

This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Stochastic modeling for vehicle platoons: 1, Dynamic grouping behavior and online platoon recognition

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1016/j.trb.2016.07.019>

PUBLISHER

© Elsevier

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at:
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Li, Baibing. 2019. "Stochastic Modeling for Vehicle Platoons: 1, Dynamic Grouping Behavior and Online Platoon Recognition". figshare. <https://hdl.handle.net/2134/22359>.

Stochastic Modeling for Vehicle Platoons (I): Dynamic Grouping Behavior and Online Platoon Recognition

Baibing Li

School of Business & Economics
Loughborough University
Loughborough LE11 3TU, United Kingdom
E-mail: b.li2@lboro.ac.uk

ABSTRACT

A vehicle platoon is a group of vehicles traveling together at approximately the same speed. Traffic platooning is an important phenomenon that can substantially increase the capacity of roads. This two-part paper presents a new approach to stochastic dynamic modeling for vehicle platoons. In part I, we develop a vehicle platoon model with two interconnected components: a Markov regime-switching stochastic process that is used to model the dynamic behavior of platoon-to-platoon transitions, and a state space model that is employed to describe individual vehicles' dynamic movements within each vehicle platoon. On the basis of the developed stochastic dynamic model, we then develop an algorithm for online platoon recognition. The proposed stochastic dynamic model for vehicle platoons also provides a new approach to vehicle speed filtering for traffic with a platoon structure.

Keywords: Dynamic grouping behavior; Markov regime switching process, State space model, Vehicle platoon, Vehicle speed, Vehicle time headway.

1. Introduction

A vehicle platoon is defined to be a group of vehicles traveling together at approximately the same speed. Traffic platooning is an important traffic phenomenon. Tightly spaced platoons will lead to savings in fuel and increased highway capacity. Clearly, when the traffic platooning structure is taken into consideration, the efficiency of traffic management can be enhanced substantially.

The importance of the research on vehicle platoons is reinforced by the rapid development on autonomous driving in the recent decade where considerable research attention has been paid to vehicle platooning formed by a number of automated vehicles that are cooperatively driven. Diakaki et al. (2015) have recently reviewed vehicle automation and communication systems, including platooning systems and cooperative following and merging systems, from a motorway traffic management perspective. Kavathekar & Chen (2011), on the other hand, have provided a survey on vehicle platooning and the relevant technical issues such as inter-vehicle communications, obstacle detection and collision avoidance; See also an interesting case study on vehicle platooning discussed in Bergenhem et al. (2012).

In the literature, the research on traffic platoons can be classified into several broad categories. First, a large body of research on traffic platoon theory focuses on platoon dispersion. In the pioneering model by Lighthill and Whitham (1955), the kinematic wave theory was used to describe the platoon traffic behavior as it travels along a link. Pacey (1956) subsequently proposed a probabilistic model for traffic platoon dispersion, upon which Robertson (1969) developed a recurrence equation to characterize the platoon dispersion phenomena. Robertson's recurrence equation involves a couple of important parameters that need to be estimated in practical applications. In recent years, many researchers, e.g., Yu (2000), Farzaneh & Rakha (2006), and Bie et al. (2012), have investigated the calibration of these platoon dispersion parameters for the traffic platoon dispersion model.

The second strand of research on vehicle platoons considers how to identify vehicle platoons and hence how to take advantages of the platoon structure in traffic management. For example, Gaur & Mirchandani (2001) investigated a method for real-time recognition of vehicle platoons. Chaudhary et

al. (2006) developed an intelligent traffic control system for identifying platoons at isolated traffic signals on signalized arterials. Jiang et al. (2006) investigated platoon-based traffic signal timing for major-minor intersection types. By minimizing the interruptions to vehicle platoon movements, Jiang et al. (2006) developed an algorithm to reduce traffic delays at intersections.

So far in the traffic literature, there is not much research done for stochastic dynamic modeling of vehicle platoons. The purpose of this two-part paper is to develop a stochastic model to describe the dynamic behavior of vehicle platoons and to investigate various platoon characteristics. In part I, we characterize vehicle platoons by both vehicle speeds and vehicle time headways so that the dynamic nature of the platoon-to-platoon transitions and within-platoon movements can be captured. In contrast, the research on vehicle platoons in the existing literature on platoon classification and recognition (e.g. Gaur & Mirchandani, 2001; Jiang et al., 2006) is solely determined by vehicle time headways, and therefore cannot reflect the platoons' dynamic nature. The developed model in this paper also provides a new probabilistic approach to online platoon recognition and online vehicle speed filtering. Then in part II in Li (2016), we will investigate statistical distribution models for some important platoon characteristics.

This paper is structured as follows. In the next section we propose a dynamic model to describe traffic platoons. Statistical inference, including model estimation, platoon recognition and speed filtering, is considered in Section 3. To illustrate the developed stochastic dynamic model for vehicle platoons, a practical example is discussed in Section 4. Finally, concluding remarks are offered in Section 5.

2. A vehicle platoon model

In this section, we develop a stochastic model to describe the dynamics of vehicle platoons. A vehicle platoon in the literature is defined to be a group of vehicles traveling together at approximately the same speed. We therefore use two microscopic traffic variables, i.e. vehicle speed and vehicle time headway, to identify vehicle platoons and to characterize the dynamic behavior of platoon-to-platoon transitions and within-platoon movements.

2.1. Vehicles and platoons

Consider a traffic flow consisting of a number of consecutive vehicles indexed by $n = 1, 2, \dots$, as illustrated in Figure 1. For any vehicle traveling alone, we follow Jiang et al. (2006) and term it as a platoon with size of one. We characterize each individual vehicle n by two microscopic traffic variables, i.e. vehicle speed v_n and vehicle time headway h_n .

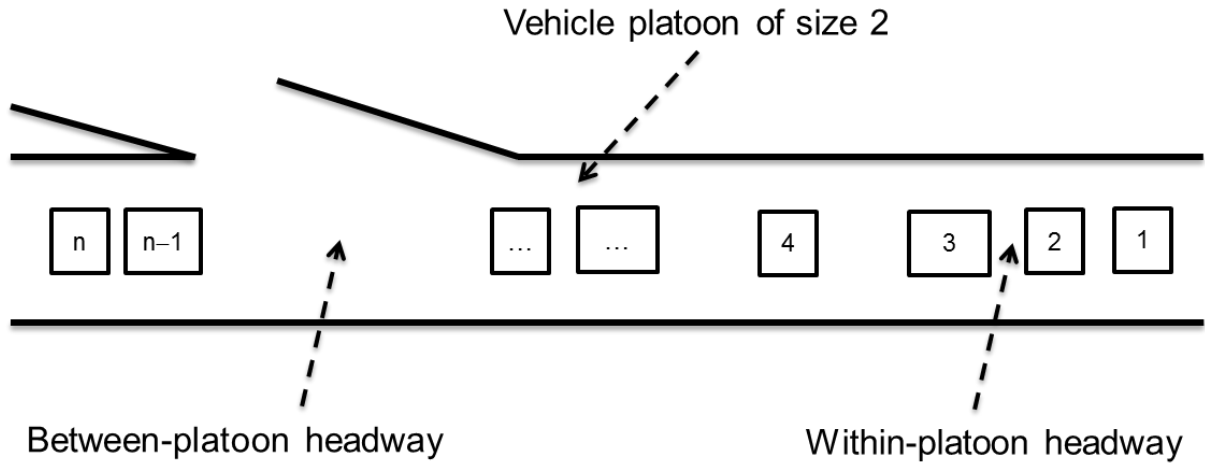


Figure 1. Illustration of vehicle platoons.

We assume that the traffic under investigation involves several velocity modes, indexed by $j \in \mathcal{M}_V = \{1, \dots, M\}$, where each velocity mode j is associated with a mean speed level μ_j and standard deviation σ_j . Without loss of generality, we restrict $\mu_1 < \dots < \mu_M$ for identifiability purposes. We use an indicator S_n to represent the velocity mode that a vehicle n is associated with; the velocity-mode indicator S_n takes a nominal level in \mathcal{M}_V , i.e. $S_n \in \mathcal{M}_V$.

Next, we turn to consider the other microscopic traffic variable, vehicle time headway. In the literature of traffic studies, vehicle time headway is shown to be in one of the two states: a ‘car-following’ mode that is associated with the vehicles following its lead vehicle, and a ‘free-speed’ mode that is associated with the vehicles traveling at a free speed. We define a headway mode

indicator $R_n \in \mathcal{M}_H = \{0,1\}$ to represent the status of the headway of a vehicle n : $R_n = 0$ representing the car-following status and $R_n = 1$ otherwise.

Now we define vehicle platoons. We use two traffic variables, vehicle speed and vehicle time headway, to characterize a vehicle platoon. If two consecutive vehicles belong to the same platoon, they must travel at the same speed level and the temporal gap between them must be small. A natural way to characterize the speed and headway statue of a vehicle n is to use a Cartesian product, $S_n \times R_n = \{(i,j)|i \in \mathcal{M}_V, j \in \mathcal{M}_H\}$: a status of $(j, 0)$ (or $(j, 1)$) indicates that vehicle n is of the car-following (or free-speed) status and is associated with speed level j . This Cartesian product $S_n \times R_n$ can be one-to-one mapped onto a set $\mathcal{M}_P = \{1, \dots, 2M\}$ by defining a vehicle platoon indicator, $G_n = S_n + MR_n \in \mathcal{M}_P$ for vehicle n . Clearly, $G_n = j \leq M$ (or $G_n = j + M$) indicates that vehicle n is of the car-following (or free-speed) status and associated with speed level $j \in \mathcal{M}_V$.

Consider a vehicle platoon $\mathbb{P}_m(j)$, indexed by m and of size L_m , that is associated with a velocity mode j ($j \in \mathcal{M}_V$). Mathematically, $\mathbb{P}_m(j)$ is defined to be a number of consecutive vehicles $m_1 + 1, \dots$, and $m_1 + L_m$ such that the following conditions (C1)-(C3) are met:

- (C1) either $\{G_{m_1} \neq j\} \cap \{G_{m_1} \neq j + M\} \cap \{G_{m_1+1} = j\}$ or $\{G_{m_1+1} = j + M\}$;
- (C2) $G_{m_1+L_m+1} \neq j$;
- (C3) $G_n = j$ for all $m_1 + 2 \leq n \leq m_1 + L_m$.

In other words, platoon $\mathbb{P}_m(j)$ is defined to be

$$\mathbb{P}_m(j) := \{\text{vehicles } m_1 + 1, \dots, \text{ and } m_1 + L_m \mid G_{m_1+1}, \dots, G_{m_1+L_m} \text{ satisfy (C1) - (C3)}\}. \quad (1)$$

Condition (C1) indicates that vehicle $m_1 + 1$ is the platoon leader of $\mathbb{P}_m(j)$, where either: (a) the gap between vehicles m_1 and $m_1 + 1$ is small but vehicles m_1 and $m_1 + 1$ belong to different velocity modes characterized by $\{G_{m_1} \neq j\} \cap \{G_{m_1} \neq j + M\} \cap \{G_{m_1+1} = j\}$; or (b) the temporal gap between vehicles m_1 and $m_1 + 1$ is large, which is characterized by $\{G_{m_1+1} = j + M\}$. On the other hand, condition (C2) indicates that the vehicle $m_1 + L_m + 1$ is the leader of a new platoon that either belongs to a velocity mode that is different from j or the temporal gap between vehicles $m_1 + L_m$ and $m_1 + L_m + 1$ is large. Finally, condition (C3) indicates that any two consecutive vehicles within the platoon must belong to the same velocity mode and the gap between the two vehicles is small.

In practice, both the velocity-mode indicator S_n and headway-mode indicator R_n (and therefore the platoon indicator G_n) are not directly observable (and hence they are latent variables); they can only be inferred using the measurements on vehicle speed v_n and vehicle time headway h_n . In the rest of this section, we develop a stochastic dynamic model, upon which we investigate stochastic recognition of vehicle platoons in Section 3.

Before concluding this sub-section, we note that in the existing literature, e.g., Jiang et al. (2006) and Chaudhary et al. (2006), vehicle platoons are identified solely based on the temporal gaps between vehicles. Clearly this implicitly assumes that vehicles' speeds are homogeneous and it is a special case of equation (1) with $M = 1$. Specifically, in this case we have $S_n \equiv 1$ and $G_n = S_n + MR_n \in \mathcal{M}_P = \{1, 2\}$. Conditions (C1)-(C3) reduce to:

$$(C1a) \quad \{G_{m_1+1} = 2\};$$

$$(C2a) \quad G_{m_1+L_m+1} = 2;$$

$$(C3a) \quad G_n = 1 \quad \text{for all } m_1 + 2 \leq n \leq m_1 + L_m.$$

Conditions (C1a) and (C3a) states that the temporal gaps at the two extreme ends of a platoon are large, whereas (C2a) requires that the temporal gaps within the platoon be small.

2.2. Vehicle time headway

In the traffic literature (e.g., Breiman et al., 1968; Cowan, 1975), there are many existing studies showing that vehicle time headways are approximately statistically independent. Therefore, throughout this paper, we follow the literature and assume that the headways $\{h_n, n = 1, 2, \dots\}$ over the entire traffic stream under investigation are approximately independent of each other. Vehicle time headway is usually modeled by the following two-component mixture distribution of headway in the literature:

$$g(h) = \theta g_0(h) + (1 - \theta)g_1(h), \tag{2}$$

where the density functions $g_0(h)$ and $g_1(h)$ correspond to the car-following and free-speed modes respectively. The parameter $\theta = \Pr\{R = 0\}$ is the prior probability of the car-following component.

Given time headway h_n , it is straightforward to apply Bayes' rule to work out the probabilities that a vehicle n belongs to the car-following mode ($R_n = 0$) and free-speed mode ($R_n = 1$) respectively:

$$r_0(h_n) = \Pr\{R_n = 0|h_n\} = \frac{\theta g_0(h_n)}{\theta g_0(h_n) + (1-\theta)g_1(h_n)}, \quad (3)$$

$$r_1(h_n) = \Pr\{R_n = 1|h_n\} = \frac{(1-\theta)g_1(h_n)}{\theta g_0(h_n) + (1-\theta)g_1(h_n)}.$$

The car-following mode of a vehicle n is solely characterized by its headway h_n (Cowan, 1975; Zhang et al., 2007). Hence, it is reasonable to assume that the car-following mode indicator is independent of the state of its lead vehicle, i.e.

$$\Pr\{R_n = i|R_{n-1}, S_{n-1}, h_n\} = \Pr\{R_n = i|h_n\} = r_i(h_n) \quad \text{for } i \in \mathcal{M}_H. \quad (4)$$

This can equivalently be written as the following Markov transition probability matrix with elements of $\Pr\{R_n = i|R_{n-1} = j, h_n\}$:

$$\mathbf{P}_H = \begin{bmatrix} r_0(h_n) & r_0(h_n) \\ 1 - r_0(h_n) & 1 - r_0(h_n) \end{bmatrix}. \quad (5)$$

In practice, there are several commonly used choices for the component distributions in (2). For example, Griffiths and Hunt (1991) considered a model with two exponentially distributed components $g_0(h)$ and $g_1(h)$. A straightforward extension of the above model is the mixture of two gamma distributions (Cowan, 1975) that includes the exponential components as its special case. In addition, a mixture of lognormal and exponential distribution is also used (Baras et al. 1979). See Zhang et al. (2007) for a further discussion. In this paper, we use the following gamma distributions to describe the headway components:

$$g_i(h) = (h - \tau)^{\alpha-1} \exp(-(h - \tau)/\lambda_i) / [\lambda_i^\alpha \Gamma(\alpha)] \quad (\text{for } h \geq \tau) \quad i \in \mathcal{M}_H, \quad (6)$$

where $\tau > 0$ is the minimum time headway, $\alpha \geq 1$ is the common shape parameter, and λ_i is the scale parameter of the distribution $g_i(h)$ with $\lambda_1 > \lambda_0$. Model (6) ensures that the transition probability in (3) is of a logit form. When $\alpha = 1$, it reduces to Griffiths and Hunt's exponential distributions. Clearly, the above model is more flexible than Griffiths and Hunt's exponential distributions: rather than to pre-specify a value of $\alpha = 1$, the optimal values of α can be estimated

using the collected data in practice. The parameters in equation (6) will be estimated in statistical inference using, e.g. the maximum likelihood method, as outlined in Section 3.

2.3. Velocity modes

In this subsection, we consider the switching process of the velocity modes. We assume that the speeds of two consecutive vehicles are dependent on each other: the speed of a lead vehicle influences the speed of its following vehicle. The inter-dependence of two consecutive vehicle platoons is modeled using a Markov switching process.

Specifically, let $p_{ij} = \Pr\{S_n = i | S_{n-1} = j\}$ denote the probability that vehicle $n - 1$ is associated with velocity mode j but vehicle n switches to mode i ($i, j \in \mathcal{M}_V$). We assume that the transition from a velocity mode j to a mode i is governed by a Markov transition probability matrix:

$$P_V = \begin{bmatrix} p_{11} & \cdots & p_{1M} \\ \vdots & \ddots & \vdots \\ p_{M1} & \cdots & p_{MM} \end{bmatrix}, \quad (7)$$

with $\sum_{i=1}^M p_{ij} = 1$ and $p_{ij} \geq 0$.

The switching process of the velocity modes can potentially be affected by the headway of individual drivers. From a driver perspective, given the average speed level μ_j of the lead vehicle, the driver needs to make a decision if he/she is willing to continue to follow the velocity level of the lead vehicle or change the speed to a different level. This, however, is subject to the feasibility of the intended maneuver characterized by the vehicle time headway. Hence, we assume that the time headway h_n of a vehicle n affects the driver's decision on the choice for velocity levels. Note that h_n is more informative than its corresponding headway mode R_n because the latter is only a binary indicator. Hence, we also assume that given the information on the current headway h_n of vehicle n , the information on the headway modes of vehicle n and its lead vehicle becomes redundant for the driver to make his/her decision. In summary, we make the following assumption:

$$\Pr\{S_n = i | S_{n-1} = j, h_n, R_n, R_{n-1}\} = \Pr\{S_n = i | S_{n-1} = j, h_n\}. \quad (8)$$

We model the discrete choice among the velocity levels by the widely used logit model with an attribute $z_n = \log(h_n - \tau)$:

$$p_{ij}(h_n) = \Pr\{S_n = i | S_{n-1} = j, h_n\} = \frac{\exp(\tilde{a}_{ij} + b_{ij}z_n)}{1 + \sum_{k \neq i} \exp(\tilde{a}_{kj} + b_{kj}z_n)} = \frac{a_{ij}(h_n - \tau)^{b_{ij}}}{1 + \sum_{k \neq i} a_{kj}(h_n - \tau)^{b_{kj}}}, \quad (9)$$

with two parameters $a_{kj} = \exp(\tilde{a}_{kj}) > 0$ and $b_{kj} \geq 0$ for $k \neq j$. We set $\tilde{a}_{ii} = 0$ and $b_{ii} = 0$ for identifiability purposes throughout the paper. The log-transformation for the attribute in (9), $\log(h_n - \tau)$, is used to ensure that the range of the attribute z_n is on the entire real-line. It also transforms the highly skewed headway distribution to an approximately symmetrical distribution.

2.4. Platoon-to-platoon switching process

On the basis of the models for the headway and velocity, we are ready to derive a Markov switching process model for the platoon indicators. Markov regime-switching processes are a very useful tool to analyze stochastic phenomena that exhibit different characteristics during different time periods. See, for example, Malyshkina et al. (2009), Malyshkina and Mannering (2010), and Xiong et al. (2014) for its applications to vehicle crash injury-severity data where the road-segment heterogeneity was taken into consideration.

Consider a vehicle n with the vehicle platoon indicator $G_n = S_n + MR_n \in \mathcal{M}_P$. Let k and l be either 0 or 1. We first note that $G_n = i + kM$ if and only if $S_n = i$ and $R_n = k$; Likewise, $G_{n-1} = j + lM$ if and only if $S_{n-1} = j$ and $R_{n-1} = l$. Hence,

$$\begin{aligned} \Pr\{G_n = i + kM | G_{n-1} = j + lM, h_n\} &= \Pr\{S_n = i, R_n = k | S_{n-1} = j, R_{n-1} = l, h_n\} \\ &= \Pr\{S_n = i | R_n = k, S_{n-1} = j, R_{n-1} = l, h_n\} \Pr\{R_n = k | S_{n-1} = j, R_{n-1} = l, h_n\}. \end{aligned}$$

From equation (8), the probability $\Pr\{S_n = i | R_n = k, S_{n-1} = j, R_{n-1} = l, h_n\}$ collapses to $\Pr\{S_n = i | S_{n-1} = j, h_n\} = p_{ij}(h_n)$. In addition, from equation (4), the probability $\Pr\{R_n = k | S_{n-1} = j, R_{n-1} = l, h_n\}$ reduces to $r_k(h_n)$. Therefore for each vehicle n and any i and $j \in \mathcal{M}_V$, we obtain

$$\begin{aligned} q_{ij}(h_n) &:= \Pr\{G_n = i | G_{n-1} = j, h_n\} = r_0(h_n)p_{ij}(h_n), \\ q_{i(j+M)}(h_n) &:= \Pr\{G_n = i | G_{n-1} = j + M, h_n\} = r_0(h_n)p_{ij}(h_n), \\ q_{(i+M)j}(h_n) &:= \Pr\{G_n = i + M | G_{n-1} = j, h_n\} = [1 - r_0(h_n)]p_{ij}(h_n), \\ q_{(i+M)(j+M)}(h_n) &:= \Pr\{G_n = i + M | G_{n-1} = j + M, h_n\} = [1 - r_0(h_n)]p_{ij}(h_n). \end{aligned}$$

Hence we have

$$q_{ij}(h_n) = q_{i(j+M)}(h_n) \quad \text{and} \quad q_{(i+M)j}(h_n) = q_{(i+M)(j+M)}(h_n) \quad \text{for any } i, j \in \mathcal{M}_V.$$

Now we can write out the Markov transition matrix for the platoon indicator G_n :

$$Q = \begin{bmatrix} q_{11} & \cdots & q_{1 \times (2M)} \\ \vdots & \ddots & \vdots \\ q_{(2M) \times 1} & \cdots & q_{(2M) \times (2M)} \end{bmatrix} = \begin{bmatrix} r_0 P_V & r_0 P_V \\ (1-r_0)P_V & (1-r_0)P_V \end{bmatrix} = P_H \otimes P_V, \quad (10)$$

where \otimes denotes the Kronecker product of two matrices P_H and P_V given by equations (5) and (7).

Clearly, if there is only one velocity level, i.e. $M = 1$, then the transition matrix (10) reduces to transition matrix (5). Correspondingly, vehicle platoons are defined by conditions (C1a)-(C3a). Dunne et al. (1968) in their pioneering study investigated a case where platoons were defined solely based on time headways without taking into account vehicle speed transitions.

The special structure on the transition matrix Q in equation (10) can be generalized. Without assuming equations (4) and (8), we can consider that the transition probability q_{ij} has the following form of semi-parametric logit function of headway h_n (Li, 2011):

$$q_{ij}(h_n) := \Pr\{G_n = i | G_{n-1} = j, h_n\} = \frac{H(\tilde{a}_{ij} + b_{ij}z_n)}{1 + \sum_{k \neq i} H(\tilde{a}_{kj} + b_{kj}z_n)} \quad \text{for any } i, j \in \mathcal{M}_P,$$

where function $H(\cdot) > 0$ is a pre-specified function; alternatively the function $H(\cdot)$ can be left unspecified and hence it is to be estimate in statistical inference. \tilde{a}_{ij} and b_{ij} are state-specific coefficients that are related to the transition from state j to i of the platoon indicator.

2.5. Within-platoon movements

To complete the stochastic dynamic modeling for vehicle platoons, we finally consider the within-platoon movements of vehicles. Recall that vehicles within platoon $\mathbb{P}_m(j)$ have theoretical mean speed μ_j of its velocity mode. In practice, however, vehicles within the same vehicle platoon usually do not travel at an identical speed level. We therefore assume that each driver adjusts his/her speed by a speed drift w_n :

$$v_n = \mu_j + w_n + \sigma_0 \varepsilon_n \quad \text{with} \quad \varepsilon_n \sim N(0,1), \quad (12a)$$

where car n is a member of a platoon $\mathbb{P}_m(j)$ with mean speed μ_j . σ_0 represents the standard deviation of the error term ε_n . The error terms ε_n and ε_m for $n \neq m$ are assumed to be mutually independent.

To characterize the dynamic behavior of within-platoon movements, we assume that the speed drift w_n follows an auto-regressive AP(p) model:

$$w_n = \sum_{k=1}^p \gamma_k w_{n-k} + \sigma_j e_n, \quad \text{with} \quad e_n \sim N(0,1) \quad (13a)$$

where the standard deviation σ_j characterizes the magnitude of the adjustment that is not from its lead vehicles' speeds. γ_k ($k = 1, \dots, p$) are coefficients. In practice, the order p of the AR process is usually small because drivers adjust their speed drift only based on a few cars in front of them. The noise terms e_n and e_m for $n \neq m$ are assumed to be mutually independent.

In the literature, Dailey (1999) used an AR(2) model to describe vehicle speeds. On the other hand, Li (2009) used a random walk model to describe the evolution of vehicle speeds. For speed drifts, however, a stationary autoregressive process that ensures effectively drifts back to the mean is a more reasonable choice. Technically, this requires that the above AR(p) process be stationary.

The above AR(p) model (13a) can be written as a state space model below:

$$W_n = AW_{n-1} + \sigma_j B e_n \quad \text{with} \quad e_n \sim N(0,1) \quad (13b)$$

where

$$A = \begin{bmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_p \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$B = [1, 0, \dots, 0]^T$, and $W_n = [w_n, w_{n-1}, \dots, w_{n-p+1}]^T$.

Equation (12a) can also be written as

$$v_n = \mu_j + CW_n + \sigma_0 \varepsilon_n \quad \text{with} \quad \varepsilon_n \sim N(0,1), \quad (12b)$$

with $C = [1, 0, \dots, 0]$. Equations (12b) and (13b) form a state space model with observation equation (12b) and state equation (13b) for all vehicles traveling within a platoon.

Combining (12b) and (13b), we can see that the state space model in this subsection stipulates that vehicle speeds follow a normal distribution within each platoon. This is consistent with many existing studies (see, e.g. Salter, 1989; May, 1990). For traffic flow across different vehicle platoons, however, the distribution of vehicle speeds is assumed to be a normal mixture in this paper. See Park et al. (2010) for a detailed discussion on mixture normal distributions for vehicle speeds. In contrast, vehicle speeds in the conventional traffic platoon dispersion models (e.g., Pacey, 1956) are usually

assumed to follow a (single) normal distribution. The assumption of normal mixture distributions for vehicle speeds clearly takes speed heterogeneity into consideration and hence it can better reflect traffic platoon dispersion. Note that the normal mixture reduces to a single normal distribution when $M = 1$.

Before concluding this section, we point out that the assumptions made in equations (4) and (8) suggest a conditional independence of the latent variables for given headway h_n . Note, however, under this conditional independence, vehicle speed v_n and vehicle time headway h_n are interdependent on each other. This is because the vehicle speed v_n is modeled by a hidden Markov chain with transition matrix $P_V = [p_{ij}(h_n)]$ in equation (9) that is conditional on headway h_n . In other words, the conditional density function of v_n depends on headway h_n ; see equation (15) for further details.

3. Statistical inference

In this section, we investigate statistical inference for the proposed model. Consider a number of consecutive vehicles, indexed by $n = 1, \dots, N$, and traveling at speed v_n with headway h_n . The stochastic model developed thus far does not address how to measure vehicles' speeds and headways.

In traffic studies, there are two commonly used methods to measure vehicles' speeds and headways. First, the vehicle speeds can be measured contemporaneously from a series of aerial photographs using the distances traveled by the vehicles in the constant time interval between adjacent photos. This method is usually expensive. Alternatively, a more practical approach is to measure the spot speeds and time headways of a line of vehicles passing the same point. In the rest of the section, we follow Chaudhary et al. (2006) and Jiang et al. (2006), and assume that the traffic data is collected using the second method.

Now based on a set of traffic data $\{v_n, h_n\}$ for $n = 1, \dots, N$, we will discuss how to identify the state-space model with Markov switching process. As it will be seen later, once the model is built, online speed filtering \hat{v}_k and online estimation of platoon indicator G_k can be carried out whenever a

new piece of information of traffic data $\{v_k, h_k\}$ becomes available for a future vehicle k , upon which inference for the corresponding platoon $\mathbb{P}_m(j)$ ($j \in \mathcal{M}_V$) can be drawn.

We first define an information set $I_n = \{v_0, h_0, \dots, v_n, h_n\}$ that contains all the information available up for the n th vehicle of interest. Let $\tilde{I}_n = \{I_n, h_{n+1}\}$ denote the information set that the headway of vehicle $n + 1$ is also known. We follow the maximum likelihood method developed by Kim and Nelson (1999, Chapter 5) to draw statistical inference. The likelihood function involves two inter-connected filters, i.e. Hamilton filter and Kalman filter, which will be discussed in details in the next two subsections.

3.1. Hamilton filter

In general, the Hamilton filter is used to deal with discrete variables. The objective of the Hamilton filtering in this section is to draw inference for G_n based on the information set I_n . This is undertaken recursively.

Specifically, suppose that we have worked out the probability of platoon indicator G_{n-1} for vehicle $n - 1$, i.e. $\Pr\{G_{n-1} = i | I_{n-1}\}$. Noting the mutual independence of h_n ($n = 1, 2, \dots$), we can calculate the following joint probability:

$$\begin{aligned} \Pr\{G_n = i, G_{n-1} = j | \tilde{I}_{n-1}\} &= \Pr\{G_n = i | G_{n-1} = j, \tilde{I}_{n-1}\} \Pr\{G_{n-1} = j | \tilde{I}_{n-1}\} \\ &= q_{ij}(h_n) \Pr\{G_{n-1} = j | I_{n-1}\}. \end{aligned} \quad (14)$$

Let $f(v_n, h_n | I_{n-1})$ denote the joint probability density function of (v_n, h_n) conditional on the information set I_{n-1} . In addition, let $f(v_n | I_{n-1}, h_n) = f(v_n | \tilde{I}_{n-1})$ be the conditional density function of v_n when h_n becomes available. From the law of total probability and (14), we obtain

$$\begin{aligned} f(v_n, h_n | I_{n-1}) &= g(h_n) f(v_n | \tilde{I}_{n-1}) \\ &= g(h_n) \sum_{j=1}^{2M} \sum_{i=1}^{2M} f(v_n | \tilde{I}_{n-1}, G_n = i, G_{n-1} = j) q_{ij}(h_n) \Pr\{G_{n-1} = j | I_{n-1}\} \end{aligned} \quad (15)$$

where the calculation of $f(v_n | \tilde{I}_{n-1}, G_n = i, G_{n-1} = j)$ will be discussed in Section 3.2. $g(h)$ is given by equation (2).

When the measurements v_n and h_n for vehicle n become available and $f(v_n|\tilde{I}_{n-1}, G_n = i, G_{n-1} = j)$ is obtained, we can apply Bayes' rule to update the joint probability for vehicle n as follows:

$$\begin{aligned}
\Pr\{G_n = i, G_{n-1} = j|I_n\} &= \Pr\{G_n = i, G_{n-1} = j|v_n, \tilde{I}_{n-1}\} \\
&= \frac{f(v_n|\tilde{I}_{n-1}, G_n=i, G_{n-1}=j) \Pr\{G_n=i, G_{n-1}=j|\tilde{I}_{n-1}\}}{f(v_n|\tilde{I}_{n-1})} \\
&= \frac{f(v_n|\tilde{I}_{n-1}, G_n=i, G_{n-1}=j) \Pr\{G_n=i, G_{n-1}=j|\tilde{I}_{n-1}\}}{\sum_{j=1}^{2M} \sum_{i=1}^{2M} f(v_n|\tilde{I}_{n-1}, G_n=i, G_{n-1}=j) \Pr\{G_n=i, G_{n-1}=j|\tilde{I}_{n-1}\}}.
\end{aligned} \tag{16}$$

Finally, we obtain the probability that the platoon indicator of vehicle n is equal to i :

$$\Pr\{G_n = i|I_n\} = \sum_{j=1}^{2M} \Pr\{G_n = i, G_{n-1} = j|I_n\}. \tag{17}$$

In practice, this probability will be used for platoon recognition. See Section 3.3 for details.

3.2. Kalman filter

To draw inference for G_n in the conditional Hamilton filter in Section 3.1, we need to calculate equation (15) that depends on the inference for individual vehicle speeds. This will be undertaken via the Kalman filter. See Simon (2006) and Li (2013) for a detailed discussion on the Kalman filter.

The Kalman filtering is carried out in a recursive manner. Specifically, for each $G_{n-1} = j$ ($j \in \mathcal{M}_P$), let $W_{n-1|n-1}^{(j)}$ be an estimate of W_{n-1} based on the information set I_{n-1} and let $P_{n-1|n-1}^{(j)}$ be the corresponding covariance matrix. In addition, throughout this section, let $l = i$ if $i \leq M$ and $l = i - M$ if $i > M$.

Suppose that, based on the information I_{n-1} up to vehicle $n-1$, the filtered speed-drift vector $W_{n-1|n-1}^{(j)}$ and its corresponding covariance matrix $P_{n-1|n-1}^{(j)}$ have been obtained. We now consider inference for vehicle n . Consider each $G_n = i$ ($i \in \mathcal{M}_P$). The one-step forecast distribution of the speed-drift vector W_n can be obtained straightforwardly using the well-known Kalman filter:

$$W_n|(\tilde{I}_{n-1}, G_n = i, G_{n-1} = j) \sim N\left(W_{n|n-1}^{(j,i)}, P_{n|n-1}^{(j,i)}\right),$$

where $W_{n|n-1}^{(j,i)} = E[W_n|\tilde{I}_{n-1}, G_n = i, G_{n-1} = j] = AW_{n-1|n-1}^{(j)}$, and

$$P_{n|n-1}^{(j,i)} = E[(W_n - W_{n|n-1}^{(j,i)})^2|\tilde{I}_{n-1}, G_n = i, G_{n-1} = j] = AP_{n-1|n-1}^{(j)}A^T + \sigma_l^2 BB^T.$$

Hence, the predictive distribution of the observed speed for vehicle n is

$$f(v_n | \tilde{I}_{n-1}, S_n = i, S_{n-1} = j) = N(\mu_t + CW_{n|n-1}^{(j,i)} P_{n|n-1}^{(j,i)} + \sigma_0^2), \quad (18)$$

which is used in the Hamilton filter, equations (15)-(17).

Applying the Kalman filter (see, e.g. Simon, 2006; Li, 2013), we can obtain the filtered speed-drift vector:

$$W_{n|n}^{(j,i)} = W_{n|n-1}^{(j,i)} + K_{n|n-1}^{(j,i)} (v_n - \mu_t - CW_{n|n-1}^{(j,i)}) \quad (19)$$

with the Kalman gain $K_{n|n-1}^{(j,i)} = P_{n|n-1}^{(j,i)} C^T / (C P_{n|n-1}^{(j,i)} C^T + \sigma_0^2)$ and its covariance matrix

$$P_{n|n}^{(j,i)} = (1 - K_{n|n-1}^{(j,i)} C) P_{n|n-1}^{(j,i)}. \quad (20)$$

Following Kim and Nelson (1999, Chapter 5), the speed-drift vector for mode $G_n = i$ is estimated by

$$W_{n|n}^{(i)} = \frac{\sum_{j=1}^M \Pr\{G_n=i, G_{n-1}=j | I_n\} W_{n|n}^{(j,i)}}{\Pr\{G_n=i | I_n\}} \quad (21)$$

$$\text{and } P_{n|n}^{(i)} = \frac{\sum_{j=1}^M \Pr\{G_n=i, G_{n-1}=j | I_n\} [P_{n|n}^{(j,i)} + (W_{n|n}^{(i)} - W_{n|n}^{(j,i)})^2]}{\Pr\{G_n=i | I_n\}}, \quad (22)$$

where $\Pr\{G_n = i, G_{n-1} = j | I_n\}$ and $\Pr\{G_n = i | I_n\}$ are given by the conditional Hamilton filter, equations (14)-(17). This process continues and the inference for vehicle $n + 1$ can be drawn when the relevant information becomes available.

It can be seen that this is an alternating algorithm for the computation of the likelihood, and the computations of Hamilton and Kalman filters are inter-connected to each other during the computation. Therefore, the statistical inference for each vehicle n is undertaken in two stages. We first update the probabilities of the platoon indicators using Hamilton filter that provides the information on the platoon indicator of each vehicle n . Next, we apply the Kalman filter to deal with the continuous speed-drift vector for a given mode.

3.3. Model estimation, speed filtering and platoon recognition

From a practical perspective, there are three major research objectives to be achieved: (a) estimation of the model parameters; (b) speed filtering; and (c) platoon recognition. We briefly discuss each of them as follows.

3.3.1. Estimation of the model parameters

The estimation of the parameters of the vehicle platoon model forms the basis of online speed filtering and online platoon recognition. The parameters to be estimated include: (a) velocity-mode parameters μ_j and σ_j ; (b) transition probability parameters a_{ij} and b_{ij} in (9); and (c) the parameters of the headway distribution in (2) and (6).

To estimate these model parameters, we first collect a number of measurements of consecutive vehicles, (v_n, h_n) ($n = 1, \dots, N$). The maximum likelihood method developed by Kim and Nelson (1999, Chapter 5) involves the Kalman filter and Hamilton filter. Note that there are $(2M) \times (2M)$ conditional posterior distributions at each iteration step (see, equations (19)-(20)); the number of the mixture components grows exponentially as the number of vehicles increases. To avoid this problem, an approximation is made in Kim and Nelson (1999, Chapter 5) so that these $(2M) \times (2M)$ conditional posterior distributions collapse into $(2M) \times 1$ distributions using equations (21) and (22). Following Kim and Nelson (1999, Chapter 5), the approximate likelihood is $\mathcal{L} = \prod_{n=1}^N f(v_n, h_n | I_{n-1})$, where $f(v_n, h_n | I_{n-1})$ is given by equation (15). Then the maximum likelihood method can be used to estimate all the unknown parameters of the model.

After the model parameters have been estimated, the Kalman filter and Hamilton filter can be used as two stand-alone algorithms for online speed filtering and online platoon recognition. These will be discussed in the next two subsections.

3.3.2. Speed filtering for traffic with a platoon structure

Speed filtering and prediction are important in traffic control and management; they are widely investigated in the literature. For example, Dailey (1999) applied the Kalman filtering method to estimate vehicle speed. Hazelton (2004) investigated speed estimation using a Monte Carlo approach. Li (2009, 2010) proposed two algorithms for the recursive estimation of vehicle speeds using a single loop detector and dual-loop detector respectively. In all these studies, the platoon structure is not taken into consideration during the speed estimation.

For the traffic with a platoon structure, the stochastic dynamic model for vehicle platoons developed in this paper provides a new approach to online estimation of vehicle speed. We outline the method below.

Given the information set I_n , we wish to estimate the speed of vehicle n . From the law of total probability, we have:

$$f(v_n | I_n) = \sum_{i=1}^{2M} \sum_{j=1}^{2M} f(v_n | I_n, G_n = i, G_{n-1} = j) \Pr\{G_n = i, G_{n-1} = j | I_n\}$$

where $f(v_n | I_n, G_n = i, G_{n-1} = j) = N(\mu_l + CW_{n|n}^{(j,i)}, P_{n|n}^{(j,i)})$ with $W_{n|n}^{(j,i)}$ and $P_{n|n}^{(j,i)}$ given by equations (19)-(20). $\Pr\{G_n = i, G_{n-1} = j | I_n\}$ is given by (16). Hence the speed of vehicle n is estimated as $\hat{v}_{n|n} = \sum_{i=1}^{2M} \sum_{j=1}^{2M} [\mu_l + CW_{n|n}^{(j,i)}] \Pr\{G_n = i, G_{n-1} = j | I_n\}$. An algorithm is outlined below.

Algorithm A (speed filtering for traffic with a platoon structure).

Initialization. Set $W_{0|0}^{(j,i)} = 0$, and $P_{0|0}^{(j,i)} = cI$ for all $i, j \in \mathcal{M}_V$, where c is a sufficiently large constant.

Let $l = i$ if $i \leq M$ and $l = i - M$ if $i > M$.

For $n = 1:N$

Computing: $\Pr\{G_n = i, G_{n-1} = j | I_n\}$ using equation (16).

One-step forecasting:

- $W_{n|n-1}^{(j,i)} = AW_{n-1|n-1}^{(j)}$;
- $P_{n|n-1}^{(j,i)} = AP_{n-1|n-1}^{(j)}A^T + \sigma_l^2 BB^T$.

Updating:

- $K_{n|n-1}^{(j,i)} = P_{n|n-1}^{(j,i)}C^T / (CP_{n|n-1}^{(j,i)}C^T + \sigma_0^2)$;
- $P_{n|n}^{(j,i)} = (1 - K_{n|n-1}^{(j,i)}C)P_{n|n-1}^{(j,i)}$;
- $W_{n|n}^{(j,i)} = W_{n|n-1}^{(j,i)} + K_{n|n-1}^{(j,i)}(v_n - \mu_l - CW_{n|n-1}^{(j,i)})$;
- $\hat{v}_{n|n} = \sum_{i=1}^{2M} \sum_{j=1}^{2M} [\mu_l + CW_{n|n}^{(j,i)}] \Pr\{G_n = i, G_{n-1} = j | I_n\}$.

Return $\hat{v}_{n|n}$.

End For.

In practical applications of the Kalman filtering, it is common practice to introduce a tuning parameter to deal with uncertainty in the future time. For this end, additional speed calibration data measured with a high-accuracy device are required. This tuning parameter is chosen in a way such that the filtered speed values are as close as possible to the additional speed calibration data.

3.3.3. Platoon recognition

Vehicle platoons can be recognized using the derived vehicle platoon indicators. Suppose that vehicle measurements (v_n, h_n) are gathered for each vehicle n . Using the conditional Hamilton filter, we can work out the probability of the platoon indicator G_n . The following rule is used for platoon recognition:

$$\text{Assign } \hat{G}_n = \operatorname{argmax}_i \Pr\{G_n = i | I_n\}, \quad (23)$$

i.e. we assign the vehicle platoon indicator \hat{G}_n for vehicle n to be an index i^* such that the corresponding probability $\Pr\{G_n = i^* | I_n\}$ attains its maximum across all the indexes in \mathcal{M}_P . Then we can follow equation (1) for platoon recognition. Specifically, for given \hat{G}_n ($n = 1, \dots, N$) from equation (23), the following algorithm identifies vehicle platoons $\mathbb{P}_m(j)$ for $m = 1, 2 \dots$

Algorithm B (online platoon recognition).

Initialization. Set $m = 0$. $\hat{G}_0 = 0$. $\mathbb{P}_0(j) = \emptyset$ (empty set).

For $n = 1:N$

If $\hat{G}_n \neq \hat{G}_{n-1}$ %start a new platoon due to a large temporal gap or different velocity modes;

- return the previous platoon $\mathbb{P}_m(j)$;
- set $j = \hat{G}_n$ when $\hat{G}_n \leq M$ or set $j = \hat{G}_n - M$ when $\hat{G}_n > M$;
- set $m = m + 1$;
- set $\mathbb{P}_m(j) = \{n\}$.

Elseif $\hat{G}_n > M$ % start a new platoon due to a large temporal gap;

- return the previous platoon $\mathbb{P}_m(j)$;

- set $j = \hat{G}_n - M$;
- set $m = m + 1$;
- set $\mathbb{P}_m(j) = \{n\}$.

Else % add into the current platoon due to the same velocity mode and a small gap;

- set $\mathbb{P}_m(j) = \mathbb{P}_m(j) \cup \{n\}$.

End If

End For.

The output of Algorithm B is a series of platoons $\mathbb{P}_m(j)$ ($m = 1, 2, \dots$). In practice, whenever a new piece of information (v_n, h_n) of vehicle n becomes available, we can use equations (17) and (23) to calculate \hat{G}_n which is further used for platoon recognition by Algorithm B.

4. A case study

In this section, we use a practical example to illustrate the proposed method.

4.1. Traffic data

Traffic measurements on vehicle speed and vehicle time headway were collected at a dual-loop station located in a three-lane freeway near Seattle during a morning off-peak time period on a normal weekday. In the following analysis we focus on the innermost lane.

The collected raw data were initially prescreened using the algorithm developed by Zhang et al. (2006) to remove erroneous records. Table 1 presents summary statistics for the vehicle speed and time headway measurements after the prescreening.

Figure 2 displays a histogram of the speed measurements. It can be seen that the speed distribution is multi-modal, involving two distinctive modes, each centering at approximately 48 mph and 60 mph respectively. In the following analysis, we set $M = 2$ for illustration purposes. In practice, the number of velocity modes M can be selected using an appropriate criterion such as AIC or use a set of validation data.

Table 1. Summary statistics of vehicle speed (mph) and time headway (s) measurements
for a sample of 1057 vehicles.

	mean	S.D.	skewness	minimum	maximum
Measured speed	57.953	5.896	-0.447	41.536	77.369
Time headway	3.496	2.970	2.204	0.490	28.100

The histogram of the time headway measurements is displayed in Figure 3. It can be seen that a majority of the vehicles had a very small headway (below 2 s) and the frequency shown in Figure 3 drops very rapidly as the headway increases. In the following analysis, we used the headway distribution (2) with gamma components (6).

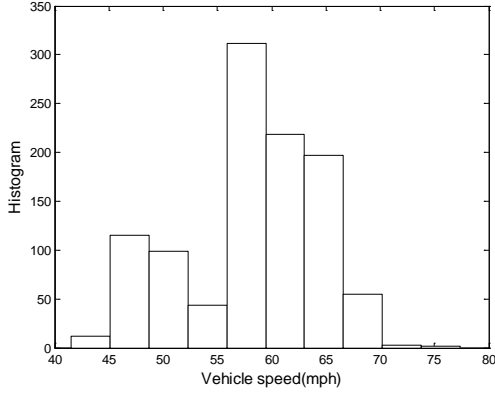


Figure 2. Histogram of speed measurements.

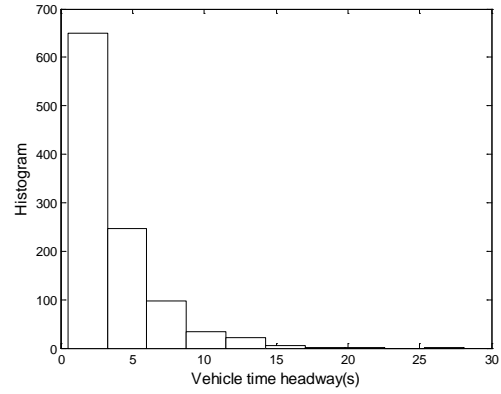


Figure 3. Histogram of vehicle time headways.

4.2. The estimated model

We considered $\mathcal{M}_V = \{1,2\}$ and $\mathcal{M}_H = \{0,1\}$. Note that when $G = S + MR$ is 1 or 3, it corresponds to the lower velocity level, whereas when G is equal to 2 or 4, it corresponds to the higher velocity level. In addition, when G is 1 or 2, the vehicles are of the car-following statue, whereas when G takes 3 or 4, the vehicles are of the free-speed status.

We applied the model developed in the previous sections to analyze the traffic data. The obtained transition matrix for the headway mode (5) is

$$P_H = \begin{bmatrix} \frac{1}{1+0.048\exp(-0.718+1.466(h-0.490))} & \frac{1}{1+0.048\exp(-0.718+1.466(h-0.490))} \\ \frac{0.048\exp(-0.718+1.466(h-0.490))}{1+0.048\exp(-0.718+1.466(h-0.490))} & \frac{0.048\exp(-0.718+1.428(h-0.490))}{1+0.048\exp(-0.718+1.466(h-0.490))} \end{bmatrix}.$$

Figure 4 depicts transition probability $r_0(h) = [1 + 0.048\exp(-0.718 + 1.466(h - 0.490))]^{-1}$. Clearly the transition probability decreases rapidly as headway increases. At approximately $h = 3$ s, the probability reduces to 0.5. Therefore, vehicles with a headway higher than this point were more likely to be of free-speed status. This is in line with the existing literature; for example, Jiang et al. (2006) used $h = 2.5$ s as a cut-off point in their empirical analysis.

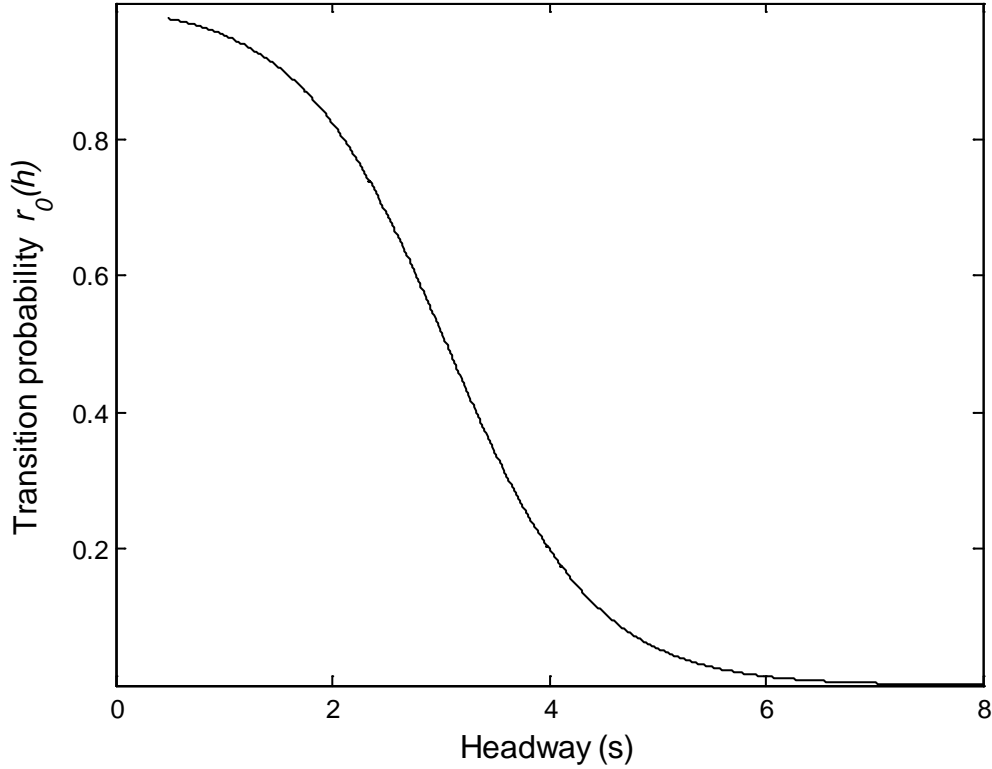


Figure 4. Transition probability $r_0(h)$ against headway h .

Next, we consider the velocity modes. For the two velocity modes, we obtained $\mu_1 = 48.660$ and $\sigma_1 = 2.087$ for lower velocity level, and $\mu_2 = 60.298$ and $\sigma_2 = 3.497$ for the higher velocity level. This indicates that the nominal speed levels of the two velocity modes were 48.7 mph and 60.3 mph respectively. Within each velocity mode, the speed variability was substantial with standard deviations of 2.1 mph and 3.5 mph respectively. The obtained velocity transition matrix (7) is

$$P_V = \begin{bmatrix} \frac{1}{1+4.842(h-0.490)^{0.093}} & \frac{0.279(h-0.490)^{0.061}}{1+0.279(h-0.490)^{0.061}} \\ \frac{4.842(h-0.490)^{0.093}}{1+4.842(h-0.490)^{0.093}} & \frac{1}{1+0.279(h-0.490)^{0.061}} \end{bmatrix}.$$

Figure 5 (left) displays the transition probability $p_{11}(h) = [1 + 4.842(h - 0.490)^{0.093}]^{-1}$ and Figure 5 (right) shows the transition probability $p_{22}(h) = [1 + 0.279(h - 0.490)^{0.061}]^{-1}$. It can be seen that the vehicle platooning behavior was very different at the two velocity levels. At the lower velocity level, the transition probability $p_{11}(h)$ decreases very rapidly. At approximately $h = 0.73$ s, the probability $p_{11}(h)$ reduces to 0.5. This suggests that when the headway of a vehicle was larger than this point, it had a higher likelihood not to follow the lead vehicle and form a platoon at the lower velocity level. In contrast, at the higher velocity level, the transition probability $p_{22}(h)$ decreases much slowly. The probability $p_{22}(h)$ reduces to 0.5 at approximately $h = 3.81$ s. Even at $h = 7.9$ s, for example, the transition probability $p_{22}(h)$ is still about 0.3. Hence, vehicles at the higher velocity level had a much higher likelihood to travel as a group.

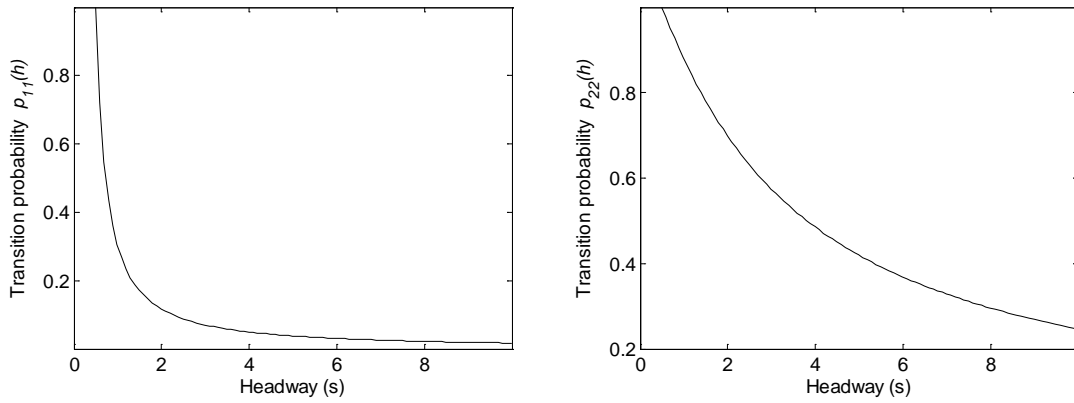


Figure 5. Probability that a vehicle remains at the lower velocity level given its lead vehicle travels at the lower velocity level (left) and probability that a vehicle remains at the higher velocity level given its lead vehicle travels at the higher velocity level (right).

Next, we applied Algorithm B for vehicle platoon recognition. The upper panel of Table 2 shows the classification of the vehicles. It can be seen that, of the total 1057 vehicles under investigation, there were 17.7% of vehicles traveling alone and 3.8% of vehicles traveling in groups at the lower velocity mode; on the other hand, there were 30.2% of vehicles traveling alone and 48.3% of vehicles traveling in groups at the higher velocity mode. Hence, there were over a half of the vehicles traveling in groups of size larger than one. This highlights the importance of research on vehicle platooning.

The lower panel of Table 2 shows that, of the total 712 platoons, there were 26.3% (or 2.7%) of the platoons traveling at the lower velocity mode that had only one vehicle (or more than one vehicle); on the other hand, there were 44.8% (or 26.2%) of platoons traveling at the higher velocity mode that had only one vehicle (or more than one vehicle). This has further reinforced that vehicle platooning behavior is very different at the different velocity levels.

Table 2. Classification of vehicles and platoons.

	Vehicle classification	
	Lower velocity mode	Higher velocity mode
Group size =1	17.7%	30.2%
Group size >1	3.8%	48.3%

	Platoon classification	
	Lower velocity mode	Higher velocity mode
Group size =1	26.3%	44.8%
Group size >1	2.7%	26.2%

Finally, we report the results on vehicle movements within the platoons. We used an AR(2) model in equation (13a). The estimated equation is:

$$w_n = 0.207w_{n-1} + 0.041w_{n-2} + \sigma_j e_n, \quad (24)$$

with $\sigma_1 = 2.087$ and $\sigma_2 = 3.497$. Equation (24) shows that the speed drift of the current vehicle n was mainly determined by the speed of the lead vehicle $n - 1$; the small coefficient 0.041 suggests that vehicle $n - 2$ (i.e. the lead vehicle of vehicle $n - 1$) does not affect vehicle n very much. The stationary process in (24) also shows a strong tendency of mean-reverting of the speed drifts within a platoon. The estimated observation equation is

$$v_n = \mu_n + w_n + 1.104\epsilon_n. \quad (25)$$

The standard deviation $\sigma_0 = 1.104$ in the observation equation (25) is relatively small in comparison with $\sigma_1 = 2.087$ and $\sigma_2 = 3.497$ in equation (24)

4.3. Model assessment

Finally, we consider the issue of model assessment. In the literature, model assessment is usually undertaken based on the comparison between the actual data and the output of a model.

Following this principle, we first compare the obtained model for the headway with the observed headways. Specifically, we split the collected headway data into 11 groups. The observed frequencies O_l ($l=1, \dots, 11$) are displayed in Table 3. The estimated headway distribution is given by

$$g(h) = 0.471g_0(h) + 0.529g_1(h), \quad (26)$$

where $g_0(h) = (h - 0.490)^{1.320} \exp(-(h - 0.490)/0.507)/[0.507^{2.320} \Gamma(2.320)]$ and $g_1(h) = (h - 0.490)^{1.320} \exp(-(h - 0.490)/1.974)/[1.974^{2.320} \Gamma(2.320)]$. On the basis of the distribution (26), we calculated expected frequencies E_l ($l=1, \dots, 11$), as displayed in Table 3.

Table 3. Observed and expected frequencies of vehicle time headway.

	Observed frequency O_l	Expected frequency E_l
$h < 1$	109	101.13
$1 \leq h < 2$	322	327.88
$2 \leq h < 3$	182	190.38
$3 \leq h < 4$	124	115.72
$4 \leq h < 5$	93	85.15
$5 \leq h < 6$	64	65.76
$6 \leq h < 7$	42	49.88
$7 \leq h < 8$	33	36.77
$8 \leq h < 10$	40	45.03
$10 \leq h < 12$	27	21.69
$12 \leq h$	21	17.60
total	1057	1057

Based on Table 3, the test statistic was calculated, $\chi^2 = \sum_{l=1}^{11} (O_l - E_l)^2 / E_l = 6.595$. The critical value with $\alpha = 0.05$ and degrees of freedom $11 - 6 = 5$ is 11.07. Therefore there is no evidence to doubt that the headway data follow the model distribution (26).

Next, we turn to consider the speed data. We assess the goodness-of-fit of the speed filtering. The dynamic model for the speed data is given by equations (24) and (25). We applied Algorithm A for speed filtering. We measured the accuracy of the speed estimates by the root square mean error (RMSE) between the observed speeds and filtered speeds. The calculated RMSE was equal to 1.410 mph. From a practical perspective, this size of the speed estimation error is reasonably small (see, e.g. Hazelton, 2004; Li, 2009).

The observed and filtered vehicle speeds are displayed in Figure 6. Overall, it can be seen that the filtered values smoothed out the random noise to large extent and the filtered speeds were able to follow the speed variation closely.

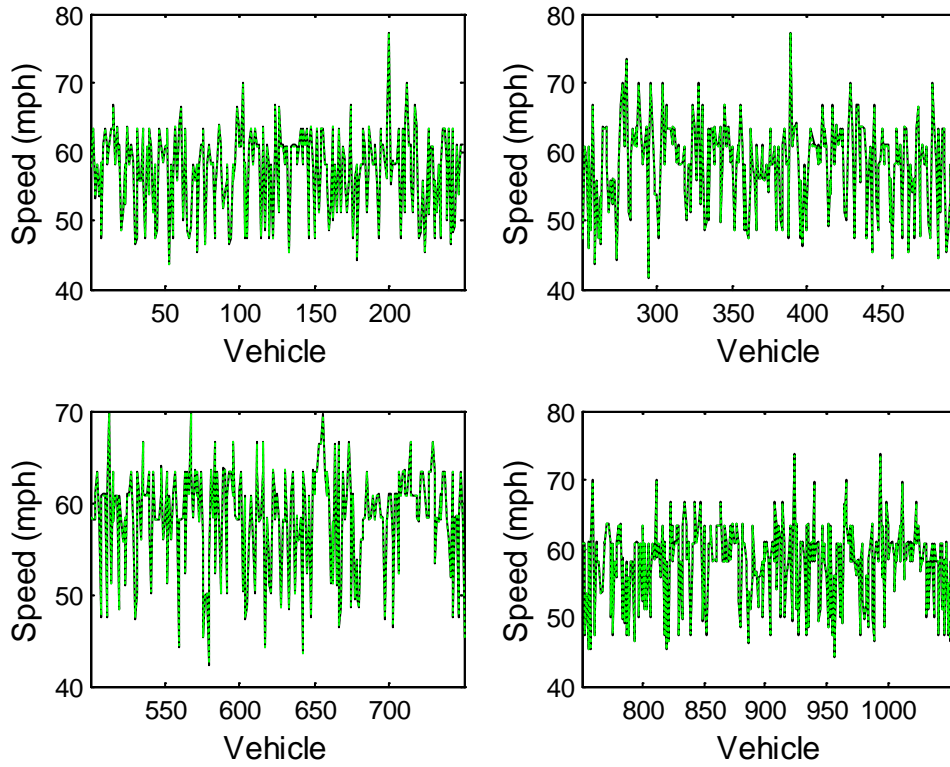


Figure 6. Speed measurements (real line —) and the corresponding filtered values (broken line ----).

The 1057 vehicles in the sample are split into four groups, where the upper left (or right) graph displays the observed and filtered speeds of the first (or second) 250 vehicles, and the lower left (or right) graph displays the observed and filtered speeds of vehicles 501 to 750 (or 751 to 1057).

5. Concluding remarks

This two-part paper investigates stochastic dynamic modeling for vehicle platoons. Part I has focused on the modeling of vehicles' dynamic grouping behavior, where a Markov regime-switching process is used to describe the platoon-to-platoon transitions, and a state space model is employed to characterize the dynamic movements within each platoon.

We have also investigated statistical inference for the proposed model, including model estimation, speed filtering and online platoon recognition. Further investigation on statistical distribution models for some important platoon characteristics will be undertaken in part II (Li, 2016).

Finally, we highlight the practical advantages of the proposed approach here. First, we note that the current platoon classification and recognition methods (e.g. Gaur & Mirchandani, 2001; Jiang et al., 2006) are solely determined by vehicle time headways, whereas the information on vehicle speeds is ignored. To better reflect the dynamic nature of traffic, the proposed method takes into account both vehicle speed and vehicle headway to identify vehicle platoons. In doing so, vehicle platoons can be better recognised in practice. Secondly, this paper has developed a new method for vehicle speed filtering for traffic with a platoon structure. This is a problem that has not been investigated yet in the literature; the existing algorithms for vehicle speed estimation (e.g., Dailey, 1999; Hazalton, 2004; Li, 2009, 2010) cannot deal with the traffic with a platoon structure. With the platooning structure taken into consideration in this paper, vehicle speeds can be estimated more accurately.

Acknowledgements

The author would like to thank the reviewers for their constructive comments that helped improve the quality of the earlier versions of the paper. The author also thanks Dr. Patikhom Cheevarunothai, Professor Yinhai Wang, and Professor Nancy Nihan for kindly providing the traffic data used in the case study.

References

- Baras, J.S., Dorsey, A.J., Levine, W.S., (1979). Estimation of traffic platoon structure from headway statistics. *Automatic Control, IEEE Transactions on* 24(4), 553-559.
- Bergenheim, C., Hedin, E., Skarin, D., (2012). Vehicle-to-vehicle communication for a platooning system. *Procedia-Social and Behavioral Sciences* 48, 1222-1233.
- Bie, Y., Liu, Z., Ma, D., Wang, D., (2012). Calibration of platoon dispersion parameter considering the impact of the number of lanes. *Journal of Transportation Engineering* 139(2), 200-207.

- Breiman L., Gafarian A. V., Lawrence R. L., Murthy V. K., (1968). An experimental analysis of single-lane time headways in freely flowing traffic. 4th Int. Sympos. Traffic. Karlsruhe.
- Chaudhary, N., Abbas, M., Charara, H., (2006). Development and field testing of platoon identification and accommodation system. Transportation Research Record: Journal of the Transportation Research Board 1978, 141-148.
- Cowan, R. J., (1975). Useful headway models. Transportation Research 9(6), 371-375.
- Dailey, D. J., (1999). A statistical algorithm for estimating speed from single loop volume and occupancy measurements. Transportation Research Part B: Methodological 33(5), 313-322.
- Diakaki, C., Papageorgiou, M., Papamichail, I., Nikolos, I., (2015). Overview and analysis of vehicle automation and communication systems from a motorway traffic management perspective. Transportation Research Part A: Policy and Practice 75, 147-165.
- Dunne, M.C., Rothery, R.W., Potts, R.B., (1968). A discrete Markov model of vehicular traffic. Transportation Science 2(3), 233-251.
- Farzaneh, M., Rakha, H., (2006). Procedures for calibrating TRANSYT platoon dispersion model. Journal of Transportation Engineering 132(7), 548-554.
- Gaur, A., Mirchandani, P., (2001). Method for real-time recognition of vehicle platoons. Transportation Research Record: Journal of the Transportation Research Board 1748, 8-17.
- Griffiths, J.D., Hunt, J.G., (1991). Vehicle headways in urban areas. Traffic Engineering and Control 32(10), 458-462.
- Hazelton, M., (2004). Estimating vehicle speed from traffic count and occupancy data. Journal of Data Science 2(3), 231-244.
- Jiang, Y., Li, S., Shamo, D. E., (2006). A platoon-based traffic signal timing algorithm for major-minor intersection types. Transportation Research Part B: Methodological 40(7), 543-562.
- Kavathekar, P., Chen, Y., (2011). Vehicle platooning: A brief survey and categorization. In ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 829-845). American Society of Mechanical Engineers.
- Kim, C.J., Nelson, C.R., (1999). State-space models with regime switching. Cambridge: The MIT Press.

- Li, B., (2009). A non-Gaussian Kalman filter with application to the estimation of vehicular speed. *Technometrics* 51(2), 162-172.
- Li, B., (2010). Bayesian inference for vehicle speed and vehicle length using dual-loop detector data. *Transportation Research Part B: Methodological* 44(1), 108-119.
- Li, B., (2011). The multinomial logit model revisited: A semi-parametric approach in discrete choice analysis. *Transportation Research Part B: Methodological* 45(3), 461-473.
- Li, B., (2013). State estimation with partially observed inputs: A unified Kalman filtering approach. *Automatica* 49(3), 816–820.
- Li, B., (2016). Stochastic modeling for vehicle platoons (II): Statistical Characteristics. Submitted to *Transportation Research Part B: Methodological*.
- Lighthill, M. J., Whitham, G. B., (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 229, No. 1178, pp. 317-345). The Royal Society.
- Malyshkina, N. V., Mannering, F. L., Tarko, A. P., (2009). Markov switching negative binomial models: an application to vehicle accident frequencies. *Accident Analysis & Prevention* 41(2), 217-226.
- Malyshkina, N. V., Mannering, F. L., (2010). Zero-state Markov switching count-data models: an empirical assessment. *Accident Analysis & Prevention* 42(1), 122-130.
- May, A. D., (1990). *Traffic Flow Fundamentals*. New Jersey: Prentice Hall.
- Pacey, G.M., (1956). The Progress of a Bunch of Vehicles released from a Traffic Signal. RRL Research Note RN/2665/GMP, Road Research Laboratory, U.K.
- Park, B. J., Zhang, Y., Lord, D., (2010). Bayesian mixture modeling approach to account for heterogeneity in speed data. *Transportation Research Part B: Methodological* 44(5), 662-673.
- Robertson, D.I., (1969). *TRANSYT - A Traffic Network Study Tool*. RRL Report LR 253, Road Research Laboratory, U.K.
- Salter, R. J., (1989). *Highway Traffic Analysis and Design*. 2nd ed. London: MacMillan.
- Simon, D., (2006). *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. New York: Wiley.

- Xiong, Y., Tobias, J. L., Mannering, F. L., (2014). The analysis of vehicle crash injury-severity data: A Markov switching approach with road-segment heterogeneity. *Transportation Research Part B: Methodological*, 67, 109-128.
- Yu, L., (2000). Calibration of platoon dispersion parameters on the basis of link travel time statistics. *Transportation Research Record: Journal of the Transportation Research Board* 1727, 89-94.
- Zhang, X., Wang, Y., Nihan, N. L., (2006). Robust algorithm for improved dual-loop detection on freeways. *Transportation Research Board Annual Meeting*, Washington DC.
- Zhang, G., Wang, Y., Wei, H. and Chen, Y., (2007). Examining headway distribution models with urban freeway loop event data. *Transportation Research Record: Journal of the Transportation Research Board*.