Fast or forced to follow: A speed heterogeneous approach to congested multi-lane bicycle traffic simulation

Paulsen, Mads; Rasmussen, Thomas Kjær; Nielsen, Otto Anker

Published in:
Transportation Research Part B: Methodological

Link to article, DOI:
10.1016/j.trb.2019.07.002

Publication date:
2019

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/.

It is a preprint of the accepted manuscript and should be cited as:

Fast or Forced to Follow: A Speed Heterogeneous Approach to Congested Multi-Lane Bicycle Traffic Simulation

Mads Paulsen, Thomas Ør Rasmussen, Otto Anker Nielsen

Department of Technology, Management and Economics, Technical University of Denmark, Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark

Abstract

Copenhagen is world-known for its large proportion of cyclists, forming a diverse group with a large variation of equipment and physical abilities. This leads to a considerable speed heterogeneity which needs to be taken into account when modelling the traffic on dedicated bicycle paths. Nevertheless, existing studies on bicycle traffic simulation have either neglected such speed heterogeneity altogether or modelled it by dividing cyclists into a few discrete classes ignoring the entirety of the speed distribution. This paper proposes an efficient bicycle traffic simulation model with continuously speed heterogeneous cyclists and corresponding congestion effects. Based on individual-specific desired speeds and headway distance preferences, the model shows realistic speed-flow relationships validated with on-site observations while being capable of delaying rapid cyclists more often than slower ones in moderate traffic flows. The scalability of the model allows it to be large-scale applicable for network loading purposes, and thus suitable for evaluating impacts of cycling related infrastructure investments.

Keywords: Bicycle traffic, Bicycle congestion modeling, Speed heterogeneity, Multi-lane traffic, Simulation,

1. Introduction

The bicycle is a very space-efficient mode of transport, why major cities across the world are starting to promote bicycling as a means of relieving congestion in urban areas. However, in places where bicycles already hold a high market share of travellers – such as Copenhagen and the Netherlands – the number of bicycles have reached a level where congestion on dedicated bicycle paths can no longer be ignored.

When modelling car traffic it is customary to consider all cars as a homogeneous mass where all cars traversing a link follow the same speed within a certain period of time. This assumption is more or less valid for car traffic as car drivers generally desire to go by the maximum allowed speed, alternatively by the highest speed possible given the congestion level. This is, however, not the case for cyclists for which higher speeds require a considerable increase in the wattage produced by the cyclist, limiting less physically fit cyclists from reaching high speeds. Furthermore,

*Corresponding author
Email address: madsp@dtu.dk (Mads Paulsen)
differences in equipment such as cargo bikes or racing bikes can significantly reduce or increase the desired speed of a cyclist.

Another difference is that overtaking is predominantly unproblematic outside rush hours for cyclists in Copenhagen as all bicycle paths are wide enough to fit at least two bicycles. However, due to the high frequency of overtakings, it may occur that the overtaking lane is occupied by a relatively slow cyclist overtaking an even slower cyclist. This makes it possible for other cyclists to cause delays to fellow cyclists even at moderate traffic intensities. Including such bicycle interactions within a fast bicycle simulation model is the primary aim of this paper.

The current literature on explicit modelling of bicycle traffic can be divided into two main categories: Heterogeneous and homogeneous modelling.

The literature on modelling of heterogeneous bicycle traffic has had a strong dominance of Cellular Automata (CA) (Wolfram, 1986) based approaches. Whereas original CA applications on traffic simulation such as Nagel and Schreckenberg (1992) only allowed simulation of homogeneous traffic, newer multi-value CA approaches have allowed simultaneous simulation of inhomogeneous traffic through discrete classes with different maximum speeds. Initially proposed for mixed traffic of cars and motorcycles in Lan and Chang (2005), it was later applied to mixed bicycle traffic by Jia et al. (2007). Since then a whole range of studies have considered mixed bicycle traffic (Gould and Karner, 2009; Zhao et al., 2013; Jin et al., 2015a; Shan et al., 2015; Zhou et al., 2015; Li and Fang, 2017; Xue et al., 2017) or mixed traffic with bicycles alongside other modes (Vasic and Ruskin, 2012; Luo et al., 2015) using similar and extended CA approaches.

Although the CA approaches are capable of modelling different classes, the discrete nature of the methodology hinders including the entirety of the speed distribution. The CA methodology partitions links into cells with a minimum size requiring fully containing any of the simulated vehicles – cyclists in our case. Since a cyclist have to stay in the same cell or move an integer number of cells per time step, different cyclist types either have to be assumed equal or differing rather considerably in order to be simulated using such approach. Thus, CA is only a suitable approach for modelling heterogeneity when dealing with traffic that can be divided into a few clearly distinguishable classes where intra-class heterogeneity can be ignored completely. This inherently restricts the potential of such models for the purpose of our study.

Three methodologies for simulation of continuous heterogeneous bicycle traffic exist in the literature. Shen et al. (2011) proposed a cell-stream flow model, Liang et al. (2012, 2018) developed psychological-physiological force models, whereas Osowski and Waterson (2017) applied a social force model. These approaches are capable of realistically simulating heterogeneous bicycle traffic in great detail but are, however, computationally infeasible for full-scale applications as explicitly stated in Osowski and Waterson (2017). Although none of the other studies touch upon computation time, the detail level of these models strongly indicates that none of them would have the computational efficiency to be suitable for traffic assignment purposes.

As a side note, heterogeneous traffic has also been modelled without focusing (solely) on bicycle traffic. Multi-class heterogeneous highway traffic was modelled in Wong and Wong (2002) using an extended Lighthill-Whitham-
Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), whereas Tang et al. (2011) applied a macroscopic following model with honk effects to a mix of bicycle and pedestrian traffic.

The second main category of bicycle traffic modelling is homogeneous simulation of bicycle traffic. This has also been explored with CA approaches (Jiang et al., 2004; Liu et al., 2008; Jiang et al., 2017; Tang et al., 2018), but has more interestingly also received some attention with alternative methodologies. Zhang et al. (2013b) combined CA with gas dynamics models for high densities to model speed-density relationships for bicycle traffic. Zhao and Zhang (2017) introduced a robust unified follow-the-leader model framework capable of modelling homogeneous traffic of either pedestrians, bicycles, or cars. Although these studies generally benefit from more advanced continuous methods, they have the drawback of completely ignoring the inherent speed heterogeneity of cyclists.

Regarding data, some studies (Mai et al., 2013; Zhang et al., 2013a; Rui et al., 2014) have been analysing single-file controlled experiments to explain cyclists’ behaviours. Additionally, some of the aforementioned bicycle models have been calibrated/validated by them (Liang et al., 2012; Xue et al., 2017; Zhao and Zhang, 2017), potentially leading to undesired biases as such experiments arguably differ from everyday bicycle traffic. First of all because overtaking is not considered in such controlled experiments. Secondly, because the cyclists have to ride in a circular or oval curve. Thirdly, since the cyclists of the controlled experiments are aware that they are participating in an experiment, and as such they might unconsciously behave differently from what they normally would in an uncontrolled environment. Collectively, these biases lead to experiments that are usually carried out using much lower speeds than the real-life counterparts they are seeking to model.

Hence, although calibrating a model using controlled experimental data is superior to no calibration, it still has some drawbacks compared to observed, uncontrolled data. Unfortunately, high quality data of observed, uncontrolled bicycle traffic is limited, although a few studies (Gould and Karner, 2009; Zhang et al., 2013b; Jin et al., 2015a; Li and Fang, 2017; Liang et al., 2018) have been able to collect such for calibrating their models.

Whereas forming fundamental diagrams require reliant aggregate data of very high traffic intensities, some characteristics of bicycling may, on the other hand, be derived by analysing bicycle traffic under low traffic intensities. Establishing a consistent bicycle simulation model based on assumptions that can be analysed at low traffic intensities is one of the aims of this study.

The study contributes to the literature by modelling explicit desired speeds for each cyclist drawn from a continuous distribution. Bicycle-to-bicycle interactions and overtaking are modelled implicitly using individualised speed-dependent headways for each cyclist and an intelligent lane choice mechanism. Furthermore, the implemented model is fast enough to be large-scale applicable, and thus has the potential to simulate the entirety of bicycle traffic within a metropolitan area.

The remainder of the paper is structured as follows. Bicycle traffic characteristics including speed and headway distance heterogeneity are analysed based on observed data in Section 2. Section 3 formulates the underlying assumptions for a bicycle simulation model capable of modelling inter-vehicular interactions of heterogeneous cyclists. Section 4 contains the design and results of small-scale experiments focusing on speed heterogeneity as well as aggre-
gate fundamental diagrams accompanied by sensitivity analyses and data validation. Finally, a discussion of model limitations as well as directions for future work are found alongside the final conclusions in Section 5.

2. Characteristics of Bicycle Traffic

This section describes some key characteristics of bicycle traffic important for the later implementation of the bicycle traffic simulation model. The analysis is based on aerial video data from COWI Ltd.\textsuperscript{1} processed using DataFromSky (RCE systems r.o., 2014). The video was shot from the tower of Frederiksberg Town Hall situated in the Copenhagen enclave Frederiksberg, with the original purpose of evaluating the use of the large parking lot next to the town hall. Still, the video data includes all east-going (towards the camera) bicycle traffic of the street Smallegade from 2pm to 5pm with a total of 1,168 cyclists. Their speed and acceleration for every cyclist have been calculated along 18 transversal gates each divided into an inner right (from the perspective of the cyclists) and an outer left part, see Figure 1.

The average flow along this stretch is less than 390 cyclists per hour why the traffic intensity is not high enough for congestion effects to be relevant. This is on purpose though, as one of the key parameters that will be estimated in the following sub-sections (desired speed) would not be possible to estimate using congested data.

Figure 1: An overview of the data area at Smallegade in Frederiksberg, Capital Region, Denmark.

\textsuperscript{1}https://www.cowi.com
2.1. Speed Heterogeneity

For each cyclist in the aerial video data the 90th percentile speed was used for determining their desired speed. Figure 2 shows the distribution of the desired speeds across cyclists, alongside best-fit estimations of 11 candidate distributions. The mean value and median of the empirical distribution are 6.29 m/s and 6.21 m/s, respectively. Furthermore, the distribution has a variance of 1.23, an excess kurtosis of 1.26, and a skewness equal to 0.55, the latter two indicating that a normal distribution may not be suitable.

![Figure 2: Distribution of observed desired speeds and best-fit estimations of candidate distributions.](image)

The data reveals relatively high speeds compared to many other studies seen in the literature from other parts of the world. Allen et al. (1998) reviews several older studies from across the world, and shows that the majority of these lie between 12 km/h (3.33 m/s) and 20 km/h (5.5 m/s). Jensen et al. (2010) analysed bicycle speeds of shared bicycles in Lyon, France, and showed an average speed of 14.5 km/h (4.03 m/s). Bernardi and Rupi (2015) measured the average speed on three segregated bicycle lanes in Bologna, Italy, to be ranging from 14.6-18.9 km/h (4.1-5.3 m/s). Li and Fang (2017) showed an average speed of 2.67 m/s for cyclists at the Nanhu campus of Wuhan University of Technology, China. Finally, Flügel et al. (2017) found the average cycling speed in Oslo, Norway to be 16.85 km/h (4.68 m/s).
Table 1: Log-likelihoods ($\ell$) as well as test statistics ($D$) and $p$-values for the Kolmogorov-Smirnov test for various candidate distributions.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\ell$</th>
<th>$D$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnson’s $S_U$</td>
<td>-1750.9</td>
<td>0.022</td>
<td>0.606</td>
</tr>
<tr>
<td>Johnson’s $S_L$</td>
<td>-1756.8</td>
<td>0.029</td>
<td>0.286</td>
</tr>
<tr>
<td>Johnson’s $S_B$</td>
<td>-1756.9</td>
<td>0.029</td>
<td>0.279</td>
</tr>
<tr>
<td>Logistic</td>
<td>-1757.7</td>
<td>0.028</td>
<td>0.310</td>
</tr>
<tr>
<td>Beta</td>
<td>-1758.4</td>
<td>0.031</td>
<td>0.200</td>
</tr>
<tr>
<td>Gamma</td>
<td>-1758.4</td>
<td>0.030</td>
<td>0.254</td>
</tr>
<tr>
<td>Hyperbolic Secant</td>
<td>-1760.9</td>
<td>0.032</td>
<td>0.181</td>
</tr>
<tr>
<td>Log-Normal</td>
<td>-1761.7</td>
<td>0.030</td>
<td>0.235</td>
</tr>
<tr>
<td>Normal</td>
<td>-1775.6</td>
<td>0.052</td>
<td>0.004</td>
</tr>
<tr>
<td>Laplace</td>
<td>-1782.5</td>
<td>0.042</td>
<td>0.032</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1851.0</td>
<td>0.077</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

However, Buch and Greibe (2015) did a study in Copenhagen, Denmark, and found an average speed of 21.7 km/h (6.02 m/s) which is reasonably close to the data obtained for this study.

In order to obtain a continuous distribution with known probability density function, we proceed by searching for a theoretical distribution that matches the empirical distribution well. This is done by assessing 11 continuous candidate distributions for which best fit estimates have been found for every parameter in their parameter set using maximum likelihood estimation. Table 1 contains the 11 distributions alongside their log-likelihood values ($\ell$) and results from the Kolmogorov-Smirnov test ($D$ and $p$).

Based on the log-likelihood, the Johnson’s $S_U$ distribution (Johnson, 1949) performs best, having a considerably higher log-likelihood than the following seven distributions which, on the other hand, do no differ much from each other. The final three distributions (Normal, Laplace, and Weibull) are clearly worse than the top eight distributions based on log-likelihood.

The Kolmogorov-Smirnov test calculates the largest vertical deviation ($D$) between the empirical and theoretical cumulative distribution function. Combined with the number of observations this number can be used to directly determine the $p$-value, i.e. the probability of the largest vertical deviation exceeding $D$, assuming that the empirical distribution was actually drawn from the theoretical distribution under consideration.

Results from the test shows that eight of the theoretical distributions do not deviate significantly from the empirical distribution when using a significance level of 5%. The Johnson’s $S_U$ distribution gives the best results, but a simpler distribution such as the logistic distribution would also be suitable according to the test. It is worth emphasising that based on the Kolmogorov-Smirnov test the normal distribution – as one of only two distributions – can be ruled out as being an appropriate choice for any significance level higher than 0.4%.

Conclusively, as the Johnson’s $S_U$ both has the highest log-likelihood and performs the best in the Kolmogorov-
2.2. Lane Behaviour

The test street is a quite moderate Copenhagen bicycle path with a width of 1.85 m and a relatively low traffic intensity in the data collection period as stated earlier. We want to assess whether cyclists in the overtaking (left) lane on average go faster than those in the inner (right) lane, despite knowing that this difference will be more evident on segments with higher traffic intensities. To do this the aerial data has been divided into cyclists riding on the right and left half of the bicycle path, respectively, and the speeds of the cyclists have been grouped by the lane at which the cyclists rode. The empirical cumulative distribution function of the speeds in each lane can be found in Figure 3. It is seen that the two empirical distributions differ, and that the left lane seems to be the faster lane.

Whether the difference in mean speed across the two lanes is statistically significant has been tested by performing a two-sample z-test for the difference between means. The resulting z-statistic exceeds 15, meaning that the difference in mean speed across lanes is strongly significant using any reasonable significance level with the corresponding p-value being less than \(10^{-50}\).

This is not to say that lane formation behaviour of cyclists only occur when overtaking. Group behaviour where cyclists ride alongside each other – as implemented in Tang et al. (2018) – is also likely to occur. However, such abreast riding can arguably be assumed not to influence other cyclists as long as the grouping behaviour stops once a faster cyclist rings his/her bell.
2.3. Headways

The capacities of bicycle paths are highly dependent on how close cyclists are willing to ride to each other: the
closer they ride, the higher the capacity. Because of this, determination of bicycle headways has already received
some attention in the literature.

According to Andresen et al. (2014) – based on single-file track experiments – the headway distance, \( d(v) \), is a
function of the speed, \( v \), having the functional form,

\[
d(v) = \theta_0 + \theta_1 v,
\]

with parameters \( \theta_0 = 1.93 \) being the sum of the average length of a bicycle \((\lambda_c = 1.73 \text{ m})\) and an additional safety
distance of \( \lambda_s = 0.2 \text{ m} \), and \( \theta_1 = 0.78 \text{ s} \) being the assumed constant headway time between the front wheel of the
following cyclist and the back wheel of the cyclist in front.

Hoogendoorn and Daamen (2016) investigated headways on a bicycle path in Delft in the Netherlands from photo-
finish data. They use the approach of Wasielewski (1979); Hoogendoorn (2005) and found that 54.1% of bicycles are
constrained by the cyclist in front of them, and that the average headway time of these cyclists is 0.784 seconds with
a standard deviation of 0.660 seconds.

In our video data we have supplementary information, as we can not only measure the headway distance but also
the speed of every cyclist. This is helpful when identifying whether a cyclist is driving freely or being constrained,
since large speed differences would indicate the former.

We filter the data such that we only consider data from cyclists that reasonably can be assumed to be following
the bicycle in front of them. Since it is difficult to keep the exact same speed as the cyclist in front, in this analysis
we allow for an absolute speed difference threshold of 0.2 m/s above which cyclists are assumed to ride freely. This
corresponds to riding freely if the difference in distance per second is larger than the desired longitudinal distance
between two stationary bicycles (0.2 metres) found in Andresen et al. (2014). Mohammed et al. (2019) identified
a cluster characterised as following cyclists revealing an average speed difference of 0.93 m/s in the raw data and
0.76 m/s after bootstrapping, indicating that the value of 0.2 m/s indeed is quite strict.

Furthermore, as in Hoogendoorn and Daamen (2016), all observations with a headway time above a certain thresh-
hold are assumed to be freely moving why such observations are also discarded. In Hoogendoorn and Daamen (2016)
the method incentivised not choosing a (too) low number, whereas doing such in this case would bias the analysis. A
threshold of 2.3 s – as opposed to 3 s in Hoogendoorn and Daamen (2016) and 5 s in Mohammed et al. (2019) – has
been used based on a flattening trend of the distribution of all headways in our dataset.

Finally, since observations of the same cyclist pair (at different gates) are highly correlated, only the median
headway distance of all accepted instances of a cyclist pair is used. The final dataset consists of 607 headway distance
observations from unique cyclist pairs. Due to the conservative thresholds of both speed differences and headway
times we are rather confident that all non-following situations have been removed by the filtration.
Besides using the speeds to filter the headway observations, it can also be used to investigate the effect the speed has on the headway distance. Figure 4 shows a scatter plot where the headway distance is plotted against the speed of the behind cyclist. Headway distances are found by multiplying the headway time with the speed of the cyclist.

In addition we fit a first and a second order polynomial as well as a square root model, see Table 2. Besides the parameter estimates and t-statistics, the table also contains the residual standard error ($\sigma_\epsilon$) and coefficient of

**Table 2**: Three linear regression models for headway distances based on speed with r-values in brackets compared to Andresen et al. (2014).

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Andresen et al. (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.234</td>
<td>-4.590</td>
<td>-4.221</td>
</tr>
<tr>
<td></td>
<td>(1.398)</td>
<td>(-1.367)</td>
<td>(-2.473)</td>
</tr>
<tr>
<td>$\sqrt{\text{Speed}}$</td>
<td>-</td>
<td>-</td>
<td>4.602</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-)</td>
<td>(6.468)</td>
</tr>
<tr>
<td>Speed</td>
<td>0.965</td>
<td>3.065</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(6.355)</td>
<td>(2.603)</td>
<td>(-)</td>
</tr>
<tr>
<td>$\text{Speed}^2$</td>
<td>-</td>
<td>-0.185</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-1.798)</td>
<td>(-)</td>
</tr>
<tr>
<td>$\sigma_\epsilon$</td>
<td>3.114</td>
<td>3.108</td>
<td>3.110</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.063</td>
<td>0.068</td>
<td>0.065</td>
</tr>
</tbody>
</table>
determination ($R^2$) for the proposed models. For comparison, the table also contains results found in Andresen et al. (2014).

The model obtained by using a first order polynomial has a steeper slope (headway time of 0.965 s) than in Andresen et al. (2014). However, the $t$-statistic of the null hypothesis that the true slope is equal to 0.780 s is 1.220, resulting in a two-sided $p$-value of 0.223 indicating that it is reasonably likely that the values from Andresen et al. (2014) could be true.

Intuitively, the headway distance should have a negative curvature as faster cyclists most likely are more experienced cyclists and as such would keep a lower headway time. In fact, this is partially supported by the second order polynomial fit which shows a negative second order term with a $t$-statistic of -1.80 corresponding to a one-sided $p$-value of 0.04. The square root model only provides a slight increase in $R^2$ compared to the first order polynomial from 0.063 to 0.068.

As the second order term in the second order polynomial is barely significant on a 5% significance level, and the slope of the first order polynomial is not significantly different from the slope of Andresen et al. (2014), the choice is really between using the square root model or the originally proposed model by Andresen et al. (2014). Although relatively weak, there is some evidence towards the headway distance having a negative curvature, why the square root model seems to be the most appropriate headway distance function. It also has a slightly higher $R^2$ than the first order polynomial despite having the same number of parameters.

2.3.1. Headway Heterogeneity

In order to account for the large heterogeneity in headway distances, the model is extended further to not only fit the mean value associated with a given speed, but also the corresponding standard deviation. As the variation of the headway distances seems to be increasing for higher speeds, see Figure 4, the standard deviation has to be modelled explicitly using a speed-dependent function. In a first model, letting $\mathcal{N}(\mu, \sigma^2)$ denote the normal distribution with mean $\mu$ and variance $\sigma^2$, we maintain normally distributed residuals and model the headway distances in the following way,

$$d(v) \sim \mathcal{N}\left(\theta_0 + \theta_1 \sqrt{v}, \left(\zeta_0 + \zeta_1 \sqrt{v}\right)^2\right) = \theta_0 + \theta_1 \sqrt{v} + \left(\zeta_0 + \zeta_1 \sqrt{v}\right) \mathcal{N}\left(0, 1^2\right). \quad (2)$$

With this formulation the mean values are kept constant, whereas the standard deviation of the residuals is a first order polynomial of $\sqrt{v}$ parameterised by $\zeta_0$ and $\zeta_1$. This model, however, has no bounds on the headways as the normal distribution has infinite tails, which may provide unrealistic headway preferences when sampling from the distribution. In order to deal with this, two additional models based on beta distributed residuals are proposed.

The generalised beta distribution, which we denote by $\mathcal{B}_g(\alpha, \beta, a, b)$, is a regular beta distribution, $\mathcal{B}(\alpha, \beta)$, that is translated and scaled to have compact support $[a, b]$. This means that for a stochastic variable $X \sim \mathcal{B}_g(\alpha, \beta, a, b)$ the standardised variable $Z_{\mathcal{B}} = \frac{X-a}{b-a} \sim \mathcal{B}(\alpha, \beta)$. Using the notation $f(x|\alpha, \beta)$ for the probability density function of $\mathcal{B}(\alpha, \beta)$
and $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$, the probability density function of the generalised beta distribution, $B_g(\alpha, \beta, a, b)$, is given by,

$$f_g(x|\alpha, \beta, a, b) = \left\{ \begin{array}{ll}
\frac{1}{b-a} \cdot f \left( \frac{x-a}{b-a} | \alpha, \beta \right) = \frac{1}{b-a} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \Gamma(\beta)} \left( 1 - \frac{x-a}{b-a} \right)^{\beta-1} & , \ x \in [a, b] \\
0, & , \ x \notin [a, b].
\end{array} \right.$$  \hspace{1cm} (3)

It is customary to require the expected value of the residuals to be 0. The easiest way to secure this, is to require the distribution to be symmetrical around 0. This can be achieved by requiring $\alpha = \beta, \alpha > 0$, and $-a = b = h, h > 0$.

Since we still want to allow larger residuals for higher expected values of $d(v)$, $h$ is parametrised by $h = \zeta_0 + \zeta_1 \sqrt{v}$.

We will denote such generalised symmetric beta distribution by $B_s(\alpha, h)$. Its probability density function, $f_s$, is given by,

$$f_s(x|\alpha, h) = f_g(x|\alpha, \alpha, -h, h) = \frac{1}{2h} \cdot f \left( \frac{x+h}{2h} | \alpha, \alpha \right), \quad \forall x \in \mathbb{R}, \quad (4)$$

Having introduced the above, three additional models focusing on modelling the residuals can be presented. Model 4 that uses the formulation from eq. (2) as well as two models using Beta-distributed residuals; A symmetric Model 5 and a potentially asymmetric Model 6 defined by eq. (5) and (6), respectively,

$$d(v) \sim \theta_0 + \theta_1 \sqrt{v} + B_s(\alpha, h) = \theta_0 + \theta_1 \sqrt{v} + 2h \left( B(\alpha, \alpha) - \frac{1}{2} \right), \quad (5)$$

$$d(v) \sim \theta_0 + \theta_1 \sqrt{v} + B_g(\alpha, \beta, -h, h) = \theta_0 + \theta_1 \sqrt{v} + 2h \left( B(\alpha, \alpha) - \frac{1}{2} \right). \quad (6)$$

For Model 4, 5, and 6 the parameters are estimated simultaneously using the L-BFGS-B algorithm (Byrd et al., 1995) in R (R Core Team, 2018). In order to secure positive standard deviations for the entire range of speeds in the headway data, $\sqrt{2} \theta_0 \geq -\theta_0$ and $\sqrt{2} \zeta_1 \geq -\zeta_0$ as well as negative values for $\theta_0$ and $\zeta_0$ were applied as constraints for the algorithm. Furthermore, Model 5 and 6 required $\alpha \geq 1$, and Model 6 additionally constrained $\beta \geq 1$.

For all of the models the parameters, their log-likelihood ($\ell$), and Akaike Information Criterion (AIC) (Akaike, 1973) are presented in Table 3. Notice that $\zeta_0$ of Model 3 is the previously implicitly estimated $\sigma_0$ from Table 2.
Table 3: Parameters, log-likelihoods ($\ell$), and Akaike Information Criteria (AIC) of four models for headway distances using different functional forms for residuals.

<table>
<thead>
<tr>
<th></th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_0$</td>
<td>-4.221</td>
<td>-4.234</td>
<td>-4.357</td>
<td>-4.539</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>4.602</td>
<td>4.602</td>
<td>4.713</td>
<td>4.840</td>
</tr>
<tr>
<td>$\zeta_0$</td>
<td>3.110</td>
<td>-4.397</td>
<td>-9.674</td>
<td>-0.841</td>
</tr>
<tr>
<td>$\zeta_1$</td>
<td>-3.109</td>
<td>6.841</td>
<td>6.959</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-</td>
<td>-</td>
<td>1.865</td>
<td>1.910</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.033</td>
</tr>
<tr>
<td>$\ell$</td>
<td>-1,549.04</td>
<td>-1,522.88</td>
<td>-1,497.87</td>
<td>-1,497.45</td>
</tr>
<tr>
<td>AIC</td>
<td>3,104.08</td>
<td>3,053.77</td>
<td>3,005.74</td>
<td>3,006.90</td>
</tr>
</tbody>
</table>

Firstly, it is seen that regarding the parameters concerned with the expected value of the headway distance, i.e. $\theta_0$ and $\theta_1$, the changes are almost indistinguishable when going from Model 3 to Model 4. When going onto Model 5 and Model 6 the changes persist to stay relatively minor.

Looking at the log-likelihood, it is seen that there is a considerable improvement when switching from a constant (Model 3) to a parameterised (Model 4) standard deviation. However, as such switch will always have a non-negative effect, it is necessary to use some kind of measure to assess whether the improvement caused by the inclusion of the $\zeta_1$-parameter is worthwhile. For nested models, as it is the case here, applying Wilks’ theorem (Wilks, 1938), commonly known as a likelihood ratio test, is a frequently used method to assess improvements in likelihood. Letting $\Lambda$ denote the likelihood ratio between Model 3 and Model 4, the test statistic can be calculated as follows,

$$D_{\Lambda} = -2 \ln \Lambda = -2 \cdot (-1549.04 + 1522.88) = 52.32.$$

Comparing that to a $\chi^2$-distribution with one degree of freedom yields a $p$-value < $10^{-12}$, showing that the improvement in likelihood is significant at any reasonable significance level.

When comparing Model 5 to Model 4, it is no longer possible to apply Wilks’ theorem as neither model is a special instance of the other model. In such cases the Akaike Information Criterion (AIC) (Akaike, 1973) can be used instead. With $k$ denoting the number of estimated parameters the AIC is defined as,

$$AIC = 2k - 2\ell,$$

and thus becomes higher when either the log-likelihood ($\ell$) decreases or when the number of parameter increases. The AIC thus discourages overfitting, and the model with the lower AIC is considered the better option. We can easily see that Model 5 has a much lower AIC (3,005.74) than Model 4 (3,053.77), why Model 5 is superior to Model 4 based on AIC. The improvement is very similar to that when going from Model 3 to Model 4.

Comparing Model 5 to Model 6 we can see that without constraining $\beta$ it is estimated to be a little larger than $\alpha$, although the differences are rather small. We can once again apply Wilks’ theorem, as Model 5 is a particular instance
of Model 6 where $\beta = \alpha$. The test statistic $D_\Lambda = 0.835$ yields a $p$-value of 0.361, meaning that the improvement in likelihood is only significant when using a significance level above 36.1%. That Model 5 seems to be the better model is also supported by the AIC which is lower for Model 5. As such, there seems to be no statistical evidence against assuming $\alpha = \beta$ when modelling our residuals with a beta distribution with support $[-h, h], h > 0$.

However, since all of the above model assessments are purely relative measures, it is sensible to test Model 5 against a baseline model. Only by doing so is it possible to evaluate whether Model 5 is actually a reasonable model, or just the best among a candidate set of models. A null model only containing the average value of headways, $\theta_0 = 6.788$ and a constant residual standard deviation $\zeta_0 = 3.211$ of the assumed normally distributed residual yields an AIC of 4,211.66. The improvement in AIC of 1,205.92 when going to Model 5 strongly suggests that adding the parameters used in Model 5 makes statistical sense.

Thus, Model 5 is chosen as the best suitable model for modelling headway distances; It has the best statistical measures, but also because it makes sense from a behavioural point of view to use one of the two bounded models (5 or 6), as using an unbounded model (3 or 4) theoretically allows extreme headway distances. This is particularly an issue when sampling from the distribution. Whereas the actual sampling procedure shall be explained in Section 3.3, for the sake of the argument Figure 5 is presented here to show how the unbounded models 3 and 4 differs fundamentally from model 5 and 6.
The figure shows a surface plot illustrating the probability of obtaining a value that is further away from the expected value in the same direction as the value on the y-axis given some speed, \( v \). In a more formal explanation, let \( D_v \) be the stochastic variable of the headway distance based on the speed \( v \) according to one of our proposed models, and let \( F_{D_v} \) be the theoretical cumulative distribution function of said stochastic variable. We then introduce the following measure, \( p \), indicating the probability of getting a value that is larger/smaller than \( d \) given \( v \),

\[
p(d|v) = \min\{F_{D_v}(d|v), 1 - F_{D_v}(d|v)\} = \frac{1}{2} \cdot \Pr\left(D_v > |d - E(D_v)| \mid v\right), \quad \forall d \in \mathbb{R}, \forall v \geq 2.
\]

It is readily seen that Model 5 and 6 are almost indistinguishable, and that they are bounded very close to the
empirical observations. This also means that sampling from such distribution will almost always provide reasonable
values, although very small negative values can occur. Model 3 and 4, on the other hand, have substantial risks of
sampling quite extreme negative as well as positive values. A probability of e.g. $10^{-5}$ might seem very small, but
since the model is meant to be large-scale applicable and able to simulate several hundred thousands of cyclists, such
probabilities will eventually come into play. Finally, it can be seen that Model 3 does not take the increasing variance
into account at all, and as such seems way off when sampling headway distances at low speeds.

Using Model 5 and the space mean speed $\bar{v}_0 = 6.104$, the expected headway time is $\frac{d_{\bar{v}_0}}{\bar{v}_0} = 0.910$ s. This
seems reasonable although slightly above the headway times of 0.78 and 0.784 s found in Andresen et al. (2014) and
Hoogendoorn and Daamen (2016), respectively.

3. Model

Whereas Section 2 was devoted to extracting certain characteristics of bicycle data from field data, this section
aims at establishing a simulation framework for heterogeneous bicycle traffic into which such findings can be used as
input. The proposed framework is based on the set of assumptions found below.

1. Every cyclist has an individual desired speed.
2. Bicycle paths can be divided into a number of pseudo-lanes.
3. Cyclists keep a headway distance to the cyclist in front of them.
4. Cyclists use the right-most pseudo-lane that can accommodate their desired speed.
   a. If no such pseudo-lane exists, they use the fastest possible pseudo-lane.
5. Cyclists stay in the same pseudo-lane for the entirety of the link.
6. Cyclists cannot enter a link if its area is fully occupied.

The remainder of this section is devoted to discussing each of the assumptions individually, before Section 3.7
summarises the input needed for the framework.

3.1. Desired Speeds

The first assumption is that every cyclist has an individual desired speed. The desired speed can be interpreted
as the speed a cyclist would ride by under free-flow conditions. In that sense it can also be understood as a personal
maximum speed for that individual.

In accordance with findings in Section 2.1, we assume that cyclists’ desired speeds follow a Johnson’s $S_U$
distribution which by denoting the inverse hyperbolic sine function by $\text{arsinh}(x) = \ln \left(x + \sqrt{x^2 + 1}\right)$ is defined as,

$$
v_c^0 \sim D = \frac{\delta}{\lambda \sqrt{2\pi \sqrt{1 + \left(\frac{v_0^e - \xi}{\lambda}\right)^2}}} e^{-\frac{1}{2} \left(\gamma + \delta \text{arsinh} \left(\frac{v_0^e - \xi}{\lambda}\right)\right)^2}, \quad v_c^0 \geq 2, \quad (9)
$$
with parameters $\gamma = -2.75$, $\xi = 3.67$, $\delta = 4.07$, and $\lambda = 3.49$. Notice that the superscript $c$ is used because any desired speed $v_c^0$ is associated with and specific to one particular cyclist, $c$.

A uniform value for each cyclist, $u_c \in U(0,1)$, can be transformed into a desired speed by the following modification of the quantile from the standard normal distribution $q_{u_c}, u_c \in [0,1]$ readily available in most programming software,

$$v_c^0 = \lambda \sinh \left( \frac{q_{u_c} - \gamma}{\delta} \right) + \xi, \quad u_c \in [0,1], \forall c \in C, \quad (10)$$

making the sampling of such values straightforward.

It is seen from eq. (9) that we truncate the distribution at $v_{min}^0 = 2 \text{ m/s}$ in order to avoid desired speed that are too low (Johnson’s $S_U$ distribution has a natural lower bound at 0). In practice such values are rejected and replaced by a resampled value. The truncation is necessary because there exists a speed below which it is virtually impossible to keep the balance on a bicycle, but also because having too low values can almost shut down links entirely with the proposed model (see Section 3.4). It is worth noting that none of the estimations of the (fully) bounded distributions (Johnson’s $S_B$ and Beta) resulted in a non-negative lower bound, meaning that the value used for truncation had to be determined manually. The value of 2 m/s seems reasonable based on our data, as no cyclists were observed riding below this speed with the minimum observation being 2.67 m/s.

In the proposed model the desired speed is constant across links for each individual. However, the framework can easily be extended to have link specific desired speeds that for instance take the gradient, surface, or wind into account.

### 3.2. Infrastructure

The second assumption requires that every link of bicycle path, $l$, has an array of pseudo-lanes, $\Psi_l$. The number of pseudo-lanes, $|\Psi_l|$, contained in this array is based on the width of the bicycle path $\omega l$. The elements of $\Psi_l$ are enumerated from right to left, i.e. from $\psi_1 l$ to $\psi_{|\Psi_l|} l$. In our model such pseudo-lanes are strictly obeyed in the sense that no interaction is assumed across pseudo-lanes.

Any function $f : \mathbb{R}^+ \rightarrow \mathbb{N}^+$ can be used for partitioning the link into lanes. We propose using a formula from an existing study (Buch and Greibe, 2015) who found the number of effective lanes – corresponding to $|\Psi_l|$ – on bicycle paths in Copenhagen to be,

$$|\Psi_l| = 1 + \left\lfloor \frac{\omega l - 0.4 \text{ m}}{1.25 \text{ m}} \right\rfloor. \quad (11)$$

The formula states that a new effective lane is obtained for every 1.25 m of width exceeding 0.4 m. Again this can easily be extended from a constant formula to also include link characteristics such as the gradient, wind, or surface type.

In the current implementation, however, the only additional attribute of a link of bicycle path is its length, $\lambda l$. This is in contrast to traditional models for car traffic where for instance the capacity of the link has to be defined explicitly.
3.3. Headway Distances

The third assumption assures that cyclists cannot overlap each other as they move through traffic. In fact, they even have to keep a minimum headway distance to the cyclist in front of them. The headway distance implemented in the model is based on the findings of Section 2.3, specifically eq. (5).

Assuming that the residuals of the distance model arise from inter-personal preferences regarding the predefined headway distance, it makes sense to have the same cyclists consistently having a relatively short or long headway distance. This can be done by rewriting eq. (5) by using the beta-distributed variable \( Z_b \sim B(\alpha, \alpha) \),

\[
d(v) \sim \theta_0 + \theta_1 \sqrt{v} + 2 \left( Z_b - \frac{1}{2} \right) \cdot (\zeta_0 + \zeta_1 \sqrt{v}), \quad v \geq 2. \tag{12}
\]

Individual values of \( \zeta_b^c \sim Z_b \) can then be drawn for each cyclist \( c \in C \) to find the cyclist-specific distance functions,

\[
d^c(v) = \theta_0^c + \theta_1^c \sqrt{v}, \quad v \geq 2, \tag{13}
\]

where \( \theta_0^c = \theta_0 + 2\zeta_0 \left( \zeta_b^c - \frac{1}{2} \right) \) and \( \theta_1^c = \theta_1 + 2\zeta_1 \left( \zeta_b^c - \frac{1}{2} \right) \) are the individual-specific distance-function parameters. In this way, the population across all cyclists \( c \in C \) approaches the distribution from eq. (12) as \( |C| \) gets large.

In order to sample from \( Z_b \), we use a rejection sampling algorithm where we for all cyclists \( c \in C \) first draw a value from a uniform proposal distribution, i.e. \( u_c \sim U(0, 1) \). We subsequently draw another uniform value \( \hat{u}_c \sim U(0, 1) \), and assign \( u_c \) to \( \zeta_b^c \) with probability \( p_a \). The acceptance probability \( p_a \) can – by denoting the probability density function of \( B(\alpha, \beta) \) by \( f \) – be found to be,

\[
p_a(u_c | \alpha, \alpha) = \frac{f(u_c | \alpha, \alpha)}{f \left( \frac{1}{2} | \alpha, \alpha \right)} = \frac{u_c^{\alpha-1} (1 - u_c)^{\alpha-1}}{\left( \frac{1}{2} \right)^{\alpha-1} (1 - \frac{1}{2})^{\alpha-1}} = (4u_c (1 - u_c))^{\alpha-1}, \quad u_c \in [0, 1]. \tag{14}
\]

If \( \hat{u}_c \geq p_a \), on the other hand, the algorithm starts over until at some point \( \hat{u}_c < p_a \) and \( \zeta_b^c \) is assigned the latest value of \( u_c \). It can be seen that the acceptance probability is exactly 1 at \( u_c = \frac{1}{2}, 0 \) at \( u_c \in [0, 1] \), and that it is positive and strictly less than 1 all other places within the support, i.e. \( p_a(u_c | \alpha, \alpha) \in ]0, 1[ \), \( \forall u_c \in ]0, 1[ \), \( \forall \alpha > 0, \forall c \in C \).

For the minimum desired speed, the corresponding expected headway distance is \( E \left[ d^c(\nu_0^{\min}) \right] = 2.31 \) metres, whereas the mean speed provides an expected distance of \( E \left[ d^c(\bar{\nu}_0) \right] = 7.29 \) metres. The number obtained by the formula is the headway distance including the length of a bicycle, which Andresen et al. (2014) found to be \( \lambda^c = 1.73 \) m on average.

The headway distance creates an inherent expected upper limit of the outgoing flow on a link as long as cyclists have finite speeds. By assuming no heterogeneity, i.e. \( \theta_0^c = \theta_0 \) and \( \theta_1^c = \theta_1 \) for all \( c \in C \), eq. (13) can be used to determine the theoretical expected maximum flow per lane at the mean speed,
meaning that in theory almost 3,016 cyclists per lane can pass a cross section of the network per hour. The heterogeneity of cyclists will most likely reduce this number noticeably in practice, though.

3.4. Lane Choice and Speed Determination

Section 3.1 described that every cyclist has a desired speed that he/she will ride by if the traffic conditions allow him/her to, whereas Section 3.2 introduced the concept of pseudo-lanes into which the traffic is assumed to be divided. Section 3.3 determined the distance needed between two consecutive cyclists in the same pseudo-lane. Combining all this allow us to specify the speed at which a cyclist will traverse a link. Apart from the desired speed of the cyclist the speed is solely a function of the arrival time and speed of the most recent cyclist that has entered the pseudo-lane that the cyclist choose to use.

As such the speed is a direct consequence of the lane choice which in the model is assumed to be performed at the moment the cyclist enters the link, see Section 3.5. A cyclist \( c \) starts considering the right-most lane \( (i=1) \), and searches leftwards (increasing \( i \)'s) in order to find a pseudo-lane \( (\psi_l^i) \) where he/she can go by his/her desired speed \( (v_c^0) \). He/she will choose the rightmost satisfactory lane; alternatively – if no lanes allow going by his/her desired speed – he/she chooses the lane with the highest possible speed \( (\psi_l^i_{\text{max}}) \).

Let \( t_l^s \) denote the time at which a cyclist reaches link \( l \). A formal formulation of the lane choice process can then be described by Algorithm 1.

\[
\begin{align*}
C_l &= |\Psi_l| \cdot \frac{3,600}{E[d(v_0)]} = |\Psi_l| \cdot 3,015.75, \quad \text{for } l \in L, \\
\end{align*}
\]

Algorithm 1 LaneChoice\((c, l, t_l^s)\)

\[
\begin{align*}
\tilde{v}^\text{max} &\leftarrow 0; \\
\text{for } i \text{ in } 1 \text{ to } |\Psi_l| &\text{ do} \\
&\text{if } \tilde{v}_l^c(\psi_l^i(t_l^s)) \geq v_0^i \text{ then return } \psi_l^i; \\
&\text{else if } \tilde{v}_l^c(\psi_l^i(t_l^s)) > \tilde{v}^\text{max} \text{ then} \\
&\quad \tilde{v}^\text{max} \leftarrow \tilde{v}_l^c(\psi_l^i(t_l^s)); \\
&\quad i^\text{max} \leftarrow i; \\
&\text{return } \psi_l^{i_{\text{max}}};
\end{align*}
\]

Here \( \tilde{v}_l^c(\psi_l^i(t_l^s)) \) denotes the speed that the cyclist \( c \) would go by, if he/she chooses pseudo-lane \( \psi_l^i \) at time \( t_l^s \). Notice that this expression is cyclist specific due to individualised desired speed and headway preferences explaining the superscript \( c \). \( \tilde{v}^\text{max} \) is the highest possible speed among the pseudo-lanes, and can be obtained at \( i^\text{max} \)'th pseudo-lane from the right, \( \psi_l^{i_{\text{max}}} \). Strictly speaking both of these should have an associated \( c \) in the notation, but this is omitted since they are only used locally.
The cyclist may choose a slower pseudo-lane, though. The speed assigned to a cyclist when entering link $l$ at time $t'_i$ denoted by $\bar{v}^c_i \left( t'_i \right)$ internally includes the lane choice described in Algorithm 1. This speed will be equal to the desired speed if at least one of the pseudo-lanes allows such speed, but will be lower if none of the pseudo-lanes can accommodate the desired speed. In both cases $\bar{v}^c_i \left( t'_i \right)$ can be described by,

$$
\bar{v}^c_i \left( t'_i \right) = \min \left\{ \max_{i \in 1, \ldots, |\Psi^l|} \bar{v}^c_i \left( t'_i \right), \bar{v}_0 \right\},
$$

(16)

with $\bar{v}^c_i \left( t'_i \right)$ still denoting the highest possible speed of pseudo-lane $\psi^l_i$ at time $t'_i$ for cyclist $c$.

Let $\tau^c(v)$ be the speed dependent headway time from cyclist $c$ to the cyclist in front of cyclist $c$. This can be derived using the headway distance function (eq. (13)),

$$
\tau^c(v) = \frac{d^c(v) - x^c}{v} = \frac{\theta_0 + \theta_1 \sqrt{v} - x^c}{v}.
$$

(17)

Notice that $v$ in eq. (17) can denote any speed of cyclist $c$, and that any of such time, link, or pseudo-lane specific speeds will have more elaborate notation. Arguments with such detailed notation will be seen frequently as input for $\tau^c(\cdot)$ in the following.

Having introduced the above, we are now ready to present an explicit functional expression for the value of $\bar{v}^c_i \left( t'_i \right)$, arising directly by finding the maximum speed enforcing the headway time of eq. (17) at the downstream end of the link. Let $t^c_{\psi^l} \left( \cdot \right)$ be the time at which the back wheel of the previous entrant of the $i$th pseudo-lane of link $l$ will leave the link\(^2\) and let $t^c_{\lambda} \left( \cdot \right)$ denote the time at which the tip of the front wheel of the current cyclist reaches the start of link $l$. Then the current cyclist can at earliest reach the downstream end of the link at $t^c_{\psi^l} \left( t'_i \right) + \tau^c \left( \bar{v}^c_i \left( t'_i \right) \right)$ as any earlier exit time would violate the headway time constraint. As speed is defined as the quotient between the distance travelled and the elapsed time, the maximum possible speed of pseudo-lane $\psi^l_i$ for cyclist $c$ can be found to be,

$$
\bar{v}^c_{\psi^l} \left( t'_i \right) = \begin{cases} 
\frac{x^c}{t^c_{\psi^l} \left( t'_i \right) - t^c_{\lambda}}, & t^c_{\psi^l} \left( t'_i \right) < t^c_{\lambda} + \tau^c \left( \bar{v}^c_i \left( t'_i \right) \right) \\
\infty, & t^c_{\psi^l} \left( t'_i \right) \geq t^c_{\lambda} + \tau^c \left( \bar{v}^c_i \left( t'_i \right) \right)
\end{cases},
$$

(18)

It is worth noting that although the headway constraint will always be enforced at the downstream end of the link, it can be violated at the upstream end of the link if the cyclist arrives shortly after the previous cyclist. However, we find it behaviourally realistic that the cyclist will accept such violation and adapt his/her speed so that the preferred headway distance is met at the downstream end of the link (the next point of evaluation, see Section 3.5) as opposed to instantaneously braking heavily to a full stop or an extremely low speed.

\(^2\)When calculating this number, it is necessary to assume that the previous cyclist will maintain the same speed at the first part (a distance of $x^c = 1.73\text{m}$) on the following link. However, the actual speed that the previous cyclist will travel on this link may differ.
As both the function value and domain boundary of \( \bar{v}_{\psi_i} (t_i') \) are functions of \( \bar{v}_{\psi_i} (t_i') \) itself, it turns out that it is beneficial to split eq. (18) into three separate cases as closed-form expressions with constant domain boundaries can be found when dealing with each case separately.

**The General Restricted Case**

When \( t_i' < t_{\bar{e}}^{\psi_i} + \tau \left( \bar{v}_{\psi_i} (t_i') \right) \) it is necessary to expand \( \tau \left( \bar{v}_{\psi_i} (t_i') \right) \) by inserting eq.(17) into eq. (18),

\[
0 = \left( t_{\bar{e}}^{\psi_i} - t_i' \right)^2 \bar{v}_{\psi_i} (t_i')^2 + \frac{2 \left( t_{\bar{e}}^{\psi_i} - t_i' \right) \left( \theta_0' - \lambda^i - \lambda^l \right) - \theta_1^2}{\theta_1^2} \bar{v}_{\psi_i} (t_i') + \frac{\left( \theta_0' - \lambda^i - \lambda^l \right)^2}{\theta_1^2}.
\]  

(19)

This is a second order polynomial of \( \bar{v}_{\psi_i} (t_i') \) with the following solution for \( \theta_1^2 + 4 \left( t_{\bar{e}}^{\psi_i} - t_i' \right) \left( \lambda^i + \lambda^l - \theta_0' \right) \geq 0 \) and \( t_{\bar{e}}^{\psi_i} \neq t_i' \),

\[
\bar{v}_{\psi_i} (t_i') = \frac{\theta_1^2 + 2 \left( t_{\bar{e}}^{\psi_i} - t_i' \right) \left( \lambda^i + \lambda^l - \theta_0' \right) - \theta_1 \sqrt{\theta_1^2 + 4 \left( t_{\bar{e}}^{\psi_i} - t_i' \right) \left( \lambda^i + \lambda^l - \theta_0' \right)}}{2 \left( t_{\bar{e}}^{\psi_i} - t_i' \right)^2}.
\]  

(20)

As mentioned, for this to be valid it is necessary but not sufficient for the radicant of the square-root to be non-negative, i.e.

\[
\theta_1^2 + 4 \left( t_{\bar{e}}^{\psi_i} - t_i' \right) \left( \lambda^i + \lambda^l - \theta_0' \right) \geq 0,
\]

which by assuming positive \( \lambda^i + \lambda^l - \theta_0' \) can be rewritten as,

\[
t_i' \leq t_{\bar{e}}^{\psi_i} + \frac{\theta_1^2}{4 (\lambda^i + \lambda^l - \theta_0')}. \quad (21)
\]

In the two following cases we will determine what happens when \( t_i' = t_{\bar{e}}^{\psi_i} \) and \( t_i' > t_{\bar{e}}^{\psi_i} + \frac{\theta_1^2}{4 (\lambda^i + \lambda^l - \theta_0')} \).

**The Singularity Case**

When \( t_i' = t_{\bar{e}}^{\psi_i} \), the main denominator of eq. (21) becomes zero. In this case we need to insert \( t_i' = t_{\bar{e}}^{\psi_i} \) directly into the first subfunction of eq. (18) which yields,

\[
\bar{v}_{\psi_i} (t_i') = \frac{\lambda^l}{\tau \left( \bar{v}_{\psi_i} (t_i') \right)} = \frac{\lambda^l}{\theta_0' + \theta_1 \sqrt{\bar{v}_{\psi_i} (t_i') - \lambda^i}}.
\]

As \( t_i' = t_{\bar{e}}^{\psi_i} \),
This can be solved straightforwardly to obtain,

\[ \tilde{v}_{c, s_l}(t_{ls}) = \left( \frac{\lambda^c + \lambda^l - \theta_0^c}{\theta_1^c} \right)^2, \quad t_{ls} = t^c_{e, s_l}. \]  

(22)

The Practically Unrestricted Case

We now turn to the cases in which eq. (21) is not fulfilled, i.e. eq. (20) have no real solutions. It turns out that if we insert the upper bound from eq. (21) into eq. (20) we can obtain a lower bound of \( \tilde{v}_{c, s_l}(t_{ls}) \) for

\[ t_{ls} > t^c_{e, s_l} + \frac{\theta_1^2}{4(\lambda^c + \lambda^l - \theta_0^c)}. \]

(23)

This can be shown to be large enough to largely exceed any reasonable desired speed for realistic values of \( \lambda^c, \lambda^l, \theta_0^c, \) and \( \theta_1^c \). In worst case, the cyclist has drawn and accepted a \( z_B = 1 \) in eq. (13), resulting in the parameters \( \theta_0^c = -14.03 \) and \( \theta_1^c = 11.55 \). Then using the smallest link length \( \lambda^l \) that will be presented in this study, 20 m, and the constant \( \lambda^c = 1.73 \), the obtained value will be,

\[ \tilde{v}_{c, s_l}(t_{ls}) = 4 \left( \frac{1.73 + 20.00 + 14.03}{11.55} \right)^2 = 38.32, \]

corresponding to approximately 138 km/h, which clearly is way beyond any reasonable desired speed for a cyclist. Thus, for any \( t_{ls} \) exceeding \( t^c_{e, s_l} + \frac{\theta_1^2}{4(\lambda^c + \lambda^l - \theta_0^c)} \), we can simply use the lower bound value obtained by eq. (23).

The situations where \( \tilde{v}_{c, s_l}(t_{ls}) \) would not be sufficiently high are if \( \theta_1^c \) is very high or if the link length, \( \lambda^l \), is very short. In such cases it would be necessary to look beyond the downstream end of the link to include the length of the headway distance. This is of course undesirable. However, having very short links or headway distance that increase immensely with increased speeds are in themselves inappropriate, why we consider this as a purely hypothetical limitation of the model. For all relevant practical purposes of the model the \( \tilde{v}_{c, s_l}(t_{ls}) \) from eq. (23) will be higher than any reasonable desired speed of a cyclist.

This means that although \( t_{ls} \) may be so late that there is no restriction on speed, i.e. \( \tilde{v}_{c, s_l}(t_{ls}) = \infty \), we do not have to address such cases when using our model, as the value obtained by using the result from eq. (23) will end up assigning the same speeds to cyclists in practice. This comes in handy as the domain of eq. (23) is given through constants, whereas the domain of the last subfunction of eq. (18) is a function of its value.

Summarising the results from the three cases, eq. (18) can be expanded to the following piecewise function,
\[
\tilde{v}_c(l_t) = \begin{cases} 
\frac{\eta_c^2 + 2(\tilde{v}^1_c + \tilde{v}^2_c)(x + x^c - \theta)}{2(\tilde{v}^1_c - \tilde{v}^2_c)} & , \quad t^1_s < t^1_c + \frac{\eta_c^2}{x(x + x^c - \theta)} \land t^1_s \neq t^1_c \\
\frac{(\tilde{v}^1_c + \tilde{v}^2_c)^2}{4} & , \quad t^1_s = t^1_c^c \\
\frac{(\tilde{v}^1_c + \tilde{v}^2_c)^2}{4} \eta_c^2 \lambda_c^c + \lambda_l^c & , \quad t^1_s > t^1_c + \frac{\eta_c^2}{x(x + x^c - \theta)} 
\end{cases}
\] (24)

Notice that contrary to eq. (18) the domains and expressions of all three subfunctions are defined solely by constants. The function can be shown to be a monotonously increasing function of \(t_s\) over the domain of its first subfunction. It is also continuous across the singularity found in \(t^1_s = t^1_c^c\). Finally it is continuous in the limit leading to the domain of the third function and constant in the remainder of the domain. As such, the function in eq.(24) is a monotonously non-decreasing function of \(t^1_s\).

The speed assigned to a cyclist, \(\tilde{v}_c(l_t)\), is used for calculating when he/she is expected to leave the link, but is actually only an upper bound for the speed. The actual speed the cyclist had on the link (\(v_c^l\)) can only be calculated retrospectively when it is known exactly when the cyclist has entered the next link, as this may occur later than expected if the next link is fully congested, see eq. (26) in Section 3.6.

3.5. Longitudinal Horizon

As it was evident from Section 3.4, the lane choices have a large influence on how a cyclist move through the network. The fifth assumption states that such choices are only made when entering a new link. This assumption is arguably the most behaviourally questionable. Normally lane choices will happen dynamically such that a cyclist will pull towards the right when finishing overtaking.

Whereas choosing a lane at the beginning of every link may be behaviourally inappropriate, it allows for a very fast implementation as the model can be implemented in an event-based fashion with events only occurring when cyclists leaves or enters a link, making the model scale extremely well.

Note that the lane choice frequency can be adjusted by transforming each link of the network into several smaller links. Although this does not make the longitudinal horizon of lane choices dynamic, it does allow it to have an upper bound as opposed to the fully link dependent horizon achieved by using the raw network.

3.6. Spatial Restrictions and Spill-Back

So far the model has a limitation on the number of cyclists that can pass over a cross section due to the headway time between cyclists. However, until now there is no guarantee that the link can actually contain every bicycle that is allocated to the link.

In order to prevent this a sixth and final assumption is added to the model. It secures that a cyclist, \(c\), cannot enter any pseudo-lane of the link, if he/she would make the occupied length/area of the link exceed the actual length/area of the link, i.e.
Can enter, \[ d^c(\tilde{v}_c(t'_l)) \leq L^l - \Omega^l \]

Is denied, \[ d^c(\tilde{v}_c(t'_l)) > L^l - \Omega^l \]  \hspace{1cm} (25)

with \( L^l = |\Psi| \lambda^l \) being the product of the number of pseudo-lanes \(|\Psi|\) and the length of the link \( \lambda^l \), \( \Omega^l \) being the space occupied by cyclists on the link, \( d^c(\tilde{v}_c(t'_l)) \) being the headway distance from eq. (13), and \( \tilde{v}_c(t'_l) \) being the speed assigned to cyclist \( c \) when entering link \( l \) at time \( t'_l \) as defined in eq. (16). By summing over \( d^c(\tilde{v}_c(t'_l)) \) for all cyclists that are currently on the link, \( \Omega^l \) can be found.

When a cyclist is denied, the queue in which he/she is positioned will be processed once a cyclist has successfully left the downstream link. This phenomenon can of course move upstream and cause spill-back over multiple links if the traffic intensity is sufficiently high.

Once the cyclist leaves a link the actual speed of a cyclist on a past link can be found by dividing the length of the link with the time spent on the link. The time waiting to enter a link counts as time spent on the previous link in this regard. This can be formulated mathematically as such: Let \( \lambda^l \) denote the length of link \( l \), \( t^c_l \) denote the time at which a cyclist is allowed entrance at link \( l \), and \( t^c_{l'} \) denote the time at which the cyclist is allowed to enter the following link, \( l' \), then the actual speed at link \( l \) is,

\[ v^c_l = \frac{\lambda^l}{t^c_{l'} - t^c_l} \]  \hspace{1cm} (26)

This implies that the speed assigned by eq. (16) is only upper bound of the actual speed the cyclists traverse the link with. The actual speed will be lower if the cyclist cannot enter the following link immediately, whereas the headway distances are always based on the provisionally assigned speed computed by eq. (16).

3.7. Model Input

One of the benefits of the proposed model is that it requires very little input. As any other model simulating a network it requires a set of links \( L \). As opposed to other models, however, for any link \( l \in L \) it only requires the width \( \omega^l \in \mathbb{R}^+ \) and length \( \lambda^l \in \mathbb{R}^+ \).

In order to have traffic to simulate it also needs demand. It requires the path – an ordered subset of \( L \) – to be known for every cyclist \( c \) in the model population \( C \). Every such path \( r^c, c \in C \) must also have an associated starting time \( t^c_r \in \mathbb{R}^+ \).

Apart from this, the model requires three non-trivial things for input. The first is a lane partitioning formula \( |\Psi| : \mathbb{R}^+ \rightarrow \mathbb{N}^+ \) which can divide any link \( l \in L \) into a positive integer value of pseudo-lanes. In our proposed implementation it is link indifferent, but it can easily be extended to include input of higher dimensions to account for additional link characteristics. The second is a desired speed distribution \( D : [0,1] \rightarrow \mathbb{R}^+ \), which can assign a desired speed \( v^c_r \) to all cyclists \( c \in C \) based on a uniformly random number. Finally, the model requires a formula to define the
headway distance $d : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, which preferably should be a function of speed. Notice that the model also supports individualised headway distance as used in this paper, but that such specification is optional.

4. Experiments

This section is devoted to assessing the capability of the model to give meaningful results when faced with some small-scale experiments. The experimental design is introduced in Section 4.1, before investigating the disaggregate and aggregate properties of the model in Section 4.2. The most questionable assumption concerning the longitudinal horizon of lane choice (see Section 3.5) is analysed in Section 4.3 where simulations with a homogeneous setup are also analysed. Finally, the model is validated against real-life observations in Section 4.4.

4.1. Experiment Design

The network used for small-scale tests is the toy network presented in Figure 6. It consists of three 100 m long links in serial ($\omega^L = [100, 100, 100]$), which are three, three, and two metres wide, respectively, i.e. $\lambda^L = [3, 3, 2]$. Using $|\Psi^l|$ as given by eq. (11), the number of pseudo-lanes are $|\Psi^L| = [3, 3, 2]$. All cyclists have to parse all three links, i.e. $r_c = (l_1, l_2, l_3), \forall c \in C$.

![Figure 6: The toy-network used for small-scale experiments of the proposed model.](image)

Such bottlenecks are quite common along major bicycle corridors in Copenhagen. Along these there are generally room for three efficient lanes, but sometimes, due to the street being especially narrow at some places, the bicycle path also has to narrow to two lanes. The bottleneck setup has been chosen because it is has been a frequently used example since Vickrey (1969), but also since it can be difficult for bicycle traffic to reach a congested regime otherwise (Agarwal et al., 2015).

In order to investigate the relationships between flows, speeds, and densities, simulations are conducted using varying levels of traffic intensity. The intensity is changed by controlling the number of cyclists ($|C|$) that have to ride through the network. A total of 199 experiments are run with in-flows on the first link ranging from $|C| = 50$ to $|C| = 10,000$ cyclists per hour with a constant increment between experiments of 50 cyclists per hour.

Every cyclist is assigned a uniformly distributed arrival time between 0 and 3600 seconds ($t^C \sim U(0, 3600), \forall c \in C$) using the same fixed seed for all experiments. Likewise, also with a fixed seed, every cyclist is given a desired speed drawn from the Johnson’s $S_U$ distribution from eq. (10), i.e. $v^C_0 \sim D, \forall c \in C$.

In all experiments the network is simulated for 3600 seconds, meaning that all cyclists will have entered the beginning of the network. There is, however, no guarantee that the network is fully emptied when the simulation period ends.
All small-scale experiments as well as the actual model implementation were carried out in Java. Concerning the computational efficiency, the experiments showed that the running time was proportional to the number of times a cyclist entered a link, using approximately 1 microsecond ($10^{-6}$ s) per such event. This has also proven to hold in pseudo-large-scale experiments on random, spatially unfounded networks. Assuming link lengths of 100 m and an average trip length of 5 km, this corresponds to simulating 1 million bicycle trips per minute, indicating that the model is indeed fast enough to be large-scale applicable in a traffic assignment model.

4.2. Results

In this section we will take a look at the results of the small scale experiments. The section is twofold with the first focusing on the speed heterogeneous properties of the model. The second part focuses on aggregate verification and which patterns are seen in the corresponding fundamental diagrams.

For all results figures, links are enumerated in increasing order from upstream to downstream, and each link keeps its own unique colour throughout the results section.

4.2.1. Speed Heterogeneous Properties

A central element in traffic simulation is the ability to delay users of the network. Figure 7 shows the proportion of slowed down cyclists on each of the three links given various outflows. A cyclist is defined as being slowed down if he/she has an actual speed strictly lower than his/her desired speed on a link. Outflows are the actual empirical outflows found in the experiments, which often will be lower than the values of $|C|$ under which the experiments were conducted.

Figure 7: Proportion of delayed cyclists on each of the three links as a function of outflow.
Firstly, it is seen that the model is in fact capable of causing delays for some cyclists under moderate traffic intensities, whereas other cyclists are not delayed. At flows of about one third of the capacity, more or less half of the cyclists will be delayed on the final narrow link. This is a feature that macroscopic models cannot include.

Secondly, it is seen that for the first and second link, the proportion of delayed cyclists increases slower for than for the third link when considering low flows. However, since the curves for the two first links are convex, they eventually catch up and also reach a point where everyone will be delayed.

Figure 8 is constructed in order to show which kind of cyclists get delayed at different traffic intensities. The plots show 1,000 actual speeds (from cyclists with ids 1 through 1,000) plotted against the desired speeds for four different experiments using different input flows ($|C|$) to the first link.

Each of the four plots represent different phases. The plot based on 1,000 cyclists represent the first phase where delays are rare and only occurs for relatively fast cyclists. The delays of this phase happens most often on the third link, as this is only has two lanes and thus a limited opportunity of overtaking. The traffic is not high enough for the bottleneck to be an issue.

The second phase can be seen in the plot based on 3,000 cyclists. Under such conditions it is quite normal for cyclists of moderate and fast desired speeds to be delayed. In this phase the delays are frequent on all three links but with the slowest speeds still being found on the last link.

In the third phase represented by 5,000 cyclists really high speeds are almost impossible to achieve. Furthermore, the bottleneck is starting to become a serious problem and many cyclist – even the ones with very low desired speeds – experience delays. Finally, the scatters are also shifted downwards a bit (slower actual speeds) compared to the scatters using 3,000 cyclists.

In the fourth phase 7,000 cyclists are trying to get through the network. This is clearly more than the network can handle, and a serious breakdown occurs on the second and the first link due to the bottleneck. This shows that spill-back effects are clearly captured by the model. Although the traffic fully breaks down to a jam speed of approximate 2.5-4 m/s before the bottleneck, the speeds on the following third link seems almost indistinguishable from what was seen using 5,000 cyclists.
Figure 8: Actual speeds of cyclists with id 1 through 1,000 as a function of desired speeds for various traffic intensities grouped by link.
To summarise the general effect of delays, the speed distribution for the third link for all experiments where $|C| \leq 7,500$ can be found in Figure 9. Values above 7,500 cyclists have not been included in the plot as no noteworthy changes occur beyond this point, probably due to the capacity being reached. It is clearly seen that the distribution becomes steeper, i.e. more homogeneous, as the traffic intensity increases. This is due to the fact that high speeds are no longer possible due to congestion, and that these cyclists have to go slower than they desire. However, even at the very large traffic intensities there is still some heterogeneity present.

Overall the model seems to have the disaggregate properties that were expected from it regarding modelling of speed heterogeneous cyclists.

4.2.2. Aggregate Results

We now move on to see how the model performs on an aggregate level. The disaggregate results of the experiments can be aggregated into well-known measures (fundamental diagrams) that are generally used for assessing traffic.
We start off by presenting the speed-flow curves which can be found in Figure 10, where the individual measurements are enriched with a best fit BPR-curve (Bureau of Public Roads, 1964). Normally the formula describes travel time, but it can be transformed to determine the speed $v_l$ of a link $l$ through the mean speed free flow travel time $\bar{v}_0$, the flow $q$, the capacity parameter $C_l$ (a soft constraint which can be exceeded), as well as the remaining two parameters $\alpha_l$ and $\beta_l$,

$$v_l(q) = \frac{\bar{v}_0}{1 + \alpha_l \left( \frac{q}{C_l} \right)^{\beta_l}}.$$  \hspace{2cm} (27)

The fits are based on the monotonically non-decreasing part of the flows, i.e. for all $|C|$'s from 50 until an increasing $|C|$ no longer provides an increase in the outflow of that link. The $\alpha_l$, $\beta_l$, and $C_l$ parameters are estimated using ordinary least squares, and the parameters for the two relevant links - the first and the second - were found to be 0.43 and 0.74 for $\alpha_l$, 5.19 and 4.99 for $\beta_l$, and 6,094.18 and 6,155.59 for $C_l$, respectively.

It is seen that on the link following the bottleneck the measurements coincide surprisingly well with the BPR function. However, on the two constrained links leading to the bottleneck the speed-flow relation cannot reasonably be assumed to follow the same form as the BPR curve. Both curves have a long slightly decreasing linear beginning after which the speeds starts to drop rapidly with a more sudden change than predicted by the BPR curves.

The maximum observed outflow of the third link (and thus also the two previous in this case) is found to be 5,606 cyclists per hour, corresponding to 2,803 cyclists per hour per pseudo-lane. This is equivalent to between 1,933 and 3,398 cyclists per hour per meter (in width), as two pseudo-lanes occur for bicycle paths with widths ranging from 29
Previous studies have also dealt with determining the maximum outflow of cyclists per hour per meter. Gould and Karner (2009) found capacities to be in the range from 1,395 to 1,641 cyclists per hour per meter through simulation. Jin et al. (2015b) and Zhou and Xu (2015) estimated capacities with mean values of 1,800 and 2,512 cyclists per hour per meter, respectively. Hoogendoorn and Daamen (2016) estimated the capacity to be 1,531 cyclists per hour per meter, but emphasised that it was underestimated due to bidirectional traffic. Liang et al. (2018) found that the capacity is just above 2,000 cyclists per hour per meter, and that the marginal effect of width is diminishing slightly.

Even though the range of our capacity estimate is quite wide, the found capacity still seems to be comparable to what is seen elsewhere in the literature, although possibly in the high end.

The flow-density relation is found in Figure 11. The concavity of the third link is barely visible, as opposed to the clear concave tendencies of the second and first links. The concave form is as expected, and it is seen that the flow of the second link greatly exceeds that of the third link until the maximum flow is reached. This will naturally lead to rapid increase in density on the second link, however without ever reaching the higher maximum density found for the third link.

Earlier studies using on-site data (Gould and Karner, 2009; Jin et al., 2015b; Liang et al., 2018) supports that the flow is a concave, increasing function of the density for low and medium densities. However, even though all three studies extrapolates these results to imply a decreasing part of the curve, the empirical evidence to support this claim is only convincing in Liang et al. (2018). Our model seems to be monotonically non-decreasing for the entire observed range of densities, seemingly contradicting the Chinese field observations. It is still formally unknown how
the real curve is supposed to look like for saturated bicycle traffic in a Danish context, although unsaturated Danish observations will be presented in Section 4.4. It does however contradict with what is already known for other modes of transport (Hoogendoorn and Knoop, 2013).

Finally, in Figure 12 we consider speed-density functions. Once again an almost straight and decreasing line is seen for the third link, whereas the other two links show some sign of negative curvature. This shows that the speed is less vulnerable at low densities, but that it is very volatile for medium to high densities. At jam density, however, the speed seems to be constant at about 3.6 m/s (12.8 km/h). This is significantly higher than in Liang et al. (2018), where the jam speed from Chinese observed data is found to be approximately 0.8 m/s. However, Gould and Karner (2009) shows a jam speed of approximately 2.2 m/s which is somewhat closer to the values obtained in this study, indicating that the results could be plausible. Especially when keeping in mind that the speeds are generally faster in Denmark as outlined in Section 2.1.

![Figure 12: Space mean speeds for each of the three links as a function of density.](image)

4.3. Sensitivity Analyses

In this section two different kinds of sensitivity analyses will be carried out. Section 4.3.1 runs simulation tests where heterogeneity is ignored, whereas Section 4.3.2 focuses on link lengths and the effect their inherent predefined longitudinal horizon of lane choices have.

4.3.1. Ignoring Heterogeneity

One of the main drivers of the proposed methodology is its capability to include heterogeneity in desired speeds as well as for headway distances. This section seeks to investigate how the model reacts when assuming strictly homogeneous cyclists instead.

31
In order to apply homogeneity, all cyclists have to be assigned the same desired speed and headway preference. The used desired speed is the space mean speed of the estimated Johnson $S_\alpha$ distribution, see Section 2.1, namely by $ar{v}_0 = 6.104 \text{ m/s}$. The corresponding headway parameters are based on Model 5 (Section 2.3.1), but whilst assuming $z_{cB} = \frac{1}{2}$, $\forall c \in C$, in eq. (12) corresponding to assigning $\theta^c_0 = \theta_0$ and $\theta^c_1 = \theta_1$ for every cyclist $c \in C$. Running the model using the input parameters presented above results in the speed-flow curve presented in Figure 13.

![Figure 13: Space mean speed as a function of outflow when assuming full homogeneity and corresponding best fit BPR curves.](image)

It can be seen that when using homogeneous input, the straight part of the speed flow curve becomes completely flat, and remains so until an average outflow of about 4,000 cyclists per hour. Above this, the congested regime is entered for the first two links, and the curvature fits very well with the best fit BPR-curves also drawn in the figure. In fact, the results obtained by the model could in this case easily and almost entirely without any loss of information be modelled through the corresponding BPR-curve. This is in contrast to the heterogeneous counterpart, Figure 10, where substituting the simulated results with a BPR curve would alter the results considerably. In the homogeneous case, the $\alpha_l$ parameters of the BPR curve for the first and the second link – the two relevant links – are 14.17 and 56.19, respectively, the $\beta_l$ parameters are 56.19 and 15.23, and the capacity parameters $C_l$ are 6,303.98 and 6,057.30. It should also be noted that even though we see the largest reduction of speed before the bottleneck in both the homogeneous and heterogeneous case, the heterogeneity actually causes speed reductions after the bottleneck too. This is an important feature of the model – and of bicycle traffic in general – that is completely ignored when assuming homogeneity.

Due to the homogeneity the cyclists can be “packed” more tightly than when dealing with heterogeneous cyclists and thus provide a higher max flow, equivalent to what is known for railway traffic for instance. Assuming homo-
geneity increases the maximum flow from 5,717 to 5,886, meaning that the realised capacity of bicycle paths would be overestimated when assuming homogeneity. It also increases the jam speed from 12.8 m/s to 14.5 m/s.

Conclusively, when ignoring heterogeneity it seems as if the model behaves very similarly to traditional flow-based models. This is a nice feature since the model then appropriately can be used on any distribution of desired speeds and headways – even when the variance of these are small.

4.3.2. Link Lengths

The results presented so far have all arisen from a setup with 100 m long links. As introduced in Section 3.5, in the proposed model the link length directly determines the longitudinal horizon for the choice of lanes which subsequently determines the speed. As an indirect consequence it also determines the frequency of lane choices and hence how often the speed can be changed. In this section, we will perform a sensitivity analysis that explores the consequences of adjusting this by repeating the experiments with links lengths of 20, 50, 200, and 500 metres, respectively.

Figure 14: Sensitivity analysis of slowed down proportion as a function of outflow for various link lengths.

We first evaluate how the proportions of delayed cyclists are affected by the link length. The relevant plots can be found in Figure 14. The model can be seen to keep its capability of delaying cyclists under moderate traffic intensities...
regardless of the link length. The curvature of the lines are influenced considerably by the link lengths though, with longer links leading to more concave functions and vice versa. The effect on the narrower third link is especially notable. Here the limited possibility of overtaking due to only having two lanes becomes paramount when the link lengths get very long, whereas the problem is minor when having a short link length. In fact, when having a very short link length of 20 m the proportion of delayed cyclists on the third link will never exceed 95 % as opposed to the other investigated link lengths.

Figure 15: Sensitivity analysis of the speed distribution for the last link as a function of cyclists inflows on the first link for various link lengths.

Figure 15 shows the effect on the speed distributions for the last link. It generally shows that the longer the link, the prompter the effect kicks in. Using a very long link length (500 m) seems to be a little too extreme, causing a dramatic speed reduction even for low traffic intensities. Reducing the links to a very short length (20 m), on the other hand, seems to invoke a more speed heterogeneous (flatter) distribution where the effect of increased traffic intensity is somewhat smaller.
Finally, we evaluate how the link length affects one of the fundamental diagrams, namely the speed-flow relation (see Figure 16). When using a link length of 20 m or 50 m the curvature of the second link is rather soft, never allowing the flow to be large enough to force the curve to be really steep. In these two cases the corresponding BPR-curves fit tremendously well, with the parameters for the second link being \((\alpha, \beta) = (0.91, 4.29)\) and \((0.91, 6.30)\) for 20 m and 50 m, respectively. For the longer links, the final part of the curves are very steep, eventually curving inwards towards the y-axis as known from speed-flow diagrams of other modes. Interestingly, when having long links (200 m and 500 m) the effect of the restricted possibility to overtake exceeds that of the bottleneck right until the flow capacity is reached. For link lengths of 50 m this happens slightly before the capacity is reached, whereas it happens considerably earlier when using really short links lengths of 20 m where the bottleneck effect is more prominent and dominates the effect of restricted overtaking for high traffic intensities.

Another interesting finding is that the maximum flow is slightly higher when using a link length of 50 or 200 metres, than when using 20 metres. This seems counter-intuitive at first, as having link lengths of 20 metres would improve the conditions for overtaking as lane choices happen much more frequently. However, the links also becomes so small, that the surplus space that is not used – because adding another cyclist to the link would make the occupied
area of the link exceed the actual area, see Section 3.6 – constitutes a considerable part of the area. This means that
the links will be poorly utilised in such cases, which most likely can explain the surprisingly lower flows using very
short links. This is supported by the fact that the maximum average density found in the experiments is much lower
for these short links (133 cyclists per lane-km), as opposed to 141, 148, 155, and 162 cyclists per lane-km for links of
length 50, 100, 200, and 500 metres, respectively.

Although the model has been found to be rather sensitive to changes in link length, it seems to behave reasonably
as long as the link lengths do not become too short or too long. In general, using \( \lambda = 50 \), \( \forall \in L \) provide results which
intuitively seem plausible.

4.4. Model Validation

In order to validate the model, on-site observations have been collected from Queen Louise’s Bridge in central
Copenhagen, using the same video-approach as in Section 2. Queen Louise’s Bridge is the street with the highest
maximum bicycle flow in Copenhagen with 3,763 cyclists between 8 a.m. and 9 a.m. in the direction towards the city
centre according to City of Copenhagen (2018). This has been the primary reason for choosing this particular street.
A single frame of the video data that covered a time period from 7.30 a.m to 8.30 a.m. on a weekday morning can be
found in Figure 17.

![Figure 17: A single frame from the video data collected at Queen Louise’s Bridge in Copenhagen.](image)

The bicycle path on the street is 3.9 metres wide, which corresponds to three efficient lanes according to eq. (11).
As opposed to our test network there is no bottleneck following Queen Louise’s Bridge. Because of this, we have
generated new simulation results using a modified network where all of the three links have three efficient lanes (the
last link only had two lanes in the test network, see Figure 6). Notice that according to eq. (11) widths of both 3.0 and
3.9 metres corresponds to three efficient lanes. This allows the results from the model to be compared directly with
the empirically collected observations.

The video data has been subdivided into time intervals of 30 s in which flows as well as average speeds and
densities are calculated. Jin et al. (2015b) uses a temporal division of the collected data of 15 seconds but states that
finding the optimal time interval is still a topic of further research. With our data we have found that using a time
interval of 30 s, as also seen in Gould and Karner (2009), seems as a reasonable interval.

Two cross sections are drawn across the bicycle path 10 and 20 metres away from the camera, respectively.
Although a sample area length of 10 metres may seem small, it is actually larger than all measurements used in Gould
and Karner (2009) and Jin et al. (2015b) which use lengths down to 2.9 m. The flow is determined by the number of
cyclists crossing the second cross section within the time interval, and adjusted to be in the unit cyclist per lane per
hour. Speeds for every cyclist are calculated as the distance divided by the time between the two cross sections. The
average speed are subsequently calculated by the harmonic mean of the speeds of all cyclists crossing the second cross
section within the time interval. For determining densities, each time interval is subdivided into slices of 0.1 s. In each
of such slices, the number of cyclists between the two cross sections is counted, and the average density of a time
interval is then defined as the average number of cyclists in the area per time slice divided by the distance between the
two cross sections and the number of lanes.
The corresponding fundamental diagrams emerging from the observed data and for the non-bottleneck simulations can be found in Figure 18. Firstly, the plots show that without the bottleneck setup it is very hard to reach the saturated regime of the curves. This is a general problem for bicycle traffic (as also mentioned in Agarwal et al. (2015)), and we see the problem here for both the simulations and for the empirical data. However, as the street from which the data is collected is the street with the highest traffic intensity for bicycles in all of Copenhagen it probably still is the best obtainable data to compare against.

The speed-flow and speed-density curves show that our simulations have speeds that are a little too high (roughly 10%) for low flows and densities compared to the observed data. This could indicate that there are some biases in either of the two video datasets that we have not corrected for, e.g. wind, gradient, or quality of the asphalt. Despite the starting point of the observations being a little too low for both the speed-flow and speed-density diagrams, we do observe a decreasing trend on the right half of the data that is very similar to that of the simulations.

Regarding the flow-density diagram we see surprisingly similar results for the empirical data and the simulations across the entire spectrum of densities. However, on this part of the curve the tendency is almost strictly linear, and
only the very beginning of a concave tendency can be seen from the data. This is in line with what is seen for low densities for empirical data of traffic that can be modelled appropriately by a BPR curve (Nielsen and Jørgensen, 2008). The simulations do have a concave form although it is far less evident than when using the bottleneck setup, see Figure 10.

Overall, although our model seems to predict a little too high speeds for freespeed-like situations the simulated results are very close to the observed data for high traffic intensities. This is in particular striking when keeping in mind that the model is calibrated using external data collected at lower traffic intensities and at a different location, i.e. it is not calibrated using the data it is compared against in contrast to most other models on bicycle traffic. This is a deliberate choice since one of the aims of this study is to model bicycle traffic at intensities exceeding what is currently observable at bicycle paths. As such, it is fully expected that our current model has some deviations relative to the empirical data. The deviations between the model and observations may later be minimised by a slight adjustment/recalibration of the input distributions, meaning that these minor discrepancies should not be considered as evidence against the proposed methodology as a whole.

It has obviously not been possible to perform a fully disaggregate evaluation of on-site observations. The desired speed can only be revealed in uncongested situations and the fundamental diagrams require congestion to be established. As such there is no straightforward way to acquire both types of information simultaneously through the available video data.

5. Conclusions and Future Work

The purpose of this paper was to develop a bicycle traffic simulation model capable of handling cyclist speed heterogeneity and its inherent interactions between cyclists. Individual desired speeds from Johnson’s $S_U$ distribution as well as individual beta-distributed headway distance parameters were assigned for each cyclist. Using simple assumptions on lane formation and choices on multi-lane bicycle paths, the simulation results of a bottleneck test network revealed that differences in desired speeds and headway distances are indeed one possible way to explain and model delays on bicycle paths. Even though the likelihood of such delays increases with the traffic intensity, delays also occur at moderate flow levels – especially for cyclists with high desired speeds due to the implicit modelling of overtaking restrictions. Although the assumptions underlying the model are simple, the resulting fundamental diagrams showed reasonable resemblance to previous empirical analyses of bicycle traffic as well as existing simulation-based analyses found in literature. Also the resulting lane capacities of the model were reasonably close to what has previously been found in the literature.

Furthermore, the model was also validated against data from the location with the highest possible bicycle traffic intensity in Copenhagen. However, it was found that under such non-bottleneck setup it is very difficult for bicycle traffic to reach a fully saturated regime. It neither happened in our model nor in the empirical data, why the model validation was only possible for the unsaturated part of the fundamental diagrams. Although the baseline speeds
were a little too high in the simulation results compared to the empirical observations, simulation model resulted in fundamental diagrams that fit the real-life observations well for moderate and high traffic intensities.

It was also shown that accounting for heterogeneity altered the fundamental diagrams by causing speed reductions at an earlier stage than when using a fully homogeneous setup. This suggests that such heterogeneity should not be ignored when modelling bicycle traffic.

The disaggregate aspects of the model that allow delays to occur under low traffic intensities could unfortunately not be validated using the acquired video data, since it is impossible to derive desired speeds of cyclists in congested situations. It could be interesting to look further into how to collect data to support such analysis in future work, for instance through providing GPS trackers to a small portion of the cyclists that are known to regularly pass through the video analysis area.

Link lengths play an important role in the proposed model, as they determine the longitudinal horizon for lane choices. Therefore, a sensitivity analysis was carried out concerning the link lengths. Overall, it seemed as if the model behaved reasonably when varying the link lengths. It did, however, also show evident changes in the curves, and that it seemed as if having too short or too long links can make the model behave undesirably. The model was relatively stable in the interval from 50 m to 100 m, though, with both values providing appropriate results.

Besides seemingly being theoretically well-founded, the proposed model also proved to scale well, with low running times increasing linearly with the number of cyclist/link interactions. This shows that the model is most likely large-scale applicable, and will be suitable for implementing in a traffic simulation for an entire metropolitan area. The current study only deals with link modelling, though. In order to have a model that also works realistically when cyclists have routes that cross each other and those of other modes, a proper intersection modelling that can be integrated with the model proposed in this study needs to be developed too.

The model can be used for various performance and policy analysis purposes. For instance, it can be used to provide insight into determining at which point the current width of bicycle paths will no longer provide sufficient capacity for the traffic demand and hence should be extended to optimise traffic flow. It can also be used to investigate the consequences of the introduction of a new type of mode on the bicycle paths, e.g. a large share of high-speed pedelecs. This would make the desired speed distribution highly bipolar, making the model even more relevant to use and general methods based on averages even more inappropriate. The model would also be able to predict to which degree the pedelecs would be able to reach their full speed potential when riding on the bicycle paths.

Most importantly, however, the theoretical foundation developed in this paper serves as an important step towards allowing bicycle traffic to be modelled at a similar level-of-detail to that of car traffic including detailed congestion patterns. This allows for the travel time of the two modes to be compared fairly in the analyses of future scenarios. Moreover, it would be possible to more accurately investigate the effects on major investments in bicycular infrastructure as a way to increase urban mobility in metropolitan areas.
Acknowledgements

The authors would like to sincerely thank Rasmus Albrink from RAW Mobility Ltd (previously working for COWI Ltd) for collecting, preprocessing, and discussing relevant aspects of the obtained video data. COWI Ltd kindly allowed us to use their data from Smallegade for analysing and estimating key model parameters from empirical observations. RAW Mobility Ltd collected the video data from Queen Louise’s Bridge used to establish the on-site fundamental diagram for congested bicycle traffic, which we are also very thankful for.

We would also like to thank the two anonymous referees for their highly constructive criticism which helped improve the quality of the paper.

References


https://www.rawmobility.dk

https://www.cowi.com


Mai, X., Lv, W., Wei, X., Song, W., Jiang, R., 2013. Analyzing the Characteristics of Unidirectional Bicycle Movement Around a Track based


