

## Inferring vehicle spacing in urban traffic from trajectory data

Jiao, Yiru; Calvert, Simeon C.; van Cranenburgh, Sander; van Lint, Hans

**DOI**

[10.1016/j.trc.2023.104289](https://doi.org/10.1016/j.trc.2023.104289)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Transportation Research Part C: Emerging Technologies

**Citation (APA)**

Jiao, Y., Calvert, S. C., van Cranenburgh, S., & van Lint, H. (2023). Inferring vehicle spacing in urban traffic from trajectory data. *Transportation Research Part C: Emerging Technologies*, 155, Article 104289. <https://doi.org/10.1016/j.trc.2023.104289>

**Important note**

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

**Copyright**

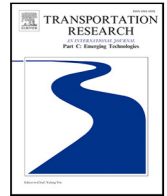
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

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

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Transportation Research Part C

journal homepage: [www.elsevier.com/locate/trc](http://www.elsevier.com/locate/trc)

## Inferring vehicle spacing in urban traffic from trajectory data

Yiru Jiao <sup>a,b,\*</sup>, Simeon C. Calvert <sup>a</sup>, Sander van Cranenburgh <sup>b</sup>, Hans van Lint <sup>a</sup><sup>a</sup> Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands<sup>b</sup> Transport and Logistics Group, Department of Engineering Systems and Services, Faculty of Technology, Policy, and Management, Delft University of Technology, The Netherlands

### ARTICLE INFO

#### Keywords:

Urban traffic  
 Vehicle interaction  
 Two-dimensional spacing  
 Fundamental Diagram

### ABSTRACT

This study presents a new method to infer the average two-dimensional (2D) spacing between interacting vehicles in urban traffic from trajectory data. In this context, 2D spacing reflects the amount of road space consumed by pairs of interacting vehicles, and is related to 2D density at the macroscopic level. Due to complex interaction and conflicts in urban traffic, the inherent assumptions in traditional traffic flow models, such as unidirectional flow and homogeneity, are often violated. Such violation challenges direct measurement of urban vehicle spacing. The proposed method addresses this challenge by focusing on the relative movement between interacting vehicles and aggregating the accumulated presence of vehicles in similar scenarios. We apply the method to a large-scale urban trajectory dataset called pNEUMA, and validate the consistency of the method through bootstrapping. By applying the method we obtain a new empirical relation between the average 2D spacing and the *relative* speeds between interacting vehicles. There are similarities between this empirical relation with the classical Fundamental Diagram of traffic flow in terms of shape and interpretation, and so we term it the “interaction Fundamental Diagram” (iFD). However, there are also key differences. The iFD does not represent steady-state (homogeneous and stationary) longitudinal behaviour; it describes the average amount of road space needed for vehicle interactions at different relative speeds. We believe these iFD relations contribute to understanding vehicle interaction in urban traffic, and can offer new insights for designing safer and more efficient urban intersections.

### 1. Introduction

Whereas on uninterrupted high-speed and high-volume traffic facilities, vehicle interactions are dominated by one-dimensional (1D, longitudinal) dynamics (e.g., car-following); in urban traffic, and particularly at intersections, vehicle interactions are both one and two-dimensional (2D), involving crossing and conflict negotiation aside from car-following and lane-changing. Vehicle interactions collectively shape traffic flow, and conversely, traffic flow influences the interactions between vehicles. Therefore, investigating vehicle interactions in urban traffic and their collective phenomena is essential to manage traffic and promote efficient, safe, and sustainable use of the shared road space.

Traffic flow modelling has predominantly focused on 1D driving, with microscopic and macroscopic models respectively examine individual vehicle interaction and aggregated traffic patterns (Hoogendoorn and Bovy, 2001). At the microscopic level, the characteristics of individual vehicles are described, such as vehicle speed, time headway, and distance headway. These characteristics are aggregated into traffic states at the macroscopic level, i.e., mean speed, flow, and density (Edie, 1963; Makigami et al., 1971). The relations between these macroscopic traffic variables are extensively studied as Fundamental Diagrams (FDs), which describe

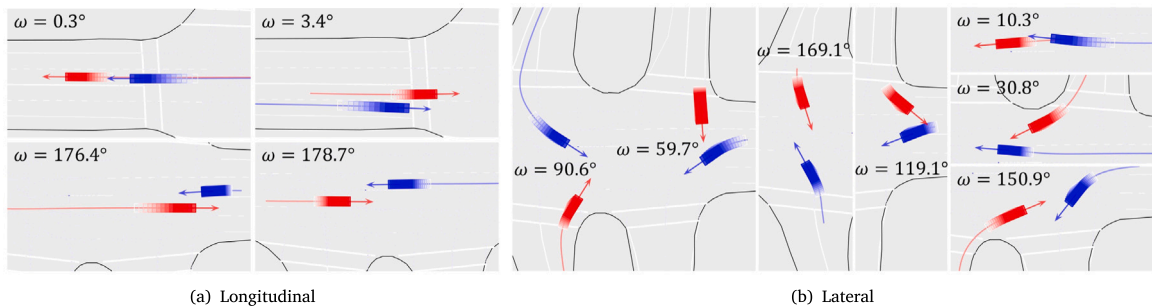
\* Corresponding author at: Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands.  
 E-mail address: [y.jiao-1@tudelft.nl](mailto:y.jiao-1@tudelft.nl) (Y. Jiao).

<https://doi.org/10.1016/j.trc.2023.104289>

Received 28 December 2022; Received in revised form 13 July 2023; Accepted 5 August 2023

Available online 23 August 2023

0968-090X/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



**Fig. 1.** Real-world examples of vehicle interactions in urban environments.  $\omega$  is the angle between moving directions of the interacting vehicles. (a) Vehicle interactions only in the longitudinal direction. (b) Vehicle interactions involving both longitudinal and lateral directions.

how steady-state traffic states evolve (e.g., Greenshields et al., 1933; Hall et al., 1992; Castillo and Benítez, 1995; Cassidy, 1998). The aggregation per se and the relations between aggregated variables rely on assumptions rooted in fluid dynamics. For instance, vehicles are assumed to move in a single direction and follow one another without disruption. The assumptions, however, are often violated in urban traffic. Accordingly, traffic flow models that account for disruptions such as lane-changes (Jin, 2010, 2013) and traffic signals (Wagner et al., 2009; Wu et al., 2011; Fourati and Friedrich, 2021) have been developed, and their corresponding FDs offer valuable information on complex real-world traffic patterns.

A gap persists concerning the assumption of unidirectional interaction, given that vehicle interaction in urban environments is multi-directional. As depicted in Fig. 1, vehicles can have various orientations on a 2D plane during their interactions in urban traffic (Zhao et al., 2022). Although pedestrian traffic is also considered 2D, it has been studied within corridors or rings where the flows remain unidirectional or bidirectional (Seyfried et al., 2005; Zhang et al., 2012; Flötteröd and Lämmel, 2015; Vanumu et al., 2017). Consequently, there is a long-standing need to identify governing quantities that can describe multi-directional vehicle interaction in urban environments and the resulted traffic phenomena.

In this study, we take the perspective of relative movement between interacting vehicles. This enables mapping the state of each specific interaction onto a consistent system to quantify the 2D spacing between the involved vehicles. By accumulating all these mapped 2D spacings, we reconstruct a spatial distribution of the vehicle interaction states, which allows us to estimate the average spacing that vehicles maintain with one another at various relative speeds. We first present the method to infer average 2D spacing between vehicles in urban traffic, and then test the method using the large-scale urban trajectory dataset pNEUMA (Barmounakis and Geroliminis, 2020). Our results reveal the characteristics of 2D vehicle spacing in different scenarios, with which we derive urban traffic states and discern their fundamental relations. We refer to the 2D spacing relations as interaction Fundamental Diagrams (iFDs), which describe how the necessary space for vehicle interactions change with their relative speeds. We believe these relations contribute towards a deeper understanding of vehicle interaction and road space use in urban traffic, which may, for example, serve traffic engineers in better and safer design of intersection layouts.

The remainder of this paper is organised as follows. Section 2 presents the methods to infer 2D vehicle spacing. Section 3 explains our experiments on the urban trajectory dataset. Section 4 presents the experiment results and findings, and Section 5 discusses further on the results. Finally, Section 6 recaps the study.

## 2. Methodology

Spacing in urban traffic is 2D as the distances drivers maintain from one another can vary depending on the direction. Two-dimensional spacing thus is equivalent to “driver space”,<sup>1</sup> which is extended for drivers from the personal space of pedestrians (Marsh and Collett, 1987; Hennessy et al., 2011). Here the driver space of a vehicle is not a determined area. Instead, we consider it as a spectrum of spacings, on which the critical spacing is where the driver can feel a rapid change in discomfort caused by proximity. Numerous studies have observed this transition from comfort to discomfort when drivers are approached by other vehicles (Taieb-Maimon and Shinar, 2001; Lewis-Evans et al., 2010; Siebert et al., 2014, 2017). Thereby, the driver space of a vehicle can be characterised by the less frequent presence of other vehicles in the surrounding, which are expected to be present more frequently if they do not cause higher levels of discomfort. Based on such violation against the expectation, we infer average 2D spacing by estimating the density of the accumulated presence of surrounding vehicles.

Fig. 2 illustrates the process proposed in this study to infer average 2D vehicle spacing. The following paragraphs will briefly explain the process. Then the specific methods will be introduced in the subsections.

<sup>1</sup> Based on the research of Gibson and Crooks (1938), a similar concept of “comfort zone” was defined by Näätänen and Summala (1974), Summala (2007), and Bärgrman et al. (2015). The comfort zone is an area within which a driver feels comfortable. Other drivers are unwilling to cross the area without extra motives and will arouse discomfort when they do cross it. For consistency, we uniformly use “driver space” to refer to this area.

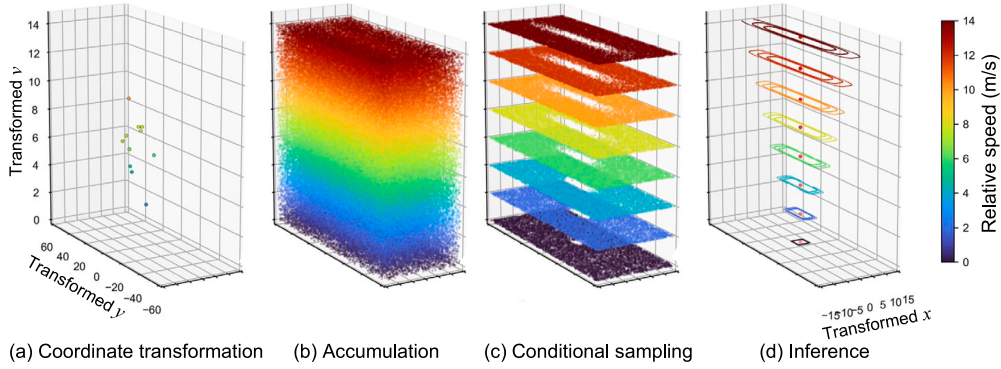


Fig. 2. Proposed framework to infer average 2D vehicle spacing from trajectory data. (a) Coordinate transformation of vehicle pairs in Fig. 1. (b) Accumulated data samples of transformed vehicle pairs. (c) Data samples that are conditioned by relative speeds, where the hollows are driver space. (d) Spacing inference from the conditioned data samples.

In Section 2.1, we firstly introduce a coordinate transformation for each pair of interacting vehicles consisting of an ego vehicle and a surrounding vehicle. The transformation emphasises the multi-directionality of the relative movement between interacting vehicles and establishes a normalised reference frame for analysing vehicle interaction. Consequently, this enables the accumulation of vehicle pairs and consistent analysis across various scenarios.

With the accumulated samples of vehicle pairs in transformed coordinates, Section 2.2 utilises driver space as a proxy to measure vehicle spacing. The intrusion into driver space can cause varying levels of discomfort, and a driver's response to such discomfort is manifested by maintaining a naturally comfortable distance from other vehicles. We interpret this discomfort response as resistance to the proximity of other vehicles, and parameterise such proximity resistance by adapting the density function of the generalised Gaussian distribution. Proximity resistance describes the comfort–discomfort transition, based on which we can quantify the spacing between vehicles in a probabilistic manner.

Section 2.3 presents an algorithm to approximate proximity resistance by estimating the density of the accumulated presence of vehicles in the surrounding of an ego vehicle. The change in proximity resistance with spacing is described as a function, which is not fixed, but varies in different scenarios. For this reason, we design our method to be scenario-conditioned with respect to variables such as speed, acceleration, or/and other traffic situations. As an illustration of this method, we use the relative speed between interacting vehicles as the basic condition in this study. Ultimately, a series of functions with different parameters in different scenarios depict the average 2D spacing between interacting vehicles.

### 2.1. Coordinate transformation

To analyse vehicle spacing more consistently, we establish a new coordinate system referring to the relative movement of interacting vehicles. For an ego vehicle  $i$  and another vehicle  $j$  in the surrounding of  $i$ , we transform their global coordinates into a local coordinate system  $C(O_i, \mathbf{v}_{ij})$ . The origin  $O_i$  is located at the position of  $i$ , and the  $y$ -axis points in the direction of  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ , the relative velocity of  $i$  to  $j$ . By doing this, all pairs of interacting vehicles share a common reference point (i.e., the ego vehicle's position) and an aligned orientation (i.e., the direction of the relative velocity of the ego vehicle to the surrounding vehicle).

Fig. 3 visually illustrates the transformation, wherein we rotate and translate the original global coordinate system. We use  $(x^{(g)}, y^{(g)})$  to denote the global position of a vehicle and  $(x^{(ij)}, y^{(ij)})$  to denote the local coordinates of the vehicle in  $C(O_i, \mathbf{v}_{ij})$ . The transformation equation is formulated as

$$\begin{bmatrix} x^{(ij)} \\ y^{(ij)} \end{bmatrix} = \begin{bmatrix} \cos \rho & -\sin \rho \\ \sin \rho & \cos \rho \end{bmatrix} \left( \begin{bmatrix} x^{(g)} \\ y^{(g)} \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \right), \quad (1)$$

where  $\rho$  represents the counterclockwise rotation angle and  $(a, b)^\top$  is the translation vector of the transformation.

The transformation parameters are solved as follows by substituting the coordinates in the global and local coordinate systems at the same position into Eq. (1). We use two positions here. One is the position of  $i$ , which is  $(x_{O_i}^{(g)}, y_{O_i}^{(g)})$  in the global coordinate system and  $(0, 0)$  in  $C(O_i, \mathbf{v}_{ij})$ . The other is the position of the head of the relative velocity vector of  $i$  to  $j$  when its tail is located at  $O_i$ . This position is  $(x_{O_i}^{(g)} + x_{\mathbf{v}_{ij}}^{(g)}, y_{O_i}^{(g)} + y_{\mathbf{v}_{ij}}^{(g)})$  in the global coordinate system and  $(0, y_{\mathbf{v}_{ij}}^{(ij)})$  in  $C(O_i, \mathbf{v}_{ij})$ , where  $y_{\mathbf{v}_{ij}}^{(ij)} = \sqrt{x_{\mathbf{v}_{ij}}^{(g)2} + y_{\mathbf{v}_{ij}}^{(g)2}}$ . As a result, the transformation parameters between the global coordinate system and  $C(O_i, \mathbf{v}_{ij})$  are

$$\begin{cases} a = -x_{O_i}^{(g)}, \\ b = -y_{O_i}^{(g)}, \\ \cos \rho = y_{\mathbf{v}_{ij}}^{(g)} / \sqrt{x_{\mathbf{v}_{ij}}^{(g)2} + y_{\mathbf{v}_{ij}}^{(g)2}}, \\ \sin \rho = x_{\mathbf{v}_{ij}}^{(g)} / \sqrt{x_{\mathbf{v}_{ij}}^{(g)2} + y_{\mathbf{v}_{ij}}^{(g)2}}. \end{cases} \quad (2)$$

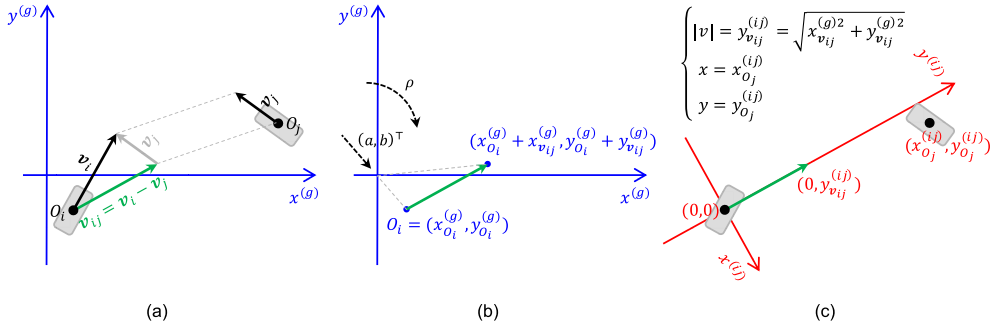


Fig. 3. Coordinate transformation. (a) An ego vehicle  $i$  and another nearby vehicle  $j$  in the global coordinate system denoted by superscript  $(g)$ , where  $O_i$  and  $O_j$  represent their positions,  $v_i$  and  $v_j$  for their velocities, and  $v_{ij}$  for the relative velocity of  $i$  to  $j$ . (b) Rotate the global coordinate system by  $\rho$  and translate  $(a, b)^T$  according to the reference point  $O_i$  and the reference orientation  $v_{ij}$ . (c) Transformed coordinates in the local coordinate system denoted by superscript  $(ij)$ .

Given that division by zero is undefined, Eq. (2) requires that  $(x_{v_{ij}}^{(g)}, y_{v_{ij}}^{(g)}) \neq (0, 0)$ , i.e.,  $v_i \neq v_j$ . This implies that if vehicle  $i$  and vehicle  $j$  are moving at the same velocity simultaneously, their relative movement cannot be transformed into a local coordinate system defined here. Then we disregard such cases as  $i$  and  $j$  are temporarily stationary relative to each other.

Finally, we transform the global position of  $j$  into  $C(O_i, v_{ij})$  as  $(x_{O_j}^{(ij)}, y_{O_j}^{(ij)})$ . Denoting  $x = x_{O_j}^{(ij)}$  and  $y = y_{O_j}^{(ij)}$ , and together with the mode of the relative velocity  $|v| = y_{v_{ij}}^{(ij)}$ , this coordinate transformation compresses the information about the relative position and velocity between vehicle  $i$  and vehicle  $j$ . For easier notation, we will use  $v$  in place of  $|v|$  in the rest of this paper. These three variables  $(x, y, v)$  encode two pieces of information. The first is the position of vehicle  $j$  relative to vehicle  $i$ . The second is the moving direction of vehicle  $i$  and  $j$ , where they move along the  $y$ -axis either towards ( $y > 0$ ) or away from ( $y < 0$ ) each other. More specifically, vehicles above the transformed  $x$ -axis are approaching the vehicle at the origin rather than being ahead of it; similarly, vehicles below the transformed  $x$ -axis are leaving it rather than being behind.

In the following sections of this paper, we will use  $(x, y, v)$  to refer to the transformed relative position and relative speed of a pair of vehicles  $i$  and  $j$ . It must be noted that this transformation makes the study specifically focus on the relative movement between vehicles. Readers are reminded that, if not specified,  $v$  refers to the relative speed between interacting vehicles rather than their absolute speeds in this study.

## 2.2. Probabilistic spacing

We consider spacing to be probabilistic to account for its variation, and we characterise it utilising the concept of driver space. Intrusion into driver space causes discomfort, which does not preclude the spatial intrusion of other vehicles, but motivates a driver to maintain a distance from others at varying levels (Camara and Fox, 2020). For an ego vehicle  $i$ , if another vehicle  $j$  is far outside the driver space of  $i$ , little discomfort is caused and the movement of  $i$  remains uninfluenced. In contrast, if vehicle  $j$  enters the driver space of  $i$  and keeps approaching, the escalating discomfort will compel the driver of  $i$  to move away (Hennessey et al., 1995; Graziano and Cooke, 2006; Summala, 2007). We therefore introduce proximity resistance to interpret the discomfort response and specify the level of intrusion-caused discomfort. With a value  $p \in (0, 1)$  denoting proximity resistance,  $p \rightarrow 0$  indicates that no discomfort is caused and  $p \rightarrow 1$  indicates extreme discomfort. In this way, any spacing in a specific scenario corresponds to a level of proximity resistance, which indicates the potential for causing discomfort.

### 2.2.1. Proximity resistance in one-dimensional spacing

We first consider proximity resistance in 1D spacing. The level of intrusion-caused discomfort has a transition from lower to higher across the critical spacing where proximity resistance changes the most rapidly. Therefore, the representation of proximity resistance should encode how far the critical spacing is and how rapid the discomfort transition is. To this end, we adapt the density function of the generalised Gaussian distribution (GGD).

The GGD is a family of symmetric probability densities that generalises the Gaussian density (Novey et al., 2010). Its function is given as  $\alpha \exp(-|\frac{x-\mu}{r}|^\beta)$ , where  $\alpha$  is the coefficient making the integral of the function 1,  $\mu$  is the centre of symmetry,  $r$  is the scale parameter, and  $\beta$  is the shape parameter (Wang et al., 2006). The family includes the density of Laplace distribution when  $\beta = 1$  and of Gaussian distribution when  $\beta = 2$ . When  $\beta \rightarrow \infty$ , GGD converges to the density of a uniform distribution on  $(\mu - r, \mu + r)$ . Thereby, with  $\beta \geq 2$ , GGD covers a continuum of densities from the Gaussian distribution to the uniform distribution.

We use  $s$  to denote 1D spacing, which is the relative positions of surrounding vehicles to an ego vehicle. Thereby, the ego vehicle is at the origin and  $\mu = 0$ . We then present proximity resistance in 1D spacing as Eq. (3). Proximity resistance is not a probability density, so we remove the coefficient  $\alpha$  to ensure  $p \in (0, 1)$ .

$$p(s|r, \beta) = \exp\left(-\left|\frac{s}{r}\right|^\beta\right), \tag{3}$$

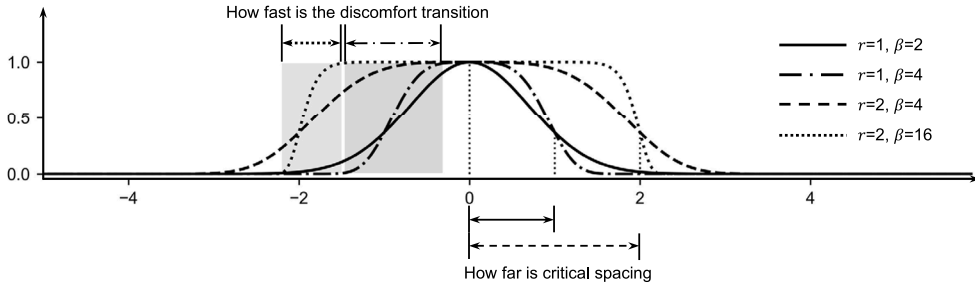


Fig. 4. Parameters  $r$  and  $\beta$  determine the change of proximity resistance in 1D spacing.

where  $r > 0$  and  $\beta \geq 2$ . As shown in Fig. 4, the scale parameter  $r$  suggests how far the critical spacing is, and the shape parameter  $\beta$  suggests how fast the discomfort transition is. This parameterisation captures the variation in proximity resistance at different relative positions, which characterises probabilistic spacing.

Vehicles oriented in different directions may cause different levels of discomfort even being equally distant. We thus further consider two components for  $r$  and  $\beta$ , respectively, to allow Eq. (3) to be asymmetric.

$$\begin{cases} r = \frac{1 + \text{sgn}(s)}{2} r^+ + \frac{1 - \text{sgn}(s)}{2} r^-, \\ \beta = \frac{1 + \text{sgn}(s)}{2} \beta^+ + \frac{1 - \text{sgn}(s)}{2} \beta^-. \end{cases} \quad (4)$$

Eq. (4) means that  $r = r^+$  and  $\beta = \beta^+$  when  $s > 0$ , and  $r = r^-$  and  $\beta = \beta^-$  when  $s < 0$ .

### 2.2.2. Proximity resistance in two-dimensional spacing

Coordinate transformation in Section 2.1 places the ego vehicle at  $(0, 0)$  and the surrounding vehicle at  $(x, y)$ , where  $x$  and  $y$  are independent<sup>2</sup> in most cases because vehicles can take any place relative to the ego vehicle. Then the proximity resistance in 2D spacing, as shown in Eq. (5), is an extension of Eq. (3).

$$p(x, y|\theta) = \exp\left(-\left|\frac{x}{r_x}\right|^{\beta_x} - \left|\frac{y}{r_y}\right|^{\beta_y}\right), \quad (5)$$

where each of  $\theta = (r_x, r_y, \beta_x, \beta_y)^\top$  has two components:

$$\begin{cases} \theta = \frac{1 + \text{sgn}(x)}{2} \theta^+ + \frac{1 - \text{sgn}(x)}{2} \theta^- \text{ for } \theta = r_x, \beta_x, \\ \theta = \frac{1 + \text{sgn}(y)}{2} \theta^+ + \frac{1 - \text{sgn}(y)}{2} \theta^- \text{ for } \theta = r_y, \beta_y. \end{cases} \quad (6)$$

Parameters in Eq. (5) have realistic explanations.  $r = \{r_x^+, r_x^-, r_y^+, r_y^-\}$  determine the positions where  $p = e^{-1}$  in different directions, and  $\beta = \{\beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-\}$  determine the rate at which  $p$  varies with  $x > 0$ ,  $x < 0$ ,  $y > 0$ , and  $y < 0$ . Hereby,  $r$  control the critical spacing resulting in higher proximity resistance, and  $\beta$  control the increase rates of proximity resistance when vehicles are moving in different directions. In this way,  $r$  and  $\beta$  together characterise 2D probabilistic spacing.

Vehicle spacing can be influenced by the movement states of the interacting vehicles, the traffic situation they are in, as well as the driving preference of their drivers. These influences are diverse and may be interrelated, making it difficult to integrate them into equations as independent variables. In this study, we consider such influences as the conditions for filtering vehicle pair samples. Therefore,  $\theta$  vary across conditioning situations.

### 2.3. Parameter inference

We infer the parameters in Eq. (5) by estimating the density of accumulated presence of surrounding vehicles. The infrequent presence of surrounding vehicles reflects how resistant drivers are to approaching one another. Conversely, the frequent presence of surrounding vehicles reflects how acceptable this approach is. We refer to the former as proximity resistance and the latter as proximity tolerance, and they complement to each other. The density of the accumulated vehicle presence around a vehicle directly corresponds to proximity tolerance, so we can indirectly infer proximity resistance as the opposite to it.

In the transformed coordinate system, a pair of vehicles consists of an ego vehicle  $i$  at  $(x_i, y_i) = (0, 0)$  and a surrounding vehicle  $j$  at  $(x_j, y_j)$ . Denote the proximity resistance between the drivers of  $i$  and  $j$  as  $p_{ij}$ , then the corresponding proximity tolerance is

<sup>2</sup> At some special scenarios such as roundabouts,  $x$  and  $y$  are correlated but can be transformed into the Frenet-Serret coordinate system to ensure their independence. In order to focus on the main idea, we do not delve into those special scenarios in this paper.



$1 - p_{ij}$ .  $1 - p_{ij}$  can be seen as the relative likelihood of a vehicle passing  $(x_j, y_j)$ , which is in proportion to the density of vehicles accumulatively passing  $(x_j, y_j)$ . Given  $n$  pairs of vehicles  $i$  and  $j$ , the likelihood of the presence of the surrounding vehicles is

$$L = \prod_{j=1}^n [1 - p_{ij}(x_j, y_j | \theta)]. \quad (7)$$

Then we can infer  $p_{ij}$  by iteratively estimating  $\theta$  until they all converge, where  $\beta$  are estimated given  $r$  and  $r$  given  $\beta$ .

The estimation of  $\beta$  aims to maximise the likelihood  $L$ , which allows for the sparsity of distant vehicles in data and prevents too slow discomfort transition. As the ego vehicle is fixed at  $(0, 0)$ , the more distant the surrounding vehicles are, the less they are recorded in data and the lower their density. In this case, a purely density-based inference would be biased towards a slower discomfort transition. Maximising  $L$  avoids this bias as vehicles farther away from the ego vehicle are at high proximity tolerance close to 1, which results in little effect on  $L$ .

The estimation of  $r$  should distinguish the critical spacing as the boundaries of driver space, within which other vehicles infrequently access. Denoting the area where  $p$  is close to 1 as  $A_{p \rightarrow 1}$ , the increase of a  $r$  (e.g.,  $r_y^+$ ) corresponds to the gradual expansion of  $A_{p \rightarrow 1}$  along the direction indicated by the  $r$  (i.e., the relative velocity direction). Holding other parameters constant,  $L$  undergoes three phases as each  $r$  increases. In the first phase,  $L$  increases slightly when almost all the surrounding vehicles are outside  $A_{p \rightarrow 1}$ . In the second phase,  $L$  decreases slowly when some surrounding vehicles are present within  $A_{p \rightarrow 1}$ . In the third phase,  $L$  decreases rapidly when more surrounding vehicles are erroneously covered by  $A_{p \rightarrow 1}$ . As a result, maximising  $L$  leads to smaller  $A_{p \rightarrow 1}$  and closer critical spacing to the ego vehicle. Alternatively to maximising  $L$ , we estimate each  $r$  by the position at which  $L$  decreases from slowly to rapidly. This position is where the second-order derivative of  $L$  with respect to  $r$  is smallest (negative), signifying the fastest change in the decreasing rate of  $L$ .

Further practical details need to be considered in the estimation. For computational convenience, we take the logarithm of  $L$ , i.e., the sum of the log-likelihood of vehicle presence. This is equivalent to  $L$  for estimating  $\theta$  since logarithms are strictly increasing functions. Eq. (5) assumes no vehicles passing through the central area of driver space. However, some vehicles indeed do so. These cases can make the estimation biased towards overly small  $r$  and  $\beta$  to keep their relative likelihood not too close to zero. To mitigate this bias, we add  $\epsilon = 10^{-4}$  to the proximity tolerance, which is thus  $1 + \epsilon - p_{ij} = 1 - (p_{ij} - \epsilon)$ . This decreases the proximity resistance between vehicles that are extremely close to each other, and increases their relative likelihood. The adjusted log-likelihood of the presence of  $n$  surrounding vehicles is then

$$\ln(L) = \sum_{j=1}^n \ln [1 + \epsilon - p_{ij}(x_j, y_j | \theta)]. \quad (8)$$

Algorithm 1 summarises the iterative inference of probabilistic spacing. Because the derivative of  $\ln(L)$  to one parameter depends on the other parameters, the algorithm is difficult to solve by hand. We thus use numerical methods for the differentiation and optimisation in the algorithm. This requires initial values for the parameters (i.e.,  $r_0$  and  $\beta_0$ ). We set  $r_0$  to the 0.1th percentiles of the surrounding vehicle positions from near to far from the ego vehicle. This heuristic setting follows the belief that few vehicles will intrude into others' driver space. Then we set  $\beta_0$  to 2 to initiate a moderate discomfort transition.

The inference is achieved when the estimates of  $r$  and  $\beta$ , i.e.,  $\hat{r}$  and  $\hat{\beta}$ , converge to either a set of values or several recurring sets of values.  $\hat{\beta}$  are estimated based on maximising  $\ln(L)$ , by which we can calculate the estimation confidence according to the numerical Hessian matrix of  $\ln(L)$ . When the inference converges to recurring sets of values, the set with the highest confidence for  $\hat{\beta}$  is taken. In addition, a maximum number of iterations is set to force a stop.

---

**Algorithm 1:** Parameter inference of 2D probabilistic spacing.

---

**Data:**  $x$  and  $y$  under particular conditions

**Result:**  $\hat{r}$  and  $\hat{\beta}$  under the particular conditions

**begin**

Initialise  $r \leftarrow r_0$  for  $r = r_x^+, r_x^-, r_y^+, r_y^-$  and  $\beta \leftarrow 2$  for  $\beta = \beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-$

**repeat**

$\hat{r} \leftarrow \operatorname{argmin}_r \frac{\partial}{\partial r} \left( \frac{\partial \ln(L)}{\partial r} \right)$  for  $r = r_x^+, r_x^-, r_y^+, r_y^-$

$\hat{\beta} \leftarrow \operatorname{argmax}_{\beta} \ln(L)$  subjected to  $\beta \geq 2$  for  $\beta = \beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-$

**until**  $\hat{r}$  and  $\hat{\beta}$  converge or maximum number of iterations reaches

---

### 3. Experiments

To demonstrate the proposed method <sup>3</sup>, we applied it on pNEUMA (Barmounakis and Geroliminis, 2020), a large-scale urban trajectory dataset. With a swarm of drones, pNEUMA recorded vehicle traces in a central business district of Athens, covering an area of 1.3 km<sup>2</sup> with over 100 km-lanes of road network and around 100 busy intersections. This area was divided into 10 subareas

<sup>3</sup> We open source our code at <https://github.com/Yiru-Jiao/DriverSpaceInference>.

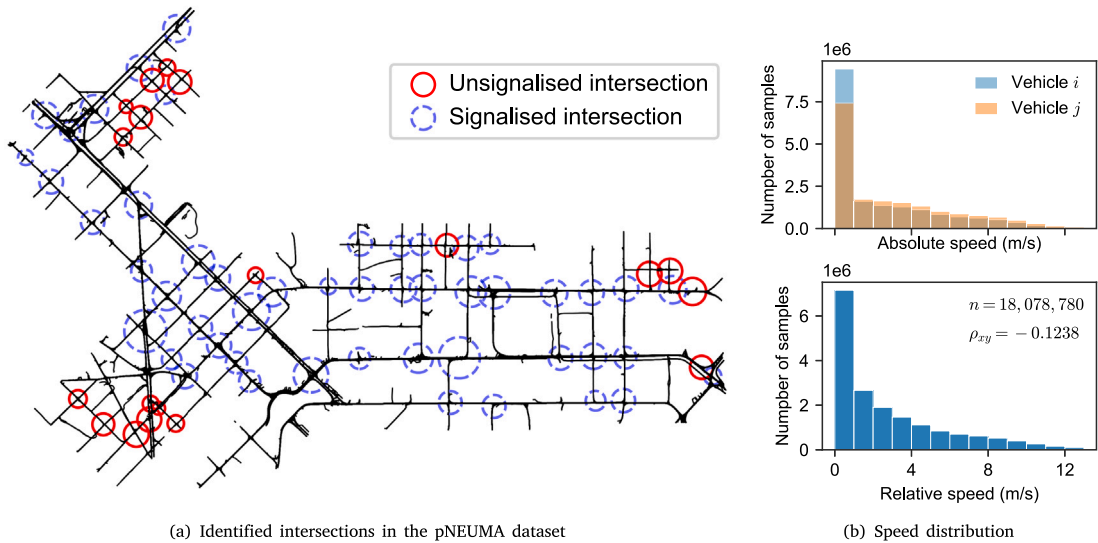


Fig. 5. Map of intersections in the dataset and vehicle sample statistics.

and each was covered by a drone. Due to limited flight time, the drones recorded 9–15 min per half hour between 8:00 and 11:00 a.m. on October 24th, 29th, 30th, and November 1st, 2018. To compare the results consistently, we use data during the latter three consecutive days where the trajectories were collected from 8:00 to 10:30 a.m. each day.

### 3.1. Preprocessing

This study pays particular attention to 2D vehicle interaction, therefore, we focused on trajectories at road intersections in the dataset. As shown in Fig. 5(a), we extracted major T-junctions and crossing intersections following the methods proposed by Cao and Krumm (2009) and Wang et al. (2014). In total 50 signalised intersections and 19 unsignalised intersections were identified. As 2D vehicle interaction is significantly reduced by signal control at signalised intersections, our experiments were focused on the unsignalised intersections. Most of these unsignalised intersections are single-lane and the recorded trajectories are insufficient to infer driver spaces for each of them, so we combined all the unsignalised intersections as an aggregated study area.

The data preprocessing exists of four steps. First, pedestrians, bicycles, and motorcycles (all are labelled “Mortorcycle” in the dataset) were excluded as they have very different movement characteristics from motor vehicles regarding speed and occupied space. Second, vehicles that never moved were also excluded as they did not interact with any others. Third, vehicles that had overlapping trajectories with another vehicle (position distance <0.5 m at every moment) were removed as these are assumed to be data errors or outliers. Such errors could be caused by the vision algorithm used to track the vehicles. Fourth, vehicle pairs were sampled at a series of moments 0.2 s apart. Seeing one in the pair as an ego vehicle and the other as a surrounding vehicle, their position coordinates were transformed according to Eq. (1) and Eq. (2). The histograms in Fig. 5(b) show the distributions of absolute speeds and relative speeds of extracted samples. The number of samples ( $n$ ) and the correlation coefficients of transformed sample positions ( $\rho_{xy}$ ) are also annotated.

### 3.2. Experiment setting

In this study, we use two conditions to specify scenarios. The first is whether the ego vehicle and the surrounding vehicle are interacting in the longitudinal direction (the ego vehicle’s heading direction) only or also in the lateral direction (perpendicular to the ego vehicle’s heading direction). We calculate the angle between the moving directions of the two vehicles and denote it as  $\omega$ . If  $\omega < 5^\circ$  or  $\omega > 175^\circ$ , they are considered to interact solely in the longitudinal direction (e.g., car-following and head-on conflict); otherwise, their interaction is considered to also involve the lateral direction (e.g., lane-changing and turning). We will refer to the two cases respectively as *Longitudinal* and *Lateral* in the following. Fig. 1 in Section 1 is referred to for some specific examples. The second condition is the relative speed between the ego vehicle and the surrounding vehicle. Samples separated by the first condition are sorted according to their relative speeds and further grouped. The grouping ensures at least 50,000 samples per group while maintaining the difference in average relative speed between groups at least 0.1 m/s apart.

We obtained 80 *Longitudinal* scenarios and 95 *Lateral* scenarios, where in each scenario an average ego vehicle was abstracted from all vehicle pairs. The inference of average spacing was performed for each scenario, which estimated series of  $\hat{r}$  and  $\hat{\beta}$ . Among



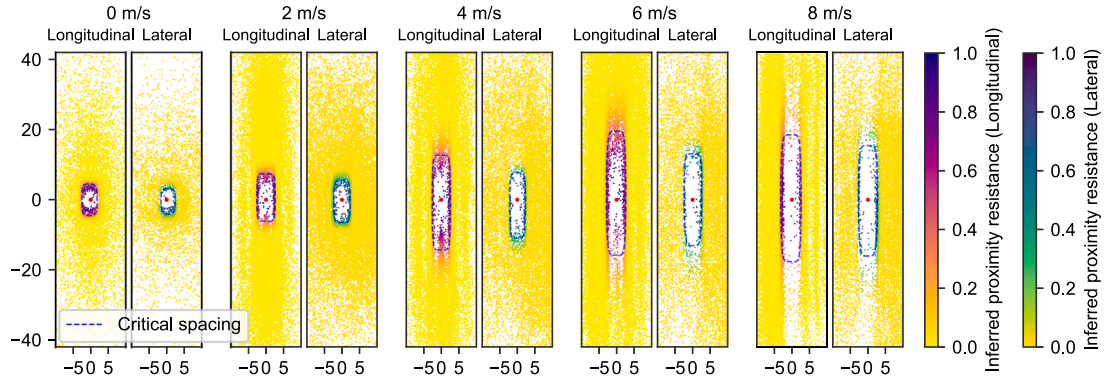


Fig. 6. Inference results in various scenarios. The averaged ego vehicles are marked with red circles at (0,0). Vehicles surrounding their ego vehicles are shown as scatters and coloured to indicate the inferred proximity resistance. The contours indicate critical spacing where proximity resistance is  $e^{-1}$ .

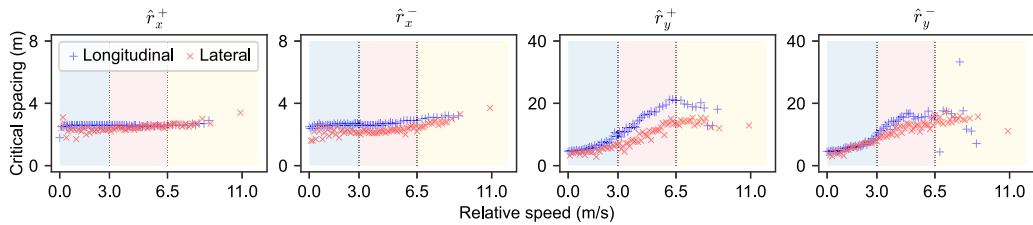


Fig. 7. Inferred critical spacing in different scenarios.  $\hat{r}_x^+$  and  $\hat{r}_x^-$  are critical spacing perpendicular to the relative velocity direction between interacting vehicles;  $\hat{r}_y^+$  and  $\hat{r}_y^-$  are that along the relative velocity direction between approaching and leaving vehicles, respectively. Two dotted lines where the respective relative speed is 3 m/s and 6.5 m/s divide each subplot into 3 cases under lower, medium, and higher relative speeds.

the estimated parameters,  $\hat{r} = \{\hat{r}_x^+, \hat{r}_x^-, \hat{r}_y^+, \hat{r}_y^-\}$  indicate the critical spacing where the transition of intrusion-caused discomfort is the most significant;  $\hat{\beta} = \{\hat{\beta}_x^+, \hat{\beta}_x^-, \hat{\beta}_y^+, \hat{\beta}_y^-\}$  depict the transition buffer from comfort to discomfort.

The values of  $\hat{r}_y^+$  and  $\hat{r}_y^-$  represent critical spacing along the direction of relative velocity when vehicles are approaching and leaving each other in a given scenario. In contrast, the values of  $\hat{r}_x^+$  and  $\hat{r}_x^-$  represent spacing along the direction perpendicular to the relative velocity, which is associated with the shape of the road section that is analysed. In *Longitudinal* scenarios,  $\hat{r}_x^+$  and  $\hat{r}_x^-$  correspond to the road widths, whereas in *Lateral* scenarios, they can be referred to the space within the intersection.  $\hat{r}_x^+$  and  $\hat{r}_x^-$  can also indicate whether the samples are evenly distributed on both sides of the  $y$ -axis. For example, if  $\hat{r}_x^- > \hat{r}_x^+$ , the sample density is significantly lower on the left side of the  $y$ -axis than on the right. To prevent this bias influencing the inference of  $\hat{r}_y$ , we set an upper limit of 7 m to  $\hat{r}_x$  in the experiments. In addition, the inference results were rejected when at least one of the  $p$ -values of  $\hat{\beta}$  is larger than 0.05, as they are statistically unreliable.

## 4. Results and findings

### 4.1. Critical spacing

Fig. 6 shows several inference results as examples. Scenarios where the average relative speed is closest to 0, 2, 4, 6, and 8 m/s are shown as five pairs of coloured scatter plots. The left side of each pair shows *Longitudinal* case and the right side shows *Lateral* case. Scatters in these plots refer to vehicle positions in the transformed coordinate system. They are coloured to indicate the inferred proximity resistance between the surrounding vehicles and the ego vehicles.

Driver spaces are the nearly empty areas with significantly fewer vehicles. These spaces are readily visible in the scatter plots, and our method properly depicts them. Their boundaries presented as the contours are the inferred critical spacing. As is seen in Fig. 6, the critical spacing increases with greater relative speeds in both *Longitudinal* and *Lateral* scenarios. However, it appears to be longer in *Longitudinal* than in *Lateral* scenarios at the same relative speed.

To further see how critical spacing changes in different scenarios, Fig. 7 displays the inferred critical spacing  $\hat{r}$  in various relative speeds. Overall, critical spacing expands along the relative velocity between interacting vehicles. This expansion is not limitless and eventually reaches a plateau after the relative velocity exceeds a certain threshold.

At relative speeds below around 3 m/s, the critical spacing along the relative velocity direction is approximately the same in *Longitudinal* and *Lateral* scenarios. However, when the relative speed is larger than 3 m/s, the critical spacing in *Longitudinal* scenarios

exceeds that in *Lateral* scenarios. This exceeding is particularly evident between approaching vehicles, as opposed to vehicles that are leaving away from each other. A possible explanation is different interaction strategies available in different scenarios. Specifically, vehicle interaction in *Longitudinal* scenarios is limited to adjusting spacing along the moving direction; while in interaction scenarios involving other directions, vehicles can make turns and thus require less space.

#### 4.2. Consistency evaluation

As is seen in Fig. 7, the inferred critical spacing do not always expand smoothly. Particularly at higher relative speeds larger than 6.5 m/s, some  $r_y^+$  and  $r_y^-$  diverge from the general curve and appear to be outliers. This entails the possibility that the proposed methods may produce inconsistent or unreliable inferences.

As the ground truth of critical spacing between vehicles is unknown, evaluating the reliability of the inference is not feasible. However, we can assess its consistency through bootstrapping. In this study,  $r$  and  $\beta$  parameterise proximity resistance change in different spacing and are considered to depend only on the specific scenario where the analysed samples lie. Therefore, the distributions of  $\hat{r}$  and  $\hat{\beta}$  should be identical for any sampling in the same scenario, which means they are pivotal. In this case, according to Davidson and MacKinnon (2000), the minimum number of bootstrapping iterations is 19 when considering a significance level of 0.05. We select 5 scenarios with relative speeds over 6.5 m/s respectively for *Longitudinal* and *Lateral* scenarios, and perform bootstrapping with 20 iterations on them. In every iteration for each scenario, we randomly select 85% of the vehicle samples with replacement, and then apply our method to infer the parameters.

Fig. 8 displays the bootstrapping results, where two sub-figures correspond to *Longitudinal* and *Lateral* scenarios. In this figure, the estimates of parameters are plotted as dots. The means of the estimates over different iterations in the same scenario are plotted as lines, and the standard deviations are plotted as shaded regions. In addition, the tables at the bottom also numerically show the statistics of bootstrapping results. To read the figure, the consistency of our inference is higher when the dots are more tightly distributed, when the shaded regions are narrower, and when the standard deviations are lower.

These plots and tables show that our method delivers consistent inferences overall. Notably in cases where the estimates appear to be outliers, such as  $\hat{r}_y^-$  in *Longitudinal* scenarios, our method still gives consistent results. Referring back to Fig. 6, one can see that the lower half of the *Longitudinal* scatter plot in the relative speed of 8 m/s does not present a clear boundary between areas of higher and lower sample densities. This unclear boundary violates the assumption that vehicles appear significantly more frequently outside of driver space than inside, and thus challenges the inference of  $\hat{r}_y^-$  as critical leaving spacing. Therefore, in contrast with potential inconsistencies in the inference method, the success in inferring driver space may depend more on the quality of data and sampling that establish clear boundaries as critical spacing.

#### 4.3. Interaction fundamental diagram

In this study, we quantify average 2D spacing in urban traffic and its relationship with relative speeds of interacting vehicles. This allows for deriving urban traffic states and their relationship, which we term as interaction Fundamental Diagram (iFD).

The accumulated samples of interacting vehicles are transformed to be in the same reference system and are conditioned at the same relative speed. As a result, we can define quasi-density  $k$  and quasi-flow  $q$  for vehicle interaction, as shown in Eq. (9). We calculate  $k$  as the inverse of  $D$ , where  $D$  is the amount of necessary space for interaction under a certain relative speed  $v$ . As a function of  $v$ ,  $k$  represents for *interaction density*. It describes space occupancy of vehicle interaction at different relative speeds. A higher interaction density means more interaction can occur within given road space, while a lower density means fewer. We then calculate  $q$  as the product of  $k$  and  $v$ , which represents for *interaction rate*. This multiplication is feasible due to coordinate transformation, after which the ego vehicles move uniformly along the  $y$ -axis. As also a function of  $v$ ,  $q$  describes time occupancy of vehicle interaction at different relative speeds. A higher interaction rate means faster interaction and is related to interaction efficiency.

$$\begin{cases} k = \frac{1}{D}, \\ q = kv. \end{cases} \tag{9}$$

The area of necessary space for vehicle interaction can vary given different levels of proximity resistance. With a proximity resistance of  $p^*$ , the dimensions of necessary space under a certain scenario can be computed by solving the inverse of Eq. (5). Eq. (10) shows the solution from inferred parameters  $\hat{r}$  and  $\hat{\beta}$ . As noted in Section 3.2,  $\hat{r}$  quantify the critical spacing and  $\hat{\beta}$  quantify the buffer across them.

$$d(p^*) = \hat{r}(-\ln(p^*))^{1/\hat{\beta}}, \tag{10}$$

where  $d = \{d_x^+, d_x^-, d_y^+, d_y^-\}$  and correspond to  $\{\hat{r}_x^+, \hat{r}_x^-, \hat{r}_y^+, \hat{r}_y^-\}$  and  $\{\hat{\beta}_x^+, \hat{\beta}_x^-, \hat{\beta}_y^+, \hat{\beta}_y^-\}$ . Then the area of necessary interaction space where the proximity resistance is smaller than  $p^*$  is computed as

$$D(p^*) = d_y^+(p^*)[d_x^-(p^*) + d_x^+(p^*)]. \tag{11}$$

Here we do not calculate the space area for vehicles leaving each other (i.e., we ignored  $d_y^-(p^*)$ ) as vehicles moving away are less restricted by one another.

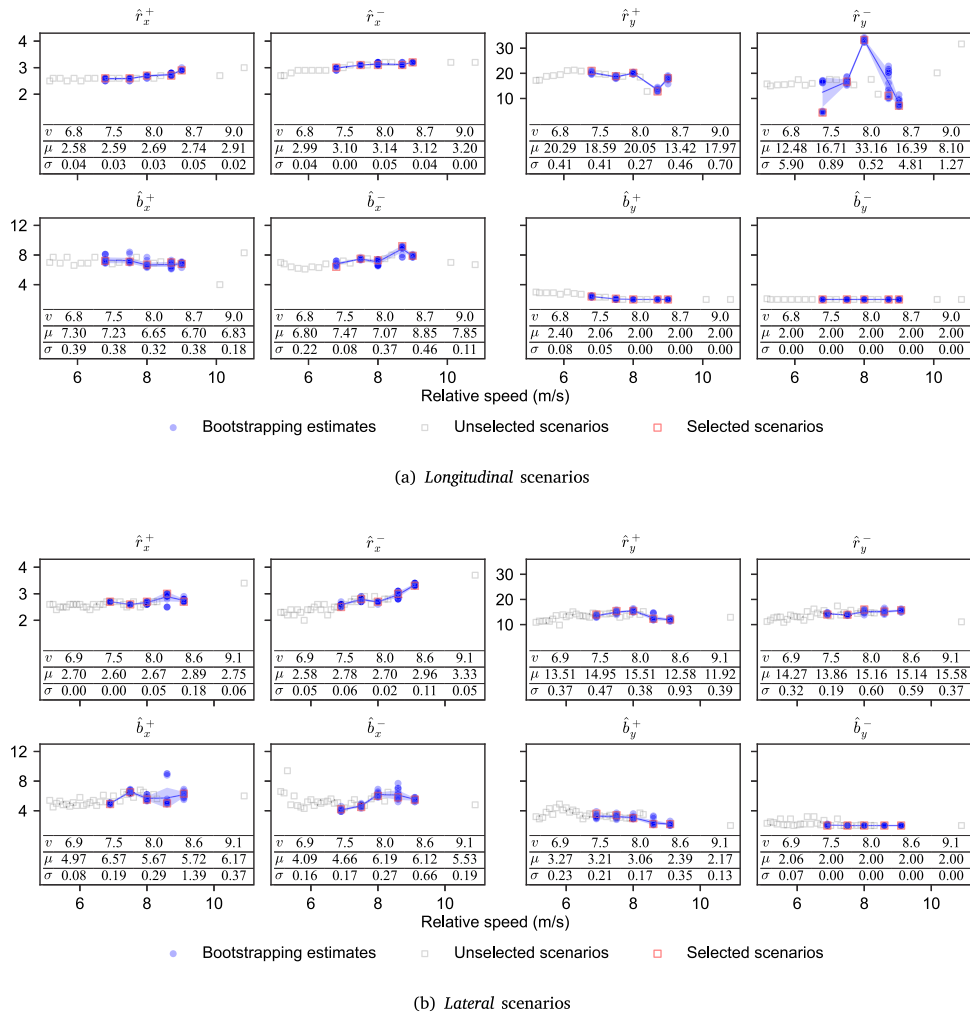


Fig. 8. Consistency evaluation of the proposed method through bootstrapping. In the bottom table of each sub-plot,  $v$  indicate the relative speeds of the selected scenarios;  $\mu$  and  $\sigma$  indicate the mean values and standard deviations of bootstrapping estimates in the selected scenarios, respectively.

Fig. 9 presents the relationships between  $k$ ,  $q$ , and  $v$  given different levels of proximity resistance. The relationships resemble FDs and so we refer to them iFDs. The blue dashed lines are iFDs at critical spacing (where proximity resistance is  $e^{-1}$ ), which mark the boundaries of frequent (less transparent) and infrequent (more transparent) occurrence of interaction states. Compared to *Longitudinal* iFDs, *Lateral* iFDs have higher interaction densities and interaction rates in various relative speeds, which suggests that lateral interactions use less road space and are more accommodated than longitudinal interactions at these unsignalised intersections.

Overall, as relative speeds increase, interaction density decreases; meanwhile, interaction rate first rises and then falls. In the initial phase of increasing interaction density, the necessary space for vehicle interaction is reduced. As a result, the limited road space can accommodate more interaction, leading to an increased interaction rate. Then in the second phase of rising interaction density, the available road space decreases faster than the reduction in necessary interaction space. Consequently, the interaction rate declines due to insufficient road space for vehicle interaction. The highest level of interaction rate represents the optimal point in the relationship, signifying where vehicle interaction is most often.

We can observe other interesting phenomena in the diagrams. To different extent in both *Longitudinal* and *Lateral* scenarios, there are noticeable drops in interaction rate for high relative speeds, with very limited increase in interaction density. This is comparable to capacity drops in longitudinal traffic flow theory, implying higher variation in driving behaviour at those particular states (Yuan et al., 2018; Calvert et al., 2018). Moreover, the diffusion of interaction states at different levels of proximity resistance suggests that stochastic spacing is a contributing factor to the scatters in empirical FDs. This is also consistent with previous studies (e.g., Nishinari et al., 2003; Geroliminis and Sun, 2011; Chen et al., 2013).

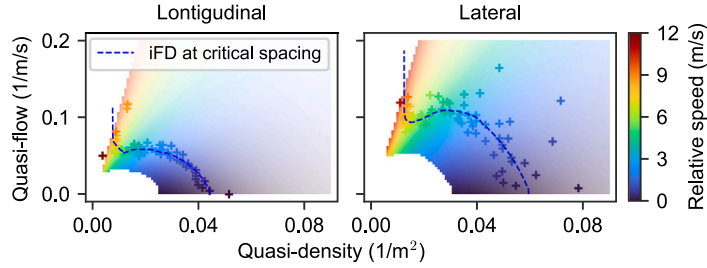


Fig. 9. Relationships between quasi-density and quasi-flow defined for urban vehicle interaction (grey scale may confuse speed and probability). These diagrams are probabilistic at different levels of proximity resistance, where the surface with varying colours and transparencies was created using smoothed curves to show the transition more clearly. The more transparent the colour, the higher the proximity resistance and the less probable vehicles are in that state.

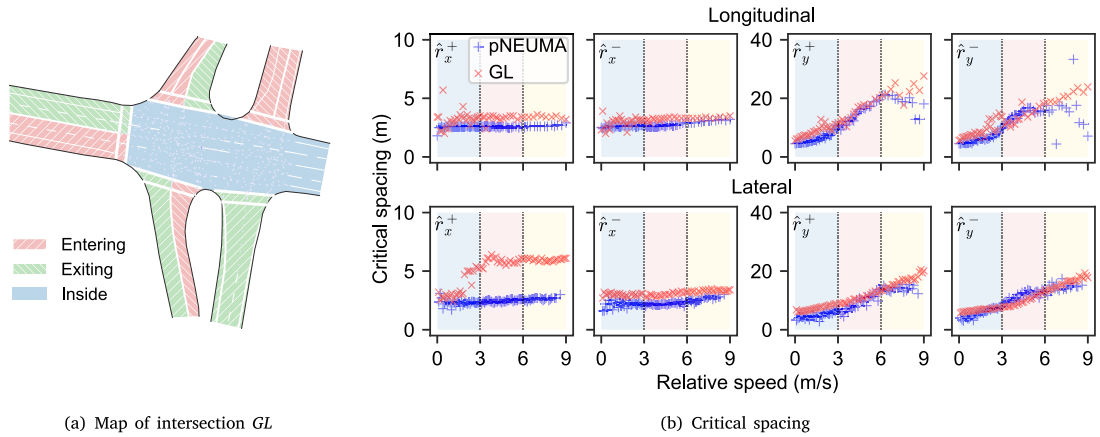


Fig. 10. Comparison of critical spacing in different scenarios between the intersection *GL* in INTERACTION and the unsignalised intersections in pNEUMA.

## 5. Discussion

### 5.1. Impact of intersection layout

In Figs. 6 and 8,  $\hat{r}_x^+$  and  $\hat{r}_x^-$  show very limited variation as relative speeds increase, and approximately correspond to road widths (around 3 m). As stated in Section 3.2, the values of  $\hat{r}_x$  are correlated with the infrastructure layout where the trajectory data are collected. Therefore, the seemingly constant  $\hat{r}_x$  are potentially related to the fact that almost all of the unsignalised intersections in the pNEUMA dataset are junctions of single-lane roads.

To investigate this conjecture further, we applied the same inference to a two-lane unsignalised intersection called *GL* in the INTERACTION dataset (Zhan et al., 2019), which was collected in the U.S. Fig. 10(a) shows a map of this intersection with the entering lane, the exiting lane, and the interior section marked. In addition, when sampling vehicle pairs, we excluded instances where one vehicle in a pair is in an entering lane while the other vehicle in the pair is in an exiting lane.

The inference results are displayed in Fig. 10(b) and are compared with the results from the unsignalised intersections in pNEUMA. As expected, in both *Longitudinal* and *Lateral* scenarios,  $\hat{r}_x^+$  and  $\hat{r}_x^-$  are larger at the intersection *GL* than at the pNEUMA intersections. This is particularly evident for  $\hat{r}_x^+$  in *Lateral* scenarios, which could be a consequence of two factors. One is the existence of yield lines at the entrances of the intersection legs, and the other is the busy entering lanes on the right side of vehicles entering the intersection on the main road.

It is necessary to underline that within the context of our methods,  $\hat{r}_x$  do not equate to “lateral spacing” in the sense of the original coordinate system. The coordinate transformation in Section 2.1 alters the view to observe the relative movement between vehicles, by simply rotating and translating the coordinate reference frame. This makes  $\hat{r}_y$  carry 2D information and be updated as vehicle interaction evolves. In Fig. 10(b),  $\hat{r}_y^+$  and  $\hat{r}_y^-$  exhibit comparable patterns at the intersection *GL* and the pNEUMA intersections. This comparability further supports and validates our methods.

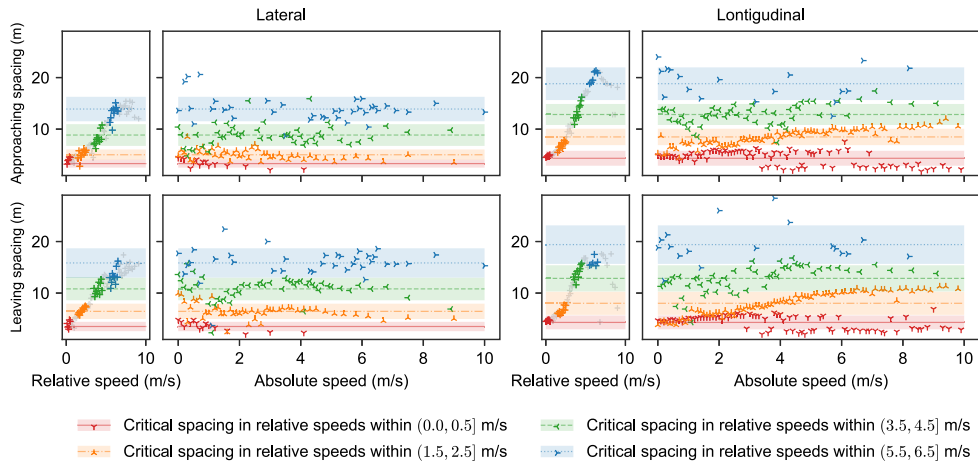


Fig. 11. Comparison of critical spacing across varying absolute speeds of the ego vehicle. For convenience of comparison, each sub-figure has a plot on the left showing the inferred critical spacing conditioned by relative speeds. The lines and shadows are mean values and standard deviations of the cases with varying absolute speeds but at similar relative speed in the main plot on the right.

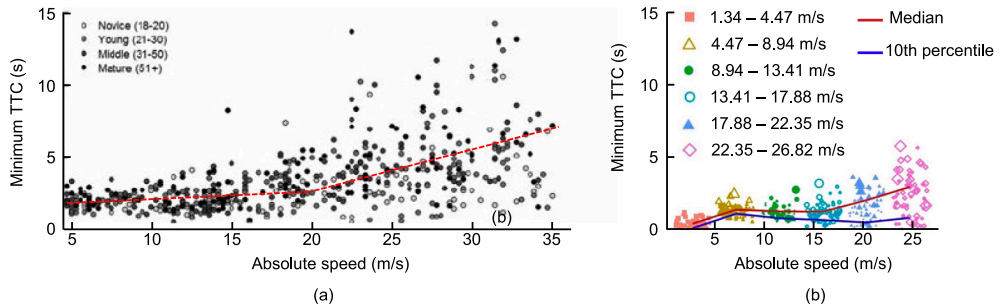


Fig. 12. The distribution of minimum Time-to-Collision with respect to absolute driving speed in the literature. The unit of speed were miles per hour in the original studies, and we converted them to meters per second. (a) Adapted from Figure 4 in Montgomery et al. (2014). (b) Adapted from Figure 8 in Kusano et al. (2015).

5.2. Impact of absolute speed

In this study, the importance of velocity differences between vehicles was emphasised for understanding 2D vehicle interaction. In light of such emphasis, our coordinate transformation omits the absolute speed, while it is often considered a crucial factor in determining vehicle spacing and interaction (Ma and Ahn, 2008; Yeo and Skabardonis, 2011).

To investigate the impact of vehicles’ absolute speeds, we examined how the spacing between vehicles varies across different absolute speeds of the ego vehicle. Fig. 11 shows the examination for both *Longitudinal* and *Lateral* scenarios under four different intervals of relative speed.

In *Lateral* scenarios, the critical spacing under the same interval of relative speed remains fairly consistent across absolute speeds ranging from 0 m/s to 10 m/s. This indicates that the impact of absolute speed is limited for *Lateral* interaction. While in *Longitudinal* scenarios there is a slight trend of increasing critical spacing with higher absolute speeds, it mainly hovers around the spacing determined by the relative speed. This observation is reasonable as we are measuring critical spacing, which, when conditioned by relative speed, is closely related to minimum Time-to-Collision (TTC). In Fig. 12, we show two plots adapted from other researchers’ empirical statistics of minimum TTC during 1D longitudinal interaction. As is clearly seen, minimum TTC is hardly influenced by absolute speeds under 20 m/s. Indeed, in urban trajectory datasets such as pNEUMA and INTERACTION, the absolute speeds of vehicles at intersections generally do not exceed 15 m/s.

We hereby suggest that the relative velocity direction between vehicles serves as a sufficient reference in the context of 2D spacing in urban traffic. The minimal impact of absolute speed on 2D critical spacing can be potentially explained from two aspects. First, we looked at vehicle interactions in congested urban traffic and the absolute speeds of vehicles are significantly slower than on highways. During such interactions, the relative velocity between vehicles plays a more critical role (Hoffmann and Mortimer, 1996; Wilmot and Khanal, 1999). Second, urban disruptions such as districts and intersections make the traffic discontinuous. Drivers may therefore be less sensitive to the changes in absolute speed in urban traffic.



## 6. Conclusion

In this study, we propose a method to infer average 2D spacing between vehicles from urban trajectory data. Our method is built upon interactions of vehicle pairs. We first transform the global coordinates of vehicle pairs into local coordinate systems based on their relative movement during interaction. The coordinate transformation establishes a normalised reference system for consistent analysis. We then accumulate all vehicle pairs together, and by estimating their spatial distribution, we infer the average 2D spacing between interacting vehicles at various relative speeds. Experiments on real-world urban trajectory data demonstrate that the inference results are consistent and behaviourally intuitive. Our inference enables the derivation of urban traffic states from the perspective of relative movement between interacting vehicles. Further, the relations between the derived traffic states are examined as interaction Fundamental Diagrams (iFDs), which describe the average amount of necessary road space for vehicle interactions at different relative speeds. Thereby, this study provides new methods and findings about vehicle interaction and traffic state estimation in cities.

Measuring 2D vehicle spacing through a driver space lens has two methodological limitations. First, the effectiveness of spacing inference is restricted by data. Driver space is delineated from the accumulated presence of vehicles. When the vehicle samples are inadequate to form an intuitively proper driver space, the spacing between vehicles may not be inferred. Second, the inferred spacing is a result of response mixture. Particularly in *Lateral* scenarios, the inference mixes ego vehicle driver's response and that of surrounding vehicle drivers. By setting the condition of static surrounding vehicles, spacing can be inferred as merely a result of discomfort response by the ego vehicle driver. Nevertheless, further research is needed to investigate this response mixture.

This study lays a methodological foundation for future research in the field of urban traffic analysis. Various conditions can be used to specify scenarios of interest in order to investigate urban traffic variations more thoroughly. For example, traffic states and iFDs may be compared across interconnected urban intersections to evaluate their service levels. Additionally, comparisons between different cities and in different time periods are also valuable for urban traffic management. Given adequate data, it would be possible to analyse average spacing between vehicles during distinct interactions, such as car-following, heading-on encounters, lane-changing, and turning. Such analysis will enhance our understanding about vehicle interaction in urban environments, and can contribute to improving traffic efficiency and reducing congestion at intersections.

## CRedit authorship contribution statement

**Yiru Jiao:** Conceptualization, Methodology, Software, Writing – original draft. **Simeon C. Calvert:** Methodology, Supervision, Writing – review & editing. **Sander van Cranenburgh:** Methodology, Supervision, Writing – review & editing. **Hans van Lint:** Conceptualization, Resources, Supervision, Writing – review & editing.

## Acknowledgements

This work is supported by the TU Delft AI Labs programme, The Netherlands. The authors would like to thank the editor and anonymous reviewers for their valuable comments and advice. Data source: pNEUMA – <https://open-traffic.epfl.ch>

## References

- Bärgman, J., Smith, K., Werneke, J., 2015. Quantifying drivers' comfort-zone and dread-zone boundaries in left turn across path/opposite direction (LTAP/OD) scenarios. *Transp. Res. Part F: Traffic Psychol. Behav.* 35, 170–184. <http://dx.doi.org/10.1016/j.trf.2015.10.003>.
- Barmounakis, E., Geroliminis, N., 2020. On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. *Transp. Res. C* 111, 50–71. <http://dx.doi.org/10.1016/j.trc.2019.11.023>.
- Calvert, S.C., van Wageningen-Kessels, F.L., Hoogendoorn, S.P., 2018. Capacity drop through reaction times in heterogeneous traffic. *J. Traffic Transp. Eng. (English Edition)* 5 (2), 96–104. <http://dx.doi.org/10.1016/j.jtte.2017.07.008>.
- Camara, F., Fox, C., 2020. Space invaders: Pedestrian proxemic utility functions and trust zones for autonomous vehicle interactions. *Int. J. Soc. Robotics* <http://dx.doi.org/10.1007/s12369-020-00717-x>.
- Cao, L., Krumm, J., 2009. From GPS traces to a routable road map. In: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, <http://dx.doi.org/10.1145/1653771.1653776>.
- Cassidy, M.J., 1998. Bivariate relations in nearly stationary highway traffic. *Transp. Res. B* 32 (1), 49–59. [http://dx.doi.org/10.1016/s0191-2615\(97\)00012-x](http://dx.doi.org/10.1016/s0191-2615(97)00012-x).
- Castillo, J., Benítez, F., 1995. On the functional form of the speed-density relationship—I: General theory. *Transp. Res. B* 29 (5), 373–389. [http://dx.doi.org/10.1016/0191-2615\(95\)00008-2](http://dx.doi.org/10.1016/0191-2615(95)00008-2).
- Chen, X., Li, Z., Li, L., Shi, Q., 2013. Characterising scattering features in flow-density plots using a stochastic platoon model. *Transportmetrica A: Transp. Sci.* 10 (9), 820–848. <http://dx.doi.org/10.1080/23249935.2013.822941>.
- Davidson, R., MacKinnon, J.G., 2000. Bootstrap tests: how many bootstraps? *Econometric Rev.* 19 (1), 55–68. <http://dx.doi.org/10.1080/07474930008800459>.
- Edie, L.C., 1963. *Discussion of Traffic Stream Measurements and Definitions*. The Port of New York Authority, New York, pp. 139–154.
- Flötteröd, G., Lämmel, G., 2015. Bidirectional pedestrian fundamental diagram. *Transp. Res. B* 71, 194–212. <http://dx.doi.org/10.1016/j.trb.2014.11.001>.
- Fourati, W., Friedrich, B., 2021. A method for using crowd-sourced trajectories to construct control-independent fundamental diagrams at signalized links. *Transp. Res. C* 130, 103270. <http://dx.doi.org/10.1016/j.trc.2021.103270>.
- Geroliminis, N., Sun, J., 2011. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transp. Res. B* 45 (3), 605–617. <http://dx.doi.org/10.1016/j.trb.2010.11.004>.
- Gibson, J.J., Crooks, L.E., 1938. A theoretical field-analysis of automobile driving. *Am. J. Psychol.* 51, 453–471.
- Graziano, M.S., Cooke, D.F., 2006. Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44 (6), 845–859. <http://dx.doi.org/10.1016/j.neuropsychologia.2005.09.009>.
- Greenshields, B.D., Thompson, J.T., Dickinson, H.C., Swinton, R.S., 1933. The photographic method of studying traffic behavior. In: *Highway Research Board Proceedings*, Vol. 13. pp. 382–399.



- Hall, F.L., Hurdle, V.F., Banks, J.H., 1992. Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways. *Transp. Res. Rec.* (1365), 12–18.
- Hennessey, M.P., Shankwitz, C., Donath, M., 1995. Sensor-based virtual bumpers for collision avoidance: configuration issues. In: Chachich, A.C., de Vries, M.J. (Eds.), *Collision Avoidance and Automated Traffic Management Sensors*, Vol. 2592. pp. 48–59. <http://dx.doi.org/10.1117/12.228920>.
- Hennessey, D.A., Howard, S., Carr, E., 2011. Driver space preference: Differences across age, gender and traffic conditions. In: Hennessey, D.A. (Ed.), *Traffic Psychology: An International Perspective*. New York, pp. 233–250.
- Hoffmann, E.R., Mortimer, R.G., 1996. Scaling of relative velocity between vehicles. *Accid. Anal. Prev.* 28 (4), 415–421. [http://dx.doi.org/10.1016/0001-4575\(96\)00005-x](http://dx.doi.org/10.1016/0001-4575(96)00005-x).
- Hoogendoorn, S.P., Bovy, P.H.L., 2001. State-of-the-art of vehicular traffic flow modelling. *Proc. Inst. Mech. Eng., Part I: J. Syst. Control Eng.* 215 (4), 283–303. <http://dx.doi.org/10.1177/095965180121500402>.
- Jin, W.-L., 2010. A kinematic wave theory of lane-changing traffic flow. *Transp. Res. B* 44 (8–9), 1001–1021. <http://dx.doi.org/10.1016/j.trb.2009.12.014>.
- Jin, W.-L., 2013. A multi-commodity lighthill-whitham-richards model of lane-changing traffic flow. *Transp. Res. B* 57, 361–377. <http://dx.doi.org/10.1016/j.trb.2013.06.002>.
- Kusano, K.D., Chen, R., Montgomery, J., Gabler, H.C., 2015. Population distributions of time to collision at brake application during car following from naturalistic driving data. *J. Saf. Res.* 54, 95–104. <http://dx.doi.org/10.1016/j.jsr.2015.06.011>.
- Lewis-Evans, B., De Waard, D., Brookhuis, K.A., 2010. That's close enough—A threshold effect of time headway on the experience of risk, task difficulty, effort, and comfort. *Accid. Anal. Prev.* 42 (6), 1926–1933. <http://dx.doi.org/10.1016/j.aap.2010.05.014>.
- Ma, T., Ahn, S., 2008. Comparisons of speed-spacing relations under general car following versus lane changing. *Transp. Res. Rec.* 2088 (1), 138–147. <http://dx.doi.org/10.3141/2088-15>.
- Makigami, Y., Newell, G.F., Rothery, R., 1971. Three-dimensional representation of traffic flow. *Transp. Sci.* 5 (3), 302–313. <http://dx.doi.org/10.1287/trsc.5.3.302>.
- Marsh, P., Collett, P., 1987. The car as a weapon. *ETC: Rev. Gen. Semant.* 44 (2), 146–151.
- Montgomery, J., Kusano, K.D., Gabler, H.C., 2014. Age and gender differences in time to collision at braking from the 100-car naturalistic driving study. *Traffic Inj. Prev.* 15, 15–20. <http://dx.doi.org/10.1080/15389588.2014.928703>.
- Näätänen, R., Summala, H., 1974. A model for the role of motivational factors in drivers' decision-making. *Accid. Anal. Prev.* 6 (3), 243–261. [http://dx.doi.org/10.1016/0001-4575\(74\)90003-7](http://dx.doi.org/10.1016/0001-4575(74)90003-7).
- Nishinari, K., Treiber, M., Helbing, D., 2003. Interpreting the wide scattering of synchronized traffic data by time gap statistics. *Phys. Rev. E* 68 (6), <http://dx.doi.org/10.1103/physreve.68.067101>.
- Novoy, M., Adali, T., Roy, A., 2010. A complex generalized Gaussian distribution— characterization, generation, and estimation. *IEEE Trans. Signal Process.* 58 (3), 1427–1433. <http://dx.doi.org/10.1109/tsp.2009.2036049>.
- Seyfried, A., Steffen, B., Klingsch, W., Boltes, M., 2005. The fundamental diagram of pedestrian movement revisited. *J. Stat. Mech. Theory Exp.* 2005 (10), P10002. <http://dx.doi.org/10.1088/1742-5468/2005/10/p10002>.
- Siebert, F.W., Oehl, M., Bersch, F., Pfister, H.-R., 2017. The exact determination of subjective risk and comfort thresholds in car following. *Transp. Res. Part F: Traffic Psychol. Behav.* 46, 1–13. <http://dx.doi.org/10.1016/j.trf.2017.01.001>.
- Siebert, F.W., Oehl, M., Pfister, H.-R., 2014. The influence of time headway on subjective driver states in adaptive cruise control. *Transp. Res. Part F: Traffic Psychol. Behav.* 25, 65–73. <http://dx.doi.org/10.1016/j.trf.2014.05.005>.
- Summala, H., 2007. Towards understanding motivational and emotional factors in driver behaviour: Comfort through satisficing. In: Cacciabue, P.C. (Ed.), *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*. London, pp. 189–207. [http://dx.doi.org/10.1007/978-1-84628-618-6\\_11](http://dx.doi.org/10.1007/978-1-84628-618-6_11).
- Taieb-Maimon, M., Shinar, D., 2001. Minimum and comfortable driving headways: Reality versus perception. *Hum. Factors* 43 (1), 159–172. <http://dx.doi.org/10.1518/001872001775992543>.
- Vanumu, L.D., Ramachandra Rao, K., Tiwari, G., 2017. Fundamental diagrams of pedestrian flow characteristics: A review. *Eur. Transp. Res. Rev.* 9 (4), <http://dx.doi.org/10.1007/s12544-017-0264-6>.
- Wagner, P., Brockfeld, E., Gartner, N.H., Sohr, A., 2009. Fundamental diagram of traffic flows on urban roads. *Transp. Res. Rec.* 2124 (1), 213–221. <http://dx.doi.org/10.3141/2124-21>.
- Wang, T., Li, H., Li, Z., Wang, Z., 2006. A fast parameter estimation of generalized Gaussian distribution. In: 2006 8th International Conference on Signal Processing, Vol. 1. <http://dx.doi.org/10.1109/icosp.2006.345546>.
- Wang, J., Rui, X., Song, X., Tan, X., Wang, C., Raghavan, V., 2014. A novel approach for generating routable road maps from vehicle GPS traces. *Int. J. Geogr. Inf. Sci.* 29 (1), 69–91. <http://dx.doi.org/10.1080/13658816.2014.944527>.
- Wilmot, C.G., Khanal, M., 1999. Effect of speed limits on speed and safety: A review. *Transp. Rev.* 19 (4), 315–329. <http://dx.doi.org/10.1080/014416499295420>.
- Wu, X., Liu, H.X., Geroliminis, N., 2011. An empirical analysis on the arterial fundamental diagram. *Transp. Res. B* 45 (1), 255–266. <http://dx.doi.org/10.1016/j.trb.2010.06.003>.
- Yeo, H., Skabardonis, A., 2011. Microscopic fundamental relationships between vehicle speed and spacing in view of asymmetric traffic theory. In: 14th International IEEE Conference on Intelligent Transportation Systems. (ITSC), pp. 1410–1414. <http://dx.doi.org/10.1109/itsc.2011.6082878>.
- Yuan, Y., Daamen, W., Goni Ros, B., Hoogendoorn, S., 2018. Investigating cyclist interaction behavior through a controlled laboratory experiment. *J. Transp. Land Use* (online) 11 (1), 833–847. <http://dx.doi.org/10.5198/jtlu.2018.1155>.
- Zhan, W., Sun, L., Wang, D., Shi, H., Clausse, A., Naumann, M., Kümmerle, J., Königshof, H., Stiller, C., de La Fortelle, A., Tomizuka, M., 2019. Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps. *arXiv* [arXiv:1910.03088](https://arxiv.org/abs/1910.03088).
- Zhang, J., Klingsch, W., Schadschneider, A., Seyfried, A., 2012. Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. *J. Stat. Mech. Theory Exp.* 2012 (02), P02002. <http://dx.doi.org/10.1088/1742-5468/2012/02/p02002>.
- Zhao, J., Knoop, V.L., Wang, M., 2022. Microscopic traffic modeling inside intersections: Interactions between drivers. *Transp. Sci.* <http://dx.doi.org/10.1287/trsc.2022.1163>.