Fundamental diagram estimation by using trajectories of probe vehicles

Toru Seo\textsuperscript{a,b,\ast}, Yutaka Kawasaki\textsuperscript{a}, Takahiko Kusakabe\textsuperscript{c}, Yasuo Asakura\textsuperscript{a}

\textsuperscript{a}Tokyo Institute of Technology, 2-12-1-M1-20, O-okayama, Meguro, Tokyo 152-8552, Japan
\textsuperscript{b}University of Michigan, 2350 Hayward St., Ann Arbor, MI 48109, The United States
\textsuperscript{c}Center for Spatial Information Science, the University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan

Abstract

The fundamental diagram (FD), also known as the flow–density relation, is one of the most fundamental concepts in the traffic flow theory. It describes the relation between equilibrated flow, density, and speed in traffic flow. Conventionally, FDs are estimated by using fixed roadside sensors such as detectors. On the other hand, there is a paucity of studies on FD estimation by using mobile sensors, namely probe vehicles. Probe vehicles enable the collection of traffic data from the entire road network, and thus they are useful sensors for large-scale traffic management. In this study, a 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. Specifically, an algorithm for estimating a triangular FD based on actual, potentially noisy traffic data is developed. The algorithm is 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.

Keywords: traffic flow theory; fundamental diagram; flow–density relation; mobile sensing; probe vehicle; connected vehicle

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 (that is sometimes referred as equilibrium or stationary traffic), i.e., traffic in which all the vehicles exhibit the same constant speed and spacing (Daganzo, 1997; Treiber and Kesting, 2013; Yan et al., 2018). 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

\textsuperscript{\ast}Corresponding author

Email addresses: t.seo@plan.cv.titech.ac.jp (Toru Seo), t.kusakabe@csis.u-tokyo.ac.jp (Takahiko Kusakabe), asakura@plan.cv.titech.ac.jp (Yasuo Asakura)

Preprint submitted to Elsevier

April 18, 2018
flow and density in actual traffic at near-steady states (Cassidy, 1998). 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 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 in general. 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. Thus, it is difficult to determine FDs on roads without sensors such as arterial roads. It is also difficult to identify the exact locations and characteristics of bottlenecks even on freeways with some sensors. The issue poses substantial difficulties for practical implementations of the aforementioned traffic science and engineering works because the installation and operational costs of roadside sensors are generally high.

**Probe vehicles** received increasing attention currently (Sanwal and Walrand, 1995; Zito et al., 1995; Herrera et al., 2010; Shladover, 2017).\(^1\) A well-known advantage of probe vehicles as a traffic monitoring tool is their significantly broad data collection range in the spatiotemporal domain when compared with roadside sensors. For example, one of the most typical applications is traffic state estimation using partial measurement data (Nanthawichit et al., 2003; Seo et al., 2017a); probe vehicles enabled the estimation of a traffic state in a wide spatiotemporal domain. However, existing studies on these types of applications of probe vehicles often assume their FDs exogenously to relate flow and density to speed measured by probe vehicles (e.g., Nanthawichit et al., 2003; Yuan et al., 2012; Wada et al., 2015; Canepa and Claudel, 2017). This is a critical limitation for application of probe vehicles because FD estimation itself is a challenging problem as discussed above. Especially, 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 the study is to propose 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, a computable algorithm for FD estimation is developed and validated by using real-world data to illustrate the empirical validity of the proposed

---

\(^1\)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
The methodology.

The rest of the paper is organized as follows. The literature on FD estimation is reviewed and originality of this study is highlighted in Section 2. Identifiability of an FD by using probe vehicle data under idealized conditions is rigorously established in Section 3. Specifically, sufficient conditions with respect to input data, FD’s functional form, and exogenous assumptions to identify an FD are discussed for general cases and explained in detail for a triangular FD case. An algorithm for estimating a triangular FD by using actual, noisy traffic data is developed in Section 4. The proposed algorithm is validated by applying it to real-world probe vehicle data in Section 5. 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 sec–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 conventional 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.

3. Identification of FD under idealized conditions

3.1. Preliminaries

3.1.1. Assumptions

We assume that traffic dynamics are described by the LWR model with a homogeneous FD and first-in first-out (FIFO) condition without a source or sink. It is expressed as follows:

\[
\frac{\partial k(t, x)}{\partial t} + \frac{\partial q(t, x)}{\partial x} = 0,
\]

(1)
\[ q(t, x) = Q(k(t, x); \theta), \]  

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 \( \theta \) denotes a vector representing the FD parameters (Lighthill and Whitham, 1955; Richards, 1956).

The functional form of an FD \( Q(k; \theta) \) is assumed as a single-valued, continuous, and piecewise differential. The functional form and parameter values of each of the piecewise differential parts of an FD are assumed as not identical to each other. Additionally, \( Q(0; \theta) = 0 \) is assumed to hold, which is obviously reasonable. Note that many of commonly used FD functional forms, such as the triangular FD (Newell, 1993), satisfy these assumptions.

A probe vehicle dataset is assumed to contain continuous trajectories of randomly sampled vehicles without any errors. It is represented as \( \left\{ X(t, n) \ | \ \forall n \in P \right\} \), where \( X(t, n) \) denotes the position that vehicle \( n \) exists at time \( t \), and \( P \) 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

\[ q(A) = \frac{\sum_{n \in N(A)} d_n(A)}{|A|}, \]  
\[ k(A) = \frac{\sum_{n \in N(A)} t_n(A)}{|A|}, \]  
\[ v(A) = \frac{\sum_{n \in N(A)} d_n(A)}{\sum_{n \in N(A)} t_n(A)}, \]

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

\[ q(A) = Q(k(A); \theta) \]  

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

Without loss of generality, the FD is defined as follows. The total number of parameters of the FD is denoted by \( g \). The number of the piecewise differential parts of the FD is \( h \), and its \( j \)-th part (labeled by the increasing order of \( k \)) includes \( g_j \) parameters and the function is denoted by \( Q_j \) (hereafter, a piecewise differential part is simply referred to as a part). Thus, \( \sum_{j=1}^{h} g_j = g = |\theta| \) holds. The \( h \)-th part includes at least one parameter, that is, the jam density \( \kappa \). An example of the FD is illustrated in Fig. 1.

A traffic state \((q, k)\) is said to belong to \( j \)-th part of an FD if \( k \in (k_{j}^{\text{min}}, k_{j}^{\text{max}}) \) holds, where \( k_{j}^{\text{min}} \) denotes the minimum or lower-bound density of \( j \)-th part and \( k_{j}^{\text{max}} \) denotes the maximum or
upper-bound density of \( j \)-th part. We assume that an observed traffic state in \( j \)-th part does not belong to the extrapolation of \((j - 1)\)-th part for the purpose of simplicity.\(^2\)

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 \((q_m(A), k_m(A))\) be probe traffic state (PTS) of probe pair \( m \) defined as

\[
q_m(A) = \frac{d_{m_+}(A)}{|A|}, \quad (5a)
\]

\[
k_m(A) = \frac{t_{m_+}(A)}{|A|}, \quad (5b)
\]

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

\[
v_m(A) = \frac{d_{m_+}(A)}{t_{m_+}(A)}. \quad (6)
\]

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

\[
q_m(A) = \frac{q(A)}{|N(A)|}, \quad (7a)
\]

\(^2\)The detailed meaning of this assumption will be discussed later in Section 3.2.
\[ k_m(A) = \frac{k(A)}{|N(A)|}, \quad (7b) \]
\[ v_m(A) = v(A), \quad (7c) \]

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

Let \(c_m - 1\) denote the number of vehicles between probe pair \(m\), and \(a_m\) 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

\[ c_m = |N(a_m)| \quad (8) \]

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

\[ q_m(a_m) = \frac{Q(c_m k_m(a_m); \theta)}{c_m} \quad (9) \]

is satisfied from Eqs. (4), (7a), and (7b). Hereafter, a PTS of pair \(m\) under the aforementioned conditions (i) and (ii) is simply referred to as \textit{steady PTS}. Equation (9) 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 \textit{probe fundamental diagram (PFD)} of \(m\). Note that the PFD of probe pair \(m\) is also expressed as follows:

\[ q_m(a_m) = Q(k_m(a_m); \theta_m). \quad (10) \]

where \(\theta_m\) denotes a pair-specific parameter vector, because the FD and PFD are similar.

In order to ensure that \(a_m\) easily satisfies the aforementioned conditions (i) and (ii), we specify \(a_m\) as follows:

\[ a_{mi} = \{ (t, x) \mid X(t, m_+) \leq x < X(t, m_-) \quad \text{and} \]
\[ -\phi(t - \tau_i) + X(\tau_i, m_+) \leq x \leq -\phi(t - \tau_i - \Delta t) + X(\tau_i + \Delta t, m_+) \}, \quad (11) \]

where \(i\) denotes an index number, \(\tau_i\) denotes a reference time for \(a_{mi}\), and \(\phi\) and \(\Delta t\) are the given parameters. An arbitrary value is acceptable for \(\tau_i\) and \(\phi\), and an arbitrary non-zero value is acceptable for \(\Delta t\) if the corresponding trajectory \(X(t, n)\) exists. The shape of \(a_m\) is illustrated in Fig. 2. The specification (11) guarantees that if traffic in \(a_m\) is steady, then distance traveled by each vehicle in \(a_m\) is equal to the others, which is \(d_{m_+}(a_{mi})\). This feature ensures that the following discussion is concise without sacrificing significant generality.

An FD is said to be \textit{identifiable} by using a given dataset (a set of traffic states or a set of PTSs) if the value of its parameter vector \(\theta\) 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 \(\theta_m\) 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

Given the aforementioned assumptions, a sufficient condition for the identification of an FD is stated as follows.

**Proposition 1.** An FD with a known functional form is identifiable if \( g_j \) or more different steady traffic state data \((q_i, k_i)\) belonging to \( j \)-th part of the FD are available for all \( j \).

**Proof.** First, let us consider determination of belongingness of traffic state data to \( j \)-th part, when belongingness of data to \( l \)-th part for all \( l < j \) is determined if \( j > 1 \). The shape of \( j \)-th part can be uniquely determined using first \( g_j \) data, as \( Q_j \) is assumed as differential. Once the \( j \)-th part is determined, the belongingness of other data belonging to \( j \)-th part (if any) are also resolved, as \( Q_j \) is assumed as differential, \( Q_l (\forall l \neq j) \) is assumed as not identical to \( Q_j \), and we assumed that a traffic state datum in \((j + 1)\)-th part does not belong to the extrapolation of \( j \)-th part. As a consequence, if \( j < h \), we know that the data whose belongingness is not determined are belonging to \((j + 1)\)-th or later parts; if \( j = h \), the belongingness of all the data is determined.

By executing the above procedure with \( j := 1 \), and subsequently iteratively execute it with \( j := j + 1 \) until \( j = h \), the belongingness of all the data is determined.

Now, the parameters of the function \( Q_j \) can be determined based on the known functional form and the given traffic state data for \( j \)-th part for all \( j \), since \( g_j \) or more different data are available for each \( j \)-th part.

This is directly related to conventional FD estimation by using detector data. Note that, in reality, it is difficult to rigorously assess whether the condition—\( g_j \) or more different steady traffic state data \((q_i, k_i)\) belonging to \( j \)-th part of the FD are available for all \( j \)—is actually satisfied based on an available dataset. However, the condition is likely to be satisfied when abundant data were collected, and it is easy to collect these types of data in today’s data age. In this proof, we used the assumption that a traffic state datum in \((j + 1)\)-th part does not belong to the extrapolation of \( j \)-th
part. This is actually a fairly reasonable assumption, because $Q_j$ is not identical to $Q_{j+1}$ and it is very rare to observe a data belonging an extrapolation of the other parts.

By substituting “FD” with “PFD of a probe pair $m$” and “traffic state $(q_i, k_i)$” with “PTS $(q_{mi}, k_{mi})$” in Proposition 1, the sufficient condition for the identification of a PFD by using probe vehicle data is stated as follows.

**Corollary 1.** An PFD of a probe pair $m$ with a known functional form is identifiable if $g_j$ or more different steady PTS data $(q_{mi}, k_{mi})$ belonging to $j$-th part of PFD are available for all $j$.

Now, we consider the identifiability of an FD when a corresponding PFD is identified.

**Proposition 2.** An FD with a known functional form is identifiable if (i) a corresponding PFD of a probe pair is identified and (ii) the value of the jam density is known.

**Proof.** Since an PFD of a pair, say $m$, is identified, the jam density of the PFD, $\kappa/c_m$, is known. Thus, the value of $c_m$ is determined by comparing $\kappa/c_m$ to the known jam density $\kappa$. Thus, the other FD parameters are determined based on Eq. (9). \qed

In fact, this FD parameter identification from known PFD parameters is very simple—a PFD is similar (in the geometric sense) to an FD with a scale of $1/c_m$ as shown in Eq. (9); thus, the FD parameters are simply obtained once $c_m$ and the PFD parameters are known.

From Corollary 1 and Proposition 2, we establish the following corollary on the identifiability of FD based on probe vehicle data:

**Corollary 2.** An FD with known functional form is identifiable if (i) $g_j$ or more different steady PTS data belonging to $j$-th part of the FD are available for all $j$ for a probe pair and (ii) the value of the jam density is known.

Besides, even if the value of the jam density is not known, some of the FD parameters are determined by following a similar approach.

**Proposition 3.** It is possible to determine parameters of an FD without the number-of-vehicle dimensions if $g_j$ or more different steady PTS data belonging to $j$-th part of the FD are available for all $j$ for a probe pair.

**Proof.** Given this condition, the PFD of a probe pair, say $m$, is identified. Since a PFD is similar (in the geometric sense) to an FD in a flow–density plane, only parameters with number-of-vehicle dimensions differ between the PFD and FD. Thus, parameters of the FD without number-of-vehicle dimensions are the same as those of the PFD. \qed

A typical example of such determinable parameters is the free-flow speed.

In summary, an FD with known functional form is identifiable if probe vehicles traveled through traffic with sufficiently wide variety of regimes (density) and the jam density is known. The first condition is realized when probe vehicle data were collected for a long time. The second condition is realized by several methods as discussed later in Section 3.4. The probe vehicle data determine the shape of an FD while the jam density determines the size of an FD.
3.3. Geometric interpretation with triangular FD case

A detailed explanation is given with a case with the triangular FD (Newell, 1993). A triangular FD is defined as

\[
Q(k; u, w, \kappa) = \begin{cases} 
uk & \text{if } k \leq \frac{uk}{u+w} \\
w(k - k) & \text{otherwise},
\end{cases}
\]  (12a)

where \(u\) denotes the free-flow speed and \(w\) denotes the backward wave speed. Equation (12a) represents a free-flowing regime, whereas Equation (12b) represents a congested regime. In this case, \(\theta\) consists of \(u, w, \text{ and } \kappa\); and \(g = 3, h = 2, g_1 = 1, \text{ and } g_2 = 2\) hold. Thus, Corollary 2 is reduced to

**Corollary 3.** 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.

Similarly, Proposition 3 is reduced to

**Corollary 4.** The free-flow speed and the backward wave speed of a triangular FD is determinable if (i) one or more steady PTS data from the free-flowing regime are available for a probe pair and (ii) two or more different steady PTS data from the congested regime are available for the probe pair.

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

Geometric interpretation of the triangular FD identification is explained using Fig. 3 as follows. Fig. 3 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 PFD are similar (in the geometric sense) with scale of \(1/c_m\) as in Eq. (9). 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 previously, traffic state \(S_m(a_{mi})\) follows the FD if region \(a_{mi}\) is steady. Similarly, PTS \(S_m(a_{mi})\) is equal to \(S_m(a_{mi})/c_m\) and follows the PFD if region \(a_{mi}\) is steady. In Fig. 3, regions \(a_{m1}\) and \(a_{m4}\) are steady; and thus their PTS are on the PFD. Conversely, regions \(a_{m2}\) and \(a_{m3}\) are non-steady; and thus their PTS are not on the PFD. Steadiness of region \(a_{mi}\) 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 \(a_{m5}\) (the gray dot in Fig. 3). In this case, the PFD is uniquely determined by selecting values of \(u, w, \text{ and } c_m\) such that all the points \(S_m(a_{m1}), S(a_{m4}), \text{ and } S(a_{m5})\) 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.
Figure 3: Illustration of triangular FD identification. Top: Probe pair and PTS in time–space diagram. Bottom: FD, PFD, and PTS on flow–density plane.
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 as known prior to the FD identification. 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 when 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 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, for the triangular FD case, 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 stated in Corollary 4. 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 implies that the method in itself is unable to consider a bottleneck. This issue could 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 multi-valued and/or non-continuous FDs such as those involving 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. 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 an estimation algorithm for a triangular FD (Newell, 1993) from actual traffic data using maximum likelihood estimation. 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 is presented in Sections 4.2 and 4.3. The triangular FD is employed owing to its simplicity, good physical interpretation of parameters, and popularity in various applications (e.g., Daganzo, 2006).

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, \( \alpha \) as
illustrated in Fig. 4. Note that $\alpha \equiv \kappa w$ holds, and thus, the characterization does not contradict the assumptions in Section 3.1.

![Figure 4: Triangular FD with its three parameters $u$, $w$, and $\alpha$. Black curve denotes FD, and red curve denotes the corresponding PFD of pair $m$.](image)

4.1. Extraction of near-steady PTSs

Suppose that multiple PTSs $S_m(a_{mi})$ for multiple $m$ and $i$ were calculated from probe vehicle data based on the definition (5) and (11). 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 3 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, $S_m(a_{mi})$ is considered as near-steady if trajectories of pair $m$ in region $a_{mi}$ satisfy a condition based on the coefficient of variation:

$$\frac{\sigma_m(a_{mi})}{\bar{v}_m(a_{mi})} \leq \theta_{\text{steady}}$$

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

In the second method, $\{S_m(a_{mi}) \mid \forall i\}$ of each $m$ is verified as to whether it is likely to satisfy the conditions in Corollary 3. Data from the probe pair $m$ are discarded if $\{S_m(a_{mi}) \mid \forall i\}$ included no $S_m(a_{mi})$ with speed lower than $u_{\text{min}}$ or no $S_m(a_{mi})$ with speed faster than $u_{\text{min}}$, where $u_{\text{min}}$ denotes a given lower bound for free-flow speed.\(^3\) This corresponds to conditions (i) and (ii) in Corollary 3. Additionally, data from probe pair $m$ are discarded if correlation between probe-flow and probe-density in $\{S_m(a_{mi}) \mid \forall i, v(a_{mi}) \leq u_{\text{min}}\}$ exceeds $\theta_{\text{corr}}$, or standard deviation in $\{\bar{v}_m(a_{mi}) \mid \forall i, v(a_{mi}) \leq u_{\text{min}}\}$ is smaller than $\theta_{\text{std}}$, where $\theta_{\text{corr}}$ and $\theta_{\text{std}}$ denotes given thresholds. This corresponds to condition (ii) in Corollary 3.

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 $I_m$.

---

\(^3\)Note that although a lower bound for free-flow speed $u_{\text{min}}$ 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.
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 $u_{\text{min}}$.

Specifically, the backward wave speed $w$ is estimated by performing a linear regression on $(k(a_{mi}), q(a_{mi}))$ for each $i \in I_m$, $m \in M$, and $v_m(a_{mi}) \leq u_{\text{min}}$; and the final estimate is given as the mean of all the regression results. Evidently, condition (ii) in Corollary 3 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 $\alpha_m$ 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 $\sigma_s$ 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

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

where $N$ denotes the probability distribution function of a normal distribution. Note that variables $u$, $\alpha_m$, $s$, and $\sigma_s$ are unknown, while the rest are known.

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

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

where $\pi_s$ 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 parameters$^5 \pi$, $u$, $w$, $\{\alpha_m\}$, $\{\sigma_s\}$ 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

$^4$Notice that it is difficult to determine whether a PTS belongs to free-flowing regime—“it is not definitely congested” does not mean “it is free-flowing.”

$^5$Hereafter, $\{\alpha_m \mid \forall m\}$ and $\{\sigma_s \mid \forall s\}$ are abbreviated as $\{\alpha_m\}$ and $\{\sigma_s\}$, respectively, for the purposes of simplicity.
(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

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

By taking its expectation of logarithm, one obtains

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

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.

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

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

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

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

s.t. $u_{\min} \leq u \leq u_{\max}$, \tag{19b}

$0 \leq \alpha_m, \forall m$, \tag{19c}

$\sum_s \pi_s = 1$, \tag{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

$$\mathcal{L}(q, k|\pi, u, \{\alpha_m\}, \{\sigma_s\}, \{\lambda\}) = \sum_{m} \sum_{i} \sum_{s} p_{mi}(s) (\ln \pi_s + \ln l(m, i, s)) - \lambda_{\pi_s} \left( \sum_s \pi_s - 1 \right)$$

---

6Hereafter, $\sum_{m \in M}, \sum_{i \in I_m}$, and $\sum_{s \in \{F, C\}}$ are abbreviated as $\sum_m, \sum_i$, and $\sum_s$, respectively, for the purposes of simplicity.
\[-\lambda_{u_{\text{min}}}(u_{\text{min}} - u) - \lambda_{u_{\text{max}}}(u - u_{\text{max}}) \]
\[+ \sum_{m} \lambda_{\alpha_{m} \text{min}} \alpha_{m} \]

(20)

where each \( \lambda \) 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:

\[u = \begin{cases} 
\tilde{u} & \text{if } u_{\text{min}} < \tilde{u} < u_{\text{max}} \\
u_{\text{min}} & \text{if } \tilde{u} \leq u_{\text{min}} \\
u_{\text{max}} & \text{if } \tilde{u} \geq u_{\text{min}}, 
\end{cases}\]

(21)

with

\[\tilde{u} = \frac{\sum_{m} \sum_{i} p_{mi}(F) q_{mi} k_{mi} / r_{m}^{2}}{\sum_{m} \sum_{i} p_{mi}(F) k_{mi}^{2} / r_{m}^{2}},\]

(22)

and

\[\alpha_{m} = \begin{cases} 
\tilde{\alpha}_{m} & \text{if } 0 < \tilde{\alpha}_{m} \\
0 & \text{if } \tilde{\alpha}_{m} \leq 0, 
\end{cases} \quad \forall m\]

(23)

with

\[\tilde{\alpha}_{m} = \frac{\sum_{i} p_{mi}(C) (q_{mi} + wk_{mi})}{\sum_{i} p_{mi}(C)}, \quad \forall m\]

(24)

and

\[\sigma_{F}^{2} = \frac{\sum_{m} \sum_{i} (p_{mi}(F) (q_{mi} - uk_{mi})^2 / r_{m}^{2})}{\sum_{m} \sum_{i} p_{mi}(F)},\]

(25)

\[\sigma_{C}^{2} = \frac{\sum_{m} \sum_{i} (p_{mi}(C) (q_{mi} - \alpha_{m} + wk_{mi})^2 / r_{m}^{2})}{\sum_{m} \sum_{i} p_{mi}(C)},\]

(26)

\[\pi_{s} = \frac{\sum_{m} \sum_{i} p_{mi}(s)}{\sum_{r \in \{F, C\}} \sum_{m} \sum_{i} p_{mi}(r)}, \quad \forall s \in \{F, C\}.\]

(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 3.

4.4. Summary of the algorithm

Input for the proposed algorithm is as follows:

- Probe vehicle trajectory data
Jam density: $\kappa$
Lower and upper bounds for free-flow speed: $u_{\text{min}}$ and $u_{\text{max}}$
Threshold for near-steady state $\theta_{\text{steady}}$
Cleaning parameter for PTSs based on Corollary 3: $\theta_{\text{corr}}$, $\theta_{\text{std}}$.
Parameter sets determining shape of time–space region $\alpha$: $\Delta m$, $\Delta t$, and $\phi$
Initial state of the EM algorithm: $p_{mi}(s)$
Threshold for the convergence of the EM algorithm: $\varepsilon$

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: $\sigma_{F}$
Standard deviation of congested flow from mean FD: $\sigma_{C}$
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 $m_{+} \in P$, $m_{-} = m_{-} - \Delta m$, $\Delta m \in \Delta M$, $t \in T$, $\Delta t \in \Delta T$, and $\phi \in \Phi$ based on the given probe vehicle data and definition (5) and (11), where $\Delta M$, $\Delta T$, and $\Phi$ 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 $\theta_{\text{steady}}$ and non-appropriate PTSs that do not satisfy Corollary 3 based on given $\theta_{\text{corr}}$ and $\theta_{\text{std}}$.

**Step 2** Backward wave speed estimation step

**Step 2.1** Estimate the backward wave speed $w$ by performing a linear regression on PTSs with $v \leq u_{\text{min}}$ 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 $\pi_{s}, u, \{\alpha_{m}\}, \{\sigma_{s}\}$ 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 $\varepsilon$, 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 an FD to a certain extent. It is possible to develop estimation algorithms for other FD functional forms by using similar methods.
Although the method requires a few input parameters, their values are determined easily (i.e., no need of fine tuning) and the estimation result would be insensitive to most parameters with the exception of $\theta_{\text{steady}}$ and $\kappa$. The parameter $\theta_{\text{steady}}$, 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 $\Delta m$, $\Delta t$, $\phi$, $\theta_{\text{corr}}$, and $\theta_{\text{std}}$ 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 3 are satisfied. It is not necessary for the upper and lower bounds for free-flow speed, $u_{\text{max}}$ and $u_{\text{min}}$, to be close to the actual free-flow speed. The upper bound $u_{\text{max}}$ 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 $u_{\text{min}}$ 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).

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.

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 (specifically, 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 and corresponded to a total of 61 days.

As probe vehicle data, we used probe vehicle data provided by Fujitsu Traffic & Road Data Service Limited. The data were collected by connected vehicles consisting of logistic trucks and vans who 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 visualized in Fig. 5 as time–space diagrams; and it indicate that some probe vehicles slowed down due to congestion approximately 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. 5); therefore, headway time in congested traffic should exceed the value.

![Figure 5: Probe vehicle trajectories on a day.](image)

In order to validate the estimation results, loop detector data provided 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. 6. A clear triangular shape is observed. 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 unexperienced 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. 7. 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 \( a \) 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, \( \theta_{\text{steady}} \), 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 PTs 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 highway with 100 km/h speed regulation. The jam density $\kappa$ was set to 100 veh/km based on Fig. 6. The initial state of the EM algorithm is $p_{mi}(F) = 1$ and $p_{mi}(C) = 0$ for $m$ and $i$ with $v_m(a_{mi}) \leq u_{\text{min}}$, and $p_{mi}(F) = 0$ and $p_{mi}(C) = 1$ for $m$ and $i$ with $v_m(a_{mi}) > u_{\text{min}}$. The convergence check parameter of the EM algorithm, $\varepsilon$, 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 PTs that appeared to satisfy the conditions in Corollary 3, and the total number of PTs 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 Tab. 1 along with other parameter settings.

The estimation result with $\theta_{\text{steady}} = 0.15$ is illustrated in Fig. 8. The solid red line denotes mean estimates for $u$ and $w$, the dashed lines represent the mean plus/minus the standard deviations $\sigma_F$ and $\sigma_C$ of the estimates, and the blue dots represent the detector data for reference. The mean estimate indicates good agreement with detector data. Additionally, the ratio of detector measurements that fall into the range between mean plus/minus standard deviation of estimates was 57.2%. The estimated free-flow speed appears slightly slower than that of detector data; and this is potentially due to the biased speed of probe vehicles (Fig. 7).

All the estimation results with different $\theta_{\text{steady}}$ are summarized in Tab. 1. Based on these results, sensitivity on $\theta_{\text{steady}}$ 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 PTs filtered by $\theta_{\text{steady}} = 0.3$; as a result, many of PTs covered both the free-flowing and congested regime that are not appropriate for the FD estimation. The case with $\theta_{\text{steady}} = 0.05$ resulted in slightly low $w$; and this potentially suggests that the estimates are slightly unstable if the number of PTs are small. With respect to the estimated $\sigma_s$, a clear tendency was obtained: namely, they tended to decrease when the $\theta_{\text{steady}}$ decreased. This is an expected result since low $\theta_{\text{steady}}$ tends to eliminate scattered PTs in the flow–density plane that typically exhibits relatively low steadiness.

| Parameter $\theta_{\text{steady}}$ | # of near-steady data | Estimation results |
|-----------------------------------|-----------------------|--------------------|
|                                   | Probe pairs | PTs | $u$ (km/h) | $w$ (km/h) | $\sigma_F$ (veh/h) | $\sigma_C$ (veh/h) |
| 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         |

Table 1: Summary of estimation results.
5.4. Discussion

The results indicated 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 $\theta_{\text{steady}}$. According to Tab. 1, the method derives fairly good estimates when $\theta_{\text{steady}}$ is less than 0.2, which is a wide range for this type of variable (i.e., coefficient of variation). Conversely, as qualitatively expected, a case with excessively high $\theta_{\text{steady}}$ results in improper estimates because it did not properly remove non-steady traffic data.

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. When compared to the detector data, the cases with $\theta_{\text{steady}}$ in $[0.1, 0.2]$ derived standard deviations similar to those obtained in detector data.

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 (as long as probe vehicles were randomly sampled). 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 the conditions (i) and (ii) of Corollary 3 (experience several traffic states) 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. 5.

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 $\theta_{\text{steady}}$ 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 (e.g., free-flow speed and backward wave speed in the case with 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.

Two future research directions are especially considerable. The first involves the development of 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 $c_m$. The second involves the incorporation 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, submitted).

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.

| Abbreviation | Description                        |
|--------------|------------------------------------|
| 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
\( d_n(A) \) Distance traveled by vehicle \( n \) in time–space region \( A \)
\( t_n(A) \) Time spent by vehicle \( n \) in region \( A \)
\( S \) Traffic state, \( S \equiv (q, k) \)
\( Q \) Density-to-flow FD function
\( Q_j \) Function of \( j \)-th piecewise differentiable part of an FD
\( g \) Number of parameters of an FD
\( g_j \) Number of parameters of \( j \)-th piecewise differentiable part of an FD
\( h \) Number of piecewise differentiable parts of an FD
\( \kappa \) Jam density
\( u \) Free-flow speed of a triangular FD
\( w \) Backward wave speed of a triangular FD
\( \alpha \) \( y \)-intercept of the congestion part of flow–density triangular FD. \( \alpha \equiv \frac{wk}{cm} \)
\( m \) Probe pair (i.e., pair of two probe vehicles)
\( m_- \) Preceding probe vehicle in pair \( m \)
\( m_+ \) Succeeding probe vehicle in pair \( m \)
\( cm \) Number of vehicles between probe vehicles in pair \( m \) plus one
\( \phi \) A parameter (angle) used to define a shape of region \( a \)
\( \Delta t \) A parameter (width) used to define a shape of region \( a \)
\( \alpha_m \) \( y \)-intercept of the congestion part of flow–density triangular PFD of pair \( m \). \( \alpha_m \equiv \frac{wk}{cm} \)
\( \theta_{\text{steady}} \) Threshold for near-steady traffic detection
\( i \) Index for datum
\( a_{mi} \) \( i \)-th time–space region between probe vehicles in pair \( m \)
\( q_m(a_{mi}) \) Probe-flow of pair \( m \) in region \( a_{mi} \)
\( k_m(a_{mi}) \) Probe-density of pair \( m \) in region \( a_{mi} \)
\( v_m(a_{mi}) \) Probe-speed of pair \( m \) in region \( a_{mi} \)
\( S_m(a_{mi}) \) PTS of pair \( m \) in region \( a_{mi} \). \( S_m(a_{mi}) \equiv (q_m(a_{mi}), k_m(a_{mi})) \)
\( q_{mi} \) Probe-flow of datum \( i \) of pair \( m \). \( q_{mi} \equiv q_m(a_{mi}) \)
\( k_{mi} \) Probe-density of datum \( i \) of pair \( m \). \( k_{mi} \equiv k_m(a_{mi}) \)
\( S_{mi} \) PTS of datum \( i \) of pair \( m \). \( S_{mi} \equiv (q_{mi}, k_{mi}) = S_m(a_{mi}) \)
\( s \) Variable representing regime of a traffic state. \( s \in \{ F, C \} \)
\( F \) Free-flowing regime
\( C \) Congested regime
\( \pi_s \) 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 \( z_{mis} \)

References

Bishop, C. M., 2006. Pattern Recognition and Machine Learning. Springer-Verlag New York, Ch. Mixture Models and EM, pp. 423–455.
Canepa, E. S., Claudel, C. G., 2017. Networked traffic state estimation involving mixed fixed-mobile sensor data using Hamilton-Jacobi equations. Transportation Research Part B: Methodological.
Cassidy, M. J., 1998. Bivariate relations in nearly stationary highway traffic. Transportation Research Part B: Methodological 32 (1), 49–59.
Cassidy, M. J., Coifman, B., 1997. Relation among average speed, flow, and density and analogous relation between density and occupancy. Transportation Research Record: Journal of the Transportation Research Board (1591), 1–6.
Chiabaut, N., Buisson, C., Leclercq, L., 2009. Fundamental diagram estimation through passing rate measurements in congestion. IEEE Transactions on Intelligent Transportation Systems 10 (2), 355–359.

Coifman, B., 2014. Revisiting the empirical fundamental relationship. Transportation Research Part B: Methodological 68, 173–184.

Coifman, B., 2015. Empirical flow-density and speed-spacing relationships: Evidence of vehicle length dependency. Transportation Research Part B: Methodological 78, 54–65.

Daganzo, C. F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological 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. Networks and Heterogeneous Media 1 (4), 601–619.

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 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. Transportation Research Record: Journal of the Transportation Research 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. In: Highway Research Board Proceedings. Vol. 14. pp. 448–477.

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. Transportation Research Part C: Emerging Technologies 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. Transportation Research Part B: Methodological 68, 205–223.

Jin, P. J., Li, J., Ran, B., 2015. The identification and empirical characterization of vehicular (Lagrangian) fundamental diagrams in multilane traffic flow. Journal of Advanced Transportation 49 (7), 769–792.

Jin, W.-L., Laval, J., 2018. Bounded acceleration traffic flow models: A unified approach. Transportation Research Part B: Methodological 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: IEEE 20th International Conference on Intelligent Transportation Systems.

Knoop, V. L., Daamen, W., 2017. Automatic fitting procedure for the fundamental diagram. Transportmetrica B: Transport Dynamics 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. Mathematical and Physical Sciences 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. Transportation Research Record: Journal of the Transportation Research Board 1855 (1), 136–142.

Nanthawichit, C., Nakatsuji, T., Suzuki, H., 2003. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. Transportation Research Record: Journal of the Transportation Research Board 1855 (1), 49–59.

Newell, G. F., 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B: Methodological 27 (4), 281–287.

Newell, G. F., 2002. A simplified car-following theory: a lower order model. Transportation Research Part B: Methodological 36 (3), 195–205.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y., 2003. Review of road traffic control strategies. Proceedings of the 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. Transportation Research Part B: Methodological 104, 256–271.
Richards, P. I., 1956. Shock waves on the highway. Operations Research 4 (1), 42–51.
Sandau, R., 2010. Status and trends of small satellite missions for earth observation. Acta Astronautica 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., 2017a. Traffic state estimation on highway: A comprehensive survey. Annual Reviews in Control 43, 128–151.
Seo, T., Kusakabe, T., 2018, submitted. Traffic state estimation using small imaging satellites and connected vehicles. Transportation Research Procedia (ISTS and IWTDCS 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: IEEE 18th International Conference on Intelligent Transportation Systems. pp. 824–830.
Seo, T., Kusakabe, T., Asakura, Y., 2017b. Calibration of fundamental diagram using trajectories of probe vehicles: Basic formulation and heuristic algorithm. Transportation Research Procedia 21, 6–17, (Selected paper from ISTS&IWTDCS, Jeju, Korea, July 7–8, 2016).
Shladover, S. E., 2017. Connected and automated vehicle systems: Introduction and overview. Journal of Intelligent Transportation Systems.
Sun, Z., Jin, W.-L., Ritchie, S. G., 2017. Simultaneous estimation of states and parameters in Newell’s simplified kinematic wave model with Eulerian and Lagrangian traffic data. Transportation Research Part B: Methodological 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. Proceedings of Infrastructure Planning 25, 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. Computer-Aided Civil and Infrastructure Engineering 26 (6), 408–419.
Wada, K., Ohata, T., Kobayashi, K., Kuwahara, M., 2015. Traffic measurements on signalized arterials from vehicle trajectories. Interdisciplinary Information Sciences 21 (1), 77–85.
Yan, Q., Sun, Z., Gan, Q., Jin, W.-L., 2018. Automatic identification of near-stationary traffic states based on the PELT changepoint detection. Transportation Research Part B: Methodological 108, 39–54.
Yuan, Y., van Lint, J. W. C., Wilson, R. E., van Wageningen-Kessels, F., Hoogendoorn, S. P., 2012. Real-time Lagrangian traffic state estimator for freeways. IEEE Transactions on Intelligent Transportation Systems 13 (1), 59–70.
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? Transportation Research Part C: Emerging Technologies 3 (4), 193–209.