of the sidewalk due to the presence of the curb, while on the side of the traffic island the surfaces are on the same level and only the surface type changes.

6.3. Face validation using simulation

Using the estimated parameters for each model, a simulation is performed, where a prediction is made for each observation in the dataset. This means that the attributes and availability conditions describing the situation at which every individual made a decision remain the same. The simulation uses the estimated model to compute all utility functions and the probabilities of each alternative. The individual's probabilities of an alternative are aggregated by averaging over all individuals to whom the corresponding alternative was available. The true (observed) choices can then be compared with the predicted (simulated) ones.

The comparison for the operational mental layer is visualised in Figure 10,

Table 2: Estimated model parameters for the operational physical layer.

Coefficient name	Coefficient value	Robust standard error	Robust t-test	Robust p-value
β_{under}	-1.01	0.05	-19.61	0.00
β_{over}	-2.04	0.10	-20.75	0.00
β_{d2Mov}	-0.40	0.04	-9.06	0.00
β_{d2Stop}	-0.24	0.05	-5.24	0.00
β_{spdMov}	-0.93	0.06	-15.73	0.00
β_{spdStop}	-0.61	0.06	-10.77	0.00
β_{step}	-2.46	0.20	-12.44	0.00
β_{offpath}	-0.83	0.09	-9.08	0.00

where the white dot at point (0,0) is the location of the 'request-green' button and the red lines indicate the boundaries of the cycle path. The observed choices show a preference for the right lane of the cycle path, which is well reproduced in the simulation results. Moreover, the choice of the first arriving cyclists to stop next to the 'request-green' button is very accurately replicated by the simulation. Another observation is that cyclists at the front of the queue are likely to stop at any lateral position on the cycle path or on the traffic island, while the longer the queue grows, the most preferable the right lane becomes, possibly due to the presence of the curb so the cyclists can rest their foot. This trend is also captured by the model; the front cells on the left lane have a higher probability than those upstream and the probability decreases with the longitudinal distance. These simulation results are, therefore, considered a good representation of reality.

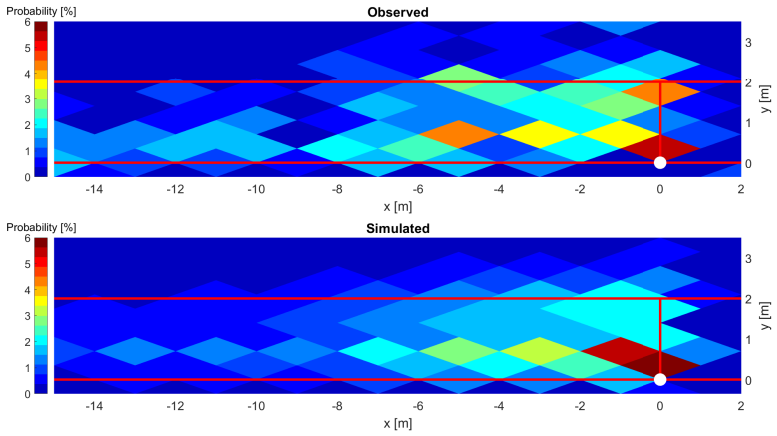


Figure 10: Probability of a diamond cell being selected as the queuing position in the observed (top) and the simulated (bottom) dataset. The white dot, point (0,0), is the location where the stop line meets the curb of the sidewalk (also the location of the 'request-green' button). The red lines indicate the boundaries of the cycle path.

The comparison for the operational physical layer is visualised in Figure 11. The pattern displayed in the two fans is similar with observed choices having less variance and thus higher probability values for no change in direction and slight deceleration, while the simulated choices are more scattered. The trend of speed reductions and small changes in the steering angle is captured well by the simulation. In order to present these results more quantitatively in a single assessment value, the positions in x and y that result from the choices of speed and angle change are calculated. The absolute percentage error made in each observation i can be computed per direction

(i.e., x and y) using the formulas

$$\begin{aligned} x_{\text{error}_i} &= \frac{|x_{\text{sim}_i} - x_{\text{obs}_i}|}{x_{\text{obs}_i}} \\ y_{\text{error}_i} &= \frac{|y_{\text{sim}_i} - y_{\text{obs}_i}|}{y_{\text{obs}_i}} \end{aligned} \tag{3}$$

The mean absolute percentage error (MAPE) in the longitudinal x direction is 4.37% and in the lateral y direction 1.79%. These values are considered low and prove that the model generates plausible results.

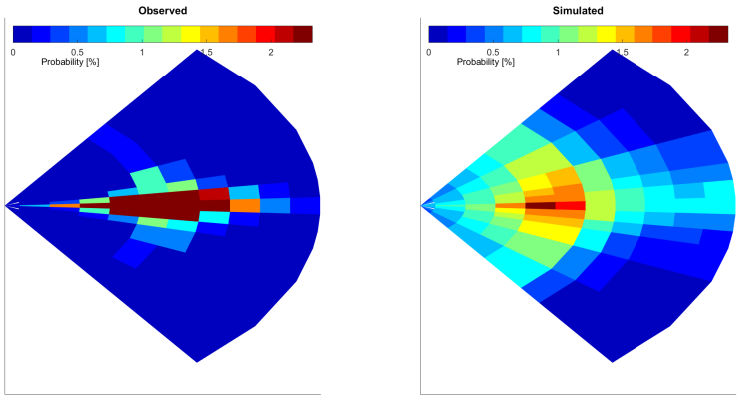


Figure 11: Probability of a combination of change in steering angle and speed to be selected in the observed (left) and the simulated (right) dataset.

7. Conclusions and recommendations

In this paper, different behavioural levels for cyclists have been defined for the first time, while focusing on the operational level. We hypothesised that this level consists of two intertwined processes, namely the path choices made within a route and the bicycle control dynamics through changes in

pedalling and steering. We put forward a novel two-layer framework to capture the tasks within the mental and physical layers of the operational level. Discrete choice theory was proposed to model each layer and the plausibility of the framework was demonstrated through an application. Using cyclist trajectory data from a signalised intersection in Amsterdam, the Netherlands, models were estimated and face validated. The models describe the behaviour of cyclists when approaching and queuing at a red traffic light, including selecting a queuing position (operational mental layer) and cycling towards it (operational physical layer). The models reveal the attributes that influence queuing behaviour.

For this specific application of the modelling framework, we found that when deciding on a queuing position, the first arriving cyclist behaves differently than the rest as there is the need to stop next to the 'request-green' button and press it. Additionally, cyclists prefer to stop on the right lane of the cycle path and upstream of the stop line. Positions on the sidewalk are less favourable than those on the traffic island, because the former are hindered by pedestrians on the sidewalk, while the latter can be attractive for left-turning cyclists. Furthermore, cyclists prefer to stop close to each other, but once a sublane becomes crowded, they prefer to go to another sublane. This disutility is traded off with their desire to stay on the right lane and once the front stopping positions are occupied, there is a trend to stop closer to the curb of the sidewalk rather than build up all sublanes equally. These results are intuitive because cyclists want to use the curb as a resting position when stopped and as an assist when accelerating. Once this intended queuing position is decided upon, the cyclists need to create a trajectory towards it,

which they do through changes in their speed and steering angle at regular time intervals with the aim of reaching that position. Based on our estimation results, cyclists behave differently towards stopped and moving cyclists, which is reasonable since stopped cyclists form an obstacle on the way and an increase of the speed difference might lead to unsafe situations that are preferably avoided. Moreover, they are attracted by their intended queuing position and deter from passing it. They additionally deter from changing surface type and, even more strongly, from stepping on and off curbs.

These findings provide valuable insights for the design of cycling infrastructure. One way to avoid long sparse queues would be to provide an elevated curb on both sides so that cyclists can use it as a resting spot. This elevation is also advisable to prevent cyclists from leaving the cycle path and interfering with pedestrian traffic. When it is not possible, changing the surface type can be an alternative measure. The reason why sparse queues should be avoided is the fact that dense queues have shorter discharge times (Goñi Ros et al. 2018), so their green phase and the cycle time of the intersection can be reduced.

The simulation results reproduce patterns observed in the empirical data and thereby demonstrate the face validity of the models. However, there is room for improvement, which could be sought in including other attributes, such as the time of day or weather conditions, or adding personal characteristics such as age, gender, bicycle type and the riding direction after the light is green. Moreover, heterogeneity between cyclists could be considered by drawing the coefficients from a distribution rather than fixing them to one value for everyone. Furthermore, the assumption with respect to the

independence of alternatives could be lifted and models that allow for correlation of alternatives, such as cross-nested or mixed-logit, could be estimated. Other modelling assumptions could be tested in future research as well, such as the vision field and the cyclists in it that are taken into account, or the anticipation and memory skills of the cyclists.

Apart from improving the currently estimated models, a future research direction is their adjustment to enable the communication of the two layers and potentially updating the intended queuing position decision. Additionally, the models could be validated on other intersections and extended for other datasets where scooters and pedestrians are present so that their effect is captured as well.

Last but not least, the generalisability of the proposed approach should be substantiated by estimating models for other choice situations in the conceptual framework. These models can then be used in microsimulations and to update the model attributes in the framework. A challenge in this process has been the shortage of cyclist trajectory data which we will tackle in future work thanks to the dataset collected through our controlled large-scale cycling experiment (Gavriilidou et al. 2019a).

Acknowledgements

This research is supported by the ALLEGRO project, which is funded by the European Research Council (Grant Agreement No. 669792) and the Amsterdam Institute for Advanced Metropolitan Solutions. The authors thank the anonymous reviewers for their careful considerations and useful comments.

References

Andresen, E., Chraibi, M., Seyfried, A., Huber, F., 2013. Basic driving dynamics of cyclists, in: Simulation of Urban MObility User Conference, Springer. pp. 18–32.

Antonini, G., Bierlaire, M., Weber, M., 2006. Discrete choice models of pedestrian walking behavior. Transportation Research Part B: Methodological 40, 667–687.

Bierlaire, M., 2016. PythonBiogeme: a short introduction. Technical Report.

Choudhury, C.F., Ben-Akiva, M., Abou-Zeid, M., 2010. Dynamic latent plan models. Journal of Choice Modelling 3, 50–70.

CROW, 2016. Ontwerpwijzer fietsverkeer.

Fukuda, D., Seo, T., Yamada, K., Yaginuma, H., Matsuyama, N., 2014. An econometric-based model of pedestrian walking behavior implicitly considering strategic or tactical decisions, in: Pedestrian and Evacuation Dynamics 2012. Springer, pp. 615–624.

Gavriilidou, A., Wierbos, M.J., Daamen, W., Yuan, Y., Knoop, V.L., Hoogendoorn, S.P., 2019a. Large-scale bicycle flow experiment: Setup and implementation. Transportation Research Record , 0361198119839974.

Gavriilidou, A., Yuan, Y., Farah, H., Hoogendoorn, S.P., 2019b. Microscopic cycling behaviour model using differential game theory, in: Traffic and Granular Flow'17. Springer, forthcoming.

Goñi Ros, B., Yuan, Y., Daamen, W., Hoogendoorn, S.P., 2018. Empirical analysis of the macroscopic characteristics of bicycle flow during the queue discharge process at a signalized intersection. Transportation Research Record .

Hoogendoorn, S.P., Bovy, P.H.L., 2004. Pedestrian route-choice and activity scheduling theory and models. Transportation Research Part B: Methodological 38, 169–190.

Huang, L., Wu, J., You, F., Lv, Z., Song, H., 2017. Cyclist social force model at unsignalized intersections with heterogeneous traffic. IEEE Transactions on Industrial Informatics 13, 782–792.

Jiang, H., Wen, T., Jiang, P., Han, H., 2013. Research on cyclists microscopic behaviour models at signalized intersection, in: 16th International Conference Road Safety on Four Continents. Beijing, China (RS4C 2013). 15-17 May 2013.

Kucharski, R., Drabicki, A., Kulpa, T., Szarata, A., 2017. Multichannel cyclist queuing behaviour at signalised cycle crossings, in: Proceedings of 6th Symposium of the European Association for Research in Transportation.

Lee, T.C., Polak, J.W., Bell, M.G., 2009. New approach to modeling mixed traffic containing motorcycles in urban areas. Transportation Research Record 2140, 195–205.

Lee, T.C., Wong, K., 2016. An agent-based model for queue formation of powered two-wheelers in heterogeneous traffic. Physica A: statistical mechanics and its applications 461, 199–216.

Li, M., Shi, F., Chen, D., 2011. Analyze bicycle-car mixed flow by social force model for collision risk evaluation, in: 3rd International Conference on Road Safety and Simulation, pp. 1–22.

Liang, X., Baohua, M., Qi, X., 2012. Psychological-physical force model for bicycle dynamics. Journal of Transportation Systems Engineering and Information Technology 12, 91–97.

Mallikarjuna, C., Rao, K.R., 2009. Cellular automata model for heterogeneous traffic. Journal of Advanced Transportation 43, 321–345.

Michon, J.A., 1985. A critical view of driver behavior models: What do we know, what should we do. Human behavior and traffic safety , 485–520.

Rasmussen, J., 1983. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE transactions on systems, man, and cybernetics , 257–266.

Shiomi, Y., Hanamori, T., Eng, M., Nobuhiro, U., Shimamoto, H., 2012. Modeling traffic flow dominated by motorcycles based on discrete choice approach, in: Proceedings of 1st LATSIS Conference.

Vasic, J., Ruskin, H.J., 2012. Cellular automata simulation of traffic including cars and bicycles. Physica A: Statistical Mechanics and its Applications 391, 2720–2729.

Yao, D., Zhang, Y., Li, L., Su, Y., Cheng, S., Xu, W., 2009. Behavior modeling and simulation for conflicts in vehicles-bicycles mixed flow. IEEE Intelligent Transportation Systems Magazine 1, 25–30.

Appendix A. Data smoothing

In this appendix, the results of smoothing with different sliding window lengths are presented, as well as of the homogenisation with different time steps. The original, smoothed and final data points are displayed in the plots of Figure A.12 for the cyclist trajectories, Figure A.13 for the cycling speed and Figure A.14 for the steering angle.

Sliding window lengths, denoted by k, of 3 and 6 are compared, which means that every data point is replaced by the mean value of its 3 or 6 surrounding frames. The values are selected based on the average frame rate of 6fps, which means that frames of 0.5 or 1 second are used for the smoothing. The only difference that can be detected in the plots is in the region of -9m in the x direction where the two cameras overlap. In the original dataset there seems to be a jump from the back camera to the front,

which is smoothed with this process. The results with $k = 6$ convert this jump to an almost continuous trajectory and are therefore favourable.

Regarding the time step, denoted by dt, values of 0.5 and 1 second are compared, as they are considered to be reasonable time intervals for a new decision to be made. Smaller values would coincide with the frame rate, while larger ones would lead to very few points per trajectory as the average trajectory duration is 7 seconds. The difference in this comparison can be observed in the plots for the speed and the angle where the larger dt is shown to be better at muting the noise in the dataset which is introduced due to the manual tracking. For this reason, the time step is chosen at 1 second.

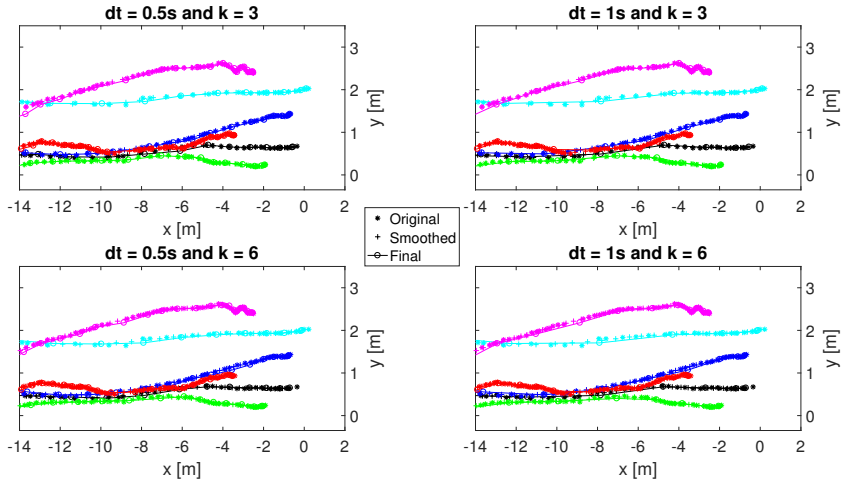


Figure A.12: Cyclist trajectories during a red light phase in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5$s and right: $dt = 1$s). The point (0,0) is the location where the stop line meets the curb of the sidewalk.

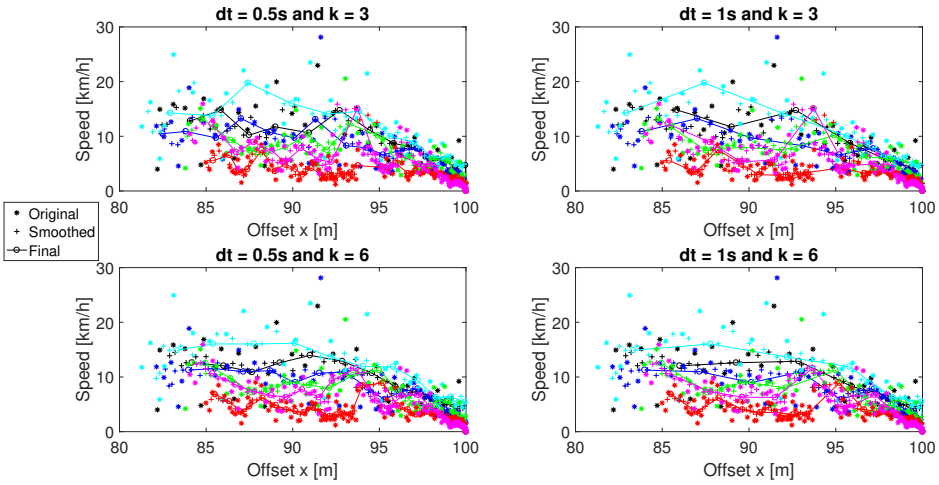


Figure A.13: Cyclist speed when approaching a red traffic light in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5$s and right: $dt = 1$s). The positions in x have been offset such that they end at the same point for all cyclists.

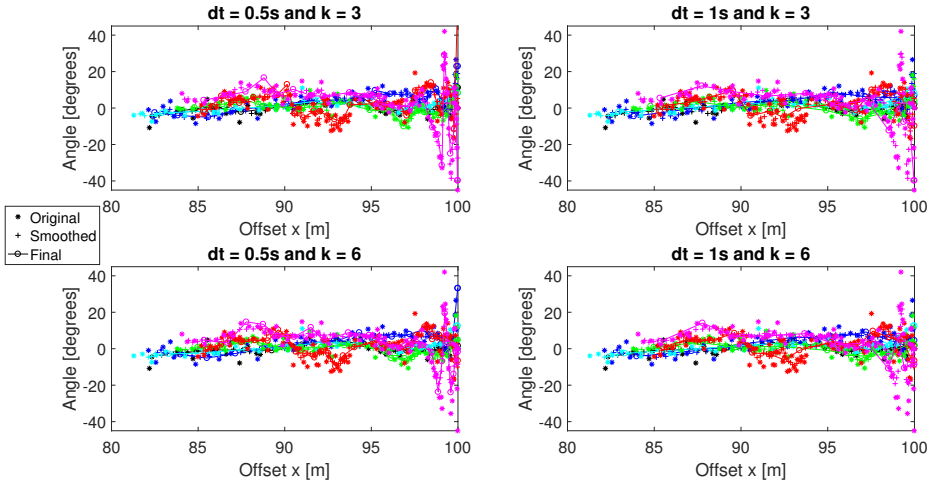


Figure A.14: Cyclist steering angle when approaching a red traffic light in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5$s and right: $dt = 1$s). The positions in x have been offset such that they end at the same point for all cyclists.