



Fundamental diagram estimation by using trajectories of probe vehicles



Toru Seo^{a,*}, Yutaka Kawasaki^b, Takahiko Kusakabe^c, Yasuo Asakura^b

^a The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

^b Tokyo Institute of Technology, 2-12-1-M1-20, O-okayama, Meguro, Tokyo, 152-8552, Japan

^c Center for Spatial Information Science, the University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba, 277-8568, Japan

ARTICLE INFO

Article history:

Received 16 April 2018

Revised 18 January 2019

Accepted 9 February 2019

Keywords:

Traffic flow theory
Fundamental diagram
Flow–density relation
Mobile sensing
Probe vehicle
Connected vehicle

ABSTRACT

The fundamental diagram (FD), also known as the flow–density relation, is one of the most fundamental concepts in the traffic flow theory. Conventionally, FDs are estimated by using data collected by detectors. However, detectors' installation sites are generally limited due to their high cost, making practical implementation of traffic flow theoretical works difficult. On the other hand, probe vehicles can collect spatially continuous data from wide-ranging area, and thus they can be useful sensors for large-scale traffic management. In this study, a novel framework of FD estimation by using probe vehicle data is developed. It determines FD parameters based on trajectories of randomly sampled vehicles and a given jam density that is easily inferred by other data sources. A computational algorithm for estimating a triangular FD based on actual, potentially noisy traffic data obtained by multiple probe vehicles is developed. The algorithm was empirically validated by using real-world probe vehicle data on a highway. The results suggest that the algorithm accurately and robustly estimates the FD parameters.

© 2019 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY license.
(<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

The *fundamental diagram (FD)*, also known as the flow–density relation, is literally one of the most fundamental concepts in the traffic flow theory. An FD describes the relation between flow and density in *steady traffic* (sometimes referred as equilibrium or stationary traffic), that is, *traffic in which all the vehicles exhibit the same constant speed and spacing* (Daganzo, 1997; Treiber and Kesting, 2013). In theory, an FD itself contains useful information on traffic features, such as the value of free-flow speed and flow capacity, and distinction between free-flow and congested regimes. Observational studies also indicated a clear relation between flow and density in actual traffic at near-steady states (Cassidy, 1998; Yan et al., 2018). Additionally, macroscopic traffic flow dynamics are modeled by combining an FD and other principles—the most widely known example is the Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956). Furthermore, FDs explain microscopic vehicle behavior to a certain extent (Newell, 2002; Duret et al., 2008; Jabari et al., 2014). Thus, FDs are utilized by a wide variety of academic and practical purposes in traffic science and engineering fields such as traffic flow

* Corresponding author.

E-mail addresses: seo@civil.t.u-tokyo.ac.jp (T. Seo), t.kusakabe@csis.u-tokyo.ac.jp (T. Kusakabe), asakura@plan.cv.titech.ac.jp (Y. Asakura).

modeling/analysis/simulation (Daganzo, 1994; 1997; 2006; Treiber and Kesting, 2013), traffic control (Papageorgiou et al., 2003), traffic state estimation (Seo et al., 2017a), and dynamic traffic assignment in network (Szeto and Lo, 2006).

In order to estimate a parametric FD in actual traffic, it is necessary to collect data from traffic, assume the functional form of its FD, and estimate the FD parameters by fitting the curve to the data. Typically, FDs are estimated by using roadside sensors (e.g., cameras and detectors) from the era of Greenshields (1935). This is a straightforward method because usual roadside sensors measure traffic count and occupancy that are closely related to flow and density, respectively, at their location.

The limitation of the roadside sensor-based FD estimation is evident: it is unable to estimate FDs where sensors are not installed. Furthermore, because the installation and operational costs of roadside sensors are generally high, it is not practically possible to deploy detectors to everywhere. Thus, it is not always possible to determine FDs on every roads, especially arterial roads and highways in developing countries. It is also difficult to identify the exact locations and characteristics of bottlenecks even on freeways with some sensors. This limitation of detectors poses substantial difficulties for practical implementations of the aforementioned traffic science and engineering works that require pre-determined FDs.

*Probe vehicles*¹ (Sanwal and Walrand, 1995; Zito et al., 1995) received increasing attention currently, because of the rise of mobile technology and connected vehicles (Herrera et al., 2010; Shladover, 2017). The notable feature of the probe vehicles is that they collect spatially continuous data from wide-ranging area. Furthermore, the data can be collected during probe vehicles' daily travel, making them a cost-effective traffic data collection tool. These are remarkable advantages compared with conventional detectors, which can collect data at limited discrete locations (Herrera and Bayen, 2010; Jenelius and Koutsopoulos, 2015; Zheng and Liu, 2017). If FDs can be estimated by probe vehicle data by leveraging their advantages, it will enable us to estimate FDs in virtually everywhere. It will further enable us to implement various traffic science and engineering works to large-scale road networks. However, there is a paucity of research on FD estimation by using probe vehicle data. Thus, systematic and computational approaches for FD estimation are desirable given the high availability of probe vehicle data.

The aim of this study is to establish methodology to estimate FDs by using probe vehicle data. In order to enable the estimation, we allow the methodology to rely on minimum exogenous assumptions such as FD's functional form and value of a parameter of the FD (i.e., the jam density that is inferred from external knowledge). Moreover, this study presents a computable algorithm for FD estimation based on data collected by multiple probe vehicles that are driving potentially noisy real-world traffic. The algorithm was applied to real-world connected vehicle data to investigate the empirical validity of the proposed methodology.

The rest of the paper is organized as follows. In Section 2, the literature on FD estimation is reviewed and originality of this study is highlighted. In Section 3, the basic idea of FD estimation using probe vehicle data is discussed, and the identifiability of a triangular FD under idealized conditions is theoretically clarified. Then, in Section 4, a computational algorithm for estimating a triangular FD by using actual, noisy traffic data is developed based on the basic idea. The method statistically estimates the most probable FD using maximum likelihood estimation based on data collected by multiple probe vehicles. In Section 5, the empirical performance of the proposed algorithm is verified by applying it to real-world probe vehicle data. Section 6 concludes this paper. The frequently used abbreviations and notations are listed in Appendix A.

2. Literature review on FD estimation

FD estimation by using detectors is extensively studied since 1935. Many approaches use traffic data aggregated for a short time period (e.g., traffic count for 30 s–5 min periods) as inputs of estimation and extract statistical relation from them (see Treiber and Kesting, 2013; Qu et al., 2017; Knoop and Daamen, 2017, and references therein). Other approaches use disaggregated data that distinguish each individual vehicle in measurement area (e.g., Duret et al., 2008; Chiabaut et al., 2009; Coifman, 2014; 2015; Jin et al., 2015). The latter approaches enable the precise identification of steady traffic states. These methodologies rely on the fact that flow and density or trajectories of all the vehicles are easily obtained by detectors. However, this is not the case for typical probe vehicles: only trajectories of randomly sampled vehicles are collected by probe vehicles. Thus, these methodologies are not applicable for FD estimation by using probe vehicle data.

There is a paucity of research on FD estimation by using probe vehicle data. A few probe vehicle-based traffic state estimation methods simultaneously estimate FDs and traffic state (e.g., Tamiya and Seo, 2002; Seo et al., 2015; Sun et al., 2017). However, they rely on additional data sources, such as detector measurement or vehicle spacing in addition to probe vehicle's positioning. Appendix in Herrera et al. (2010) mentioned an idea for manual inference of an FD by using probe vehicle data under special conditions. However, it was not formulated and not computable. Their idea involves manually detecting a shockwave between a saturated free-flowing traffic and congested traffic, deriving the backward wave speed of a triangular FD by manually measuring the speed of the shockwave, and subsequently calculating the other FD parameters by using a given jam density. However, detection of a shockwave, detection of saturated traffic, and identification of the speed of shockwave are not trivial tasks in actual traffic.

¹ In this study, the term "probe vehicle" denotes a vehicle that continuously measures and reports its position and time (i.e., spatiotemporal trajectory) by using global navigation satellite systems, mobile phones, connected vehicle technologies, or other similar on-vehicle systems.

3. Identification of FD by using probe vehicles under idealized conditions

In this section, we mathematically investigate identifiability of FDs by using probe vehicle data under idealized conditions, such as homogeneous traffic described by the LWR model and accurate measurement.

3.1. Preliminaries

3.1.1. Assumptions

We assume that traffic dynamics are described by the LWR model (Lighthill and Whitham, 1955; Richards, 1956) with a homogeneous FD and first-in first-out (FIFO) condition without a source or sink. It is expressed as

$$\frac{\partial k(t, x)}{\partial t} + \frac{\partial q(t, x)}{\partial x} = 0, \quad (1)$$

$$q(t, x) = Q(k(t, x); \theta), \quad (2)$$

where $q(t, x)$ and $k(t, x)$ denote flow and density, respectively, at time–space point (t, x) , Q denotes a flow–density FD, and θ denotes a vector representing the FD parameters.

The functional form of an FD $Q(k; \theta)$ is assumed as to be *triangular* (Newell, 1993), which is expressed as

$$Q(k; u, w, \kappa) = \begin{cases} uk & \text{if } k \leq \frac{w\kappa}{u+w} \\ w(\kappa - k) & \text{otherwise,} \end{cases} \quad (3a) \quad (3b)$$

where u denotes the free-flow speed, w denotes the backward wave speed, and κ denotes the jam density. Thus, θ consists of u , w , and κ . Eq. (3a) represents a *free-flowing regime*, whereas Eq. (3b) represents a *congested regime*. The triangular FD is employed owing to its simplicity, good agreement with actual phenomena, and popularity in various applications (e.g., Daganzo, 2006). Note that it is easy to generalize the proposed methodology to incorporate single-valued, continuous, piecewise/fully differential FDs; see Appendix B.

A probe vehicle dataset is assumed to contain continuous trajectories of randomly sampled vehicles without any errors. It is represented as $\{X(t, n) | \forall n \in \mathbf{P}\}$, where $X(t, n)$ denotes the position that vehicle n exists at time t , and \mathbf{P} denotes a set of all the probe vehicles. The penetration rate of probe vehicles is unknown.

3.1.2. Definitions

Let *traffic state* in a time–space region be defined by Edie's generalized definition (Edie, 1963), namely,

$$q(\mathbf{A}) = \frac{\sum_{n \in \mathbf{N}(\mathbf{A})} d_n(\mathbf{A})}{|\mathbf{A}|}, \quad (4a)$$

$$k(\mathbf{A}) = \frac{\sum_{n \in \mathbf{N}(\mathbf{A})} t_n(\mathbf{A})}{|\mathbf{A}|}, \quad (4b)$$

$$v(\mathbf{A}) = \frac{\sum_{n \in \mathbf{N}(\mathbf{A})} d_n(\mathbf{A})}{\sum_{n \in \mathbf{N}(\mathbf{A})} t_n(\mathbf{A})}, \quad (4c)$$

where \mathbf{A} denotes a time–space region, and $q(\mathbf{A})$, $k(\mathbf{A})$, and $v(\mathbf{A})$ represent flow, density, and speed, respectively, in the time–space region \mathbf{A} , $\mathbf{N}(\mathbf{A})$ denotes a set of all the vehicle in \mathbf{A} , $d_n(\mathbf{A})$ and $t_n(\mathbf{A})$ denote distance traveled and time spent, respectively, by vehicle n in \mathbf{A} , and $|\mathbf{A}|$ denotes the area of \mathbf{A} . If traffic in \mathbf{A} is steady,

$$q(\mathbf{A}) = Q(k(\mathbf{A}); \theta) \quad (5)$$

is satisfied by the definition of the LWR model (2).

Let us consider a *probe pair*, that is, a pair of arbitrary two probe vehicles. A pair is denoted by m , and the preceding probe vehicle in the pair m is denoted by m_- whereas the following probe vehicle is denoted by m_+ . Note that arbitrary and unknown number of vehicles may exist between a probe pair.

Let $(q_m(\mathbf{A}), k_m(\mathbf{A}))$ be *probe traffic state (PTS)* of probe pair m defined as

$$q_m(\mathbf{A}) = \frac{d_{m_+}(\mathbf{A})}{|\mathbf{A}|}, \quad (6a)$$

$$k_m(\mathbf{A}) = \frac{t_{m_+}(\mathbf{A})}{|\mathbf{A}|}, \quad (6b)$$

following Edie's definition. Similarly, probe-speed is also defined as

$$v_m(\mathbf{A}) = \frac{d_{m_+}(\mathbf{A})}{t_{m_+}(\mathbf{A})}. \quad (7)$$

The PTS denotes flow and density of probe vehicle m_+ only in time–space region \mathbf{A} . If (i) traffic in \mathbf{A} is steady, (ii) distance traveled by each vehicle in \mathbf{A} is equal to the others, and (iii) probe vehicle m_+ exists in \mathbf{A} , then the following relations are satisfied:

$$q_m(\mathbf{A}) = \frac{q(\mathbf{A})}{|\mathbf{N}(\mathbf{A})|}, \tag{8a}$$

$$k_m(\mathbf{A}) = \frac{k(\mathbf{A})}{|\mathbf{N}(\mathbf{A})|}, \tag{8b}$$

$$v_m(\mathbf{A}) = v(\mathbf{A}), \tag{8c}$$

where $|\mathbf{N}(\mathbf{A})|$ denotes the size of $\mathbf{N}(\mathbf{A})$ (i.e., number of the vehicles in \mathbf{A}).

Let $c_m - 1$ denote the number of vehicles between probe pair m , and \mathbf{a}_m denote a time–space region whose boundary is on trajectories of m_+ and m_- (the region includes m_+ but not m_-). Since source/sink and FIFO-violations are absent, c_m is constant and

$$c_m = |\mathbf{N}(\mathbf{a}_m)| \tag{9}$$

always holds irrespective of shape, time, and location of \mathbf{a}_m . If (i) traffic in \mathbf{a}_m is steady and (ii) distance traveled by each vehicle in \mathbf{a}_m is equal to the others, then

$$q_m(\mathbf{a}_m) = \frac{Q(c_m k_m(\mathbf{a}_m); \theta)}{c_m} \tag{10}$$

is satisfied from Eqs. (5), (8a), and (8b). Hereafter, a PTS of pair m under the aforementioned conditions (i) and (ii) is simply referred to as *steady PTS*. Eq. (10) means that steady PTSs follow a relation that is similar (in the geometric sense) to the FD with a scale of $1/c_m$ on a flow–density plane. Hereafter, this relation is referred to as *probe fundamental diagram (PFD)* of m . Note that the PFD of probe pair m is also expressed as

$$q_m(\mathbf{a}_m) = Q(k_m(\mathbf{a}_m); \theta_m), \tag{11}$$

where θ_m denotes a pair-specific parameter vector, because the FD and PFD are similar. Because of the similarity, θ_m consists of free-flow speed of the PFD, u , backward wave speed of the PFD, w , and jam density of the PFD, κ/c_m .

In order to ensure that \mathbf{a}_m easily satisfies the aforementioned conditions (i) and (ii), we specify \mathbf{a}_m as follows:

$$\mathbf{a}_{mi} = \{(t, x) \mid X(t, m_+) \leq x < X(t, m_-) \text{ and} \\ -\phi(t - \tau_i) + X(\tau_i, m_+) \leq x \leq -\phi(t - \tau_i - \Delta t) + X(\tau_i + \Delta t, m_+)\}, \tag{12}$$

where i denotes an index number, τ_i denotes a reference time for \mathbf{a}_{mi} , and ϕ and Δt are the given parameters. An arbitrary value is acceptable for τ_i and ϕ , and an arbitrary non-zero value is acceptable for Δt if the corresponding trajectory $X(t, n)$ exists. The shape of \mathbf{a}_m is illustrated in Fig. 1. The specification (12) guarantees that if traffic in \mathbf{a}_m is steady, then distance traveled by each vehicle in \mathbf{a}_m is equal to the others, which is $d_{m_+}(\mathbf{a}_{mi})$. This feature ensures that the following discussion is concise without sacrificing significant generality.

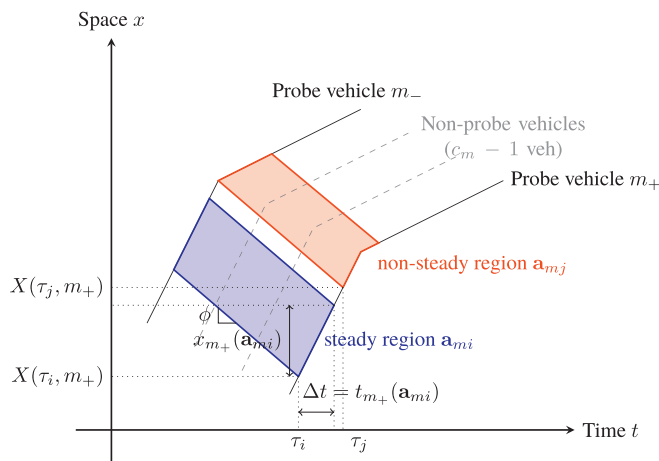


Fig. 1. Shape of the time–space region \mathbf{a}_m .

An FD is said to be *identifiable* by using a given dataset (a set of traffic states or a set of PTSs) if the value of its parameter vector θ is determined based on the dataset. Similarly, a PFD of certain probe pair m is said to be identifiable by using the given dataset (a set of PTSs of m) if the value of its parameter vector θ_m is determined based on the dataset.

For the readers' convenience, important abbreviations and notations used throughout this paper are summarized in [Appendix A](#).

3.2. Identifiability of FD based on probe vehicle data

It is mathematically obvious that following proposition holds true:

Proposition 1. *A triangular FD is identifiable if (i) one or more steady traffic state data from the free-flowing regime are available and (ii) two or more different steady traffic state data from the congested regime are available.*

Given the relation between a FD and a corresponding PFD of a probe pair shown in [Eq. \(10\)](#), following corollary can be obtained from [Proposition 1](#):

Corollary 1. *A triangular PFD is identifiable if (i) one or more steady PTS data from the free-flowing regime are available for a particular probe pair and (ii) two or more different steady traffic state data from the congested regime are available for the same probe pair.*

This means that the value of the free-flow speed and the backward wave speed can be determined by appropriate probe vehicle data.

According to [Eq. \(10\)](#) again, a PFD is similar (in the geometric sense) to a corresponding FD with a scale of $1/c_m$ on a flow–density plane. Therefore, once a PFD is identified and the value of the jam density is given, the corresponding FD can be identified. By combining this fact and [Corollary 1](#), following corollary can be obtained:

Corollary 2. *A triangular FD is identifiable if (i) one or more steady PTS data from the free-flowing regime are available for a probe pair, (ii) two or more different steady PTS data from the congested regime are available for the probe pair, and (iii) the value of the jam density is known.*

Notice that it is not necessary for an analyst to know whether a specific PTS belongs to free-flowing or congested regime beforehand. The belongingness is determined as a result of the identification of the FD.

3.3. Geometric interpretation of FD identification based on probe vehicle data

Geometric interpretation of the triangular FD identification based on [Corollary 2](#) is explained using [Fig. 2](#). [Fig. 2](#) shows traffic flow consisting of free-flow traffic and congested traffic as a time–space diagram (top) and a flow–density plane (bottom). In the time–space diagram, probe vehicles consisting of pair m are denoted by solid lines, non-probe vehicles are denoted by dashed gray lines, and a shockwave is denoted by an arrow. In the flow–density plane, dots indicate traffic states or PTSs, solid lines indicate transition of traffic states; the dashed line denotes the FD, and the thick dashed line denotes the PFD. Notice that the FD and the PFD are similar (in the geometric sense) with scale of $1/c_m$ as in [Eq. \(10\)](#). The blue area (top) and dot (bottom) indicate a steady region with free-flow traffic, the green ones indicate non-steady regions, and the red ones indicate steady region with congested traffic.

As explained, traffic state $S_m(\mathbf{a}_{mi})$ follows the FD if region \mathbf{a}_{mi} is steady. Similarly, PTS $S_m(\mathbf{a}_{mi})$ is equal to $S_m(\mathbf{a}_{mi})/c_m$ and follows the PFD if region \mathbf{a}_{mi} is steady. In [Fig. 2](#), regions \mathbf{a}_{m1} and \mathbf{a}_{m4} are steady; and thus their PTS are on the PFD. Conversely, regions \mathbf{a}_{m2} and \mathbf{a}_{m3} are non-steady; and thus their PTS are not on the PFD. Steadiness of region \mathbf{a}_{mi} can be inferred by using trajectories of probe vehicles. The necessary conditions for the steadiness are that the two probe vehicles' speeds are time-invariant, and they are equal to each other.

Suppose that the probe vehicles m_- and m_+ traveled to another congested region, say \mathbf{a}_{m5} (the gray dot in [Fig. 2](#)). In this case, the PFD is uniquely determined by selecting values of u , w , and κ/c_m such that all the points $S_m(\mathbf{a}_{m1})$, $S(\mathbf{a}_{m4})$, and $S(\mathbf{a}_{m5})$ are on the PFD $q = Q(k; u, w, \kappa/c_m)$. In other words, a triangle that satisfy following conditions can be determined uniquely: its vertex is at point (0,0), one of its edges is on line $q = 0$, and its another two edges pass the blue point, red point, and gray point. Thus, the FD is identified from the probe vehicle data and the jam density.

3.4. Discussion

The proposed method identifies all the parameters of an FD with the exception of its jam density by using probe vehicle trajectory data. The value of the jam density is required to be known prior to the FD identification, as it relates an FD and a PFD. There are several methods to obtain the value: common knowledge and remote sensing would be especially useful at this moment.

The jam density is considerably time- and space-independent variable compared to the other FD parameters such as capacity. Thus, it is reasonable to set the jam density based on a value that is observed at nearby locations. Alternatively, remote sensing (e.g., image recognition) from satellites is used to measure the jam density directly. Although remote sensing

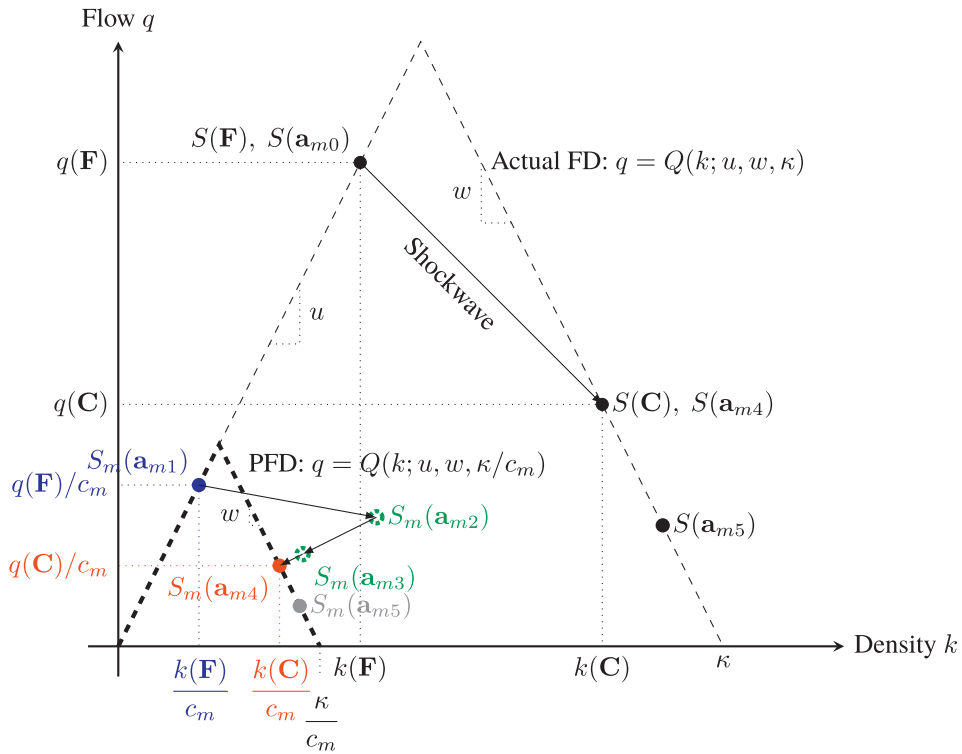
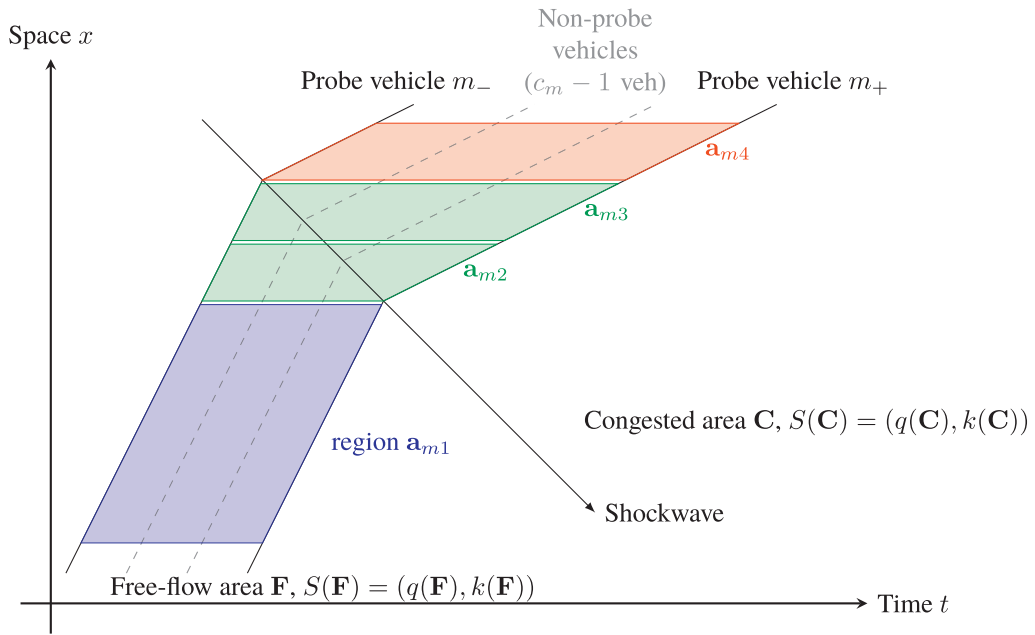


Fig. 2. Illustration of triangular FD identification. Top: Probe pair and PTS in time–space diagram. Bottom: FD, PFD, and PTS on flow–density plane.

does not provide us time-continuous traffic state variables, such as flow and speed (because the time interval between two measurements is extremely long, few hours to few days at least), it measures space-continuous density accurately (McCord et al., 2003). Thus, the jam density could be inferred as an upper limit of such measured density.

Besides, the proposed method identifies the values of the free-flow speed and the backward wave speed even if the value of the jam density is unknown as explained by Corollary 1. Since these two variables are important to characterize traffic flow, this property would be useful for various purposes such as traffic speed reconstruction and travel time estimation (Treiber et al., 2011).

A limitation of the proposed method is that a homogeneous FD is assumed. It means that the method in itself is unable to consider a bottleneck. This issue can be resolved by extending the method: for example, it is possible to estimate section-dependent FDs by applying the proposed method to each section iteratively. Another limitation is that the proposed method does not consider phenomena related with multi-valued and/or non-continuous FDs such as a capacity drop.

4. Estimation algorithm for actual traffic

Actual traffic data are not exactly consistent with the theoretical LWR model, and thus it is not possible to directly apply the methodology described in Section 3 to actual data—all the data will be considered as non-steady, and PTSs reported by different probe pairs may not derive the same PFD and FD. The inconsistency is mainly due to two sources. First, various traffic phenomena, such as the existence of non-equilibrium flow, bounded acceleration, heterogeneity among vehicles, FIFO violation in multi-lane sections, and the resulting stochasticity of macroscopic features, are not considered by the LWR model (Jabari et al., 2014; Jin et al., 2015; Coifman, 2015; Qu et al., 2017; Jin and Laval, 2018). Second, data always include measurement noise.

In order to account for this challenge, this section presents a statistical estimation algorithm for a triangular FD from actual traffic data collected from multiple probe pairs. Specifically, a method of extracting near-steady traffic state data from probe vehicle data is presented in Section 4.1. Then, an estimation method for free-flow speed and backward wave speed based on near-steady traffic state data collected by multiple probe pairs, which may exhibit noisy and heterogeneous behavior, is presented in Sections 4.2 and 4.3.

In order to simplify the mathematical representation, the triangular FD is characterized by free-flow speed, u , backward wave speed, w , and y -intercept of the congestion part of FD, α , as illustrated in Fig. 3. Note that $\alpha \equiv \kappa w$ holds, and thus, the characterization does not contradict the assumptions in Section 3.1.

4.1. Extraction of near-steady PTSs

Suppose that multiple PTSs $S_m(\mathbf{a}_{mi})$ for multiple m and i were calculated from probe vehicle data based on the definition (6) and (12). This set of PTSs is referred to as *raw PTS set*.

It is necessary to extract PTSs that are near-steady, meaning that they nearly satisfy the conditions in Corollary 2 from the raw PTS set. To do this, two filtering methods are employed. The first method is based on the linearity of the trajectories of two probe vehicles in pair m . Specifically, $S_m(\mathbf{a}_{mi})$ is considered as near-steady if trajectories of pair m in region \mathbf{a}_{mi} satisfy a condition based on the coefficient of variation:

$$\frac{\sigma_m(\mathbf{a}_{mi})}{\bar{v}_m(\mathbf{a}_{mi})} \leq \theta_{\text{steady}}, \quad (13)$$

where $\sigma_m(\mathbf{a}_{mi})$ and $\bar{v}_m(\mathbf{a}_{mi})$ denote the standard deviation and mean, respectively, of the instantaneous speed of probe vehicles in pair m in region \mathbf{a}_{mi} , and θ_{steady} denotes a given threshold.

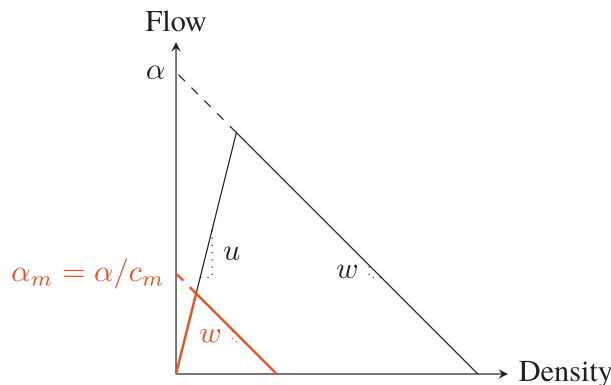


Fig. 3. Triangular FD with its three parameters u , w , and α . Black curve denotes FD, and red curve denotes the corresponding PFD of pair m . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In the second method, $\{S_m(\mathbf{a}_{mi}) | \forall i\}$ of each m is verified as to whether it is likely to satisfy the conditions in [Corollary 2](#). Data from the probe pair m are discarded if $\{S_m(\mathbf{a}_{mi}) | \forall i\}$ included no $S_m(\mathbf{a}_{mi})$ with speed lower than u_{\min} or no $S_m(\mathbf{a}_{mi})$ with speed faster than u_{\min} , where u_{\min} denotes a given lower bound for free-flow speed.² This corresponds to conditions (i) and (ii) in [Corollary 2](#). Additionally, data from probe pair m are discarded if correlation between probe-flow and probe-density in $\{S_m(\mathbf{a}_{mi}) | \forall i, v(\mathbf{a}_{mi}) \leq u_{\min}\}$ exceeds θ_{corr} , or standard deviation in $\{\bar{v}_m(\mathbf{a}_{mi}) | \forall i, v(\mathbf{a}_{mi}) \leq u_{\min}\}$ is smaller than θ_{std} , where θ_{corr} and θ_{std} denotes given thresholds. This corresponds to condition (ii) in [Corollary 2](#).

Hereafter, a set of PTS data obtained by the above procedure is referred to as *near-steady PTS set*. A set of probe pairs that collected elements in the near-steady PTS set is denoted as M , and a near-steady PTS set obtained by pair m is denoted as I_m .

4.2. Backward wave speed estimation

Backward wave speed is first estimated separately from the free-flow speed. This is because it is easy to determine whether a PTS datum definitely belongs to congested regimes based on the lower bound of the free-flow speed u_{\min} .

Specifically, the backward wave speed w is estimated by performing a linear regression on $(k(\mathbf{a}_{mi}), q(\mathbf{a}_{mi}))$ for each $i \in I_m, m \in M$, and $v_m(\mathbf{a}_{mi}) \leq u_{\min}$; and the final estimate is given as the mean of all the regression results. Condition (ii) in [Corollary 2](#) is essential to estimate w appropriately by the method.

4.3. Free-flow speed estimation by using EM algorithm

The free-flow speed u is subsequently estimated. To do this, an iterative algorithm categorized as expectation–maximization (EM) algorithm ([Dempster et al., 1977](#)) is constructed. In this algorithm, free-flow u is estimated in conjunction with the y -intercept value α_m and other variables, such as standard deviation of free-flowing and congested part of an FD, based on the near-steady PTS set.

4.3.1. Likelihood function

Assume that a probe-flow of a steady PTS belonging to regime s (free-flowing, $s = F$, or congested, $s = C$) follows a normal distribution in which the mean is the true PFD and standard deviation is $r_m \sigma_s$, where σ_s denotes a standard deviation of traffic state from an FD, r_m denotes a pair-specific normalizing parameter defined as \bar{q}_m/\bar{q} , \bar{q}_m denotes the mean of probe-flow of pair m , and \bar{q} denotes the mean of probe-flow of all of the probe pairs. Thus, likelihood of regime s of PTS datum i of pair m is expressed as

$$l(m, i, s) = \begin{cases} f(q_{mi} - uk_{mi}, r_m \sigma_s), & \text{if } s = F, \\ f(q_{mi} - (\alpha_m - wk_{mi}), r_m \sigma_s), & \text{if } s = C, \end{cases} \quad (14)$$

where $f(x, \sigma)$ denotes a value of the probability distribution function of a normal distribution with mean 0 and standard deviation σ at x . Note that variables u, α_m, s , and σ_s are unknown, while the rest are known.

The likelihood of the mixture distribution for all the near-steady PTS set is given by

$$L(q, k | \pi, u, \{\alpha_m | \forall m\}, \{\sigma_s | \forall s\}) = \prod_{m \in M} \prod_{i \in I_m} \sum_{s \in \{F, C\}} \pi_s l(m, i, s), \quad (15)$$

where π_s denotes the contribution of state s in the mixed distribution ([Bishop, 2006](#)). The intuitive meaning of L is a likelihood of observing PTS (q_{mi}, k_{mi}) set when parameters³ $\pi, u, w, \{\alpha_m\}, \{\sigma_s\}$ are given.

Our problem is to estimate the latent parameters $\pi, u, \{\alpha_m\}, \{\sigma_s\}$ from the observed near-steady PTS (q_{mi}, k_{mi}) set. This can be solved by applying an EM algorithm. If the complete dataset ([Bishop, 2006](#)) is available (i.e., latent variable z_{mis} that is 1 if S_{mi} belongs to regime s , 0 otherwise is given), then [Eq. \(15\)](#) is transformed to

$$L(q, k | \pi, u, \{\alpha_m\}, \{\sigma_s\}, z) = \prod_{m \in M} \prod_{i \in I_m} \sum_{s \in \{F, C\}} \pi_s^{z_{mis}} l(m, i, s)^{z_{mis}}. \quad (16)$$

By taking its expectation of logarithm, one obtains⁴

$$E[LL(q, k | \pi, u, \{\alpha_m\}, \{\sigma_s\}, \{p_{mi}(s)\})] = \sum_m \sum_i \sum_s p_{mi}(s) (\ln \pi_s + \ln l(m, i, s)), \quad (17)$$

where $p_{mi}(s)$ denotes an expectation of z_{mis} . The EM algorithm determines the value of latent parameters $\pi, u, w, \alpha, \sigma$, and p by maximizing [Eq. \(17\)](#) with an iterative procedure termed as E-steps and M-steps.

² Note that although a lower bound for free-flow speed u_{\min} is used in this procedure, it is not necessary to know the precise value of free-flow speed before estimation. This point is discussed later.

³ Hereafter, $\{\alpha_m | \forall m\}$ and $\{\sigma_s | \forall s\}$ are abbreviated as $\{\alpha_m\}$ and $\{\sigma_s\}$, respectively, for the purposes of simplicity.

⁴ Hereafter, $\sum_{m \in M}, \sum_{i \in I_m}$, and $\sum_{s \in \{F, C\}}$ are abbreviated as Σ_m, Σ_i , and Σ_s , respectively, for the purposes of simplicity.

4.3.2. E-step

In the E-step, $p_{mi}(s)$ is updated under given $\pi_s, u, w, \alpha_m, \sigma$. It is expressed as

$$p_{mi}(s) = \frac{\pi_s l(m, i, s)}{\sum_{r \in \{F, C\}} \pi_r l(m, i, r)}, \quad \forall s \in \{F, C\}, \quad \forall i \in I_m, \quad \forall m \in M. \quad (18)$$

4.3.3. M-step

In the M-step, $\pi_s, u, \{\alpha_m\}, \{\sigma_s\}$ is updated such that the likelihood (17) is maximized under given $p_{mi}(s)$. Specifically, it is expressed as

$$\max_{\pi, u, \{\alpha_m\}, \{\sigma_s\}} E[LL(q, k | \pi, u, \{\alpha_m\}, \{\sigma_s\}, \{p_{mi}(s)\})] = \sum_m \sum_i \sum_s p_{mi}(s) (\ln \pi_s + \ln l(m, i, s)), \quad (19a)$$

$$\text{s.t. } u_{\min} \leq u \leq u_{\max}, \quad (19b)$$

$$0 \leq \alpha_m, \quad \forall m, \quad (19c)$$

$$\sum_s \pi_s = 1, \quad (19d)$$

where u_{\min} and u_{\max} are non-negative constants. In order to solve problem (19), the method of Lagrange multiplier is adopted. The Lagrangian for the problem is defined as

$$\begin{aligned} \mathcal{L}(q, k | \pi, u, \{\alpha_m\}, \{\sigma_s\}, \{p_{mi}(s)\}, \{\lambda\}) &= \sum_m \sum_i \sum_s p_{mi}(s) (\ln \pi_s + \ln l(m, i, s)) \\ &\quad - \lambda_{\pi_s} (\sum_s \pi_s - 1) \\ &\quad - \lambda_{u_{\min}} (u_{\min} - u) - \lambda_{u_{\max}} (u - u_{\max}) \\ &\quad + \sum_m \lambda_{\alpha_m \min} \alpha_m \end{aligned} \quad (20)$$

where each λ with subscript denotes a Lagrangian multiplier corresponding to each constraint in problem (19).

By adopting the standard procedure that exploits the Karush–Kuhn–Tucker condition (Karush, 1939; Kuhn and Tucker, 1951), this problem is analytically solved. The solution is summarized as follows:

$$u = \begin{cases} \tilde{u} & \text{if } u_{\min} < \tilde{u} < u_{\max} \\ u_{\min} & \text{if } \tilde{u} \leq u_{\min} \\ u_{\max} & \text{if } \tilde{u} \geq u_{\max}, \end{cases} \quad (21)$$

with

$$\tilde{u} = \frac{\sum_m \sum_i p_{mi}(F) q_{mi} k_{mi} / r_m^2}{\sum_m \sum_i p_{mi}(F) k_{mi}^2 / r_m^2}, \quad (22)$$

and

$$\alpha_m = \begin{cases} \tilde{\alpha}_m & \text{if } 0 < \tilde{\alpha}_m \\ 0 & \text{if } \tilde{\alpha}_m \leq 0, \end{cases} \quad \forall m \quad (23)$$

with

$$\tilde{\alpha}_m = \frac{\sum_i p_{mi}(C) (q_{mi} + w k_{mi})}{\sum_i p_{mi}(C)}, \quad \forall m \quad (24)$$

and

$$\sigma_F^2 = \frac{\sum_m \sum_i (p_{mi}(F) (q_{mi} - u k_{mi})^2 / r_m^2)}{\sum_m \sum_i p_{mi}(F)}, \quad (25)$$

$$\sigma_C^2 = \frac{\sum_m \sum_i (p_{mi}(C) (q_{mi} - \alpha_m + w k_{mi})^2 / r_m^2)}{\sum_m \sum_i p_{mi}(C)}, \quad (26)$$

$$\pi_s = \frac{\sum_m \sum_i p_{mi}(s)}{\sum_{r \in \{F, C\}} \sum_m \sum_i p_{mi}(r)}, \quad \forall s \in \{F, C\}. \quad (27)$$

Notice that the right-hand side of Eq. (22) is indefinite if $p_{mi}(F) = 0$ holds for all m, i ; in other words, it is not possible to determine the free-flow speed if data are not observed from the free-flowing regime. This corresponds to condition (i) in Corollary 2.

4.4. Summary of the algorithm

Input for the proposed algorithm is as follows:

- Trajectories of multiple probe vehicles
- Jam density: κ
- Lower and upper bounds for free-flow speed: u_{\min} and u_{\max}
- Threshold for near-steady state θ_{steady}
- Cleaning parameter for PTSs based on [Corollary 2](#): θ_{corr} , θ_{std} .
- Parameter sets determining shape of time–space region \mathbf{a} : Δm , Δt , and ϕ
- Initial state of the EM algorithm: $p_{mi}(s)$
- Threshold for the convergence of the EM algorithm: ε

Main output of the proposed algorithm is as follows:

- Free-flow speed: u
- Backward wave speed: w
- Standard deviation of free-flowing flow from mean FD: σ_F
- Standard deviation of congested flow from mean FD: σ_C
- Number of vehicles between the probe pairs: c_m

The entire procedure of the proposed algorithm is summarized as follows:

Step 1 Near-steady PTS set construction step

Step 1.1 Construct raw PTS set by calculating PTSs for all $m_+ \in \mathbf{P}$, $m_- = m_+ - \Delta m$, $\Delta m \in \Delta M$, $t \in T$, $\Delta t \in \Delta T$, and $\phi \in \Phi$ based on the given probe vehicle data and definition (6) and (12), where ΔM , ΔT , and Φ denote sets of given parameter values.

Step 1.2 Construct near-steady PTS set by removing specific PTSs from the raw PTS set. Specifically, remove non-steady PTSs based on [Eq. \(13\)](#) with given θ_{steady} and non-appropriate PTSs that do not satisfy [Corollary 2](#) based on given θ_{corr} and θ_{std} .

Step 2 Backward wave speed estimation step

Step 2.1 Estimate the backward wave speed w by performing a linear regression on PTSs with $v \leq u_{\min}$ in the near-steady PTS set.

Step 3 Free-flow speed estimation step

Step 3.1 Define the likelihood function (17). Give an initial state on $p_{mi}(s)$.

Step 3.2 Update π_s , u , $\{\alpha_m\}$, $\{\sigma_s\}$ by executing the M-step.

Step 3.3 Update $p_{mi}(s)$ by executing the E-step.

Step 3.4 Check the convergence: if the change rate of all the parameters is lower than a given threshold ε , then it converges. If it converges, output the current parameter values and halt the algorithm. If not, go to Step 3.2.

4.5. Discussion

The proposed method estimates mean free-flow speed and backward wave speed of a triangular FD along with other auxiliary variables. Two of the auxiliary variables correspond to the standard deviation of flow from the mean FD; therefore, the method also captures the stochasticity of the FD. The method statistically estimates an FD based on data collected by multiple pairs of probe vehicles whose behavior may be heterogeneous as in general traffic. As long as the number of probe vehicles are sufficient and they are randomly sampled, the estimation results can be considered as a representative of the whole traffic.

Although the method requires a few input parameters, their values can be determined easily (i.e., no need of fine tuning) and the estimation result would be insensitive to most parameters with the exception of θ_{steady} and κ . The parameter θ_{steady} , threshold for near-steady state identification, must be important for the proposed method, since it is relevant to the original definition of the FD. This issue is investigated by performing a sensitivity analysis in [Section 5.3](#). The jam density could be determined reasonably as discussed in [Section 3.4](#); moreover, even if the jam density is not known, the free-flow speed and the backward wave speed are estimated as discussed in [Section 3](#). The other parameters are not essential. Any combinations of Δm , Δt , ϕ , θ_{corr} , and θ_{std} are acceptable as long as it produces sufficient number of near-steady PTSs from a given probe vehicle dataset such that the conditions (i) and (ii) of [Corollary 2](#) are satisfied. It is not necessary for the upper and lower bounds for free-flow speed, u_{\max} and u_{\min} , to be close to the actual free-flow speed. The upper bound u_{\max} is only required to eliminate non-realistic local optima with unrealistically high u (e.g., over 1000 km/h) as discussed subsequently in this section. The lower bound u_{\min} is required to determine congested traffic speed that exhibits significantly slower speed than the free-flow speed; therefore low values (e.g., 60 km/h for a highway) are acceptable.

The objective function of the problem (16) is not necessarily convex nor unimodal. Thus, local optima that does not represent an FD may exist. For example, if the upper limit for free-flow speed (19b) does not exist, obvious local optima are obtained at u with unrealistically high speeds, such as 1000 km/h, which classifies all the PTSs belonging to congested regime. The constrain (19b) is useful to eliminate such local optima.

The proposed method estimates u and w separately although it is possible to formulate a joint estimation problem of u , w , and other auxiliary variables as a single optimization problem similar to (19) (Seo et al., 2017b; Kawasaki et al., 2017). However, the separate estimation approach is adopted for the sake of empirical performance, because it was found that a joint estimation approach is difficult to solve stably due to numerous number of local optima, originated from fluctuations in actual traffic (Kawasaki et al., 2017). Note that u cannot be estimated by linear regression as in w ; because it is not possible to extract data from free-flow regimes without knowing a precise value of u in prior.

Reliability of an estimate of the proposed method can be assessed by a bootstrap method (c.f., Bishop, 2006), in which the method is repeatedly executed with re-sampled data in Monte Carlo simulation manner. Note that standard, simpler criteria such as the likelihood ratio and the Akaike information criteria cannot be applied to assess the reliability of the method with different parameter values, especially θ . This is because both of the number of model parameters and the number of samples depend on the value of θ .

It must be noted that the proposed near-steady PTS extraction method is not theoretically perfect, since it only considers speed of probe vehicles. In fact, as long as only probe vehicle data is used, it is not possible to exactly identify steady state; because behavior of non-probe vehicle does not necessarily influences probe vehicles. Therefore, from practical perspective, it would be important to have an acceptable extraction method. As we will see in Section 5, the proposed extraction method is acceptable in terms of its empirical performance.

The relation between the proposed method and the manual inference method proposed by Appendix in Herrera et al. (2010) is as follows. They share a few of the ideas: that is, they both estimate triangular FD parameters, namely u and w , based on probe vehicle data and a given jam density. Meanwhile, essential differences exist. First, the proposed method is based on the result of more general identifiability of an FD as discussed in Section 3. Thus, the proposed method does not rely on the detection of a shockwave nor the detection of saturated free-flowing traffic (which are used by Herrera et al. (2010)), which are special phenomena and difficult to automatically detect by using probe vehicles. Furthermore, the proposed method is a computable algorithm where discrepancies between the LWR theory and actual traffic are considered based on the steadiness of traffic and the noise terms.

5. Empirical validation

In this section, validation results of the estimation algorithm based on real-world data are presented.

5.1. Datasets

We used traffic data collected at approximately 4 km length section in an inter-city highway near Tokyo, Japan (i.e., the section between Kawasaki Junction and Yokohama–Aoba Interchange in Tomei Expressway). In this section, on-ramps, off-ramps, and merge/diverging sections are absent. The number of lanes is three, and the speed regulation was 100 km/h. According to the statistics, congestion occurs frequently during weekends and holidays mainly due to tourists in this section. The data collection duration was weekends and holidays from April 2016 to September 2016, 61 days in total.

As probe vehicle data, we used data collected by Fujitsu Traffic & Road Data Service Limited. The data were collected by connected vehicles mainly consisting of logistic trucks and vans that were connected to the Internet for fleet management purposes. The data included map-matched location information of each vehicles for every 1 s. Probe vehicle data collected on a day (April 9, 2016) are illustrated in Fig. 4 as time–space diagrams; and it indicates that some probe vehicles slowed down due to congestion between 9 and 11 a.m. The number of trips performed by probe vehicles during the aforementioned period was 23 103 trips. This corresponds to mean headway time between two consecutive probe vehicles 3.8 min. However, many of probe vehicles avoided congested time periods (this is confirmed by Fig. 4); therefore, headway time in congested traffic should exceed the value.

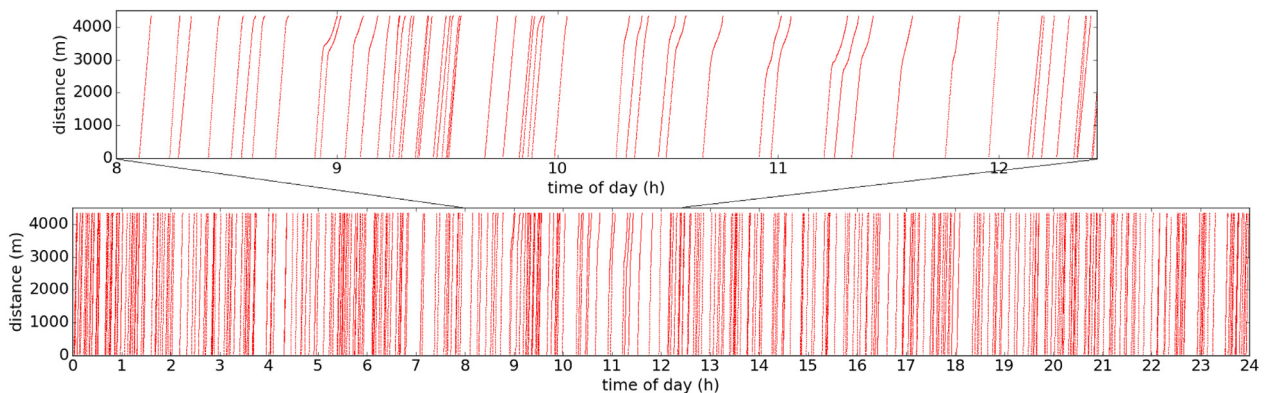


Fig. 4. Probe vehicle trajectories on a day.

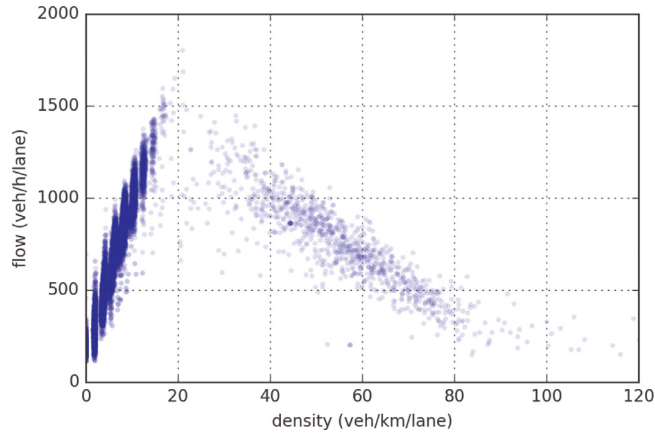


Fig. 5. Reference flow–density relation based on the detector data.

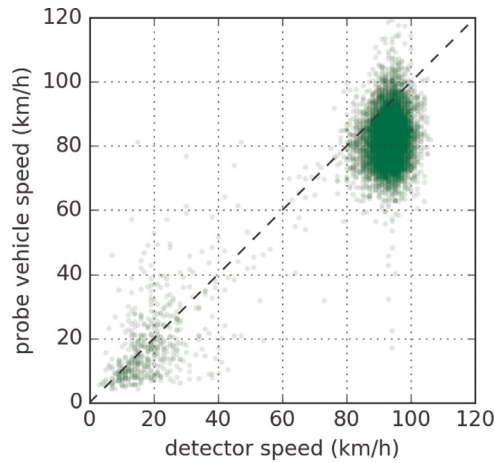


Fig. 6. Comparison between speed measured by probe vehicles and detector.

In order to validate the estimation results, loop detector data collected by Central Nippon Expressway Company Limited were used as a reference. The detector is located in the middle of the section. It measures vehicle type-specific flow and occupancy for every 5 min. Density was obtained from the data by using the method proposed by Cassidy and Coifman (1997). The flow–density plot based on the detector data is shown in Fig. 5. A clear triangular shape is observed; this is consistent with the existing empirical results (Cassidy and Coifman, 1997; Yan et al., 2018). According to the plot, the flow capacity was approximately 1500 veh/h and the jam density was approximately 100 veh/km. These values are slightly lower than those of usual highway. This might be due to the property of drivers during the period in which most of them were potentially inexperienced drivers (i.e., sightseeing travelers in weekends) and/or possible miss detection by the loop detector. Nevertheless, the values of the free-flow speed and the backward wave speed seems reasonable; therefore, this data can be considered as useful to validate the proposed method.

A comparison between speed measured by probe vehicles and detector for each time period is shown in Fig. 6. Essentially, the probe vehicle data exhibits a tendency similar to that of the detector data. However, in the high-speed regime, probe vehicle data tend to be slightly slower than the detector data. The bias is potentially because the probe vehicles consisted of logistic trucks and vans that are typically slower than average cars. From these results, we conclude that the probe vehicle data and detector data were sufficiently consistent in terms of speed measurement.

5.2. Parameter setting

The time–space regions \mathbf{a} were constructed based on all the combinations $(\Delta m, \Delta t, \phi)$ of the following values: $\Delta m \in \{1, 2, 3, 4\}$, $\Delta t \in \{1, 2, 3, 4\}$ (s), $\phi \in \{5, 10, 15, 20, 30\}$ (km/h). The threshold for near-steady state identification, θ_{steady} , was selected from 0.005 to 0.3. We performed a sensitivity analysis on this parameter by selecting each of them. The cleaning parameters for PTSs were set as $\theta_{\text{corr}} = -0.8$ and $\theta_{\text{std}} = 1$ (km/h). The upper and lower bound for free-flow speed were set as $u_{\text{min}} = 60$ (km/h) and $u_{\text{max}} = 120$ (km/h), respectively, and they are considered as reasonably mild constraints for a

Table 1
Summary of estimation results.

Parameter	# of near-steady data		Estimation results			
	Probe pairs	PTSs	u (km/h)	w (km/h)	σ_F (veh/h)	σ_C (veh/h)
θ_{steady}						
0.005	57	1 257	83.62	17.77	99.94	140.53
0.010	94	2 824	83.65	16.26	97.87	138.93
0.050	478	40 695	82.56	14.32	84.98	100.02
0.100	1172	229 892	81.94	16.02	166.82	181.30
0.150	1762	437 337	81.20	17.44	192.58	211.43
0.200	2344	654 732	80.55	17.63	197.48	191.81
0.300	3442	1 141 629	78.17	31.23	274.98	529.18

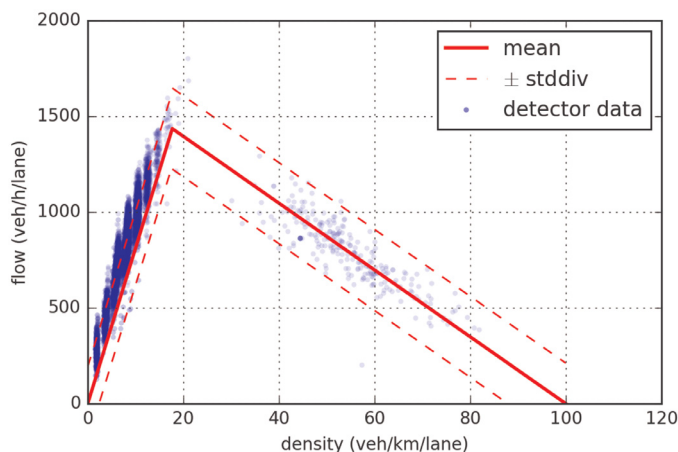


Fig. 7. Estimated FD and near-steady detector data. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

highway with 100 km/h speed regulation. The jam density κ was set to 100 veh/km based on Fig. 5. The initial state of the EM algorithm is $p_{mi}(F) = 1$ and $p_{mi}(C) = 0$ for m and i with $v_m(\mathbf{a}_{mi}) \leq u_{\min}$, and $p_{mi}(F) = 1$ and $p_{mi}(C) = 0$ for m and i with $v_m(\mathbf{a}_{mi}) > u_{\min}$. The convergence check parameter of the EM algorithm, ε , was set to 0.01.

5.3. Estimation results

First, an estimation result under a moderate setting, $\theta_{\text{steady}} = 0.15$, is presented. Given the parameter setting, 1762 probe pairs produced multiple near-steady PTSs that appeared to satisfy the conditions in Corollary 2, and the total number of PTSs was 437 337. The estimated values were free-flow speed $u = 81.2$ (km/h) and backward wave speed $w = 17.44$ (km/h). The values of estimates are summarized in Table 1 along with other parameter settings.

The estimation result with $\theta_{\text{steady}} = 0.15$ is illustrated in Fig. 7. The solid red line denotes mean estimates for u and w , the dashed lines represent the mean plus/minus the standard deviations σ_F and σ_C of the estimates, and the blue dots represent near-steady traffic states extracted from the detector data (by using a method similar to Cassidy (1998)) for reference. The mean estimate indicates good agreement with detector data. The estimated free-flow speed appears slightly slower than that of detector data. This is likely due to the biased speed of probe vehicles.⁵ Nevertheless, it is noteworthy that many of detector-measured traffic states fall within the standard deviation.

All the estimation results with different θ_{steady} are summarized in Table 1. Based on these results, sensitivity on θ_{steady} is investigated. With respect to the estimated u and w , similar results were obtained in all cases with the exception of $\theta_{\text{steady}} = 0.3$ and $\theta_{\text{steady}} = 0.05$. The case with $\theta_{\text{steady}} = 0.3$ resulted in low u and excessively high w . This is due to the low steadiness in the PTSs filtered by $\theta_{\text{steady}} = 0.3$; as a result, many of PTSs covered both the free-flowing and congested regime, making them inappropriate for the FD estimation. The case with $\theta_{\text{steady}} = 0.05$ resulted in slightly low w ; and this suggests that the estimates are slightly unstable if the number of PTSs are small. With respect to the estimated σ_s , a clear tendency was obtained: namely, they tended to decrease when the θ_{steady} decreased. This is an expected result since low θ_{steady} tends to eliminate scattered PTSs in the flow–density plane that typically exhibits relatively low steadiness.

⁵ According to Fig. 6, mean speed of probe vehicles in free-flowing regime is roughly 83 km/h (which is similar to the estimated u), whereas detectors measurement is roughly 95 km/h.

5.4. Discussion

The results indicate that the proposed method estimates reasonable values of the free-flow speed and the backward wave speed, depending on the value of threshold for the steadiness θ_{steady} . According to Table 1, the method derives fairly good estimates when θ_{steady} is less than 0.2, which is a wide range for this type of variable, namely, coefficient of variation. Conversely, as qualitatively expected, a case with excessively high θ_{steady} results in improper estimates because it did not properly remove non-steady traffic data. The estimated mean FDs are slightly biased compared to the detector data, likely due to biases in probe vehicles' driving behavior. This is one of the limitation of probe vehicle data. Meanwhile, Appendix B suggested that the proposed methodology can estimate unbiased FD if probe vehicles are not biased.

In addition to the mean FD, the proposed method also estimates standard deviation in the FD, implying that it can capture stochasticity of the FD. The standard deviation in free-flow regime is almost always lower than that of the congested regime; this is a reasonable result considering the nature of traffic flow.

With respect to the required amount of probe vehicle data, the results suggest that probe vehicle data with a few minutes of mean headway collected for few months are sufficient to estimate the FD. If the penetration rate is lower than this case, the FD estimation is possible by collecting the data for a longer period. On the contrary, if data collection duration was too short, we expect that the method may not estimate an appropriate FD, as probe vehicle data may not cover various traffic states. An analyst needs choose an appropriate data collection duration to ensure that probe vehicle data does not violate conditions (i) and (ii) of Corollary 2. With respect to the section length required to estimate an FD, the results suggest that the length of a few kilometers is sufficient for the employed data. It is difficult to precisely derive the minimum requirement because it depends on traffic conditions and the amount of data. Generally, if traffic congestion occurs frequently or the amount of probe vehicle data is large, then some of the probe vehicles are likely to satisfy conditions (i) and (ii) of Corollary 2 in a short road section. For this dataset, the minimum would be less than a kilometer, because some of the probe vehicles experience several traffic states at one kilometer length segment from the downstream end of this section as illustrated in Fig. 4.

The optimization problem may involve unrealistic local optima as discussed in Section 4.5. However, in the empirical analysis, the proposed method always yielded solution with physically reasonable results with the exception of the excessively high θ_{steady} case. This is potentially due to the separated estimation approach and the constraint on free-flow speed that we employed.

6. Conclusion

A theory and an algorithm for fundamental diagram (FD) estimation by using trajectories of probe vehicles is presented. They estimate FD parameters (i.e., free-flow speed and backward wave speed of a triangular FD) based on the jam density and time-continuous trajectories of randomly sampled vehicles. The algorithm is developed to estimate a reasonable FD from actual noisy traffic data. The algorithm was empirically and quantitatively validated based on real-world probe vehicle data on a highway. The results suggested that the algorithm accurately and robustly estimates triangular FD parameters and captures the stochasticity of the FD.

The following future research directions are considerable. First, it is important to develop a method that considers spatially heterogeneous FDs such that a bottleneck is endogenously detected. This is possible by extending the method to estimate location-dependent u and w and global c_m . Second, sophistication of the near-steady PTS extraction method is valuable. For example, consideration of duration of near-steady state would be useful. However, too strict criteria will result in the shortage of data; thus, careful design is required. Third, extension to other FD functional forms is desirable. Appendix B presents preliminary results. Finally, it is considerable to incorporate of a jam density inference method based on remote sensing by using satellites (c.f., Section 3.4), in order to make the method independent of exogenous knowledge. For this purpose, the “small satellites” that attracted significant attention recently (Sandau, 2010) will be useful owing to their flexible and low-cost sensing capability. The latter direction is now being investigated by the authors (Seo and Kusakabe, 2018).

Acknowledgments

The probe vehicle data were provided by Fujitsu Traffic & Road Data Service Limited and the detector data were provided by Central Nippon Expressway Company Limited. The study was financially supported by the Japan Society for the Promotion of Science (KAKENHI Grant-in-Aid for Young Scientists (B) 16K18164 and Scientific Research (S) 26220906). The authors would like to express their sincere appreciation for the help.

Appendix A. List of abbreviations and notations

Important abbreviations and notations that are used throughout this paper are listed below.

FD	Fundamental diagram
LWR	Lighthill–Whitham–Richards
FIFO	First-in first-out
PTS	Probe traffic state
PFD	Probe fundamental diagram
t	Time
x	Location
q	Flow
k	Density
v	Speed
$d_n(\mathbf{A})$	Distance traveled by vehicle n in time–space region \mathbf{A}
$t_n(\mathbf{A})$	Time spent by vehicle n in region \mathbf{A}
S	Traffic state. $S \equiv (q, k)$
Q	Density-to-flow FD function
κ	Jam density
u	Free-flow speed
w	Backward wave speed
α	y -intercept of the congestion part of flow–density triangular FD. $\alpha \equiv wk$
m	Probe pair(i.e., pair of two probe vehicles)
m_-	Preceding probe vehicle in pair m
m_+	Succeeding probe vehicle in pair m
c_m	Number of vehicles between probe vehicles in pair m plus one
ϕ	A parameter (angle) used to define a shape of region \mathbf{a}
Δt	A parameter (width) used to define a shape of region \mathbf{a}
α_m	y -intercept of the congestion part of flow–density triangular PFD of pair m . $\alpha_m \equiv wk/c_m$
θ_{steady}	Threshold for near-steady traffic detection
i	Index for datum
\mathbf{a}_{mi}	i th time–space region between probe vehicles in pair m
$q_m(\mathbf{a}_{mi})$	Probe-flow of pair m in region \mathbf{a}_{mi}
$k_m(\mathbf{a}_{mi})$	Probe-density of pair m in region \mathbf{a}_{mi}
$v_m(\mathbf{a}_{mi})$	Probe-speed of pair m in region \mathbf{a}_{mi}
$S_m(\mathbf{a}_{mi})$	PTS of pair m in region \mathbf{a}_{mi} . $S_m(\mathbf{a}_{mi}) \equiv (q_m(\mathbf{a}_{mi}), k_m(\mathbf{a}_{mi}))$
q_{mi}	Probe-flow of datum i of pair m . $q_{mi} \equiv q_m(\mathbf{a}_{mi})$
k_{mi}	Probe-density of datum i of pair m . $k_{mi} \equiv k_m(\mathbf{a}_{mi})$
S_{mi}	PTS of datum i of pair m . $S_{mi} \equiv (q_{mi}, k_{mi}) = S_m(\mathbf{a}_{mi})$
s	Variable representing regime of a traffic state. $s \in \{F, C\}$
F	Free-flowing regime
C	Congested regime
π_s	Contribution of state s in the mixed distribution
z_{mis}	1 if S_{mi} belongs to regime s , 0 otherwise
$p_{mi}(s)$	Expectation of z_{mis}

Appendix B. Cases with general continuous FDs

In this appendix, the applicability of the proposed methodology to FDs with general continuous functional forms is presented.

B1. Identifiability

We overwrite an assumption in Section 3.1.1 to make the methodology more general. Hereafter, the FD is assumed to be continuous.

Following the same idea described in Section 3, we have following proposition on the identifiability of continuous, differentiable FDs based on probe vehicle data:

Proposition 2. *A continuous, differentiable FD whose number of parameters is b is identifiable if (i) b or more different steady PTS data are available for a probe pair and (ii) the value of the jam density is known.*

Proof. The corresponding PFD of the probe pair is identifiable from the said PTS data; this is evident from the identifiability condition of continuous, differentiable functions. By comparing the jam density of the PFD and that of the FD, the value of the FD parameters can be derived. \square

B2. Algorithm

Now we consider a computational algorithm of FD estimation based on potentially noisy probe vehicle data as in Section 4. Following the same idea described in Section 4, an FD estimation algorithm is constructed as follows:

Step 1 Near-steady PTS set construction step. Construct a near-steady PTS set by Step 1.1 of the algorithm shown in Section 4.4.

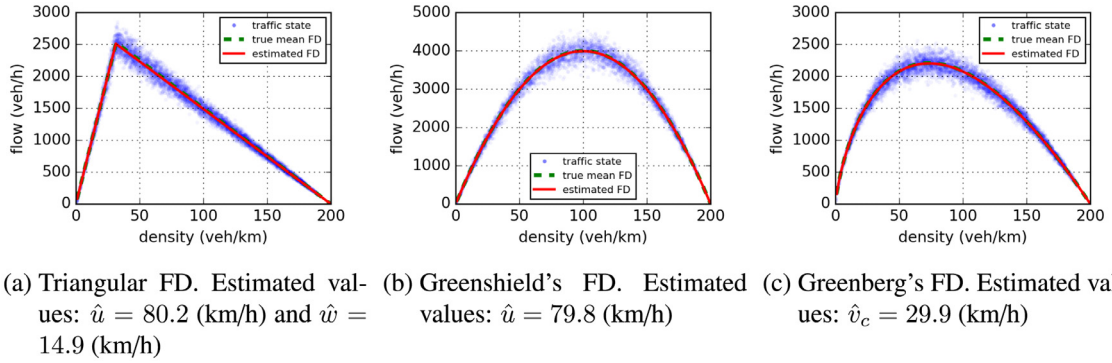


Fig. B1. Simulation-based FD estimation results.

Step 2 FD parameters estimation step. Solve the following residual minimization problem:

$$\min_{\theta, \{c_m\}} \sum_{m \in M} \sum_{i \in I_m} (c_m q_{mi} - Q(c_m k_{mi}; \theta))^2, \tag{B.1a}$$

$$\text{s.t. } c_m \geq 0, \quad \forall m, \tag{B.1b}$$

$$\text{physical constraints on } \theta, \tag{B.1c}$$

where θ denotes the FD parameters vector, (q_{mi}, k_{mi}) denotes a near-steady PTS datum i of probe pair m , and c_m denotes number of vehicles between a probe pair m plus one. Note that we use Eq. (B.1b) as a heuristic constraint on c_m , although true physical constraint is $c_m \geq 1$. This is because of a certain property of the optimization problem: this setting greatly increases the opportunity of finding physically reasonable optima by gradient methods. For the reason, see Seo et al. (2017b).

Although the expression of this algorithm is simpler than the proposed method for the triangular FD case in Section 4, the former may be more difficult to solve, as the latter leverages the piecewise linearity of the triangular FD. Development of an efficient algorithm for general FDs is out of scope of this appendix. Instead, in this appendix, problem (B.1) is solved by a quasi-Newton method with multiple initial values, which is general but potentially computationally costly.

B3. Simulation-based evaluation

In order to evaluate the method, synthetic data generated by traffic simulation was used. First, a functional form and parameters for an underlying FD was assumed: we used either of a triangular FD, Greenshield's FD ($Q_{GS}(k) = uk(1 - k/\kappa)$), or Greenberg's FD ($Q_{GB}(k) = v_c k \ln(\kappa/k)$, where v_c is a parameter) for this illustration purpose. The values of parameters were $u = 80$ (km/h), $w = 15$ (km/h), $\kappa = 200$ (veh/km), and $v_c = 30$ (km/h). Then, traffic states in 1 km homogeneous road section with various demand and supply conditions were generated by the underlying FD. A flow-dependent perturbation was added to the generated traffic states to simulate system and observation noises. Probe vehicles that record their location in every 1 seconds were randomly sampled with penetration rate of 5% from the generated ground truth traffic, so that 50 probe pairs were constructed. Finally, the FD estimation problem (B.1) was solved using L-BFGS-B algorithm (Byrd et al., 1995) with 30 random initial values.

Estimation results are summarized in Fig. B.1. It is clear that estimated FDs almost perfectly coincide with the corresponding true mean FDs. From these results, we conclude that the proposed methodology can estimate some continuous FDs in addition to the triangular FDs. Additionally, the proposed methodology can estimate an unbiased triangular FD if probe vehicle data was not biased. These findings complement the findings of the case study with real probe vehicle data in Section 5.

References

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer-Verlag, New York.

Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* 16 (5), 1190–1208.

Cassidy, M.J., 1998. Bivariate relations in nearly stationary highway traffic. *Transp. Res. Part B Methodol.* 32 (1), 49–59.

Cassidy, M.J., Coifman, B., 1997. Relation among average speed, flow, and density and analogous relation between density and occupancy. *Transp. Res. Record J. Transp. Res. Board* 1591, 1–6.

Chiabaut, N., Buisson, C., Leclercq, L., 2009. Fundamental diagram estimation through passing rate measurements in congestion. *IEEE Trans. Intell. Transp. Syst.* 10 (2), 355–359.

Coifman, B., 2014. Revisiting the empirical fundamental relationship. *Transp. Res. Part B Methodol.* 68, 173–184.

Coifman, B., 2015. Empirical flow-density and speed-spacing relationships: evidence of vehicle length dependency. *Transp. Res. Part B Methodol.* 78, 54–65.

Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transp. Res. Part B Methodol.* 28 (4), 269–287.

- Daganzo, C.F., 1997. Fundamentals of Transportation and Traffic Operations. Pergamon Oxford.
- Daganzo, C.F., 2006. On the variational theory of traffic flow: well-posedness, duality and applications. *Netw. Heterogeneous Med.* 1 (4), 601–619.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* 39 (1), 1–38.
- Duret, A., Buisson, C., Chiabaut, N., 2008. Estimating individual speed–spacing relationship and assessing ability of newell's car-following model to reproduce trajectories. *Transp. Res. Record J. Transp. Res. Board* 2088 188–197.
- Edie, L.C., 1963. Discussion of traffic stream measurements and definitions. In: Almond, J. (Ed.), *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, pp. 139–154.
- Greenshields, B.D., 1935. A study of traffic capacity. *Highway Res. Board Proc.* 14, 448–477.
- Herrera, J.C., Bayen, A.M., 2010. Incorporation of lagrangian measurements in freeway traffic state estimation. *Transp. Res. Part B Methodol.* 44 (4), 460–481.
- Herrera, J.C., Work, D.B., Herring, R., Ban, X.J., Jacobson, Q., Bayen, A.M., 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: the mobile century field experiment. *Transp. Res. Part C Emerg. Technol.* 18 (4), 568–583.
- Jabari, S.E., Zheng, J., Liu, H.X., 2014. A probabilistic stationary speed–density relation based on newell's simplified car-following model. *Transp. Res. Part B Methodol.* 68, 205–223.
- Jenelius, E., Koutsopoulos, H.N., 2015. Probe vehicle data sampled by time or space: consistent travel time allocation and estimation. *Transp. Res. Part B Methodol.* 71, 120–137.
- Jin, P.J., Li, J., Ran, B., 2015. The identification and empirical characterization of vehicular (lagrangian) fundamental diagrams in multilane traffic flow. *J. Adv. Transp.* 49 (7), 769–792.
- Jin, W.L., Laval, J., 2018. Bounded acceleration traffic flow models: a unified approach. *Transp. Res. Part B Methodol.* 111, 1–18.
- Karush, W., 1939. Minima of functions of several variables with inequalities as side conditions. Master's Thesis. Department of Mathematics, University of Chicago.
- Kawasaki, Y., Seo, T., Kusakabe, T., Asakura, Y., 2017. Fundamental diagram estimation using GPS trajectories of probe vehicles. In: *Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems*.
- Knoop, V.L., Daamen, W., 2017. Automatic fitting procedure for the fundamental diagram. *Transp. B Transp. Dyn.* 5 (2), 133–148.
- Kuhn, H.W., Tucker, A.W., 1951. Nonlinear programming. In: Neyman, J. (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 481–492.
- Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. a theory of traffic flow on long crowded roads. *proceedings of the royal society of london. series a. Mathemat. Phys. Sci.* 229 (1178), 317–345.
- McCord, M.R., Yang, Y., Jiang, Z., Coifman, B., Goel, P.K., 2003. Estimating annual average daily traffic from satellite imagery and air photos: empirical results. *Transp. Res. Record J. Transp. Res. Board* 1855 (1), 136–142.
- Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part i: general theory. *Transp. Res. Part B Methodol.* 27 (4), 281–287.
- Newell, G.F., 2002. A simplified car-following theory: a lower order model. *Transp. Res. Part B Methodol.* 36 (3), 195–205.
- Papageorgiou, M., Diakaki, C., Dinopolou, V., Kotsialos, A., Wang, Y., 2003. Review of road traffic control strategies. *Proc. IEEE* 91 (12), 2043–2067.
- Qu, X., Zhang, J., Wang, S., 2017. On the stochastic fundamental diagram for freeway traffic: model development, analytical properties, validation, and extensive applications. *Transp. Res. Part B Methodol.* 104, 256–271.
- Richards, P.L., 1956. Shock waves on the highway. *Oper. Res.* 4 (1), 42–51.
- Sandau, R., 2010. Status and trends of small satellite missions for earth observation. *Acta Astronaut* 66 (1), 1–12.
- Sanwal, K.K., Walrand, J., 1995. Vehicles as Probes. California Partners for Advanced Transit and Highways (PATH). Working Paper.
- Seo, T., Bayen, A.M., Kusakabe, T., Asakura, Y., 2017. Traffic state estimation on highway: a comprehensive survey. *Annu. Rev. Control* 43, 128–151.
- Seo, T., Kusakabe, T., 2018. Traffic state estimation using small imaging satellites and connected vehicles. *Transp. Res. Procedia* 34, 4–11. (ISTS and IWTDSCS 2018, 4–6 August 2018, Matsuyama, Japan).
- Seo, T., Kusakabe, T., Asakura, Y., 2015. Traffic state estimation with the advanced probe vehicles using data assimilation. In: *Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 824–830.
- Seo, T., Kusakabe, T., Asakura, Y., 2017. Calibration of fundamental diagram using trajectories of probe vehicles: basic formulation and heuristic algorithm. In: *Proceedings of the Transportation Research Procedia* 21, 6–17. ISTS&IWTDSCS, Jeju, Korea. 7–8, 2016.
- Shladover, S.E., 2017. Connected and automated vehicle systems: introduction and overview. *J. Intell. Transp. Syst.*
- Sun, Z., Jin, W.L., Ritchie, S.G., 2017. Simultaneous estimation of states and parameters in Newell simplified kinematic wave model with Eulerian and lagrangian traffic data. *Transp. Res. Part B Methodol.* 104, 106–122.
- Szeto, W.Y., Lo, H.K., 2006. Dynamic traffic assignment: properties and extensions. *Transportmetrica* 2 (1), 31–52.
- Tamiya, K., Seo, T., 2002. Q-V Configuration for urban street using probing car data. In: *Proceedings of the Infrastructure Planning*, 25, pp. 599–602. (in Japanese).
- Treiber, M., Kesting, A., 2013. *Traffic Flow Dynamics: Data, Models and Simulation*. Springer Berlin Heidelberg.
- Treiber, M., Kesting, A., Wilson, R.E., 2011. Reconstructing the traffic state by fusion of heterogeneous data. *Comput. Aided Civ. Infrastruct. Eng.* 26 (6), 408–419.
- Yan, Q., Sun, Z., Gan, Q., Jin, W.L., 2018. Automatic identification of near-stationary traffic states based on the PELT changepoint detection. *Transp. Res. Part B Methodol.* 108, 39–54.
- Zheng, J., Liu, H.X., 2017. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transp. Res. Part C Emerg. Technol.* 79, 347–362.
- Zito, R., D'Este, G., Taylor, M.A.P., 1995. Global positioning systems in the time domain: how useful a tool for intelligent vehicle-highway systems? *Transp. Res. Part C Emerg. Technol.* 3 (4), 193–209.