Dynamic modeling of trip completion rate in urban areas with MFD representations

Paper number: 18-06192

Dynamic modeling of trip completion rate in urban areas with MFD representations

Raphaël Lamotte (corresponding author)
School of Architecture, Civil and Environmental Engineering
Urban Transport Systems Laboratory
Ecole Polytechnique Fédérale de Lausanne (EPFL)
GC C2 389, Station 18, 1015 Lausanne, Switzerland

Mikhail Murashkin
School of Architecture, Civil and Environmental Engineering
Urban Transport Systems Laboratory
Ecole Polytechnique Fédérale de Lausanne (EPFL)
GC C2 384, Station 18, 1015 Lausanne, Switzerland

Anastasios Kouvelas
School of Architecture, Civil and Environmental Engineering
Urban Transport Systems Laboratory
Ecole Polytechnique Fédérale de Lausanne (EPFL)
GC C2 390, Station 18, 1015 Lausanne, Switzerland

Nikolas Geroliminis
School of Architecture, Civil and Environmental Engineering
Urban Transport Systems Laboratory
Ecole Polytechnique Fédérale de Lausanne (EPFL)
GC C2 383, Station 18, 1015 Lausanne, Switzerland

For Presentation Only at the 97th Annual Meeting Transportation Research Board
Washington, D.C. January 2018

Special call by AHB45 committee on "Advancing theory and application of large-scale urban traffic network models"
6344 words + 6 figures = 7844 words

November 16, 2017

ABSTRACT

A large part of the research utilizing the concept of Macroscopic Fundamental Diagram (MFD) relies heavily on an approximation of the trip completion rate derived for steady states. This model, referred to hereafter as the PL model, only requires the knowledge of the MFD, the average trip length and the current accumulation. An alternative approach, known as the trip-based model, allows to determine the trip completion rate exactly (provided that the road network is governed by an MFD) when the trip length distribution and inflow history are known. This work investigates the soundness of the PL approximation under time-varying inflow by comparing it to the more complex trip-based model. The trip length distribution is shown to be an important determinant of the accuracy of the PL model, not only via its mean but also via its coefficient of variation. The PL approximation is exact when trip length follows an exponential distribution, and relatively good when the coefficient of variation is close to 1. Other coefficients of variation lead to the emergence of hysteresis phenomena, whose properties depend on whether the coefficient of variation is smaller or greater than 1. A third type of model (the M model) is proposed to address the cases where the PL model does not provide sufficient accuracy. The M model has a dynamic behavior very similar to the one of the trip-based model, but it is described by an ordinary differential equation, thus being more suitable for control purposes. Despite their differences in accuracy, the PL and M models are found to perform equally well when integrated in a model predictive control framework.

INTRODUCTION

Different methods have been proposed over the last 50 years to describe the average speed and average link flow at the zone level (1, 2, 3). Today, the most popular method relies on the concept of Macroscopic Fundamental Diagram (MFD), also known as Network Fundamental Diagram. These diagrams describe the effects of congestion by relating together two of the three spatially aggregated variables describing the state of road network: average speed, production, and accumulation of vehicles (i.e. the number of vehicles in the zone).

When Daganzo (4) re-introduced these diagrams, he associated them with the concept of Network Exit Function (NEF), thereby highlighting their potential for regional traffic flow management. A NEF is a function expressing the outflow of the zone, i.e. the trip completion rate (from the zone view-point, trips may be completed either by parking inside the zone or by crossing the perimeter). However, because it depends only on some constants and on the current accumulation, the NEF introduced by Daganzo (4) has come to be known as another MFD. To avoid any confusion, we will refer to this specific NEF as the outflow MFD or PL model (this abbreviation refers to its mathematical expression), while the relations between accumulation and speed or production will be referred to as speed MFD and production MFD. An important particularity of the outflow MFD is that even though it is dedicated to study the dynamics of congestion, its formula was derived using a steady-state assumption (4).

More recently, several authors studying the departure time choice problem at the city scale introduced an alternative description of the congestion dynamics, based on the speed MFD but avoiding the steady state approximation (5, 6, 7, 8). This so-called “trip-based" model is computationally more demanding than the PL model and may cause intractability. Yet, it also provides a sounder treatment of propagation phenomena, avoiding some artifacts associated with the PL model, such as the temporary reduction of experienced travel time that can follow a demand surge (9, 10).

These developments naturally raised several questions: what are the fundamental differences between the PL model and the trip-based model? Under which conditions do they perform similarly? Which model is most suitable for control applications? The recent paper by Mariotte et al. (10) analyses some of these issues, but focuses primarily on the consequences of a non-stationary inflow with homogeneous trip length, in line with Daganzo (4). The present paper breaks with this line of research and investigates the impact of trip length heterogeneity.

Under an exponential distribution of trip length (coefficient of variation $\sigma/L = 1$), the trip-based model and the PL model are shown to be equivalent. For coefficients of variations close to 1, the discrepancy between the two models remains quite small, such that the PL model may describe the congestion dynamics more accurately than a trip-based model applied with a distribution whose coefficient of variation would have been poorly estimated. The cases $\sigma/L << 1$ and $\sigma/L >> 1$ are found to have entirely different dynamic properties.

We also introduce a third type of NEF which reproduces extremely well the behavior of the trip-based model at a much lower computational cost. This is achieved by keeping track of the average remaining distance to be traveled. By comparison, the trip-based model keeps track of the distance remaining to be traveled by each individual user, while the PL model does not keep any record of traveled distance. Such a model offers valuable intuition and represents an attractive trade-off for control applications. An example of integration in a Model-Predictive Control (MPC) framework is proposed.

DIFFERENT MODELS

The speed (production) and outflow MFDs

On one hand, the speed and production MFDs of road networks are analogous to the Fundamental Diagram (FD) of road sections. The three network variables equivalent for speed, flow and density are the average speed v [$\mathrm{km\,h^{-1}}$], the production or network flow P [veh $\mathrm{km\,h^{-1}}$], and the accumulation of vehicles inside the zone n [veh]. While the three variables are related by $P(t) \triangleq n(t)v(t)$, the MFD provides another relationship describing the effect of congestion, for instance $P = \mathcal{P}(n)$, or $v = \mathcal{V}(n)$. Although the existence of such a relation should not be considered as universal, it has been observed in multiple cities with relatively little scatter (11, 12, 13, 14, 15, 16, 17), using local measurements from e.g. inductive loop detectors and/or probe data. The importance of scatter has been related to the heterogeneity across links (18) and clustering algorithms have been developed to identify homogeneous regions suitable for MFD analyses (19). The speed MFD $\mathcal{V}$ is generally assumed to be continuous and strictly decreasing from the free-flow speed v_{f} to 0 on an interval $[0, n_{\text{jam}}]$. The corresponding production MFD follows a unimodal curve, reaching its maximum at the so-called critical accumulation.

On the other hand, traffic flow management at the regional level relies on the conservation equation

$$\dot{n}(t) = I(t) - O(t), \tag{1}$$

where $n(t)$ denotes the accumulation of vehicles inside the zone at time t, $\dot{n}$ its time derivative, $O(t)$ the outflow rate and $I(t)$ the inflow rate (note that trips may start either inside the zone, or by crossing the perimeter). While the inflow rate is often exogenous, the outflow rate is estimated via a NEF. The outflow MFD provides a simple expression for this outflow, which can be derived as follows. Consider a system in steady state, where vehicles enter the network at a constant rate I. Let L denote the average trip length of entering users and $M(t)$ the total distance that remains to be traveled by all users in the network at time t. The evolution of M is governed by the following equation:

$$\dot{M}(t) = I(t)L - n(t)v(t). \tag{2}$$

Since $\dot{n} = 0$ and $\dot{M} = 0$ in steady state, $O = I = \frac{nv}{L} = \frac{P}{L}$. Daganzo (4) postulates that this result still holds approximately as long as the production MFD exists and the inflow varies slowly enough, so that $O(t)$ can be approximated as $O_{PL}(t) \triangleq \mathcal{O}(n(t)) \triangleq \mathcal{P}(n(t))/L$ (hence the appellation "PL model"). This assumption was given some support by Geroliminis and Daganzo (11), who observed using both loop detector data and taxi data that the ratio of production over trip completion rate remained approximately constant over their observation period for the city center of Yokohama, Japan. The objective of this work is to further investigate the conditions under which this approximation is reasonable and to propose alternative NEFs more suitable under time-varying conditions.

The trip-based model

The trip-based model (TB) of outflow derives directly from the existence of a speed MFD and vehicle conservation, without requiring the steady state assumption. One way to introduce it starts from the simple observation that a user with trip length l_0 that entered at time t_0 should exit after traveling l_0, i.e. after a delay τ_0 satisfying

$$\int_{t_0}^{t_0+\tau_0} \mathcal{V}(n(u))\,\mathrm{d}u = l_0. \tag{3}$$

Among the users that entered the network at time s, the proportion that is still in the network at the time $t > s$ is given by $1 - F\left(\int_s^t \mathcal{V}(n(u))\,\mathrm{d}u\right)$, where $F(\cdot)$ is the cumulative distribution function (cdf) of trip length corresponding to the trip-generating process (the corresponding pdf is denoted f).[1] Assuming that the flow $I(s)$ entering the zone is known for all times $s < t$ and that $I(s) = 0$ for all times $s < 0$, the accumulation at time t is

$$n(t) = \int_0^t I(s)\left(1 - F\left(\int_s^t \mathcal{V}(n(u))\,\mathrm{d}u\right)\right)\mathrm{d}s. \tag{4}$$

[1] Note that in full generality F might be time dependent. It is assumed constant in this paper as real world observations generally exhibit low variability in time.

By differentiating Eq. (4), we obtain an expression that has the same form as Eq. (1), but where the outflow is described by:

$$O_{\mathrm{TB}}(t) = \mathcal{V}(n(t)) \int_0^t I(s) f\left(\int_s^t \mathcal{V}(n(u))\,\mathrm{d}u\right) \mathrm{d}s. \tag{5}$$

While this expression is in general difficult to solve analytically, it can be easily implemented in an event- and agent-based simulation (see Section 4.2). In two cases however, the TB model turns out to be identical to the PL model. First, since both the TB and PL models are based on the speed MFD, they are trivially equivalent in steady state. This is true regardless of the trip length distribution. Second, both models are equivalent when the trip length follows an exponential distribution (with constant coefficient). Indeed, an exponential distribution is characterized by the fact that for all $l \in \mathbb{R}^+$, $f(l) = (1 - F(l))/L$, where L is the mean of the distribution. By replacing f in Eq. (5) and by combining it with Eq. (4), one obtains that $O_{\mathrm{TB}}(t) = \mathcal{V}(n(t))n(t)/L = O_{\mathrm{PL}}(t)$, regardless of the inflow variations. This result illustrates the well-known "memory-less" property of the exponential distribution.

It is worth mentioning as well that the outflow described by Eq. (5) is always positive but is not bounded above. While the current practice consists in bounding the outflow by the receiving capacity of the neighboring regions (plus some internal capacity corresponding to internal trips), it is our view that that the physical boundary capacity is actually very large (especially for internal trips) and rarely binding, such that it can be ignored altogether. Note however that this statement applies only to the instantaneous outflow. Inflows and outflows exceeding the capacity of the outflow MFD cannot be sustained on the long term.

The M model

The PL and TB models might be considered as two extreme ways of modeling the outflow. While the TB model keeps track of the distance remaining to be traveled by each single user, the PL model does not keep any record of past events. We now propose a trade-off between these two extreme alternatives. The M model summarizes all past events into the average distance remaining to be traveled.

In steady state, the average distance remaining to be traveled is simply given by $L^* = \int_0^{+\infty} g(l) \frac{l}{2}\,\mathrm{d}l$, where g is the pdf of the trip length distribution among all users present in a snapshot. Since users remain in the network for a duration proportional to their trip length, $g(l)$ is proportional to $f(l)l$. Imposing that $\int_0^{+\infty} g(l)\,\mathrm{d}l = 1$ implies that $g(l) = \frac{f(l)l}{L}$. Thus

$$L^* = \int_0^{+\infty} g(l) \frac{l}{2}\,\mathrm{d}l = \int_0^{+\infty} f(l) \frac{l^2}{2L}\,\mathrm{d}l = \frac{L^2 + \sigma^2}{2L}. \tag{6}$$

We propose then to account for variations in the remaining distance to be traveled via the following alternative model:

$$O_{\mathrm{M}}(t) = \frac{n(t)v(t)}{L}\left(1 + \alpha\left(\frac{M(t)}{n(t)L^*} - 1\right)\right) = \frac{n(t) + \alpha\left(\frac{M(t)}{L^*} - n(t)\right)}{L} v(t), \tag{7}$$

where $M(t)$ represents the total remaining distance to be traveled by all users (as in Eq. (2)).

We show first that all three models (TB, PL and M) are equivalent in the steady state or when trip length follows an exponential distribution. Since we already showed the equivalence between the TB and PL models for these two cases, we only need to show that the M model is also equivalent to the PL model. We do so by demonstrating that $M(t) = n(t)L^*$ always holds in these two cases.

In the steady state, $M(t) = n(t)L^*$ holds by definition of L^*. Then, if trip length follows an exponential distribution, $\dot{M}(t) = I(t)L - n(t)v(t) = \left(I - \frac{n(t)v(t)}{L}\right)L = \dot{n}(t)L$. Taking $M(0) = n(0) = 0$, we have $M(t) = n(t)L$. Since for the exponential distribution $L^* = L$, we obtain the desired result.

Let us now consider the general case. If $M(t) > n(t)L^*$, the average distance to be traveled is larger than in the steady state, so one might expect that $O_{\mathrm{TB}}(t) < \frac{n(t)v(t)}{L}$. Similarly, if $M(t) < n(t)L^*$, one might expect that $O_{\mathrm{TB}}(t) > \frac{n(t)v(t)}{L}$. This suggests that the parameter α in Eq. (7) should be negative. In practice, we suggest to use $\alpha = -3$. The choice of this value is motivated by the fact that, if trip length follows a gamma distribution with pdf $f(l) = \frac{4l}{L^2}e^{\frac{-2l}{L}}$, then the M model is equivalent to the TB model.

In order to obtain that result, we first write $M(t)$ as

$$\begin{aligned} M(t) &= \int_0^{+\infty} l^* \int_0^t I(s) f\left(l^* + \int_s^t \mathcal{V}(n(u))\,\mathrm{d}u\right) \mathrm{d}s\,\mathrm{d}l^* \\ &= \int_0^t I(s) \int_0^{+\infty} l^* f\left(l^* + \int_s^t \mathcal{V}(n(u))\,\mathrm{d}u\right) \mathrm{d}l^*\,\mathrm{d}s, \end{aligned} \tag{8}$$

where l^* denotes the distance remaining to be traveled by individual users. Note then that if $f(l)$ satisfies

$$f(l) = \frac{1}{L}\left((1 - F(l)) + \alpha\left(\frac{1}{L^*}\int_0^{+\infty} l^* f(l^* + l)\,\mathrm{d}l^* - (1 - F(l))\right)\right), \tag{9}$$

then by replacing $M(t)$ and $n(t)$ in Eq. (7) by their expressions taken respectively from Eq. (8) and (4), we obtain that $O_{\mathrm{M}}(t) = O_{\mathrm{TB}}(t)$.

We now prove that the gamma distribution $f(l) = \frac{4l}{L^2}e^{\frac{-2l}{L}}$ satisfies equation (9) for $\alpha = -3$. First, note that

$$\begin{aligned} 1 - F(l) &= \int_0^{+\infty} f(l + l^*)\,\mathrm{d}l^* = \int_0^{+\infty} \frac{4(l + l^*)}{L^2} e^{\frac{-2(l+l^*)}{L}}\,\mathrm{d}l^* \\ &= \left(\frac{4l}{L^2}\int_0^{+\infty} e^{\frac{-2l^*}{L}}\,\mathrm{d}l^* + \int_0^{+\infty} \frac{4l^*}{L^2} e^{\frac{-2l^*}{L}}\,\mathrm{d}l^*\right) e^{\frac{-2l}{L}} = \left(\frac{2l}{L} + 1\right) e^{\frac{-2l}{L}}. \end{aligned}$$

Besides, note that

$$\int_0^{+\infty} l^* f(l^* + l)\,\mathrm{d}l^* = \left(l\int_0^{+\infty} \frac{4l^*}{L^2} e^{\frac{-2l^*}{L}}\,\mathrm{d}l^* + \int_0^{+\infty} l^* \frac{4l^*}{L^2} e^{\frac{-2l^*}{L}}\,\mathrm{d}l^*\right) e^{\frac{-2l}{L}} = (l + L)e^{\frac{-2l}{L}}.$$

By combining these two equations with Eq. (9) and by taking into account that for the considered distribution $\frac{\sigma}{L} = \frac{1}{\sqrt{2}}$ and $L^* = \frac{3L}{4}$, Eq. (9) boils down to

$$4l = 2l + L + \alpha\left(\frac{4}{3}(l + L) - (2l + L)\right),$$

which is always true for $\alpha = -3$.

Thus, the M model and the TB model are equivalent (i) in steady state regardless of the trip length distribution, (ii) under time-varying conditions when trip length follows an exponential distribution ($\sigma/L = 1$) and (iii) under time-varying conditions when trip length follows a gamma distribution with $f(l) = \frac{4l}{L^2}e^{\frac{-2l}{L}}$ (in that case $\sigma/L = 1/\sqrt{2}$), provided that $\alpha = -3$. Furthermore, it is shown in the following section that the M model with $\alpha = -3$ produces results that are similar to those of the TB model for other distributions. Thus, setting α to -3 should be considered as a good rule of thumb, even though other values might work better for specific cases.

COMPARISON AND SENSITIVITY ANALYSIS

Description of key factors and metric of accuracy

Due to the lack of real world data, the numerical applications reported hereafter were made with σ/L varying between 0 and 1.2. Larger coefficients of variations are not considered as they would require rather fat tails and large maximum trip length (which would be typically much larger than the zone width). We introduce hereafter two families of distributions allowing us to vary σ continuously while keeping the same average trip length L:

- uniform distributions of the type $U\left[L-\sqrt{3}\sigma, L+\sqrt{3}\sigma\right]$ are used to obtain ratios $\sigma/L \in \left(0, 1/\sqrt{3}\right]$;
- mixtures of uniform distributions of the type $w_1 U[0,L] + w_2 U\left[0, L+\frac{3\sigma^2}{L}\right]$, where $w_1 = 1-\frac{L^2}{3\sigma^2}$, $w_2 = \frac{L^2}{3\sigma^2}$ are used to obtain $\sigma/L > 1/\sqrt{3}$.

The other key factor analyzed hereafter is inflow variability. Its importance as a determinant of the accuracy of the PL model has been known since the introduction of this model (4). Here, we obtain different levels of inflow variability by changing the time-scale of the inflow variations (via the coefficient h in the inflow equation (11)).

In order to quantify the differences in predicted dynamic behavior, we introduce the following metric

$$\xi_{\text{A/B}} = \frac{\int |n_{\text{A}}(t) - n_B(t)| \, \mathrm{d}t}{\int |n_{\text{B}}(t) - n_s| \, \mathrm{d}t}, \tag{10}$$

where n_A and n_B denote the accumulations obtained over time with two different models A and B, B acting as the reference, while n_s denotes the steady state accumulation (corresponding to the inflow $(1-p_{\text{peak}})N/T$ - see Eq. (11)). The bounds of the integral are chosen to exclude any warm-up period and focus on the effects of the perturbation. The metric ξ can be interpreted as the average error in accumulation over time, normalized by the excess in vehicles hours traveled due to the demand peak with the reference model. For the sake of readability, the models being compared are indicated in plain text rather than as subscripts in the remainder of this article.

Implementing the trip-based model in simulations

The system whose evolution is described by (4) has been recognized as analytically intractable with general inflow functions and trip length distributions (5). It can however be solved exactly for the case of discrete users (7, 9, 10). The solving procedure takes as input a speed MFD $\mathcal{V}(n)$ and a population, described by three vectors of length N (which is the number of agents): a first vector contains the departure times of all agents, a second contains the trip lengths and a third contains the weights (this third vector can also be omitted if all agents have the same weight). Then, the algorithm proceeds event by event, in a chronological order, by keeping track of time t and of the cumulative distance traveled by a fictive user since the first departure (which we take as the origin of time): $x(t) = \int_0^t v(u) \, \mathrm{d}u$. An event is the departure or arrival of an agent. When a departure occurs (denote t_{dep} the departure time, l the trip length and w the weight), we increase the current accumulation n by the weight of that agent (w) and add to the list of exits that an agent with weight w should exit when x will be equal to $x(t_{\text{dep}}) + l$. When an arrival occurs, we simply decrease the accumulation n by the weight of the user exiting. Once an event is processed, we identify the next event by comparing the next departure time with the next arrival, given that accumulation and speed remain constant between two events.

Considering discrete users also raises issues related to stochastic processes. These, although very interesting, are beyond the scope of the present paper. To circumvent these issues and approximate the deterministic solution of Eq. (4) for a continuum of users, we used both a large number of agents (2×10^6)

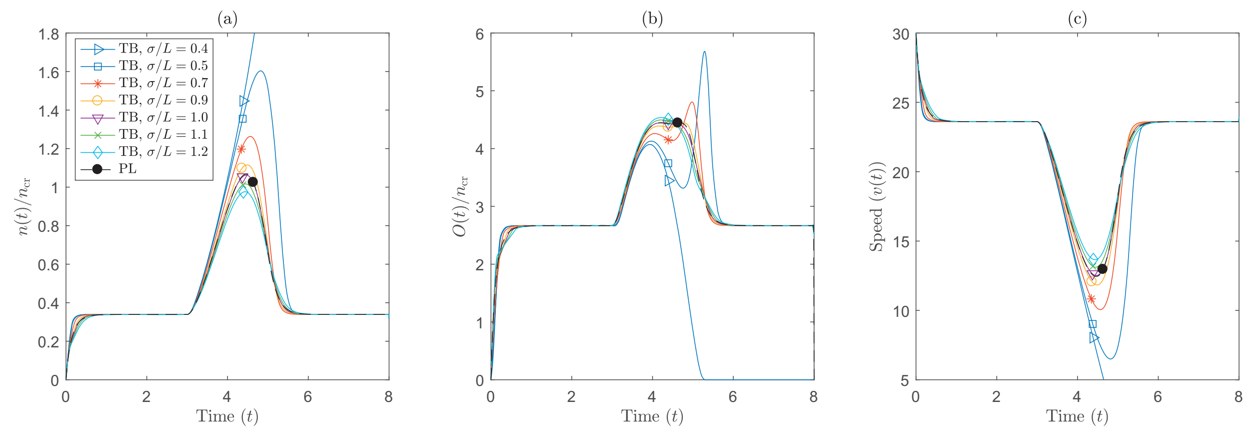

FIGURE 1 Time series of accumulation, outflow and speed with several ratios σ/L. Accumulation and outflow are normalized by the critical outflow.

and an ad hoc discretization scheme. While a continuum of users would usually be best represented by a large number of different trip lengths, generating 2×10^6 well-distributed values of trip length and associating them randomly to departure times would introduce strong temporal variations in the average trip length and standard deviation of the trip generating process. To avoid this phenomenon, we only generated a small number (1000) of representative trip lengths, and associated them with a one-to-one mapping to batches of 1000 users, taken by increasing departure times.

We summarize hereafter the different numbers and functions used for simulations. Given a total demand N (in units of vehicles), the inflow is given for all $t \in [0,T]$ by:

$$I(t) = \begin{cases} (1-p_{\text{peak}})\frac{N}{T} + p_{\text{peak}}\frac{N\pi}{2h}\cos(\frac{\pi}{h}t), & \text{if } t \in [T/2-h/2, T/2+h/2] \\ (1-p_{\text{peak}})\frac{N}{T}, & \text{otherwise.} \end{cases} \tag{11}$$

The simulation duration T is chosen to ensure that the system has enough time to stabilize before and after the peak. The parameter p_{peak} is the proportion of the total inflow that is part of the perturbation and h corresponds to the width of the demand peak. Small values of h correspond to highly variable inflows. In the reference scenario, $h = 2\,\text{h}$. The average trip length in all simulations is $3\,\text{km}$, the average speed when the network is empty was taken to be $30\,\text{km}\,\text{h}^{-1}$ and the speed MFD has the form:

$$\mathcal{V}(n) = \begin{cases} 30(1-n/n_{\text{jam}})^2, & \text{if } n \in [0, n_{\text{jam}}] \\ 0, & \text{otherwise.} \end{cases}$$

Note that the production is maximized for the critical accumulation $n_{\text{cr}} = n_{\text{jam}}/3$. With the PL model, the corresponding maximum outflow is $C = \frac{30}{L}\left(\frac{2}{3}\right)^2 n_{\text{cr}} = \beta n_{\text{cr}}$, where $\beta = \frac{40}{9} \simeq 4.44\,\text{h}^{-1}$. We will, somehow abusively, refer to C as the capacity of the network considered. Rather than specifying some values for the demand (via p_{peak} and N) or for the capacity (via n_{jam}), we specify the demand relatively to the capacity. The size of the inflow perturbation is set identical to the jam accumulation ($Np_{\text{peak}} = n_{\text{jam}}$) for all simulations and the steady-state inflow is set equal to some percentage of the capacity (60% in the reference scenario, i.e. $(1-p_{\text{peak}})\frac{N}{T} = 0.6C = 0.6\left(\frac{40}{9}\right) n_{\text{cr}} \simeq 2.67 n_{\text{cr}}$).

Gridlock and hysteresis

The time series of accumulation, outflow and speed obtained with the trip-based model for trip length distributions corresponding to ratios of σ/L ranging from 0.4 to 1.2 are represented in Fig. 1, along with the

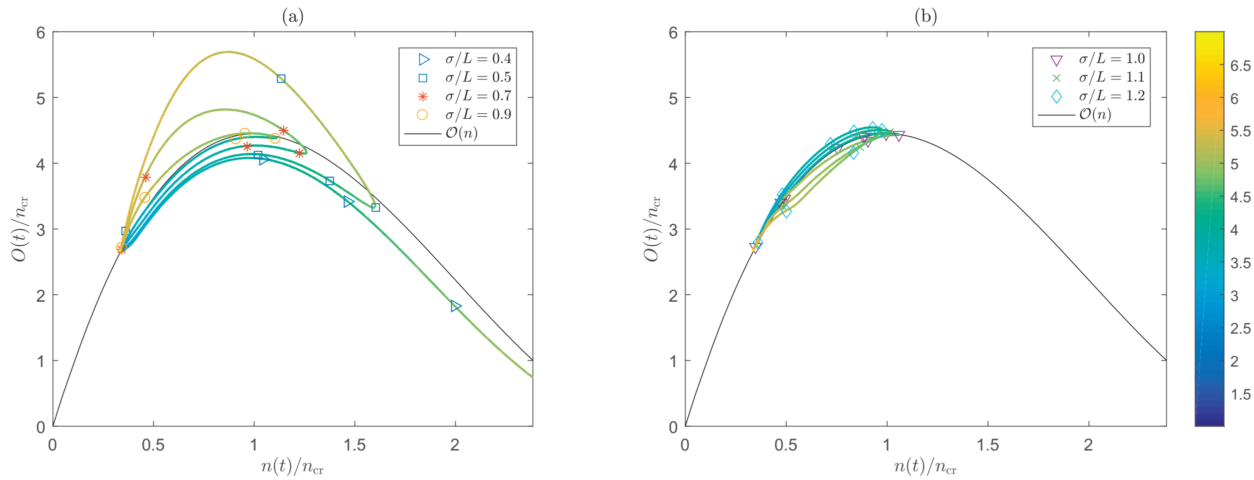

FIGURE 2 State trajectories obtained with the trip-based for several ratios σ/L, represented in the accumulation-outflow space together with the outflow MFD. The points corresponding to the warm-up period ($t < 1$) and to the end of the simulation ($t > 7$) are not represented. The color indicates the occurrence time.

time series obtained with the PL model (which only depends on L). In the three sub-figures, the time series obtained with the PL model are strikingly similar to those obtained with the trip-based model and $\sigma/L = 1$. This observation is to be considered in conjunction with the equivalence result established for exponential distributions (which also have $\sigma/L = 1$). These results also confirm the fundamental role played by the first and second moments of the trip length distribution.

Note also that the existence of congestion might lead to differences in the steady state. As smaller coefficients of variations correspond to more severely congested situations, the differences between different coefficients of variations are amplified by congestion. For $\sigma/L = 0.4$, the system reaches gridlock (and is trapped in this state).

The Fig. 2a and Fig. 2b represent the trajectory of the system in the space of the outflow MFD for the cases $\sigma/L < 1$ and $\sigma/L \geq 1$, respectively. Starting from the steady state, the system describes a loop in the outflow-accumulation space. This loop is counter-clockwise for $\sigma/L \leq 0.9$ and clockwise for $\sigma/L \geq 1.1$. For $\sigma/L = 1$, the trajectory follows the outflow MFD curve closely but actually describes an elongated "8". The same patterns could actually also be observed under free-flow conditions but they are amplified here by the severe congestion.

Sensitivity analysis

We now provide some statistics regarding the sensitivity of the different models to variations in h (inflow variability) and in the coefficient of variation σ/L. In order to highlight the generality of the obtained results, we provide these statistics both under congested conditions and under free-flow conditions, i.e. when the speed is not influenced by accumulation. In that case, Eq. (4) boils down to:

$$n(t) = \int_0^t I(s)\left(1 - F\left((t-s)v_{\mathrm{f}}\right)\right) \mathrm{d}s.$$

Since we only consider trip length distributions that are mixtures of uniform distributions, F is piece-wise linear and the entire integral can be solved easily for a large range of inflow functions, including sinusoidal functions.

Fig. 3 summarizes the values of the accuracy metric ξ with the TB model as a reference for a wide

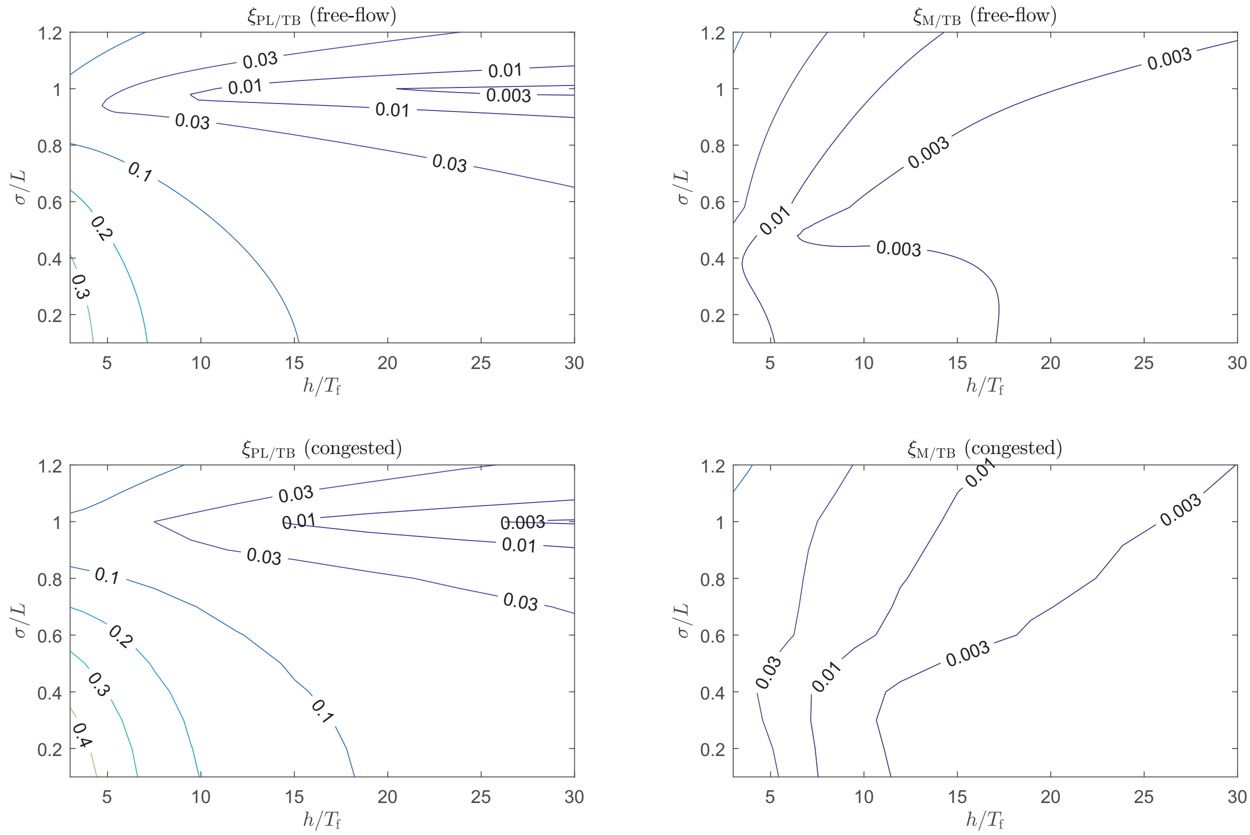

FIGURE 3 Error between the different models (PL vs TB in the first column, M vs TB in the second column) under free-flow (first row) and congested conditions (second row).

range of inflow variability (h, normalized with the average free-flow travel time T_{f}), coefficient of variation, and for both the PL and M models. The subplots on the first row are for free-flow conditions and those on the second row for congested conditions.

In both the congested and free-flow cases, the error between the PL and TB models is maximized when the inflow varies very rapidly and when the coefficient of variation of trip length is close to 0. In such cases, the error can exceed 30%. For more reasonable inflows however, the error is dramatically reduced for the entire range of coefficients of variations. When $h/T_{\mathrm{f}} > 18$ for instance, the error between the PL and TB models is systematically smaller than 10%. The error between the PL and TB models seems to be minimized for a value of σ/L around 1, which is in line with the result obtained for the exponential distribution. In comparison, the M model produces much smaller errors (less than 1% for a wide range of scenarios) and is less sensitive to the value of σ/L.

Imperfectly estimated coefficient of variation

Overall, the previous results suggest that the dynamics of the TB model depend heavily on the trip length distribution considered. We investigate in this section the consequences of using a complex model (TB) but with the wrong trip length distribution. More specifically, we use only (mixtures of) uniform distributions of the type described in Section 4.1 but we consider different coefficients of variation.

After running 13 scenarios based on the TB model with trip length distributions having different coefficient of variations, this metric was computed for all the 13^2 possible pairs of time series of accumulation and results were summarized in Fig. 4. When the two trip length distributions are identical, the error is of

σ/L in the reference trip-based model	σ/L considered for the trip-based model (approximation): 0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	PL
0	0.0%	3.8%	11.6%	19.0%	24.8%	29.2%	32.6%	35.6%	38.2%	40.6%	42.9%	45.3%	47.7%	43.2%
0.1	4.0%	0.0%	8.1%	15.8%	21.8%	26.4%	29.9%	33.0%	35.8%	38.4%	40.9%	43.5%	46.2%	41.3%
0.2	13.1%	8.8%	0.0%	8.4%	14.9%	19.9%	23.8%	27.3%	30.5%	33.6%	36.6%	39.7%	42.7%	36.9%
0.3	23.4%	18.7%	9.1%	0.0%	7.2%	12.6%	17.0%	21.0%	24.8%	28.4%	31.9%	35.4%	38.8%	32.0%
0.4	33.0%	27.9%	17.6%	7.7%	0.0%	5.9%	10.7%	15.2%	19.4%	23.5%	27.4%	31.2%	35.0%	27.4%
0.5	41.2%	35.8%	24.8%	14.4%	6.3%	0.0%	5.2%	10.0%	14.6%	18.9%	23.1%	27.2%	31.2%	23.1%
0.6	48.3%	42.7%	31.2%	20.4%	11.9%	5.4%	0.0%	5.2%	10.0%	14.5%	19.0%	23.3%	27.5%	18.9%
0.7	55.2%	49.3%	37.4%	26.4%	17.8%	11.0%	5.4%	0.0%	5.1%	9.8%	14.5%	19.0%	23.4%	14.4%
0.8	61.7%	55.6%	43.5%	32.4%	23.6%	16.7%	10.8%	5.3%	0.0%	5.0%	9.9%	14.6%	19.2%	9.8%
0.9	67.9%	61.8%	49.7%	38.5%	29.6%	22.4%	16.4%	10.6%	5.2%	0.0%	5.0%	9.9%	14.7%	5.0%
1	74.3%	68.2%	56.0%	44.7%	35.6%	28.3%	22.1%	16.2%	10.6%	5.2%	0.0%	5.1%	10.0%	1.3%
1.1	80.8%	74.7%	62.6%	51.2%	41.9%	34.4%	28.0%	21.9%	16.1%	10.6%	5.2%	0.0%	5.1%	5.5%
1.2	87.6%	81.6%	69.4%	57.8%	48.3%	40.7%	34.1%	27.7%	21.8%	16.1%	10.6%	5.2%	0.0%	10.8%

FIGURE 4 Comparison of relative errors ξ when considering two scenarios based on the trip-based model with different trip length distributions and when using the PL model as an approximation of the trip-based model. The peak duration h was increased from 2 to 2.15 compared to the scenario described in Section 4.2, such that setting $\sigma/L = 0$ does not lead to gridlock.

course null. Note also that the table is not exactly symmetric as the denominator in the calculation of $\xi_{\text{A/B}}$ is not the same as in the calculation of $\xi_{\text{B/A}}$. For comparison, we also evaluated the metric ξ when the reference is the TB model with one of the 13 possible trip length distribution and the approximation is the PL model. For a coefficient of variation $\sigma/L = 0.7$, the PL model achieves an error $\xi = 14.4\%$. The trip-based model achieves smaller errors only when the ratio σ/L is relatively close to its true value ($\sigma/L \in [0.5, 0.9]$). The worst accuracy is obtained when considering a trip-based model with homogeneous trip length, in which case the error reaches 55.2%. This suggests that the PL model may actually perform relatively well in many realistic cases and that considering a more complex model (like the TB model) might be counter-productive if the second moment of the trip length distribution cannot be estimated with a sufficient accuracy.

CONTROL

In this section, we investigate the usage of the presented models for control purposes. The concept of perimeter control is utilized, where the inflow to an urban area (zone) is restricted in order to achieve minimization of total delays. For this purpose, we define a control variable $u(t)$ that acts on the boundary of the zone and restricts the inflow, based on the current measured accumulation $n(t)$. Vehicles that cannot enter the zone right away are stored in a so-called "virtual queue", whose size at time t (in units of vehicles) is denoted $\mathrm{VQ(t)}$. For simplicity, we assume the entire inflow comes from neighboring regions, so that it can be fully controled at the perimeter. Thus, when perimeter control is active on the boundary, equation (1) becomes

$$\dot{n}(t) = u(t)\tilde{I}(t) - O(t), \tag{12}$$

with

$$\tilde{I}(t) = \begin{cases} I(t) & \text{if } VQ(t) = 0, \\ c & \text{if } VQ(t) > 0, \end{cases}$$

where c denotes the boundary capacity and $u(t) \in [0, 1]$. The controller can for instance restrict all the users from entering the network by setting $u(t) = 0$ in cases of strong gating, or can allow everyone to enter ($u(t) = 1$) in cases of light traffic conditions, i.e. low accumulation values inside the zone of interest. Vehicles that are not allowed to enter by the control policy are stored in a virtual queue at the boundary of

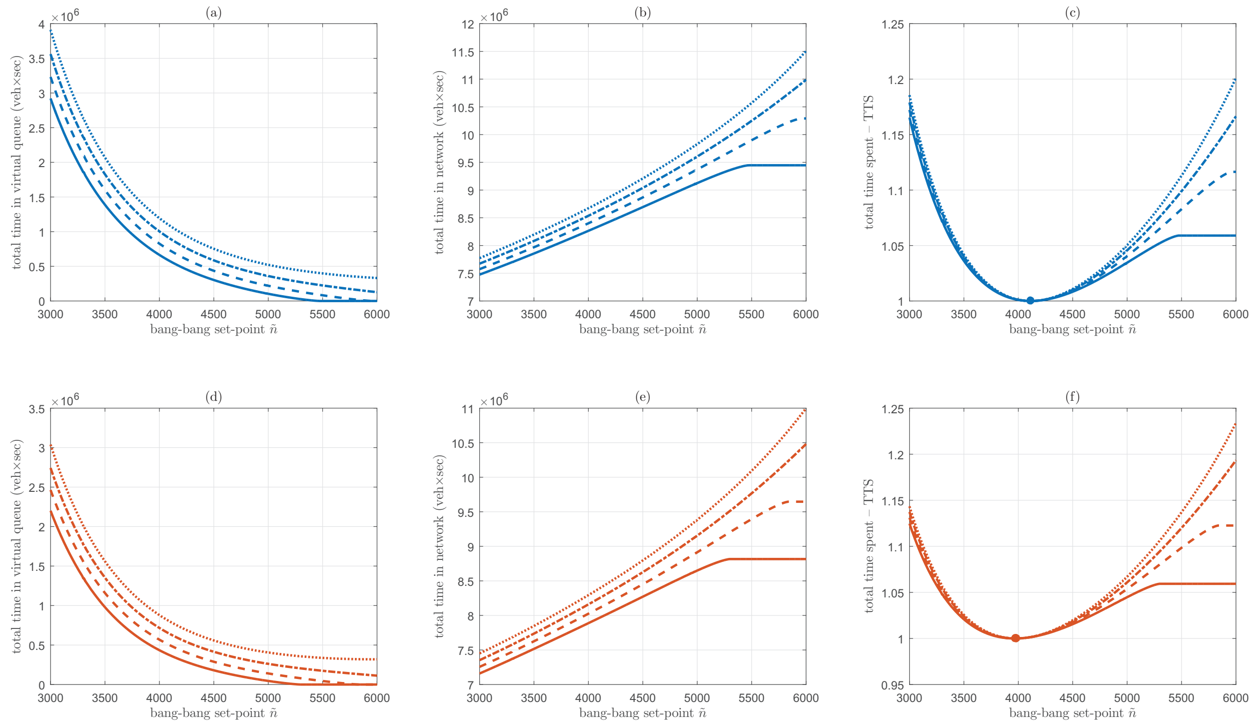

FIGURE 5 Analysis of different set-points for the bang-bang control policy for models PL and M.

the network. We model the vehicle accumulation in the virtual queue and study its contribution to the overall system delay.

Bang-bang policy

One control policy that is presented in Daganzo (4) is the so-called bang-bang policy, where the control actions switch from minimum to maximum according to the measured accumulation $n(t)$ inside the zone. The objective of this policy is to keep the region accumulation as close as possible to a set point $\tilde{n} = n_{\text{cr}}$ for most of the time by closing the "gates" to the area. This policy is proven to be optimal for the PL model, in the sense that if we operate around the critical accumulation n_{cr} the outflow of the system over time is maximized, and no matter the accumulation of the virtual queue, the total delay is also minimized (see Daganzo (4) for the complete proof).

Here we investigate different values for the set-point $\tilde{n}$ and we apply the bang-bang control to the presented models in order to assess the performance of the system. In order to compute the total delay we consider the time spent in the network by all vehicles plus the time spent in the virtual queue waiting to enter the network. The summation of the two gives the total time spent (TTS) which is the criterion for comparing different strategies. Fig. 5 presents the numerical experiments for a range of $\tilde{n}$ from 3000 to 6000 vehicles, whereas the critical point of the MFD that maximizes flow is $n_{\text{cr}} = 4110$. For all this range the bang-bang policy is activated in the PL and M model and the figure presents the evolution of the integral of accumulations in the network and virtual queue. The PL model achieves the minimum TTS at a point very close to the actual critical accumulation, whereas the M model minimizes the TTS for an accumulation about 150 vehicles (3%) smaller than the critical (Fig. 5(c)). Moreover, this point is not constant for the M model and moves slightly to the left or right depending on the demand variations. This discrepancy between the models is due to the different dynamic equations and the resulting hysteresis on the MFD that affects the optimal point for the bang-bang policy.

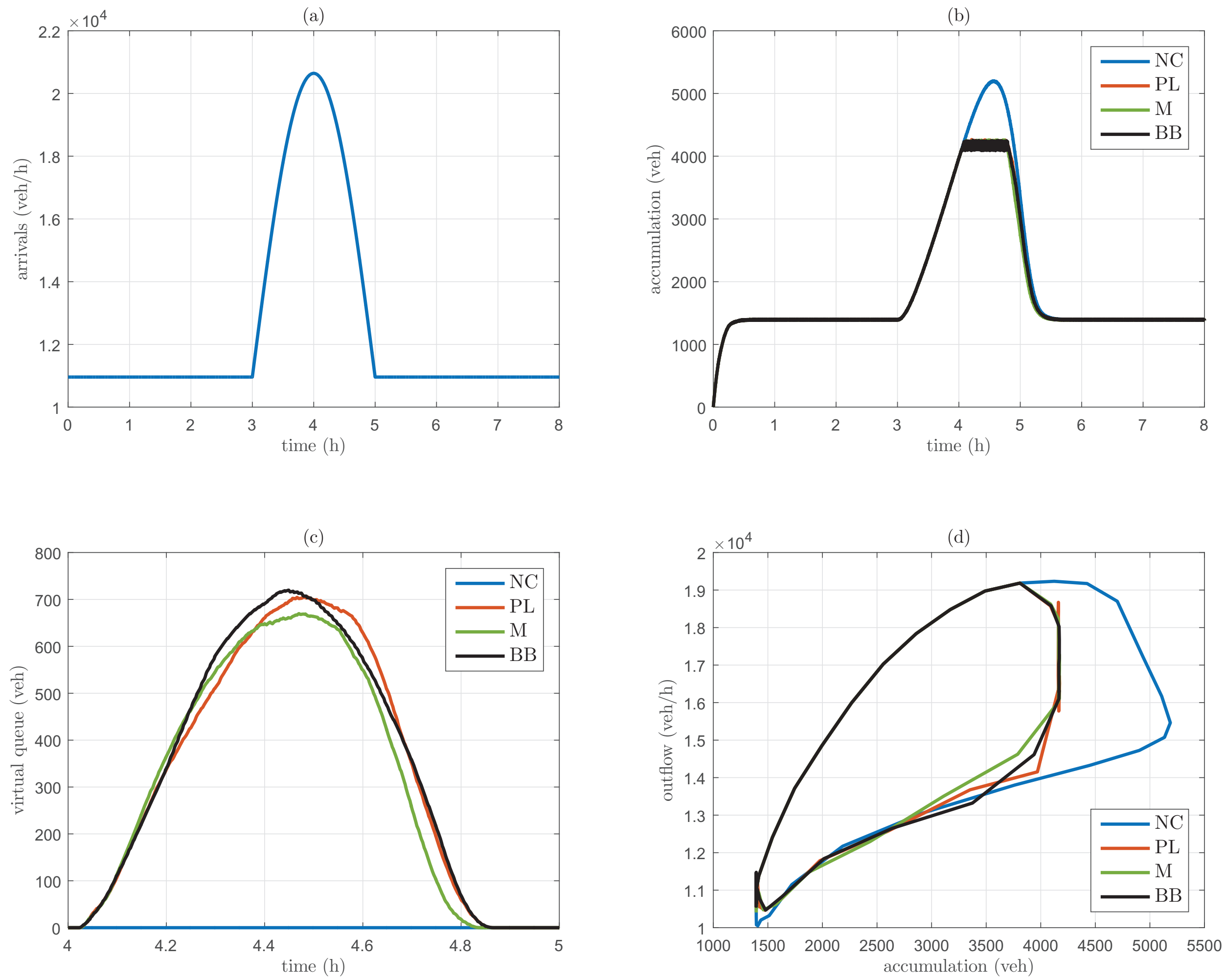

FIGURE 6 Simulation of no control (NC) case and application of MPC utilizing the different models (PL and M).

Model predictive control (MPC)

Another usage of the presented models could be for real-time control of urban areas based on the concept of model predictive control. In that case, given a current state of the network $n(t)$ and the trajectory of the future inflow $I(t)$ for a given prediction horizon (e.g. up to $t + T_\mathrm{p}$), one can use the model and solve an optimization problem to minimize delays for this horizon. Such a finite horizon optimal control problem has an objective function of the form

$$\mathcal{J} = \min \int_t^{t+T_\mathrm{p}} n(t)\,\mathrm{d}t + \int_t^{t+T_\mathrm{p}} \mathrm{VQ(t)}\,\mathrm{d}t, \tag{13}$$

and is subject to the constraints that all the variables should follow explicitly the dynamic equations of the considered model.

Fig. 6 presents some simulation results for MPC by utilizing the models presented in the previous sections. The TB model is used as the plant (real process) for the experiments and the models PL and M as the prediction models for the MPC. We simulate the inflow of Fig. 6(a) which brings the system to a steady state, then we have the peak hour and we finish again with the same steady state. Fig. 6(b) presents the trajectories of accumulations for all the studied scenarios. First of all, we have the no control (NC) case where no controller is applied to the trip-based plant. Then, we have the MPC cases with both models and the bang-bang (BB) policy that regulates the accumulation around the actual critical point of the unimodal production MFD curve. It should be noted that for these experiments we use the same discretization as described in the section with the simulation comparison. The two models have very similar performance to the BB policy (with slight improvement that is not visible in the plot), resulting in reduced delays for the system compared to NC (area between the blue curve and the other curves). Note that in the total delay of the control experiments one needs also to include the area that is computed by the integral of the vehicles in the virtual queue which is depicted in Fig. 6(c). Nevertheless, the control examples have still significant improvements compared to NC, as the order of magnitude of the vehicles in the virtual queue is much lower that the accumulation inside the zone of interest. Finally, Fig. 6(d) represents the outflow curves of all instances. We can observe the hysteresis of the outflow MFD as well as the difference between NC and the other cases, that do not reach states that reduce the zone outflow, as they are regulated at a point that maximizes the outflow (we get many states around this point during the simulation).

From the results presented in this figure we can conclude that the PL model gives very similar results to the M model when used for the MPC horizon. Although the M model is much closer to the trip-based plant (as demonstrated in the simulation comparison section), when utilized for model predictive control the two models exhibit similar results. The PL model has been used in previous works for perimeter control purposes (see e.g. Kouvelas et al. (20) for a multi-region system), and the analysis presented here supports the conjecture that it is not an inadequate model for future prediction and control.

CONCLUSIONS

This investigation on the influence of the trip length distribution on the trip completion rate under variable inflow provides a better understanding of the conditions under which the PL model represents a sound approximation. It is shown in particular that the PL model describes well realistic situations with large trip length heterogeneity (σ/L close to 1) but performs rather poorly when trip length is homogeneous (which is the most common assumption in the literature). In the latter case, the M model introduced in this paper can capture most of the complexity of the TB model, at a considerably smaller cost. The simple PL model seems however to be sufficient for control strategies with feedback.

REFERENCES

[1] Godfrey, J. W., The mechanism of a road network. *Traffic Engineering and Control*, Vol. 11, No. 7, 1969, pp. 323–327.

[2] Herman, R. and I. Prigogine, A two-fluid approach to town traffic. *Science*, Vol. 204, No. 4389, 1979, pp. 148–151.

[3] Mahmassani, H., J. C. Williams, and R. Herman, Investigation of network-level traffic flow relationships: some simulation results. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 971, 1984, pp. 121–130.

[4] Daganzo, C. F., Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, Vol. 41, No. 1, 2007, pp. 49–62.

[5] Arnott, R., A bathtub model of downtown traffic congestion. *Journal of Urban Economics*, Vol. 76, 2013, pp. 110–121.

[6] Fosgerau, M., Congestion in the bathtub. *Economics of Transportation*, Vol. 4, No. 4, 2015, pp. 241–255.

[7] Daganzo, C. F. and L. J. Lehe, Distance-dependent congestion pricing for downtown zones. *Transportation Research Part B: Methodological*, Vol. 75, No. 0, 2015, pp. 89–99.

[8] Lamotte, R. and N. Geroliminis, The morning commute in urban areas with heterogeneous trip lengths. *Transportation Research Procedia*, Vol. 23, 2017, pp. 591–611, presented at the 22nd International Symposium on Transportation and Traffic Theory in Chicago, Illinois.

[9] Lamotte, R. and N. Geroliminis, The Morning Commute in Urban Areas: Insights from Theory and Simulation. In *Transportation Research Board 95th Annual Meeting*, 2016.

[10] Mariotte, G., L. Leclercq, and J. A. Laval, Macroscopic urban dynamics: Analytical and numerical comparisons of existing models. *Transportation Research Part B: Methodological*, Vol. 101, 2017, pp. 245–267.

[11] Geroliminis, N. and C. F. Daganzo, Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, Vol. 42, No. 9, 2008, pp. 759–770.

[12] Buisson, C. and C. Ladier, Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record: Journal of the Transportation Research Board*, , No. 2124, 2009, pp. 127–136.

[13] Ji, Y., J. Luo, and N. Geroliminis, Empirical Observations of Congestion Propagation and Dynamic Partitioning with Probe Data for Large-Scale Systems. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2422, 2014, pp. 1–11.

[14] Tsubota, T., A. Bhaskar, and E. Chung, Macroscopic fundamental diagram for Brisbane, Australia: empirical findings on network partitioning and incident detection. *Transportation Research Record: Journal of the Transportation Research Board*, , No. 2421, 2014, pp. 12–21.

[15] Wang, P. F., K. Wada, T. Akamatsu, and Y. Hara, An empirical analysis of macroscopic fundamental diagrams for Sendai road networks. *Interdisciplinary Information Sciences*, Vol. 21, No. 1, 2015, pp. 49–61.

[16] Loder, A., L. Ambühl, M. Menendez, and K. W. Axhausen, Empirics of multi-modal traffic networks – Using the 3D macroscopic fundamental diagram. *Transportation Research Part C: Emerging Technologies*, Vol. 82, 2017, pp. 88–101.

[17] Lamotte, R., R. Meyer, and N. Geroliminis, La congestion en ville: une approche globale. In *Big Data et politiques publiques dans les transports* (A. de Palma and S. Dantan, eds.), Economica, 2017, chap. 10.

[18] Geroliminis, N. and J. Sun, Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological*, Vol. 45, No. 3, 2011, pp. 605–617.

[19] Saeedmanesh, M. and N. Geroliminis, Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. *Transportation Research Procedia*, Vol. 23, 2017, pp. 962–979, papers Selected for the 22nd International Symposium on Transportation and Traffic Theory Chicago, Illinois, USA, 24-26 July, 2017.

[20] Kouvelas, A., M. Saeedmanesh, and N. Geroliminis, Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. *Transportation Research Part B: Methodological*, Vol. 96, 2017, pp. 26–45.