Analysis of Bicycle Headway Distribution, Saturation Flow and Capacity at a Signalized Intersection using Empirical Trajectory Data

Yuan, Yufei; Goni Ros, Bernat; Poppe, Mees; Daamen, Winnie; Hoogendoorn, Serge

DOI
10.1177/0361198119839976

Publication date
2019

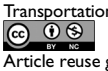
Document Version
Final published version

Published in
Transportation Research Record

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–12



DOI: 10.1177/0361198119839976
journals.sagepub.com/home/trr


Analysis of Bicycle Headway Distribution, Saturation Flow and Capacity at a Signalized Intersection using Empirical Trajectory Data

Yufei Yuan[1], Bernat Goñi-Ros[1], Mees Poppe[1], Winnie Daamen[1], and Serge P. Hoogendoorn[1]

Abstract

Predicting the bicycle flow capacity at signalized intersections of various characteristics is crucial for urban infrastructure design and traffic management. However, it is also a difficult task because of the large heterogeneity in cycling behavior and several limitations of traditional capacity estimation methods. This paper proposes several methodological improvements, illustrates them using high-resolution trajectory data collected at a busy signalized intersection in the Netherlands, and investigates the influence of key variables of capacity estimation. More specifically, it shows that the (virtual) sublane width has a significant effect on the shape of the headway distribution at the stop line. Furthermore, a new method is proposed to calculate the saturation headway (a key variable determining capacity), which excludes the cyclists initially located close to the stop line using a distance-based rule instead of a fixed number (as is usually done in practice). It is also shown that the saturation headway is quite sensitive to the sublane width. Moreover, a new, empirically based method is proposed to identify the number of sublanes that can be accommodated in a given cycle path, which is another key influencing variable. This method yields considerably lower estimates of the number of sublanes than traditional methods, which rely solely on the (available) cycle path width. Finally, the authors show that methodological choices such as the sublane width and the method used to estimate the number of sublanes have a considerable effect on capacity estimates. Therefore, this paper highlights the need to define a sound methodology to estimate bicycle flow capacity at signalized intersections and proposes some steps to move toward that direction.

The ability to predict the bicycle flow capacity of cycle paths at signalized intersections is crucial for urban infrastructure design and traffic management. Capacity (maximum hourly rate at which bicycles can pass the stop line of cycle paths) depends on various key variables, such as the geometric characteristics of cycle paths, the interference with other traffic modes (e.g., car traffic), the traffic control schemes, and the cyclist headway distributions. Estimation of the latter is a difficult task because of the large heterogeneity in cycling behavior. Previous studies on headway distribution models mainly focus on motorized vehicles, and can be split into two categories. The first states that all the headways can be modeled from a single distribution. Examples are the normal distribution (*1*), the exponential distribution (*2*), the gamma distribution (*3*), and the lognormal distribution (*4*). The second category consists of composite headway models. In this case, the total headway distribution is made up of two different headway distributions, resulting from whether the vehicle/bicycle is following the vehicle/bicycle in front or moving freely. Examples of composite headway models are a semi-Poisson model (*5*), a generalized queueing model (*6*), and a distribution-free approach (*7*). Applications of headway distribution models on cyclist data are rare. A distribution-free (or rather, non-parametric) model estimation approach following the theory of composite headways has been performed on a cyclist intersection in the Netherlands (*8*).

Another key variable of capacity estimation is the saturation flow, which is often calculated as the inverse of the average headway observed between cyclists and

[1]Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Transport and Planning, Delft, The Netherlands

Corresponding Author:
Address correspondence to Yufei Yuan: y.yuan@tudelft.nl

their leaders once the queue is moving in a stable manner, multiplied by the number of cyclist sublanes that the cycle path can accommodate. There are multiple examples of empirically derived saturation flows, and various values of saturation flows have been reported. A study in the Netherlands reported 3,000–3,500 cyclists per hour for a 0.78 m-wide cyclist lane (*9*). An empirical analysis in California and Colorado suggested a maximum flow rate of 4,500 cyclists per hour on a 2.43 m-wide cycle path, and it concluded that cyclist saturation headway started at the sixth cyclist in a queue (*10*). A case study in Beijing, China found 1,836–2,088 cyclists per hour for a 1.25 m-wide path (*11*). A 2 m-wide cycle-track in Santiago de Chile was reported to have a saturation flow of 4,657 cyclists per hour, whereas for Tavistock Square in London the saturation flow was equal to 4,320 cyclists per hour for a path with the width of 1 m (*12*). Another point of interest related to saturation flow calculation is the cycle path width. Usually, cyclists can form multiple queues as the "sublane" of a bicycle path is not fixed; this is different from the lane usage of motorized vehicles. In the literature, many different sublane widths have been suggested, for example, 0.78 m in the Netherlands (*9*), 1.00 m in Germany (*13*) and even 1.60 m in Norway (*14*). However, there is a lack of a unified paradigm to compute saturation flow rate and thus the bicycle flow capacity. Besides, there are several limitations in traditional calculation methods. For instance, the calculation of the number of sublanes in Botma and Papendrecht was solely based on the available cycle path width, and ignored leader–follower relations in queue discharge processes (*9*); the calculation of saturation headway in Raksuntorn and Khan suffered from the problem of headways close to zero, because leader–follower pairs were not clearly identified (*10*).

Generally, the capacity of a cycle path at a signalized intersection (C) is calculated as follows (*9*, *10*):

$$C = q_s \cdot \frac{T_e}{T} \tag{1}$$

where:

$$q_s = \varphi \cdot \frac{1}{h_s} \tag{2}$$

and:

$$T_e = T_G - T_L + T_{YG} \tag{3}$$

In Equations 1–3: q_s denotes the saturation flow; h_s denotes the saturation headway between leader–follower pairs; T denotes the total cycle time; T_e is the effective green time; T_G is the total green time; T_L is the start-up lost time; T_{YG} is the period of the yellow phase in which cyclists still pass the stop line; and φ is the number of virtual sublanes in a given cycle path. This paper proposes a series of methodological improvements on capacity estimation, illustrates them using high-resolution trajectory data collected at a busy intersection in the Netherlands, and investigates the influence of key variables. More specifically, it shows that the (virtual) sublane width has a significant effect on the shape of the headway distribution at the stop line. Furthermore, a new method is proposed to calculate the saturation headway h_s (a key variable determining capacity), which excludes the cyclists initially located close to the stop line using a distance-based rule instead of a fixed number (as is usually done in practice), and the authors show that the saturation headway is quite sensitive to the sublane width. Moreover, a new, empirically based method is proposed to identify the number of sublanes φ available to cyclists in a given cycle path, which is another key variable influencing capacity. The authors show that this method yields considerably lower estimates of the number of available sublanes than traditional methods, which rely solely on the (available) cycle path width. Finally, it is shown that methodological choices such as the values of the sublane width, as well as the method used to estimate the number of available sublanes, have a considerable impact on capacity estimates.

The rest of this paper is structured as follows. the next section first describes the data set used in this study, and then illustrates the calculations of several key variables of capacity estimation. The results of the analyses are then presented. The final section contains the conclusions of this study and some suggestions for future research.

Data and Methods

This section describes the empirical trajectory data set analyzed in this study, and then it describes the method used to measure cyclist headways at the stop line as well as the methods proposed to estimate the saturation headway, the saturation flow, the start-up lost time and thus the capacity of cycle paths at signalized intersections. For clarification and easy reference purposes, all the variables introduced in this work and their definitions are listed in Table 1.

Data Set Characteristics

The data set analyzed in this study consists of the trajectories of 691 cyclists moving along a cycle path that leads to a signalized intersection in Amsterdam (the Netherlands). The cycle path is 2 m wide, unidirectional, and segregated from car traffic. Traffic demand for this cycle path is relatively high in peak hours, and consists of people riding both bicycles and scooters. In this data set the percentage of scooters is very low; hereafter we

Table 1. Variables and Their Definitions (in Alphabetical Order of the Notation Names)

Notations	Definitions
C	capacity of a cycle path at a signalized intersection
d	distance to the stop line, $d \in D$
d_s	distance threshold leading to the most accurate estimate of the saturation headway
D	a set containing various consecutive values of d
$\bar{h}_c(d)$	average headway of cyclists closer to the stop line than d
$\bar{h}_f(d)$	average headway of cyclists farther from the stop line than d
$\bar{h}_f^E(d)$	piecewise linear function fitting to the $(d, \bar{h}_f(d))$ points
h_i	headway of cyclist i
h_s	saturation headway between leader–follower pairs
i	cyclist index, $i = 1, 2, \ldots, N$. It gives the rank of the cyclists regarding the time when they pass cross-section x_{ref}
j	index of the leader of cyclist i ($j<i$)
k	index of queue discharge period
$M_c(d)$	a set containing the indices i of all cyclists whose initial longitudinal position x_i is higher than $x_{ref} - d$ for a given value of distance d
$M_f(d)$	a set containing the indices i of all cyclists whose initial longitudinal position x_i is lower than $x_{ref} - d$ for a given value of distance d
N	total number of cyclists in a queue
$N_c^k(d)$	number of cyclists standing closer to the stop lane than d in a queue discharge period k
$N_c^E(d)$	linear function fitting to the $(d, N_c^k(d))$ data points of all the selected discharge periods
q_s	saturation flow
R_λ^k	number of cyclists who are in each position λ of a chain in each queue discharge period k
t_c	headway increment because of the start-up reaction and acceleration
t_i	time instant when cyclist i passes cross-section x_{ref}
t_G	time instant when the green phase starts
T	total cycle time
T_e	effective green time
T_G	total green time
T_L	start-up lost time
T_{YG}	period of the yellow phase where cyclists still pass the stop line
w	width of a virtual sublane for each cyclist
W	physical width of a cycle path
W_a	available cycle path width
W_u	actually used cycle path width
x_i	longitudinal position of cyclist i
x_{ref}	position of a reference cross-section along the roadway used to calculate headways
$y_i(t_i)$	lateral displacement of cyclist i at time instant t_i
λ	index of cyclist position in a leader–follower chain, $\lambda = 1, \cdots, \Lambda$
Λ^k	total number of chain positions in each queue discharge period k
φ	number of virtual sublanes in a given cycle path
φ_E	empirical number of sublanes
φ_T	theoretical number of sublanes
$\mathcal{K}_N$	kernel distribution function, which is a non-parametric function
$\mathcal{N}(\mu_N, \sigma_N)$	normal distribution function
μ_N and σ_N	mean and standard deviation of the headways
$\mathcal{X}(\mu_X, \sigma_X)$	lognormal distribution function
μ_X and σ_X	mean and standard deviation of the natural logarithm of the headways

refer to these as *cyclists* in both cases with an assumption that there is no effect on the analysis results. Access to the intersection is regulated by a traffic light.

The trajectories were derived from top-view video images using the methodology described in Goñi-Ros et al. (*15*). The video images were obtained using two cameras mounted on a 10 m pole, which was placed next to the cycle path (more details about this type of installation can be found in Duives [*16*]). The two camera views can be seen in Figure 1 in (*15*). The two camera views combined covered 20 m of cycle path upstream of the traffic light. The frame rate varied between 5 and 10 fps during the video recordings.

Traffic along this cycle path stretch was recorded between 12:45 and 19:00 h on June 6, 2016. By looking at the video footage, the queue discharge periods that

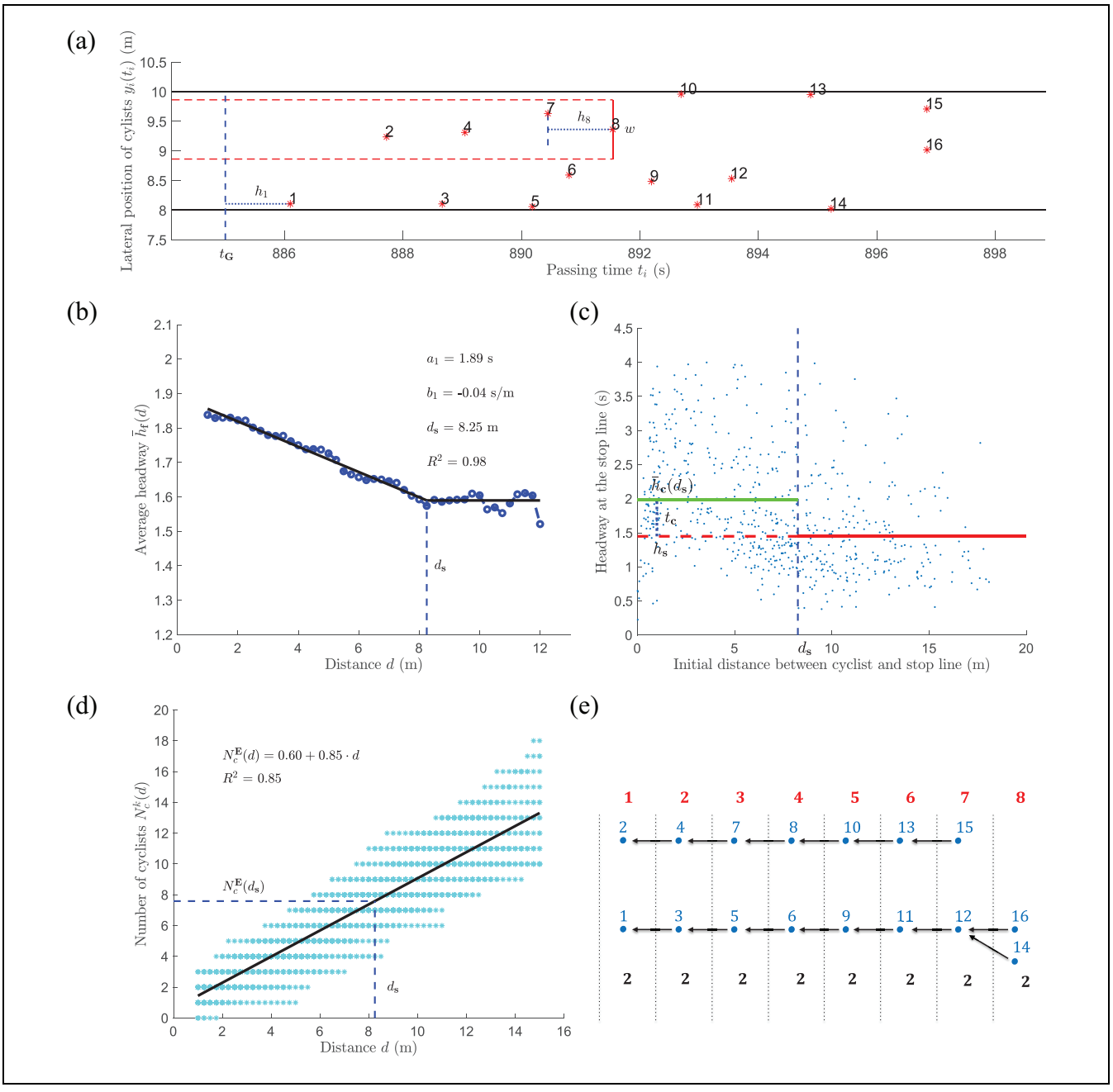


Figure 1. Illustration of the methods used to calculate different variables. (*a*) Illustration of the method used to calculate the time headways (w = 1.0 m). The numbers assigned to (t_i, y_i) points correspond to the cyclist indices i. Green phase starts from time t_G; (*b*) Example of piecewise-linear function fit to determine the distance threshold d_s (w = 1.0 m); (*c*) Illustration of the method used to calculate the saturation headway (h_s) and the headway increment t_c, once the distance threshold d_s has been identified (w = 1.0 m); (*d*) Linear relation between N_c^k (d) and d. Light blue markers denote the number of cyclists standing closer to the stop line than d in each queue discharge period k. (e) Illustration of the method used to empirically determine the number of sublanes (based on the data for the same queue discharge period shown in Figure 1[*a*]). Blue numbers correspond to cyclist identification indices (i); red numbers correspond to the cyclist position within the chain (λ); and black numbers correspond to the R_λ^k values. In this case, $\varphi_E^k = 2$.

met the following criteria were selected (see also Goñi-Ros et al. [*15*]): (a) the queue is formed by at least seven cyclists; (b) no more than two cyclists within the queue are riding scooters; (c) all cyclists move toward the stop line and pass it without getting out of the path; (d) no pedestrian crosses the cycle path during the queue discharge process; (e) the discharge process is not affected by downstream traffic conditions; and

(f) all cyclists pass the stop line before the end of the green phase.

The total number of selected queue discharge periods was 57, with an average queue size of 12.12 cyclists, 2% of which were scooters (see also Goñi-Ros et al. [*15*]). For every selected period, the trajectories of all cyclists forming part of the queue were derived using a procedure that comprises six steps: (1) video clip decomposition; (2) manual cyclist tracking; (3) height transformation; (4) orthorectification and scaling; (5) time coding; and (6) trajectory merging. Please refer to Goñi-Ros et al. (*15*) for a full description of this procedure.

Definition of Headway of a Cyclist

In road traffic, the (gross) time headway—or simply *headway*—of a vehicle n is defined as the difference between the time when a certain part (e.g., rear bumper) of the preceding vehicle in the same lane (i.e., vehicle $n-1$) passes a reference cross-section along the roadway (x_{ref}) and the time when the same part of vehicle n passes that cross-section (*17*). Thus defined, the value of the headway is a direct consequence of the longitudinal acceleration behavior of the driver, which is mostly influenced by the movement of the preceding vehicle in the same lane, and can be used to calculate the capacity of a roadway (*18*). In bicycle traffic, however, the definition of headway is less straightforward, as cyclists can generally not be allocated to well-defined lanes. This makes it difficult to identify the preceding cyclist that constrains the longitudinal acceleration of a given cyclist.

To overcome this issue, headways can be defined based on the concept of *virtual sublane*, as proposed, for example, by Hoogendoorn and Daamen (*8*) and Botma and Papendrecht (*9*). For a given queue discharge period, let us assign an index i to every cyclist in the queue, where the index gives the rank of the cyclists regarding the time they pass cross-section x_{ref} ($i = 1, 2, \ldots, N$, where N is the total number of cyclists in the queue). Furthermore, let t_i denote the time instant when cyclist i passes cross-section x_{ref}, and let $y_i(t_i)$ denote the lateral displacement of cyclist i at time instant t_i. Then, the following rule can be used to determine leader j (see also Figure 1*a*). The leader of cyclist i is the cyclist with the largest index $j<i$ for which:

$$|y_j(t_j) - y_i(t_i)| \leqslant \frac{w}{2} \tag{4}$$

where $w>0$ denotes the width of a virtual sublane for each cyclist.

Parameter w corresponds to the minimum width of the space right in front of the cyclist that needs to be empty so that he/she can move forward in a safe manner. The value of w is around 75–80 cm in Hoogendoorn and Daamen (*8*) and Botma and Papendrecht (*9*), 100 cm in Brilon et al. (*13*), and 160 cm in Allen at al. (*14*). This difference between studies are partially as a result of the different bicycle points used to define lateral positions y_i and y_j (e.g., center or end of handlebar) in each study. Another reason is the lack of knowledge on the size of the lateral influencing area of cyclists (swerving and swaying), which leads to different assumptions on the value of w.

Here, the leader of cyclist i is defined as the cyclist that passed cross-section x_{ref} earlier than cyclist i closer in time, within the virtual sublane used by cyclist i (which has a width equal to w). Once cyclist i has been assigned a leader j, his/her headway h_i can be calculated as follows (see also Figure 1*a*):

$$h_i = t_i - t_j \tag{5}$$

Note that using the method described above, it is not possible to assign a leader to some cyclists, particularly those who are initially located close to the stop line (because no cyclist passes the stop line before them within their virtual lane). For these cyclists, it is assumed that t_j in Equation 5 is the time instant when the green phase starts (t_G), as shown in Figure 1*a* for the calculation of h_1.

Headway Distribution Modeling

Using the method described in above, the headways of cyclists at the intersection stop line (i.e., x_{ref} was defined as the cross-section of the stop line) in the 57 selected queue discharge periods were calculated. As will be discussed later, the headways at the stop line are relevant for the estimation of saturation headway, start-up lost time, and thus cycle path capacity.

A histogram of the headways of all cyclists who have a leader (all following headways) was built. Three different functions were fitted to the headway data to determine the shape of the distribution: (a) normal distribution function, $\mathcal{N}(\mu_{\text{N}}, \sigma_{\text{N}})$; (b) lognormal distribution function, $\mathcal{X}(\mu_{\text{X}}, \sigma_{\text{X}})$; and (c) kernel distribution function, $\mathcal{K}_{\text{N}}$. The first and second functions are well-known parametric functions, where parameters μ_{N} and σ_{N} correspond to the mean and standard deviation of the headways, and parameters μ_{X} and σ_{X} correspond to the mean and standard deviation of the natural logarithm of the headways. Instead, $\mathcal{K}_{\text{N}}$ is a non-parametric function, which has been fit using the MATLAB function fitdist(x,'kernel') with a normal smoothing function and default parameter values.

Note that a critical influencing factor of the headway calculation method (see "Definition of Headway of a Cyclist") is the virtual sublane width (w). To investigate the sensitivity of the headway distribution to the virtual

sublane width, the authors calculated the headways, built the histograms and fit the distribution functions for various values of w.

Calculation of Saturation Headway

Similarly to road traffic (*19*), the *saturation headway* for cyclists can be defined as the average headway between cyclists and their leaders once the queue is moving in a stable manner. Like in road traffic, the first group of cyclists in a queue generally have longer headways than cyclists located more upstream (*10*). The main reason is that they generally lose time reacting to the traffic light (*15*) and they pass the stop line when still accelerating to their desired speed, whereas the other cyclists pass the stop line at more or less constant speed. For this reason, in practice, the headways of the first group of cyclists are excluded when calculating the saturation headway (*10*).

As mentioned above, when calculating the saturation headway, the basic assumption is that once the queue moves in a stable manner, the average headway stays constant, whereas the average headway of the first group of cyclists is longer. Therefore, the main challenge is to determine which headways need to be excluded from the calculation. For example, Raksuntorn and Khan analyzed the discharge process on a 2.43 m-wide cycle path and estimated that the headways of the first five cyclists in the queue need to be excluded (*10*); note, however, that they did not calculate the headways based on the concept of virtual sublane (see "Definition of Headway of a Cyclist"), so their headways are not necessarily the difference in passing times between follower–leader pairs. Here, the authors propose to exclude not a fixed number of cyclists (like in Raksuntorn and Khan [*10*]), but those cyclists who are initially closer than a certain distance to the stop line (d_s, distance threshold). The reason is that the authors deem the initial distance to the stop line to influence the speed of cyclists at the stop line more than their ranking regarding their initial proximity to the stop line. In other words, beyond a certain initial distance d_s, cyclists are assumed to have enough space to accelerate to their desired speed before crossing the stop line, and only then they are included in the calculation of the saturation headway.

To determine the distance threshold d_s leading to the most accurate estimate of the saturation headway, a model-based approach was used (see also Figure 1*b*). Let $M_c(d)$ and $M_f(d)$ denote two disjoint sets containing the indices i of all cyclists (in all selected discharge periods) whose initial longitudinal position x_i is higher and lower than $x_{ref} - d$, respectively, for a given value of distance d. Here, the headways of cyclists who did not have a leader were excluded. Firstly, a set D containing various consecutive values of d was defined, and the average headways contained in sets $M_c(d)$ and $M_f(d)$ calculated for all values of d contained in set D. Thus, the average headway of cyclists closer to the stop line than d and the average headway of cyclists farther from the stop line than d (which can be denoted by $\bar{h}_c(d)$ and $\bar{h}_f(d)$, respectively) were obtained for all the specified values of d. Secondly, the authors visualized the results in a plot where the axes were d and the average headway in set $M_f(d)$ (see Figure 1*b*). A visual inspection of this plot shows that the average headway of cyclists initially located farther from the stop line than d decreases more or less linearly with d and becomes quite stable after a certain value of d. Thirdly, the value of d that leads to the most accurate estimate of the saturation headway (i.e., d_s) was identified by fitting the following piecewise linear function to the $(d, \bar{h}_f(d))$ points (note that the data points for $d > 12$ m were not used to fit the piecewise linear function, because for $d > 12$ m the number of elements (cyclists) in set $M_f(d)$ is too low and the average headway $\bar{h}_f(d)$ varies too much for different values of d). (see also Figure 1*b*):

$$\bar{h}_f^E(d) = \begin{cases} a_1 + b_1 \cdot d & \text{if } d < d_s \\ A_1 & \text{if } d \geq d_s \end{cases} \quad (6)$$

which has three parameters: a_1, b_1 and d_s, and $A_1 = a_1 + b_1 \cdot d_s$.

Once the distance threshold d_s has been derived, the saturation headway (h_s) can be defined as the average headway of all cyclists initially located farther from the stop line than d_s (see also Figure 1*c*):

$$h_s = \bar{h}_f(d_s) \quad (7)$$

Note that the distance threshold d_s may be different for different values of w. The authors calculated the saturation headway using the headway data derived using different virtuals sublane widths (w) to investigate the sensitivity of the saturation headway to the value of w. In all cases, the piecewise linear model (Equation 6) fits the $(d, \bar{h}_f(d))$ data points reasonably well (R^2 higher than 0.96).

Calculation of Start-Up Lost Time

As discussed in "Calculation of Saturation Headway", the first group of cyclists in the queue (more specifically, those located closer to the stop line than d_s), generally experience a start-up reaction and acceleration time, which causes them to have longer headways at the stop line on average than the following cyclists. For the latter, the influence of start-up reaction and acceleration on their headways at the stop line is minimal, so their headway is shorter (h_s, on average).

The average headway of the first group of cyclists in the queue (both with and without a leader) can be decomposed into two elements: the saturation headway (h_s) and a headway increment t_c resulting from the start-up reaction and acceleration, which can thus be defined as follows (see also Figure 1*c*):

$$t_c = \bar{h}_c(d_s) - h_s \tag{8}$$

Similar to the work in Raksuntorn and Khan (*10*), the start-up lost time (T_L) can then be defined as the headway increment (t_c) multiplied by the expected number of cyclists standing closer to the stop line than d_s before the traffic light turns green ($N_c^E(d_s)$):

$$T_L = N_c^E(d_s) \cdot t_c \tag{9}$$

Variable $N_c^E(d_s)$ in Equation 9 is determined using a linear model estimated by fitting the following linear function to the $(d, N_c^k(d))$ data points of all the selected discharge periods k (see Figure 1*d*):

$$N_c^E(d) = a_2 + b_2 \cdot d \tag{10}$$

where: a_2 and b_2 are constants; d denotes an element (distance value) of set D (see "Calculation of Saturation Headway"); and k is an index identifying the queue discharge period.

As shown in Figure 1*d*, variable $N_c^E(d_s)$ is obtained by substituting d in Equation 10 by the value of d_s, which is derived as explained in "Calculation of Saturation Headway". The start-up lost time assuming different virtual sublane widths (w) was calculated to investigate the sensitivity of T_L to the value of w.

Calculation of the Number of Sublanes

The number of virtual sublanes (φ) can be calculated either theoretically (which is the traditional approach, used for example in Botma and Papendrecht (*9*)) or based on empirical observations as proposed in this work. Theoretically, φ could be calculated as follows (*9*):

$$\varphi_T = \frac{W_a}{w} \tag{11}$$

where W_a is the available cycle path width, which is the path width that is actually used by cyclists if we look at their positions on the ground (W_u) plus $w/2$ on each side (so plus w in total):

$$W_a = W_u + w \tag{12}$$

From this definition of available cycle path width, it follows that in some sites W_a may not be equal to the physical width of the cycle path (W). For example, in this study's site, the cycle path is segregated from both the road and the sidewalk, so cyclists tend to use the full width of the cycle path ($W_u \simeq W$); therefore, $W_a > W$. Instead, in the sites analyzed in Raksuntorn and Khan (*10*), the cycle paths are not segregated from the road, so cyclists do not use the full width of the cycle path ($W_u < W$), which implies that the difference $W_a - W$ is probably lower and maybe even negative.

This theoretical definition of φ may overestimate the number of virtual sublanes that can be accommodated in a given cycle path. An alternative method to calculate φ is to determine the *actual* number of virtual sublanes observed in empirical data ($\bar{\varphi}_E$). For this purpose, the authors propose the following innovative approach. For a given discharge period k, the leader–follower chains (as shown in Figure 1*e*) are built. Then, the number of cyclists R_λ^k who are in each position λ of a chain are counted, where $\lambda = 1, \cdots, \Lambda^k$. For example, $\lambda = 1$ means the first position in a chain, and in Figure 1*e* there are two cyclists who are in the first position of a chain, so $R_1^k = 2$. The total number of chain positions (Λ^k) differs among queue discharge periods.

The number of virtual sublanes of a given discharge period can then be defined as follows:

$$\varphi_E^k = \frac{\sum_{\lambda=1}^{\Lambda^k} R_\lambda^k}{\Lambda^k}, \tag{13}$$

and the number of virtual sublanes ($\bar{\varphi}_E$) is the average value of φ_E^k in all queue discharge periods.

The saturation flow (q_s in Equation 2) and the capacity of the cycle path (C in Equation 1) were calculated assuming different sublane width (w) and using different estimates of the number of sublanes φ (theoretical and empirical methods). The traffic control scheme parameters were set to $T_G = 20$ s, and $T_{YG} = 4$ s and $T \in \{60, 120\}$ s.

Results

This section presents the results of the previously described analyses, which have been performed using the empirical trajectory data set described in the same section.

Headway Distribution Functions and Influence of Virtual Sublane Width

The histograms shown in Figure 2, *a–d* provide an overview of the empirical headway distributions at the stop line with different virtual sublane widths (w). These figures show that the change in virtual sublane width has a significant influence on the shape of the headway distribution. In general, headways decrease as the width of

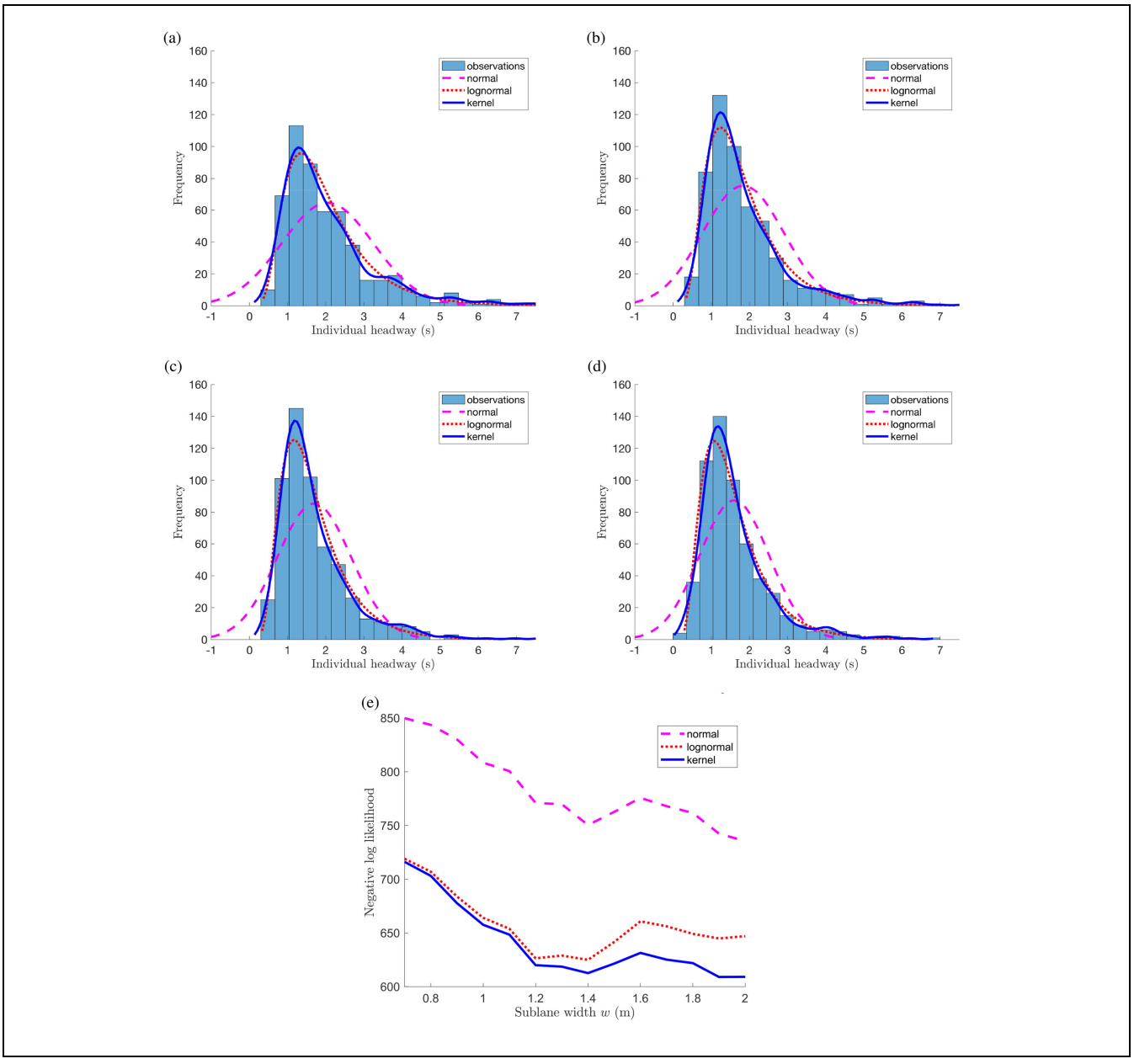


Figure 2. Headway distributions and negative log likelihood with three distribution functions with respect to different sublane widths. Note only the headways of cyclists who have a leader (following headways) have been included in this analysis. (*a*) Headway distribution with w = 0.80 m; (*b*) headway distribution with w = 1.00 m; (*c*) headway distribution with w = 1.20 m; (*d*) headway distribution with w = 1.40 m; (*e*) negative log likelihood of three distribution functions with different sublane widths.

virtual sublane increases: the average headway decreases from 2.16 to 1.28 s if w increases from 0.7 to 2.0 m (see Table 2). This makes sense, because with greater values of w cyclists can be assigned to leaders that are laterally more distant and pass the stop line closer in time (see "Definition of Headway of a Cyclist"). However, the headway varies considerably among cyclists: the range between first (Q1) and third quantiles (Q3) is 0.80–1.36 s wide depending on the sublane width (see Table 2). As expected, the kernel distribution function fits the headway histogram best, followed by the lognormal distribution which also fits the data reasonably well (see Figure 2, *a–e*). Besides, it is observed that the average headway of cyclists who have a leader is generally smaller than the one derived from all cyclists (with and without a leader) (see Table 2).

Table 2. Values of Different Variables with Respect to the Sublane Width w

	Virtual sublane width w (m)													
	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Headway of cyclists that have a leader (s)														
Mean	2.16	2.05	1.93	1.84	1.79	1.70	1.66	1.61	1.54	1.49	1.44	1.37	1.32	1.28
Median	1.82	1.70	1.61	1.54	1.50	1.43	1.41	1.37	1.33	1.30	1.26	1.19	1.18	1.16
1st Quantile (Q1)	1.29	1.21	1.16	1.12	1.11	1.08	1.07	1.02	0.98	0.94	0.91	0.85	0.82	0.79
3rd Quantile (Q3)	2.65	2.49	2.35	2.20	2.17	2.03	2.00	1.96	1.83	1.79	1.74	1.67	1.63	1.59
Headway of all cyclists (s)														
Mean	2.30	2.19	2.07	1.98	1.91	1.83	1.78	1.73	1.65	1.59	1.55	1.47	1.43	1.38
Median	2.00	1.85	1.76	1.68	1.65	1.58	1.54	1.50	1.43	1.40	1.33	1.29	1.26	1.22
Q1	1.38	1.30	1.24	1.20	1.17	1.12	1.11	1.08	1.02	0.98	0.94	0.88	0.86	0.84
Q3	2.85	2.73	2.56	2.45	2.37	2.23	2.19	2.14	2.01	1.98	1.92	1.83	1.80	1.73
Normal distribution $\mathcal{N}$														
Parameter μ_N	2.16	2.05	1.93	1.84	1.79	1.70	1.66	1.61	1.54	1.49	1.44	1.37	1.32	1.28
Parameter σ_N	1.26	1.20	1.13	1.07	1.03	0.96	0.95	0.91	0.91	0.91	0.89	0.87	0.83	0.82
Lognormal distribution $\mathcal{X}$														
Parameter μ_X	0.63	0.57	0.52	0.47	0.45	0.40	0.37	0.34	0.28	0.24	0.20	0.13	0.09	0.05
Parameter σ_X	0.52	0.52	0.51	0.51	0.50	0.50	0.51	0.52	0.55	0.59	0.60	0.63	0.65	0.67
Distance threshold d_s (m)														
	8.25	7.75	8.25	8.25	8.00	7.75	7.75	7.00	6.75	7.00	6.50	6.75	7.00	6.75
Saturation headway h_s (s)														
Mean	1.72	1.66	1.52	1.45	1.42	1.40	1.36	1.34	1.29	1.26	1.21	1.16	1.12	1.10
Median	1.57	1.43	1.32	1.23	1.25	1.22	1.21	1.20	1.18	1.15	1.11	1.08	1.03	1.03
Q1	1.09	1.08	0.98	0.94	0.95	0.94	0.94	0.93	0.91	0.89	0.86	0.81	0.77	0.75
Q3	2.20	2.15	1.97	1.77	1.72	1.66	1.61	1.60	1.54	1.52	1.46	1.43	1.38	1.36
Start-up lost time T_L (s)														
Mean	3.15	3.18	3.86	4.04	4.05	3.69	3.77	3.66	3.40	3.17	3.22	3.07	3.18	2.96
Median	3.26	3.92	4.12	4.46	4.18	4.01	3.93	3.89	3.49	3.29	3.28	3.14	3.09	2.89
Q1	3.10	2.65	3.26	3.41	3.03	2.73	2.66	2.63	2.36	1.97	1.95	1.44	1.61	1.36
Q3	3.97	3.76	4.76	5.49	5.71	5.60	5.63	5.15	4.79	4.41	4.47	4.35	4.38	4.12
Theoretical number of sublanes (φ_T)														
	3.86	3.50	3.22	3.00	2.82	2.67	2.54	2.43	2.33	2.25	2.18	2.11	2.05	2.00
Empirical number of sublanes (φ_E)														
Mean ($\bar{\varphi}_E$)	2.18	2.04	1.93	1.86	1.82	1.72	1.68	1.63	1.59	1.52	1.51	1.49	1.45	1.46
Median	2.17	2.00	1.83	1.82	1.80	1.67	1.63	1.63	1.50	1.40	1.38	1.38	1.33	1.30
Q1	1.83	1.75	1.69	1.63	1.63	1.40	1.40	1.31	1.23	1.13	1.13	1.13	1.11	1.09
Q3	2.50	2.27	2.18	2.03	2.00	1.91	1.86	1.83	1.83	1.80	1.76	1.71	1.68	1.70
Saturation flow q_s (cyc./h)														
With $\varphi = \varphi_T$	8060	7610	7645	7456	7133	6866	6741	6532	6495	6449	6476	6528	6579	6574
With $\varphi = \bar{\varphi}_E$	4561	4435	4573	4626	4617	4424	4452	4376	4437	4347	4497	4613	4653	4800
Capacity C (cyc./h)														
$\varphi = \varphi_T$, T = 120s	1401	1320	1283	1240	1186	1162	1137	1107	1115	1120	1121	1139	1142	1153
$\varphi = \varphi_T$, T = 60s	2801	2641	2566	2480	2371	2324	2273	2214	2230	2239	2243	2278	2283	2305
$\varphi = \bar{\varphi}_E$, T = 120s	793	770	768	769	767	749	751	742	762	755	779	805	807	842
$\varphi = \bar{\varphi}_E$, T = 60s	1585	1539	1535	1539	1535	1497	1501	1483	1523	1509	1558	1609	1615	1683

Distance Threshold, Saturation Headway and Start-Up Lost Time

The distance threshold d_s used to exclude the first group of cyclists of the queue when calculating the saturation headway (see "Calculation of Saturation Headway") decreases slightly with the increasing sublane width (see Figure 3*a* and Table 2). For sublane widths between 0.7 and 2.0 m, d_s is estimated to be 6.5–8.25 m, which implies the exclusion of the first 6.1–7.6 cyclists, according to the linear model shown in Figure 1*d*. This is a slightly higher number than in Raksuntorn and Khan (*10*), in which the authors estimated that the first five cyclists needed to be excluded, although the available width of the cycle path analyzed in their study (*10*) is slightly lower than in this study's site.

The saturation headway decreases with the increasing sublane width (see Figure 3*b* and Table 2), as expected. More specifically, the saturation headway decreases from 1.72 to 1.10 s if w increases from 0.7 to 2.0 m.

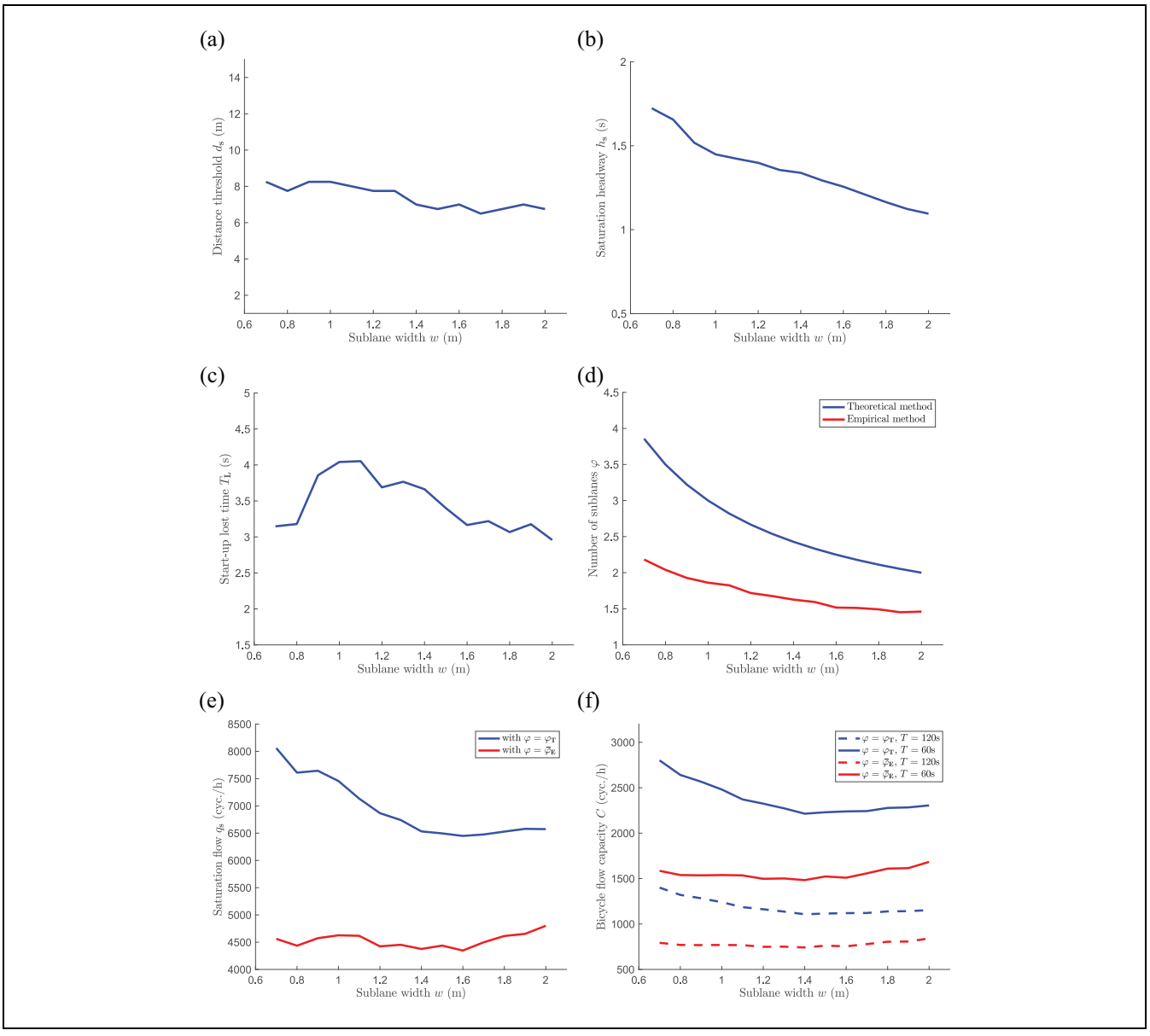


Figure 3. Relations of distance threshold, saturation headway, start-up lost time, number of sublanes, saturation flow and capacity with respect to the sublane width chosen to calculate the headways at the stop line. (*a*) relation between sublane width and distance threshold; (*b*) relation between sublane width and saturation headway; (*c*) relation between sublane width and start-up lost time; (*d*) relation between sublane width and number of sublanes; (e) relation between sublane width and saturation flow; (*f*) relation between sublane width and capacity (T_G = 20 s, and T_{YG} = 4 s).

Furthermore, h_s is a highly stochastic variable (the range between Q1 and Q3 of h_s is 0.61–1.11 s depending on the value of w (see Table 2)). For values of w between 1.00 and 1.40 m (which can be considered as reasonable values of w), the saturation headway equals 1.34–1.45 s.

The start-up lost time, instead, seems to increase for w = [0.7, 1.1] m and decrease for w = [1.1, 2.0] m (see Figure 3*c*). For values of w between 1.00 and 1.40 m, T_L is within the range of 3.66–4.04 s. This variable is highly stochastic, as the range between Q1 and Q3 of T_L is 0.87–2.87 s depending on the value of w (see Table 2). Thus, the value of T_L is again slightly higher than the T_L value reported in Raksuntorn and Khan (*10*), which was 2.5 s, as more cyclists (1.1–2.6 more) are included in the first cyclist group of the queue in the present study.

Saturation Flow and Capacity Estimation

Clearly, the proposed empirical method to calculate the average number of virtual sublanes $\bar{\varphi}_E$ (see "Calculation of the Number of Sublanes") yields considerately lower

estimates than the theoretical method, especially for lower values of w (see Figure 3*d*). The estimates obtained using the empirical method are between 1.46 and 2.18 virtual sublanes (depending on the value of w), 27–44% less than the theoretical method. This may be explained by the fact that cyclists do not always sufficiently utilize/occupy the space of cycle paths when they form leader–follower pairs during queue-formation processes. Note, however, that φ_E is a stochastic variable, as the range between Q1 and Q3 of φ_E is 0.37–0.67 depending on the value of w (see Table 2).

The fact that the observed number of sublanes is much lower than the theoretically assumed number of sublanes has enormous implications for the calculation of the saturation flow (and thus the capacity). As shown in Figure 3*e* and Table 2, the saturation flow calculated using the mean empirical estimates of φ ($\bar{\varphi}_E$) and saturation headway (h_s) is around 4350–4800 cyc./h depending on the value of w (values similar to those observed by Raksuntorn and Khan (*10*) and Seriani et al. (*12*) in cycle paths of similar widths), 27–43% lower than the saturation flow calculated using the theoretical estimates of φ (φ_T). Also, the saturation flow changes less with w if the empirical estimate of φ is used (see Figure 3*e*). This has crucial implications for the estimation of capacity (as shown in Figure 3*f*, and Table 2).

Conclusions

The ability to predict the bicycle flow capacity at signalized intersections of various characteristics is crucial for urban infrastructure design and traffic management. However, it is also a difficult task because of the large heterogeneity in cycling behavior as well as several limitations of traditional capacity estimation methods (e.g., the ignorance of leader–follower relations). This study presented an improved methodology to estimate the saturation flow and the capacity of cycle paths at signalized intersections. More specifically, this methodology includes a method to measure headways between leader–follower pairs using the concept of virtual sublane, an improved method to calculate the saturation headway and the start-up lost time (using a distance-based rule), as well as a new empirically based method to estimate the number of virtual sublanes that the cycle path can accommodate. Estimates of these variables were derived from cyclist trajectory data of a busy intersection in Amsterdam (the Netherlands). One of the main findings is that saturation headway, start-up lost time, and number of virtual sublanes are highly stochastic variables. This needs to be taken into account when interpreting the saturation flow and capacity estimates calculated using the means of these variables. For a reasonable range of w (which is considered to be 1.00–1.40 m), the authors estimated (for the site in Amsterdam) that the saturation headway is 1.34–1.45 s depending on the value of w, the start-up lost time is 3.66–4.04 s, the number of virtual sublanes (measured using the empirical method) is 1.63–1.86 (whereas the theoretical number is 2.43–3.00), and the saturation flow is 4,376–4,626 cyc./h. The saturation flow estimates are significantly lower (27–43% lower) than if the theoretical method is used to calculate the number of virtual sublanes ($\varphi_T = W_a/w$), as is usually done in practice.

Further research is necessary to identify the lateral influencing area of cyclists and to define the virtual sublane width based on empirical evidence. Moreover, it is necessary to analyze more sites with various cycle path widths (W_a) to establish an empirical relation between number of virtual sublanes and W_a, which would allow more accurate prediction of the capacity of cycle paths of various width and other characteristics to be made, and also for optimal widths of cycle paths to be derived. This is very useful for cycle path design and traffic management. Alternative methods to measure the number of sublanes can also be explored. In this work, no distinction was made between scooters and different types of bicycles (e.g., electric bicycles, and human-powered bicycles including regular bicycles, cargo bicycles, racing bicycles). Note, however, that the percentage of scooters in this data set was very low. As future work, the influence of the presence of scooters, and different composition of bicycle types on the calculation of bicycle headway and capacity should be investigated. Also the influence of traffic composition should take into account different traffic conditions. This is not trivial, as traffic operations are different for regular and saturated flow conditions. For the latter, conflicts caused by speed differences among road users would not be so different dependent on the composition, as the flow is fully packed in this situation.

Acknowledgments

This research was supported by the ALLEGRO project (Unravelling slow mode traveling and traffic: with innovative data to create a new transportation and traffic theory for pedestrians and bicycles), which is funded by the European Research Council (Grant Agreement No. 669792), and the Amsterdam Institute for Advanced Metropolitan Solutions.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: BG-R, YY, MP; data collection: BG-R, WD; analysis and interpretation of results: YY, BG-R; manuscript preparation: YY, BG-R, MP; comments to draft manuscript: WD, SPH. All authors reviewed the results and approved the final version of the manuscript.

References

1. Adams, W. F. Road Traffic Considered as a Random Series. *Institution Civil Engineers*, 1936.
2. Cowan, R. J. Useful Headway Models. *Transportation Research*, Vol. 9, No. 6, 1975, pp. 371–375.
3. Zhang, G., Y. Wang, H. Wei, and Y. Chen. Examining Headway Distribution Models with Urban Freeway Loop Event Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 1999: 141–149.
4. Tolle, J. E. The Lognormal Headway Distribution Model. *Traffic Engineering and Control*, Vol. 8, No. 8, 1971. https://trid.trb.org/view/117094.
5. Buckley, D. J. A Semi-Poisson Model of Traffic Flow. *Transportation Science*, Vol. 2, No. 2, 1968, pp.107–133.
6. Branston, D. Models of Single Lane Time Headway Distributions. *Transportation Science*, Vol. 10, No. 2, 1976, pp.125–148.
7. Hoogendoorn, S. P. Unified Approach to Estimating Free Speed Distributions. *Transportation Research Part B: Methodological*, Vol. 39, No. 8, 2005, pp.709–727.
8. Hoogendoorn, S. P., and W. Daamen. Bicycle Headway Modeling and Its Applications. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2587: 34–40.
9. Botma, H., and H. Papendrecht. Traffic Operation of Bicycle Traffic. *Transportation Research Record: Journal of the Transportation Research Board*, 1991. 1320: 65–72.
10. Raksuntorn, W., and S. I. Khan. Saturation Flow Rate, Start-Up Lost Time, and Capacity for Bicycles at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1852: 105–113.
11. Liu, X., L. D. Shen, and F. Ren. Operational Analysis of Bicycle Interchanges in Beijing, China. *Transportation Research Record: Journal of the Transportation Research Board*, 1993. 1396: 18–21.
12. Seriani, S., R. Fernandez, and E. Hermosilla. Experimental Study for Estimating Capacity of Cycle Lanes. *Transportation Research Procedia*, Vol. 8, 2015, pp.192–203.
13. Brilon, W., M. Grossmann, and H. Blanke, Verfahren für die Berechnung der Leistungsfähigkeit und Qualität des Verkehrsablaufes auf Straßen, *Forschung Straßenbau und Straßenverkehrstechnik*, No. 669, 1994. (In German)
14. Allen, D. P., N. Rouphail, J. Hummer, and J. Milazzo. Operational Analysis of Uninterrupted Bicycle Facilities. *Transportation Research Record: Journal of the Transportation Research Board*, 1998. 1636: 29–36.
15. Goñi-Ros, B., Y. Yuan, W. Daamen, and S.P. Hoogendoorn. Empirical Analysis of the Macroscopic Characteristics of Bicyce Flow during the Queue Discharge Process at a Signalized Intersection. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672(36); 51–62.
16. Duives, D. C. *Analysis and Modelling of Pedestrian Movement Dynamic at Large-scale Events*. T2016/15. TRAIL Thesis Series, 2016.
17. Daganzo, C. F. *Fundamentals of Transportation and Traffic Operations*, Pergamon-Elsevier, Oxford, UK, 1997.
18. Shao, C., and X. Liu. Estimation of Saturation Flow Rates at Signalized Intersections. *Discrete Dynamics in Nature and Society*, No. 2012, 2012, pp. 1–9.
19. Transportation Research Board. *Highway Capacity Manual–Sixth Edition: A Guide for Multimodal Mobility Analysis*. Transportation Research Board of the National Academies of Science, Washington, D.C., 2016.

The Standing Committee on Bicycle Transportation (ANF20) peer-reviewed this paper (19-04329).