Bi-modal macroscopic traffic dynamics in a single region

Mahendra Paipuri, Ludovic Leclercq

Bi-modal Macroscopic Traffic Dynamics in a Single Region

Mahendra Paipuri[a,*], Ludovic Leclercq[a]

[a] *Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT, F-69518, Lyon, France.*

Abstract

Bi-modal or 3D-MFD relates the accumulation of cars, buses to total production at the network level. The current work provides a detailed discussion of extended MFD-based models namely, accumulation-based and trip-based models that accounts for bi-modal flows through 3D-MFD. In addition, delay accumulation-based models, also known as exit-flow models in classical traffic flow theory, are revisited. Fundamental modeling differences between different MFD-based models are illustrated using a benchmark test case. A new FIFO-based entry flow function is also proposed in order to address the inconsistencies of the conventional entry flow function in the context of the 3D-MFD case. A novel weak internal FIFO discipline is proposed to circumnavigate the violation of internal FIFO order during network unloading in the delay accumulation-based model. MFD-based models are verified using the solutions of micro-simulations performed on an idealized grid network. The importance of separating the 3D-MFD into partial car and bus 3D-MFDs is highlighted. Moreover, it is also shown that partial bus 3D-MFD should be further split when dedicated bus lanes are present in the network to account for unequal mean speeds between different bus lanes.

Keywords: 3D Macroscopic Fundamental Diagram, Bi-modal, MFD-based Modeling, FIFO, Micro-simulation, Delay accumulation-based model

2010 MSC: 00-01, 99-00

1. Introduction

Ever since the existence of a well-defined Macroscopic Fundamental Diagram (MFD) was established for the city of Yokohama, Japan (Geroliminis and Daganzo, 2008), there has been a rapid development in the MFD-based models to describe the traffic dynamics at large scale networks. MFD relates the density of vehicles to mean flow in the network. This relation was first introduced by Godfrey (1969) and later revisited by Mahmassani *et al.* (1984) based on simulation studies. The first formulation of a urban model based on MFD was proposed by Daganzo (2007). The main advantage of this modeling framework is its ability to predict the aggregate traffic state dynamics with a relatively lower computational cost. These MFD-based simulators have been

*Corresponding author

Email addresses: mahendra.paipuri@univ-eiffel.fr (Mahendra Paipuri), ludovic.leclercq@univ-eiffel.fr (Ludovic Leclercq)

found promising in several applications like traffic state estimation (Knoop and Hoogendoorn, 2014; Yildirimoglu and Geroliminis, 2014; Kavianipour *et al.*, 2019), perimeter control (Keyvan-Ekbatani *et al.*, 2012; Haddad and Mirkin, 2017; Ampountolas *et al.*, 2017; Mohajerpoor *et al.*, 2019), congestion pricing (Gu *et al.*, 2018) and cruising for parking (Cao and Menendez, 2015; Leclercq *et al.*, 2017), *etc.*

Urban networks usually comprise multiple modes like cars, buses and bicycles, *etc.*, sharing the network infrastructure. Previous works, *e.g.*, (Boyac and Geroliminis, 2011; Chiabaut *et al.*, 2014; Loder *et al.*, 2017), suggest that the buses and cars affect the network dynamics in different ways. Most of the MFD-based simulators proposed in the literature usually restrict to single-mode or 2D-MFD formulation, *i.e.*, relation between the density of *all* vehicles and mean flow of *all* vehicles in the network. Hence, a single-mode 2D-MFD considers the cars and buses to be alike and it cannot account for the influence of different mode shares on the network dynamics and performance. Geroliminis *et al.* (2014) is first to address this issue and proposed a bi-modal or 3D-MFD for the area of downtown San Francisco based on micro-simulations. This so-called 3D-MFD for bi-modal traffic data relates the accumulation of cars, buses to the total mean flow in the network. Ortigosa *et al.* (2015) analyzed 3D-MFDs for the cities of Zurich and San Francisco using micro-simulations to study the effect of dedicated bus lanes on the network performance.

The first empirical study of 3D-MFD is proposed by Loder *et al.* (2017) for the city of Zurich. They conclude that adding a public transport bus to the network has a more negative impact on the mean speed of cars compared to adding a car. More recently, Loder *et al.* (2019) proposed a new functional form for 3D-MFD based on the structure and topology of car and bus network. Huang *et al.* (2019) investigated the existence of 3D-MFD using the GPS data of private cars, taxis and public buses for the city of Shenzhen in China. While there is plenty of literature around 2D-MFD and numerous applications proposed using the network dynamics resolved by 2D-MFD, most of the works proposed based on 3D-MFD focus on network performance and capacity characterization. There has been little attention in the area of extension of a general MFD modeling framework to 3D-MFD settings. This work aims to fill this gap by proposing the MFD modeling framework founded on 3D-MFDs and study the model properties.

There are primarily two different types of MFD-based models proposed in the literature namely, accumulation-based and trip-based models. Daganzo (2007) proposed the accumulation-based model in the framework of a single reservoir system. This framework is extended to consider multiple trip lengths in Geroliminis (2009); Yildirimoglu *et al.* (2015). Despite the fact that the accumulation-based is relatively simple to resolve and computationally less demanding, it suffers from a few drawbacks as highlighted in Mariotte *et al.* (2017). Hence, the trip-based approach gained significant attention in the recent past as it can address few issues of the accumulation-based model. The idea of the trip-based model was proposed by Arnott (2013) and it was revisited by Daganzo and Lehe (2015); Lamotte and Geroliminis (2016); Mariotte *et al.* (2017); Leclercq *et al.* (2017). The modeling of congestion spill-backs in the context of multi-reservoir systems

is proposed in Mariotte and Leclercq (2019). However, Leclercq and Paipuri (2019) showed that no model is perfect and a hybrid model bridging both accumulation-based and trip-based approaches gives more consistent results both in free-flow and congestion regimes. In addition to traditional MFD-based models, the accumulation-based model with outflow delay is also considered in the present work. This model is proposed and studied at the link scale by Friesz *et al.* (1989); Daganzo (1995); Astarita (1996). It is revisited in the current work at the network scale to address some drawbacks of the accumulation-based model. In terms of physical interpretation, It can be considered as the time-continuous variant of the trip-based model. Recently, Haddad and Zheng (2018) proposed a perimeter control model based on the delay accumulation-based model. However, they simplified the model by assuming a constant average delay. Zhong *et al.* (2018) proposed another perimeter control strategy founded on the delay accumulation-based model. However, the authors have not investigated the stability issues that the model can pose during network congested scenarios.

Ampountolas *et al.* (2017) used the accumulation-based model using 3D-MFD in the context of perimeter control. However, their work focuses on control with a fixed composition of bi-modal traffic. Recently, Dakic *et al.* (2019) used the accumulation-based model for macroscopic modeling in the context of perimeter control using a bi-modal MFD. To the knowledge of the authors, the current work is the first to propose the complete modeling frameworks for not only the accumulation-based model but also the trip-based and the delay accumulation-based models considering a 3D-MFD. One of the aims of this work is to provide a complete review of different model properties under transition and steady state conditions while resolving traffic dynamics. Two different test cases are built to study the properties and limitations of the different models. The solutions of the MFD-based simulators are compared to a reference solution that is computed by solving the system of linear hyperbolic equations, where the flux is defined by 3D-MFD (Leclercq *et al.*, 2015). This approach explicitly considers the traffic state evolution during the distance traveled in the reservoir. It guarantees that delays (wave propagation) from one perimeter to another (for instance, inflow to outflow and *vice-versa*) are consistent and thus, enable to assess whether a modeling approach provides consistent solution during transitional periods. This model is referred to as the continuum space-time model in the present work.

The remainder of the paper is organized as follows: Section 2 provides the details of the functional form of 3D-MFD assumed in this work, a brief introduction to different MFD-based models used. A detailed investigation of the delay accumulation-based model is also provided in this section. Section 3 investigates the different MFD-based models with numerical examples. This section demonstrates the essential differences between different MFD-based modeling frameworks along with the limitations, if any, for different models. Several novel modeling techniques are proposed to address the limitations. Section 4 presents verification with micro-simulation results using a grid network setting. Finally, Section 5 gives a brief overview of the conclusions made in this work.

Table 1: Coefficient values used to compute the 3D-MFD.

Cars		Buses	
Coefficient	Value	Coefficient	Value
$\beta_{c,0}$	15	$\beta_{b,0}$	15
$\beta_{c,c}$	-0.015	$\beta_{c,b}$	-0.003
$\beta_{b,c}$	-0.3	$\beta_{b,b}$	-0.06

2. Methodology: 3D-MFD and MFD-based models

2.1. Functional form of 3D-MFD

For the sake of simplicity, the functional form proposed in Loder *et al.* (2017) is used in the present work. This is not restrictive as other functional forms like exponential function proposed by Geroliminis *et al.* (2014), can also be used. The mean speed of the cars, v_c, is expressed as a linear function of the accumulation of cars, n_c, and the accumulation of buses, n_b as follows,

$$v_c(n_c, n_b) = \beta_{c,0} + \beta_{c,c}\, n_c + \beta_{b,c}\, n_b. \tag{1}$$

The constant of the function, $\beta_{c,0}$, corresponds to the free-flow speed of the cars. The coefficients $\beta_{c,c}$ and $\beta_{b,c}$ represent the marginal effect of each mode on the car mean speed, *i.e.*, the amount by which the free flow speed of the cars is reduced by adding a vehicle of each mode. Similarly, the speed of buses is given as,

$$v_b(n_c, n_b) = \beta_{b,0} + \beta_{c,b}\, n_c + \beta_{b,b}\, n_b, \tag{2}$$

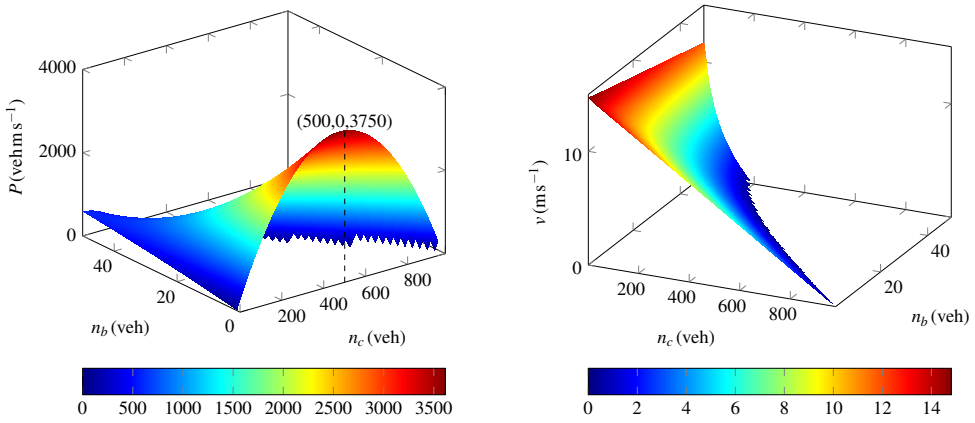
where $\beta_{c,b}$ and $\beta_{b,b}$ represent the marginal effect of each mode on bus mean speed. Now, the total production in the network isthe sum of the production of cars, P_c, and production of buses, P_b. As the production is defined as the product of accumulation and mean speed, the total production can be expressed as,

$$P = n_c\, v_c + n_b\, v_b. \tag{3}$$

From the eq. (3), P can be expressed as $v\,(n_c + n_b)$, where v is the mean speed of all vehicles in the network. Using the definition of total production and eq. (3), the mean vehicular speed in the network can be written as,

$$v = v_c\, \frac{n_c}{n_c + n_b} + v_b\, \frac{n_b}{n_c + n_b}. \tag{4}$$

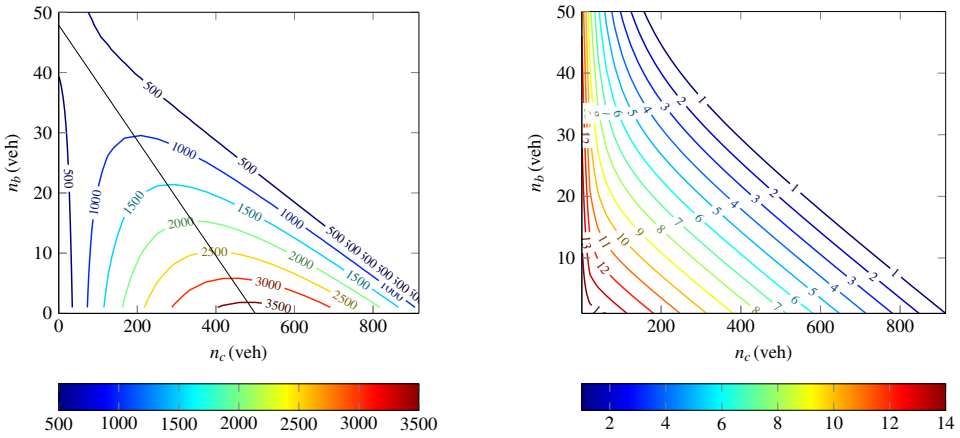
Table 1 presents the values of coefficients used in the present work and correspondingly, Fig. 1 shows production MFD and velocity MFD surfaces obtained. Unless otherwise stated, the remainder of the work uses coefficient values presented in Table 1 for modeling 3D-MFD. As shown in Fig. 1a, when $n_b = 0$ the critical accumulation of vehicles is 500 veh and maximum production is 3 750 veh m s^{-1}. It can be noticed that maximum network production occurs at zero bus accumulation and it decreases as the number of buses increases in the network. Moreover, the



(a) Production MFD surface.

(b) Velocity MFD surface.

Figure 1: Production and velocity 3D-MFD surfaces.



(a) Production MFD contour plot.

(b) Velocity MFD contour plot.

Figure 2: Production and velocity 3D-MFD contour plots.

critical accumulation, n_{cr}, is no longer a constant value but depends on the partial accumulations of buses and cars in the network. Fig. 2 presents the contour lines for both production and velocity 3D-MFDs. The black line in the production 3D-MFD corresponds to the line of critical accumulation. It means for any given bus accumulation, the critical line represents the maximum available production in the network. Mathematically, the equation of the line of critical accumulation can be expressed as $\beta_{c,0} + 2\,\beta_{c,c}\,n_c + (\beta_{c,b} + \beta_{b,c})\,n_b = 0$.

2.2. *Continuum space-time model*

This model has been first introduced by Leclercq *et al.* (2015) for network region with multiple paths of different lengths. It is the only one among MFD-based models that can consider explicitly the internal travel distances with different traffic scenarios. While being computationally demanding, it properly reproduces wave propagation between the region perimeter and therefore, this solution is used as the reference in the present work.

The system of hyperbolic conservation equations for each mode, m, can be represented as,

$$\frac{\partial k_m}{\partial t} + \frac{\partial k_m v(\mathbf{K})}{\partial x} = 0 \quad \forall m, \tag{5}$$

where k_m is the density of mode m and $v(\mathbf{K})$ is the velocity defined by 3D-MFD, which is the function of all densities, *i.e.*, $\mathbf{K} = [k_1, \ldots, k_m]^T$. In the special case of $m = 1$ and flux computed by fundamental diagram, eq. (5) becomes Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956). The model presented in eq. (5) can be interpreted as multi-modal or multiclass traffic flow models. Different multiclass models proposed in literature are well documented in Fan and Work (2015). Since all vehicles travel at the same speed given by velocity MFD in the present model, this model is referred to as homogeneous multiclass model. The strict hyperbolicity of the homogeneous multiclass models can be proven when $m \leq 2$ (Benzoni-Gavage and Colombo, 2002; Keyfitz and Kranzer, 1980). System (5) can be re-written as,

$$\frac{\partial \mathbf{K}}{\partial t} + \mathbf{A}\frac{\partial \mathbf{K}}{\partial x} = \mathbf{0}, \tag{6}$$

where $\mathbf{A}$ is the Jacobian matrix and for the case of $m = 2$, it is given as follows,

$$\mathbf{A} = \begin{bmatrix} v + k_c \frac{\partial v}{\partial k_c} & k_b \frac{\partial v}{\partial k_c} \\ k_c \frac{\partial v}{\partial k_b} & v + k_b \frac{\partial v}{\partial k_b} \end{bmatrix}. \tag{7}$$

It can be shown that the Eigen values of the Jacobian $\mathbf{A}$ are real and distinct if $\frac{\partial v}{\partial k_c} < 0$ and $\frac{\partial v}{\partial k_b} < 0$. The functional form of 3D-MFD proposed earlier fulfills this condition.

A single reservoir system with two different trips, one for cars and one for buses, is considered in this work. The numerical resolution of the system (5) using MFD as flux function is discussed in-detail in Leclercq *et al.* (2015). A similar approach is used here, albeit, 3D-MFD is used to define the flux function and hence, numerical scheme details are omitted. A constant CFL number of 0.5 is used along with an adaptive time stepping scheme in all computations. The trip length is discretized into 200 cells and a ghost cell is used on either end of the trip length to enforce boundary conditions.

2.3. Accumulation-based model

Essentially, the accumulation-based model is the simplification of the continuum space-time model, where the space derivative of the flux is neglected. The resulting conservation equation (Daganzo, 2007) to resolve the reservoir dynamics is given as,

$$\frac{\mathrm{d}n_m}{\mathrm{d}t} = q_{m,\mathrm{in}}(t) - q_{m,\mathrm{out}}(t) \quad \forall m, \tag{8}$$

where n_m is the space-averaged accumulation in the reservoir for mode m, $q_{m,\mathrm{in}}(t)$ is the total effective inflow and $q_{m,\mathrm{out}}(t)$ is the total effective outflow. The main advantage of this model (8) compared to the continuum space-time (5) is that this system is well-defined for any number of modes, m. As the system (8) comprises of only Ordinary Differential Equations (ODE), the

numerical resolution is straightforward and simple. On the other hand, the accumulation-based model accounts for the traveled distance incorrectly inside the reservoir, which may lead to inaccuracies in network dynamics during transition periods (Mariotte *et al.*, 2017).

The theory of single reservoir and multi-reservoir accumulation-based model is discussed in-detail in Mariotte *et al.* (2017); Mariotte and Leclercq (2019). The following equations are the direct extension of the stated previous works to a bi-modal or 3D-MFD. Hence, at the reservoir entry, the effective inflow is defined as,

$$q_{m,\text{in}}(t) = \min\left(\lambda_m(t), I_m(n_m,n)\right), \tag{9}$$

where $\lambda_m(t)$ is the demand for mode m and $I_m(n_m,n)$ is the entry supply function, which depends on the reservoir state. The entry supply function is defined as per the following relation,

$$I_m(n_m,n) = \begin{cases} \dfrac{n_m}{n}\dfrac{P_{cr}(n_m,n)}{L_m} & \text{if } n \le n_{cr}(n_m,n) \\ \dfrac{n_m}{n}\dfrac{P(n_m,n)}{L_m} & \text{otherwise.} \end{cases} \tag{10}$$

In eq. (10), n is the total vehicular accumulation, *i.e.*, $n = \sum_m n_m$, (n_{cr}, P_{cr}) are the critical accumulation and production, respectively and L_m is the trip length of mode m. Following the definition of entry supply function (10), it can be shown that $\sum_m I_m = \dfrac{P_s(n_m,n)}{L}$, which is the flow constraint, *i.e.*, sum of flows on all routes cannot exceed reservoir capacity. Here, $P_s(n_m,n)$ and L are entry supply production and average trip length (Geroliminis, 2009), respectively and they are defined as follows,

$$P_s(n_m,n) = \begin{cases} P_{cr}(n_m,n) & \text{if } n \le n_{cr}(n_m,n) \\ P(n_m,n) & \text{otherwise} \end{cases} \quad ; \quad L = \frac{n}{\sum_m \frac{n_m}{L_m}}. \tag{11}$$

The inflow supply function can be expressed in terms of entry supply production as $I_m(n_m,n) = \dfrac{n_m}{n}\dfrac{P_s(n_m,n)}{L_m}$. In addition to flow, another constraint can be defined based on production, *i.e.*, $\sum_m L_m I_m = P_s$, which states that the sum of entering production of each route should not exceed reservoir total production capacity. The entry supply function definition (10) is valid, when all routes (or modes) are in congestion. However, when only fraction of the routes are in congestion, say car route is congested in reservoir and buses operate in free flow, reservoir capacity maybe under-utilized by using (10). Hence, in this case, to ensure that the reservoir capacity is used fully, the following constraint on production is defined,

$$L_m q_{m,\text{in}} = \begin{cases} L_m \lambda_m & \text{if } L_m \lambda_m < L_m I_m(n_m,n) \\ \max\left(\sum_{n\neq m}(P_s(n_m,n) - L_n \lambda_n), \dfrac{n_n}{n} P_s(n_m,n)\right) & \text{otherwise.} \end{cases} \tag{12}$$

By definition (12), the sum of entering production is equal to the total available production in the reservoir, *i.e.*, $\sum_m L_m q_{m,\text{in}} = P_s(n_m,n)$.

Similarly, the effective outflow is given as,

$$q_{m,\text{out}}(t) = \min\left(\mu_m(t), O_m(n_m,n)\right), \tag{13}$$

where $\mu_m(t)$ is the supply restriction, if exists, and $O_m(n_m,n)$ is the outflow demand function and it can be defined as,

$$O_m(n_m,n) = \begin{cases} \frac{n_m}{n}\frac{P(n_m,n)}{L_m} & \text{if } n \leq n_{cr}(n_m,n) \\ \frac{n_m}{n}\frac{P_{cr}(n_m,n)}{L_m} & \text{otherwise.} \end{cases} \tag{14}$$

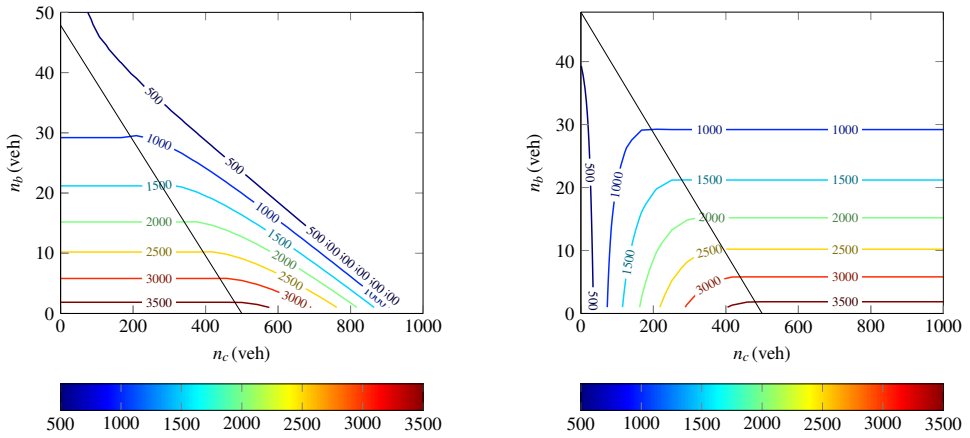
The details of the formulation of entry supply and outflow demand functions are provided in Mariotte and Leclercq (2019) and the references within and hence, not discussed here. In the definition of the outflow demand function, the outflow share for each mode is computed based on the ratio of partial accumulation of each mode to the total accumulation. However, for accurate prediction of dynamics, the ratio of vehicles of each mode to the total number of vehicles of all modes *that are close to exit* of the reservoir must be considered. Since this information is unavailable in the accumulation-based framework (due to absence of space dimension), the ratio of vehicles in the whole reservoir is used to compute partial outflow demand. This assumption has implications on the resulting dynamics, but this is the closest approximation that can be made in the context. The important point worth making here is that the critical accumulation and critical production in eqs. (10) and (14) are not constant values as in the case of 2D-MFD, but decreasing functions depending on the accumulation of each mode, n_m. This has an important implication in the modeling, which will be demonstrated in numerical examples. Analogous to entry supply production in inflow computations, a new variable called exit demand production, $P_d(n_m,n)$, can be defined as,

$$P_d(n_m,n) = \begin{cases} P(n_m,n) & \text{if } n \leq n_{cr}(n_m,n) \\ P_{cr}(n_m,n) & \text{otherwise,} \end{cases} \tag{15}$$

and similarly, outflow demand functions can be expressed in terms of exit demand production as $O_m(n_m,n) = \frac{n_m}{n}\frac{P_d(n_m,n)}{L_m}$.

The entry supply and exit demand functions discussed till now are a conventional way of treating the entry and the exit flows in the MFD-based framework. They are developed and studied in the context of conventional 2D-MFD. However, it will be demonstrated in numerical examples that this type of treatment, especially at the entry of the reservoir, has limitations when 3D-MFD is considered. Hence, to address this issue, a new entry flow function based on First-In-First-Out (FIFO) discipline is proposed in this work. This type of entry flow has not been explored in the context of MFD-based modeling. Note that other formulations of entry flow functions that consider higher flow constraint than the MFD maximum capacity and/or higher critical accumulation have been proposed in the literature. They are not discussed here however, it is worth noting that FIFO discipline at the entry can be extended to any of these functional forms and address the limitations of the 3D-MFD modeling framework. This will be discussed in the subsequent sections after demonstrating the shortcomings of the conventional entry flow function.

Figure 3 shows the entry supply and exit demand functions graphically on contour plots.



(a) Entry supply production function.

(b) Exit demand production function.

Figure 3: Contour plots of entry supply and exit demand production functions.

Note that these are analogous to entry and exit functions of a cell in the Cell Transmission Model (CTM) (Daganzo, 1994). A first-order forward Euler scheme with a time step, Δt, of 1 sec is used in the present work to numerically resolve the accumulation-based model.

2.4. Trip-based model

Now considering the trip-based formulation, mathematically it can be expressed as,

$$L_m = \int_{t-\tau_m(t)}^{t} v(n(s))\,\mathrm{d}s \qquad \forall m. \tag{16}$$

Consider $\tau_m(t)$ is the travel time of a user of mode m, who entered the reservoir at time t. The speed at each time instant depends on the total accumulation, n, in the reservoir, which is given by velocity 3D-MFD. Hence, the area under the speed-time curve between the times, $t - \tau_m(t)$ and t gives the total travel distance, which is trip length L_m. The significant modeling difference in the trip-based model compared to the accumulation-based model is that the former considers the traveled distance explicitly. This results in the resolution of more accurate traffic dynamics during the transition compared to the accumulation-based model. This is crucial for bi-modal flows as buses and cars have different travel distances in the same region. This will be evident in the numerical results presented subsequently, where the trip-based results are closer to the reference solution. Event-based resolution proposed in Mariotte *et al.* (2017); Leclercq *et al.* (2017) is used in the present work. The trip starting times of all vehicles are created based on the demand for each mode. Based on the current accumulation and trip length of each mode, the time at which the leading vehicle in the queue leaves the reservoir is computed. Similarly, the time at which the next vehicle enters the reservoir is known *a priori* from the demand. Depending on which event, entry or exit, happens first, the next event and time instance are updated. The entry supply, exit demand are applied to the first vehicle in the waiting list to enter the reservoir and the leading vehicle in the queue to exit the reservoir, respectively.

Congestion propagation in the trip-based model is modeled by using the same principles as in the accumulation-based framework. For instance, for a mode m the entry supply time for a vehicle N is imposed by changing the time at which the vehicle can enter the reservoir. This can be expressed as follows,

$$t^{N}_{m,\text{entry supply}} = t^{N-1}_{m,\text{entry}} + \frac{1}{I_m(n_m,n)} \quad \forall m, \tag{17}$$

where $t^{N-1}_{m,\text{entry}}$ is the entry time of previous vehicle $N-1$ for mode m and $t^{N}_{m,\text{entry supply}}$ is the time at which, vehicle N can enter the reservoir. As stated in the context of the accumulation-based model, the FIFO-based entry flow function is proposed in the framework of the trip-based model as well. As will be seen in the discussion, implementing FIFO-based entry in the trip-based framework is more straightforward compared to the accumulation-based model.

2.5. Delay accumulation-based model

The principal hypothesis of the delay accumulation-based model is that the outflow is delayed in the reservoir by the order of travel time at any time instance, t. This model was first introduced by Ran *et al.* (1993); Friesz *et al.* (1989) in the context of link level traffic dynamics and later used by Haddad and Zheng (2018); Zhong *et al.* (2018) in the context of MFD-based modeling. Consider vehicles that enter the reservoir at time t, at an inflow of $q_{m,\text{in}}(t)$ for mode m. They leave the reservoir at time $t+\tau_m(t)$, where $\tau_m(t)$ is the travel time inside the reservoir for the mode m. Under the assumptions of vehicle conservation and FIFO rule, vehicles of mode m that enter at time t must be equal to vehicles that leave the reservoir at time $t+\tau_m(t)$. Mathematically, it can be expressed as,

$$\int_{-\infty}^{t} q_{m,\text{in}}(s)\,\mathrm{d}s = \int_{-\infty}^{t+\tau_m(t)} q_{m,\text{out}}(s)\,\mathrm{d}s. \tag{18}$$

Differentiating the eq. (18) with respect to t and rearranging yields,

$$q_{m,\text{out}}(t+\tau_m(t)) = \frac{q_{m,\text{in}}(t)}{1+\dfrac{\mathrm{d}\tau_m(t)}{\mathrm{d}t}}. \tag{19}$$

Since, $\dfrac{\mathrm{d}\tau_m(t)}{\mathrm{d}t}$ is not known *a priori*, it can be computed using chain rule as follows,

$$\frac{\mathrm{d}\tau_m(t)}{\mathrm{d}t} = \sum_{i=\{c,b\}} \frac{\partial \tau_m}{\partial n_i}\frac{\mathrm{d}n_i(t)}{\mathrm{d}t} \equiv \sum_{i=\{c,b\}} \frac{\partial \tau_m}{\partial n_i}\left(q_{i,\text{in}}(t) - q_{i,\text{out}}(t)\right). \tag{20}$$

The travel time function can be obtained from the velocity MFD, *i.e.*, $\tau_m(n_m,n) = \dfrac{L_m}{v(n_m,n)}$. As the trip length of each mode, L_m, is constant and $v(n_m,n)$ is a well-defined analytical function, a non-linear travel time function is obtained. Moreover, this obtained travel time function is well-defined and continuous with respect to accumulation. It is evident that as the accumulation reaches jam accumulation and travel time tends to infinity. This poses stability problems for cases with severe supply restrictions. The entry and exit flow functions in this framework are treated in the same way as in the case of the accumulation-based model and hence, details are omitted.

Notice that the conventional accumulation-based model uses production MFD as the input to the model, whereas in the delay accumulation-based model velocity MFD is embedded in the travel time function.

For the case of linear travel time function with only one mode, *i.e.*, $\tau = \alpha + \gamma n$, where α is the free-flow travel time and constant inflow rate of q_{in}, it is proven in Carey and McCartney (2002), the delay outflow at any time t_n can be expressed as,

$$\bar{q}_{out} = q_{in} \frac{\sum_{i=0}^{n-1} (q_{in}\,\gamma)^i}{\sum_{i=0}^{n} (q_{in}\,\gamma)^i}, \tag{21}$$

where $\bar{q}_{out}$ is the delayed outflow at time $t + \tau_m(t)$. Expanding the expression (21) for each time step results in the series of discontinuous jumps with outflow remaining constant within each segment. Hence, if the inflow demand is constant, delayed outflow reaches the steady state by finite jump discontinuities. If the travel time function is non-linear, the outflow profile is more complex within the discontinuous jumps. This phenomenon has implications during congestion spill-backs, *i.e.*, when the entry flow is limited by the reservoir capacity, these finite jumps can become unbounded and consequently, diverging the numerical solution.

From eq. (19), it is evident that outflow at time $t + \tau_m(t)$ is non-negative if and only if $\frac{\mathrm{d}\tau_m(t)}{\mathrm{d}t} > -1$. Physically, this restriction translates to strict FIFO behavior inside the reservoir. However, it is already shown in Carey and McCartney (2002) that certain inflow profiles, mainly decreasing demand functions, may result in $\frac{\mathrm{d}\tau_m(t)}{\mathrm{d}t} \leq -1$ and hence, violation of internal FIFO. By observing eq. (20), it is clear that $\frac{\partial \tau_m}{\partial n_i}$ is always positive as travel time function is always increasing. Hence, $\frac{\mathrm{d}\tau_m(t)}{\mathrm{d}t}$ is negative if and only if $(q_{i,\mathrm{in}}(t) - q_{i,\mathrm{out}}(t)) < 0$, *i.e.*, during unloading. Daganzo (1995) proved that internal FIFO is preserved in this model if and only if travel time function is a linear. In a simplified case of 2D-MFD, if travel time function is assumed to be linear, the corresponding production MFD can be expressed as $\frac{Ln}{\alpha + \gamma n}$, where L is trip length. Evidently, this type of production MFD cannot be used in practical applications, where real traffic data tends to follow piece-wise parabolic or piece-wise linear relationship between mean flow and accumulation. On the other hand, using a non-linear (for example parabolic) MFD results in non-linear travel time function, which can violate the internal FIFO discipline during unloading. This can be regarded as a strong limitation in this modeling framework. However, this work addresses this restriction by proposing a new concept: *weak internal FIFO discipline*. This is achieved by re-constructing the outflow cumulative curve locally whenever there is a violation in the internal FIFO discipline. The principle of weak internal FIFO discipline is discussed in-detail in the numerical examples of the following section.

All the above stated MFD-based models assume a well-defined MFD for the network under consideration. However, the link heterogeneity in the urban networks introduces a considerable amount of scatter in the MFD (Ramezani *et al.*, 2015). The heterogeneity in the link density can cause hysteresis loops, which are typically characterized by network performance during loading and unloading phases. It is shown that considering the hysteresis phenomenon is important for

accurate resolution of traffic dynamics in Paipuri *et al.* (2019) in the context of 2D-MFD using micro-simulation study of a real network. The stated work can be extended to the case of 3D-MFD by appropriately calibrating the MFD surfaces by taking the hysteresis loops into account.

3. Bi-modal MFD model solutions and extensions

3.1. Test cases description

Test cases in the present work assume a reservoir with two modes namely, cars and buses. Production and velocity 3D-MFDs computed based on the coefficients presented in Table 1 are used in all examples, unless otherwise stated. The trip lengths of cars and buses, L_c and L_b, are assumed to be 1 000 m and 2 000 m, respectively. A total simulation time of 10 000 sec is considered and the first 1 000 sec is warm-up period.

This work aims to study the MFD-based models during transitional regimes for different scenarios. MFD-based models are known to behave well when the inflow varies slowly in time as this complies with the steady state approximation upon which the definition of MFD is based in a reservoir. As steep varying inflows are common in the network loadings, it is necessary to consider such conditions to see how the traffic dynamics are resolved in the transitional phase. The current work is restricted to the limit case of step demand function with sudden increase and later, sudden decrease of inflow. This magnifies the potential shortcomings during transitional regimes, which in turn enables them to address. Two demand scenarios are considered: a low demand case, where traffic states are confined to the free-flow side of 3D-MFD. The fundamental differences between modeling approaches during the transition regime are demonstrated using this example. The second case consists of higher demand, where congestion spill-back reaches the entry of the reservoir. The limitations of each model are studied under this scenario and appropriate solutions are discussed.

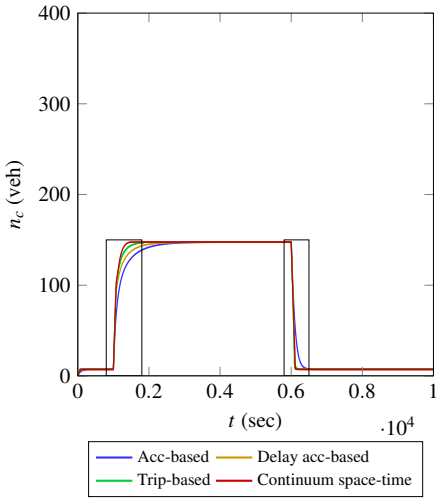
3.2. Low demand free-flow scenario

Firstly, a free-flow scenario is considered to demonstrate the essential differences between different MFD-based models. A step demand case is assumed for the inflow profile according to the following definition,

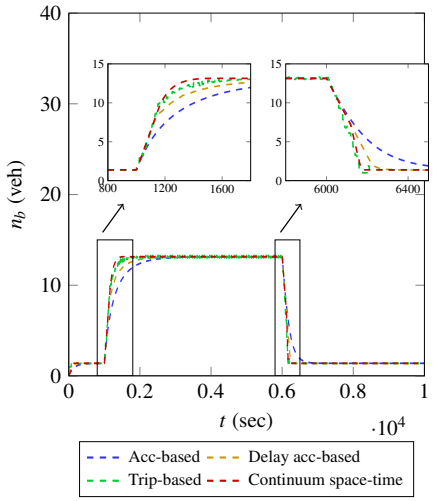
$$[\lambda_c,\, \lambda_b]^T = \begin{cases} [0.1,\, 0.01]^T & \text{if } 0 < t \leq 1000 \text{ or } 6000 < t \leq 10000 \\ [1.3,\, 0.06]^T & \text{otherwise,} \end{cases} \tag{22}$$

where λ_c and λ_b are demands for cars and buses in veh s^{-1}, respectively. These flow values are chosen in such a way that the accumulations of cars and buses stay on the left hand side, *i.e.*, in the free-flow regime, of the critical accumulation line on the MFD surface shown in Fig. 2a.

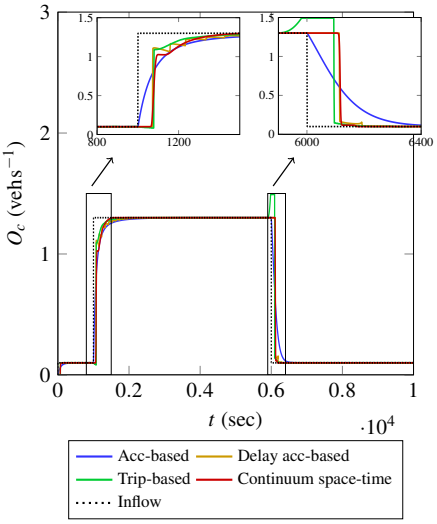
Figure 4 presents the evolution of accumulation, outflow and mode share ratio with time for cars and buses. In this case, no spill-backs are observed in the reservoir and inflow follows the demand profile. Therefore, plots of inflow evolution with time are omitted in this case. Figs. 4a



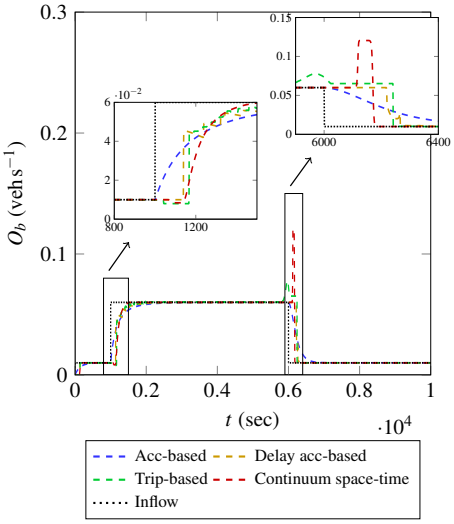
(a) Accumulation *vs.* time for cars.



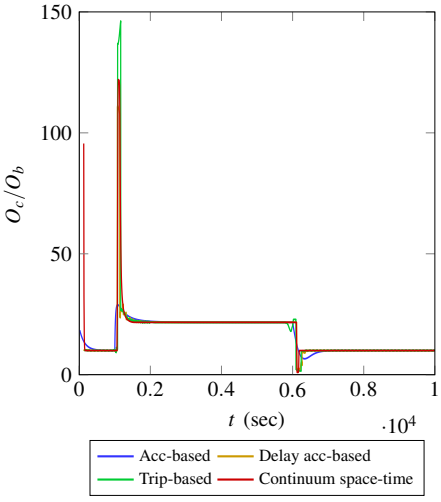
(b) Accumulation *vs.* time for buses.



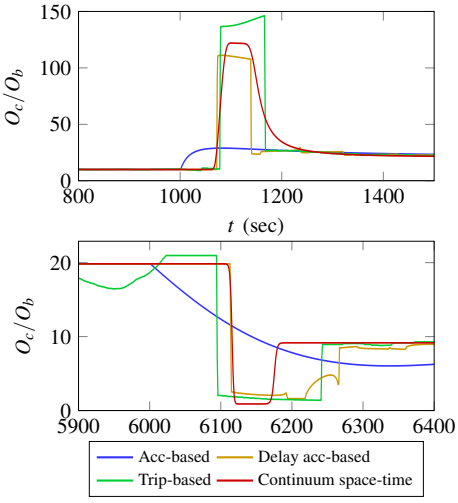
(c) Outflow *vs.* time for cars.



(d) Outflow *vs.* time for buses.



(e) Modeshare ratio *vs.* time.



(f) Details of evolution of mode share at transition periods. Top plot is at demand surge and bottom one is at demand drop.

Figure 4: Low demand free-flow scenario: Evolution of accumulation, outflow and mode share with time for cars and buses for all MFD-based models considered.

Table 2: Low demand free-flow scenario: Relative errors in $\mathscr{L}_2$ norm for accumulation and outflow for MFD-based models with respect to continuum space-time model.

Model	Accumulation		Outflow	
	Cars	Buses	Cars	Buses
Accumulation-based	0.0688	0.0927	0.0729	0.1552
Delay accumulation-based	0.0293	0.0434	0.0238	0.1360
Trip-based	0.0162	0.0219	0.0638	0.1447

and 4b show the evolution of accumulation with time for cars and buses, respectively. According to the assumed demand profile (22), there is a sharp increase in the demand at $t = 1\,000\,\text{sec}$ and correspondingly, a sharp decrease at $t = 6\,000\,\text{sec}$. Although all the models presented in the plots reach the same equilibrium, it is interesting to observe the transition region to study the fundamental differences between models. Consider the demand surge at time $t = 1\,000\,\text{sec}$ highlighted in Figs. 4a and 4b. Since the continuum space-time model is the reference solution, it is evident from the plot that the trip-based model follows the reference solution more closely followed by the delay accumulation-based model and finally, the classical accumulation-based model is most diffusive or least accurate. Since there is no explicit space variation term in the accumulation-based model, only space averaged solution is obtained at each time step, which results in a more diffused solution. On the other hand, the trip-based model accounts for the wave propagation inside the reservoir to a reasonable accuracy, which can be observed from the plot. These results simply confirm existing observations with the uni-modal formulation of these two MFD-based models (Mariotte *et al.*, 2017; Mariotte, 2018).

Delay accumulation-based is in between these two models, *i.e.*, this model cannot account for wave propagation like trip-based, however, the delay in the outflow introduces an average effect of the wave propagation. The essential difference between trip-based and delay accumulation-based is that the former can account for the variation of travel time during a trip, whereas the latter cannot account for this variation within a trip. The delay (or travel time) is computed at the instance vehicle enters the reservoir and it does not change during the trip. The same applies during the demand drop, *i.e.*, at $t = 6\,000\,\text{sec}$, albeit, the transition to the equilibrium is rather fast and hence, the differences in models are less obvious. However, the trend can be observed clearly in the case of buses in Fig. 4b, where the trip-based is more accurate and the accumulation-based is less accurate. This shows that an accurate representation of delays between perimeters is even more crucial for bi-modal flows as buses and cars have usually different trip lengths inside the reservoir.

Table 2 presents the relative errors in $\mathscr{L}_2$ norm for the MFD-based models with respect to the reference solution, *i.e.*, continuum space-time model. The solution of each MFD model is computed at different time steps and so, in order to estimate the errors, all the solutions are

projected on to a reference equidistant time intervals. The projection operation, in this case, is done by a simple linear interpolation and hence, additional errors introduced are of the same order. It is clear from the relative error values that the trip-based gives most accurate while the classical accumulation-based model is least accurate when the evolution of accumulation is considered. Error in the accumulation for the delay accumulation-based is in-between the trip-based and the accumulation-based, rather closer to the trip-based model.

Figures 4c and 4d present the evolution of outflow with time for different MFD-based models considered for cars and buses, respectively. The inflow demand is shown in dotted lines in the plot, where demand surge and drop can be observed. At $t = 1\,000\,\text{sec}$, it is clear from the plots that the outflow in the accumulation-based model reacts immediately to an increase in inflow. On the contrary, both trip-based model and delay accumulation-based model produce a sharp increase in the outflow around $t = 1\,100\,\text{sec}$, where the difference in time instants *i.e.*, is the time at which inflow is increased and actual time where outflow increase is noticed, is travel time inside the reservoir. Again, this confirms the observations with classical MFD-based models about the important differences between different approaches.

In the delay accumulation-based, a step wise increase in the outflow is observed and this phenomenon is in accordance with eq. (21) and discussed in-detail in Carey and McCartney (2002). The reason for non-constant steps in the increase of outflow in the delay accumulation-based model is due to the non-linearity in travel time function. Similarly, during the demand drop at $t = 6\,000\,\text{sec}$, the outflow in the accumulation-based model starts to decrease gradually, whereas the trip-based and the delay accumulation-based produce a sharp decrease after accounting for the travel time inside the reservoir. An increase in the outflow in the trip-based model can be observed just before the demand drop. This is due to the causality effect that is discussed in-detail in Leclercq and Paipuri (2019). In brief, the causality effect is noticed when demand drop occurs, which results in a decrease of accumulation inside the reservoir and an increase in the mean speed of all vehicles. Thus, the mean speed of vehicles close to exit also increases, which results in an increase of outflow before equilibrium is reached. A small increase in the outflow in the buses is noticed at the end of the demand peak period. This is due to the fact that cars finish their trip faster than buses owing to the smaller trip lengths. Hence, demand drop occurs in cars sooner than buses, which makes more outflow capacity available for buses and hence, an increase in outflow for buses is noticed. Now considering the relative error values in outflow, it can be noticed that the trip-based model has a larger error than the delay accumulation-based model. This is again due to the causality effect before the end of the demand peak period. Since the delay accumulation-based model cannot account for the variation of travel time or mean speed during the trip, the new vehicles entering the reservoir have no effect on existing downstream vehicles in the reservoir. Hence, the causality effect is less observed in the delay accumulation-based model and produces a more accurate solution with respect to the reference solution.

Finally, differences between models during the transition period can be clearly noticed by

monitoring the mode share ratio, *i.e.*, the ratio of outflow of cars to buses, which is an important factor for multi-modal simulations. Fig. 4e presents the mode share ratio evolution with time. The equilibrium values of ratio of outflows computed from inflow (22) during demand surge and warm-up periods are 21.67 and 10, respectively. During demand surge, the ratio increases sharply to a value higher than 100 in all models except the accumulation-based, stays at that value for a certain time and then converges to the equilibrium state. This sharp increase is due to the fact that the outflow of cars increases prior to the buses (owing to the shorter trip length of cars) and therefore, the ratio of outflows during this period should be 130 from the inflow profile (22). Fig. 4f shows that the continuum space-time model estimates this mode share ratio of 130, while the trip-based and the delay accumulation-based models over-predicts and under-predicts, respectively. The inaccuracies in the trip-based model are due to the rounding-off and interpolation errors. In the trip-based model, the outflow is re-constructed from the exit times of each vehicle and this re-construction procedure introduces some rounding-off errors. On the other hand, the outflow in the delay accumulation-based model increases through finite discontinuous jumps and tends to equilibrium through these series of jumps. Hence, a small under-prediction is observed owing to inaccuracies introduced by these jumps. Once the outflow of buses reacts to the inflow surge, the ratio tends to the equilibrium ratio of 21.67. The exact opposite behavior is observed during the demand drop, where the outflow of cars decreases first, followed by buses. However, the accumulation-based model is unable to predict this phenomenon and it is evident from Fig. 4f that during both demand drop and demand surge phases, the ratio of outflows gradually converges to equilibrium ratio in the accumulation-based model. The accumulation-based formulation not only fails to predict the dynamic evolution of mode share but also takes a long time to converge to the equilibrium ratio. This is a significant drawback for bi-modal traffic simulations. This can be noticed during the demand drop period, where at the end of 6400 sec, the accumulation-based model still has not reached the equilibrium ratio.

3.3. High demand congested-flow scenario

This scenario is chosen in such a way that congestion appears inside the reservoir because of high demand and spill-back reaches the entry of the reservoir limiting the inflow. This is achieved by increasing the inflow demand of buses relative to the demand of cars. A similar step demand case as in the previous example is considered as follows,

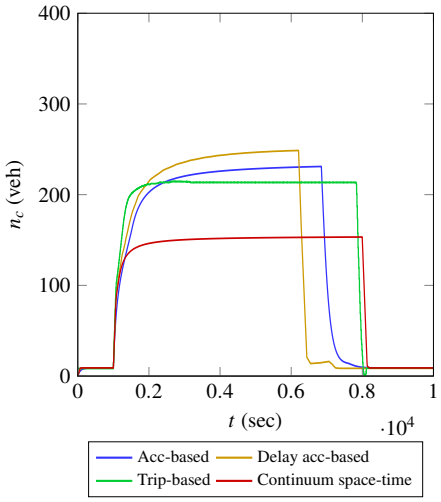
$$[\lambda_c,\, \lambda_b]^T = \begin{cases} [0.1,\, 0.01]^T & \text{if } 0 < t \leq 1000 \text{ or } 6000 < t \leq 10000 \\ [1.3,\, 0.13]^T & \text{otherwise.} \end{cases} \tag{23}$$

In this case, the inflow demand of buses is increased to $0.13\,\mathrm{veh\,s^{-1}}$ from $0.06\,\mathrm{veh\,s^{-1}}$ in the free-flow scenario. As the number of buses increase in the reservoir, the capacity of the network reduces according to the assumed 3D-MFD and therefore, congestion spill-backs appear in the reservoir.

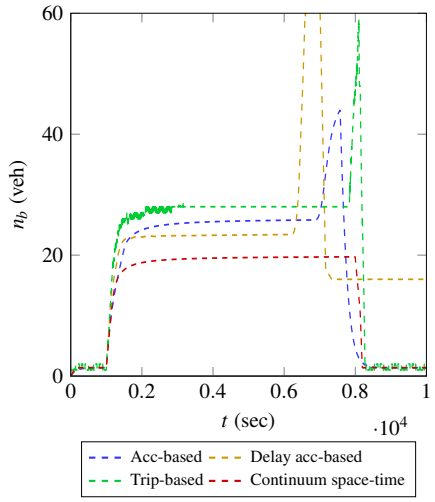
Figure 5 shows the evolution of accumulation, inflow and outflow for the different MFD-based models along with the reference solution. It is clearly evident that all models are not only different during transition regimes, but also, importantly, they reach different steady states on the 3D-MFD plane. This example illustrates the essential differences in the modeling framework between a conventional 2D-MFD and a bi-modal 3D-MFD.

Firstly, consider the solution of the continuum space-time model. The inflow reaches the given demand of $1.3\,\mathrm{veh\,s^{-1}}$ and almost immediately drops to an equilibrium value of $1.15\,\mathrm{veh\,s^{-1}}$. In the continuum space-time model, the space dimension is discretized into a finite number of cells and the inflow dynamics are evolved based on the solution in the first cell of the left hand boundary. As soon as the inflow of cars and buses increases after $t = 1\,000\,\mathrm{sec}$, the density of cars and buses in the first cell reaches the maximum allowable value based on the 3D-MFD and the inflow adapts accordingly to sustain at this equilibrium value. On the contrary, in the case of other MFD-based models, the inflow adapts based on the average density of the whole reservoir. Hence, the inflow demand of $1.3\,\mathrm{veh\,s^{-1}}$ sustains for a longer period of time before congestion spill-back reaches the entry. In other words, the entry flow function works based on the density value of the first cell in the continuum space-time model, whereas it works based on the average density value in the whole reservoir in the MFD-based models. It can be concluded that it is not possible to compare the continuum space-time model with MFD-based results in the case of congestion spill-backs. Therefore, the results of the continuum space-time model are omitted in the rest of the discussion for the present scenario.

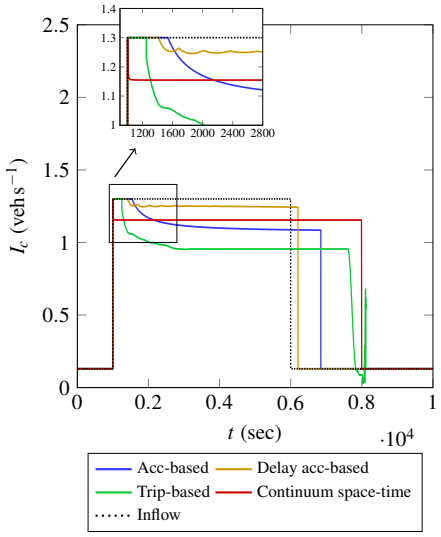
Now consider the other MFD-based models. As explained in Section 2, all three modeling frameworks use the same entry flow function defined in eq. (10) along with constraint (12). However, as stated earlier the critical production in the case of 3D-MFD is not a constant value and it depends on the partial accumulation of cars and buses. On the contrary, the total maximum production in the case of 2D-MFD is always the same irrespective of the partial accumulations on each route. From the previous test case, it is concluded that different MFD-based models have different behaviors during transition periods. Consequently, the ratio of accumulation of cars to buses evolves differently during demand surge. Fig. 6a presents the ratio of buses to total accumulation at the demand surge period. The oscillations in the trip-based model are due to the discreteness of the numerical scheme. It is clear that the ratio evolves differently for different models owing to their conceptual frameworks. As this ratio is directly used in the entry flow function (10), the inflow assumes different values and hence, this difference in the partial accumulations results in different equilibria on the 3D-MFD plane for different models. This phenomenon is absent when a conventional 2D-MFD is used as the critical accumulation is always the same irrespective of the partial accumulations. Observe Fig. 6, which presents the evolution of accumulations of cars and buses on the MFD contour surface. The black line indicates the critical production, *i.e.*, the peak of any parabolic 2D-MFD for a given bus accumulation. The three MFD-based models reach different equilibria states on the surface.



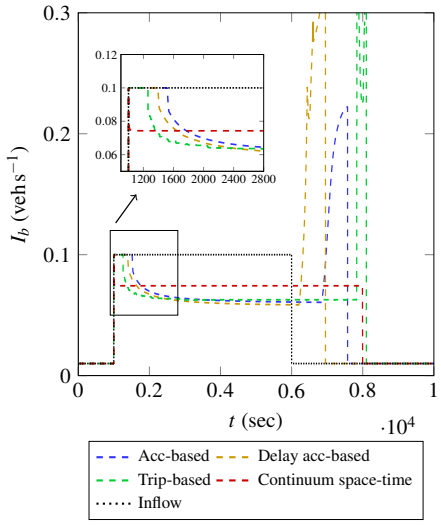
(a) Accumulation *vs.* time for cars.



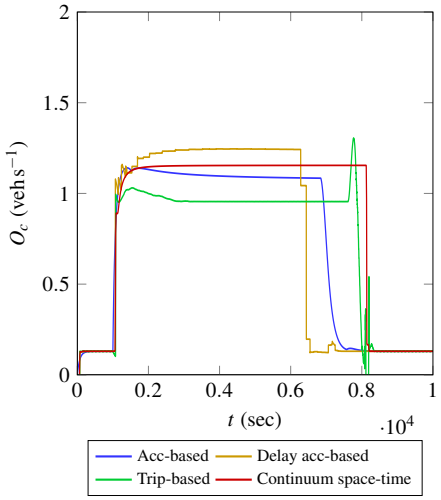
(b) Accumulation *vs.* time for buses.



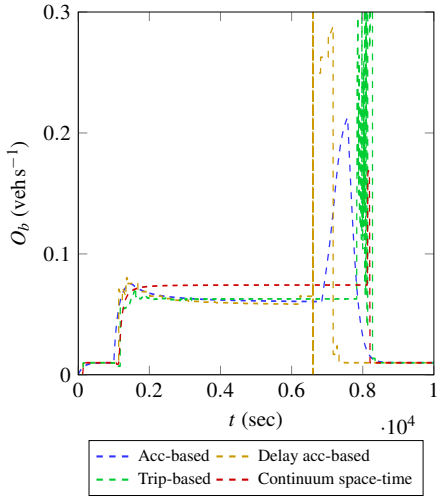
(c) Inflow *vs.* time for cars.



(d) Inflow *vs.* time for buses.

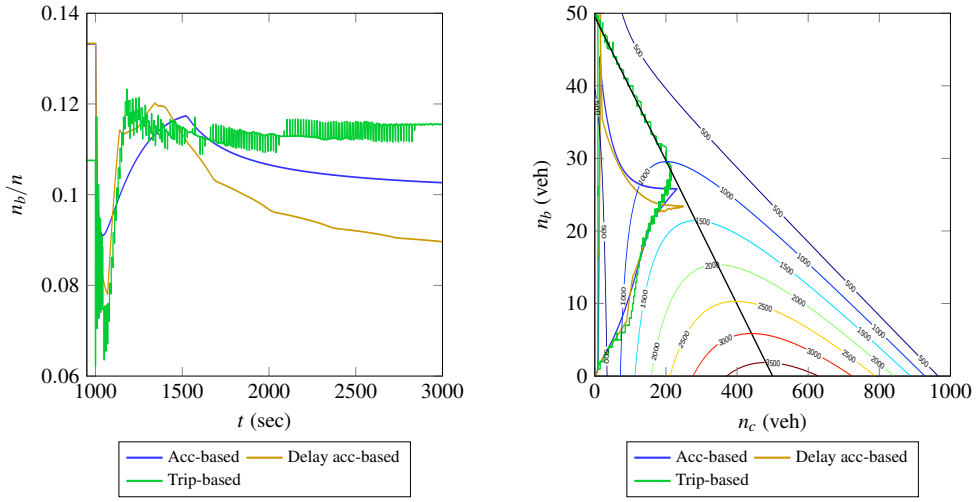


(e) Outflow *vs.* time for cars.



(f) Outflow *vs.* time for buses.

Figure 5: High demand congested-flow scenario: Evolution of accumulation, inflow and outflow with time for cars and buses for all MFD-based models considered.



(a) Bus share ratio *vs.* time for cars.

(b) Evolution of cars and buses accumulation on 3D-MFD plane.

Figure 6: High demand congested-flow scenario: Evolution of bus share and evolution of partial accumulations on MFD plane for all MFD-based models considered.

Another inference worth making in this scenario is the evolution of traffic dynamics during the period of demand drop. It can be noticed from Figs. 5c and 5d that the steady state inflow values sustain for a longer period even after demand drop, *i.e.*, $t = 6\,000$ sec. This is due to the queue discharging phenomenon that happens during congestion. The entry flow function keeps two queues for each mode at the entry of the reservoir and the rate at which the queues can be discharged depends on the partial accumulations (10). According to the present entry flow function, the queue of the buses takes longer to discharge than that of cars. Hence, when the queue of the cars disappears, the entry flow function, due to the constraint on total available production (12), allocates the maximum available production (or capacity) to buses. This is evident from a huge increase in the inflow of the buses for MFD-based models before the end of the queue discharge. It can be observed in the MFD plane as well, where the accumulations tend to go in an upward direction, *i.e.*, going in the direction of increasing bus accumulation and decreasing car accumulation. This is possible only if there are different entry points for cars and buses in the network (separate networks), but not if all routes are mixed, which is often the case in reality. In the case of mixed routes, it is highly unlikely that one mode of the vehicles will have a priority over the other while discharging queues in a congested scenario. When spill-backs reach the entry, the vehicles start to queue in the order of their arrival at the perimeter and enter the reservoir in the same order. Hence, this type of entry flow function violates the FIFO rule at the entry of the reservoir due to inconsistent queue discharge of each mode.

To summarize the discussion above, the entry flow function presented in eq. (10) has implicit dependence of maximum capacity on the mode share ratio. The differences in the evolution of mode share ratio between different models result in different capacities and consequently,

different equilibria. This can be addressed by introducing FIFO discipline at the entry, which guarantees that the current share of inflows between modes matches the demand and it is no longer depends on the internal mode share in the reservoir. This assumption goes well with physical rationale. When vehicles start to queue at the entry of a homogeneous border, which is always the case in the MFD framework, there is no reason for one mode (in absence of dedicated lanes) to be more competitive than the other, *i.e.*, bypass the queue faster. Therefore, when buses and cars are mixed in a congested inflow, they should experience the same delay to reach the reservoir entry, which FIFO discipline enforces in the proposed entry flow function.

FIFO-based entry flow function

The principal idea behind the FIFO-based entry supply function is to build a global queue instead of a local queue per trip at the entry of the reservoir. For example, in the present case of having two modes (or trip lengths) inside the reservoir, the FIFO-based entry flow function uses only one global queue rather than two for each mode. In the case of the trip-based model, implementing the FIFO-based entry flow function is straight-forward. Based on the departure times of the cars and buses, if the congestion spill-back reaches the entry of the reservoir, the vehicles queue at the entry in the order of their departure times, irrespective of the mode. Then, the entry supply time for the leading vehicle in each queue can be defined as follows,

$$t^{N}_{\text{entry supply}} = t^{N-1}_{\text{entry}} + \frac{1}{I_F(n_m, n)} \quad \forall N, \tag{24}$$

where I_F is the FIFO-based entry supply function. The capacity available for entry flow should not exceed the reservoir's and therefore, the entry supply function, I_F, can be computed as follows,

$$I_F(n_m, n) = \frac{P_s(n_m, n)}{L}, \tag{25}$$

where P_s and L are defined in eq (11). FIFO-based entry flow ensures that maximum capacity is always utilized at the entry by forming a global queue. However, implementing this FIFO-based entry flow function in the framework of the accumulation-based models is not trivial.

The present work proposes the design of FIFO-based entry for the accumulation-based models based on the partial cumulative curves, which is based on the work of Chevallier and Leclercq (2008). Fig. 7 shows graphical representation of FIFO-based entry system. In the plot, N^i_c and N^i_b are inflow cumulative curves for cars and buses and likewise, N^d_c and N^d_b are demand cumulative curves for cars and buses, respectively. The total demand and inflow cumulative curves are denoted by N^d and N^c, respectively. Consider the evolution of the system is known until time t^n, for the next time step, t^{n+1}, the inflow for each mode is computed based on eqs. (9) and (10). Let the inflow values "predicted" at time t^{n+1} be $q^{n+1,p}_{c,\text{in}}$ and $q^{n+1,p}_{b,\text{in}}$ for cars and buses, respectively. Assuming Δt to be constant time step, the predicted partial cumulative curves at time t_{n+1} can be

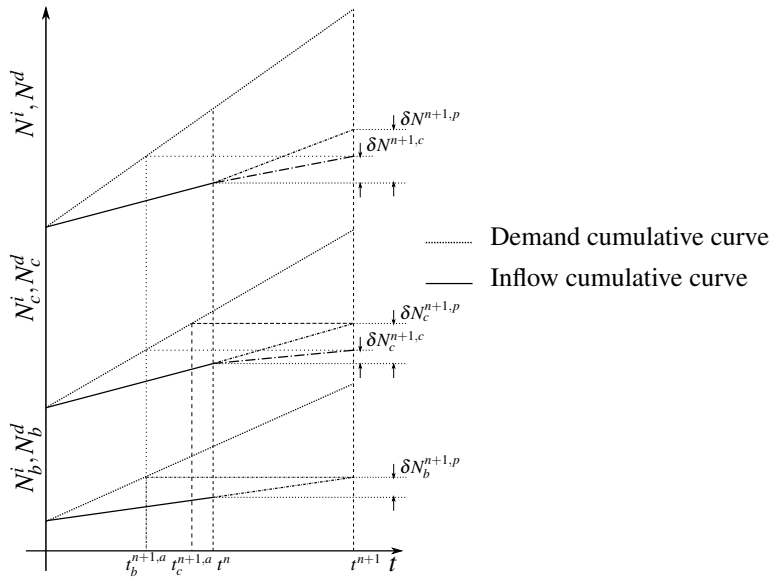


Figure 7: FIFO-based entry supply function: Graphical representation of implementation of FIFO-based entry supply function. N is the cumulative curve, subscripts c, b denote cars and buses, respectively and superscripts d, i correspond to demand and inflow, respectively. The superscripts p and c on the incremental cumulative vehicles denote predicted and corrected values, respectively. Finally, superscript a on time denotes the arrival (or creation) time of the vehicle.

expressed as,

$$N_c^{i,n+1,p} = N_c^{i,n} + \delta N_c^{n+1,p} = N_c^{i,n} + q_{c,\text{in}}^{n+1,p}\,\Delta t, \tag{26a}$$

$$N_b^{i,n+1,p} = N_b^{i,n} + \delta N_b^{n+1,p} = N_b^{i,n} + q_{b,\text{in}}^{n+1,p}\,\Delta t. \tag{26b}$$

It is evident from the graph that the actual inflow is lower than the demand, which results in the formation of a queue at the entry of the reservoir. Hence, the predicted number of vehicles $\{N_c^{i,n+1,p}, N_b^{i,n+1,p}\}$ at time t^{n+1} had been created at time $\{t^{n+1} - t_c^{n+1,a}, t^{n+1} - t_b^{n+1,a}\}$, respectively. In other words, cars and buses have a delay (or waiting time) of $\{t^{n+1} - t_c^{n+1,a}, t^{n+1} - t_b^{n+1,a}\}$, respectively. If the FIFO rule were to hold true, the ratio of the number of vehicles at time t^{n+1} should be the same as the time of the creation of each mode, *i.e.*, vehicles for each mode created at the same time should experience the same delay before entering the reservoir. The mode that experiences maximum delay is chosen and flow values are "corrected" to obey FIFO discipline in the queue. In the graph, it is clear that the buses have more delays according to the predicted inflow values. It is possible to obtain the cumulative number of cars at time $t_b^{n+1,a}$ and the corrected inflow is $\dfrac{\delta N_c^{n+1,c}}{\Delta t}$. Now, it can be verified graphically that the ratio of the number of vehicles at time t^{n+1} and time of arrival $t_b^{n+1,a}$ is the same. Similarly, the total inflow can also be corrected by setting the total number of vehicles at time t^{n+1} to be equal to total vehicle count at time $t_b^{n+1,a}$. In a nutshell, the proposed FIFO-based entry supply can be briefed as follows,

- Compute the predicted inflow values based on the entry supply function (10).
- Calculate the delay for each mode by estimating at which time instance the demand cumu-

lative curve has the same number of vehicles as inflow cumulative curve at time t^{n+1} for each mode.

- The corrected inflow values correspond to the mode which experiences a maximum delay. Estimate the value of the cumulative number of vehicles at the time of the maximum delay for two modes.

- Finally, compute the corrected inflow values from the cumulative curve value at time t^n and the new corrected cumulative number of vehicles at time t^{n+1}.

This prediction-correction strategy can be applied to both conventional accumulation-based model and delay accumulation-based models. Consider a case of the constant demand of λ_c, λ_b for cars and buses, respectively. As assumed earlier, let the predicted cumulative number of vehicles for each mode be $\{N_c^{i,n+1,p}, N_b^{i,n+1,p}\}$. The arrival time for each mode can be computed as,

$$t_c^{n+1,a} = \frac{N_c^{i,n+1,p}}{\lambda_c}, \tag{27a}$$

$$t_b^{n+1,a} = \frac{N_b^{i,n+1,p}}{\lambda_b}, \tag{27b}$$

owing to a simple constant demand. Depending on the arrival times of each mode, the inflow based on FIFO discipline at time t^{n+1} can be expressed as follows,

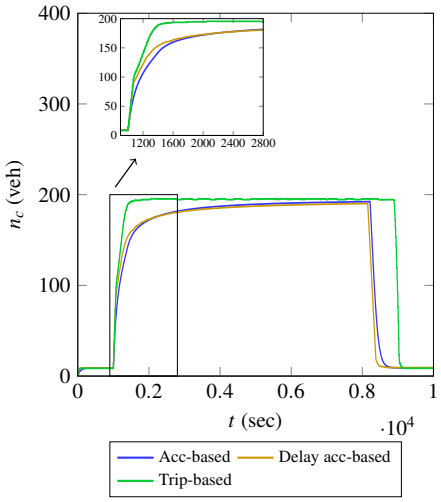
$$q_{\text{in}}^{n+1} = \begin{cases} \left[\dfrac{N_c^{i,n+1,p} - N_c^{i,n}}{\Delta t}, \quad \dfrac{N_c^{i,n+1,p}\,\dfrac{\lambda_b}{\lambda_c} - N_c^{i,n}}{\Delta t} \right]^T & \text{if } t_c^{n+1,a} = \min(t_c^{n+1,a}, t_b^{n+1,a}), \\ \left[\dfrac{N_b^{i,n+1,p}\,\dfrac{\lambda_c}{\lambda_b} - N_c^{i,n}}{\Delta t}, \quad \dfrac{N_b^{i,n+1,p} - N_c^{i,n}}{\Delta t} \right]^T & \text{if } t_b^{n+1,a} = \min(t_c^{n+1,a}, t_b^{n+1,a}). \end{cases} \tag{28}$$

Note that the solution presented in eq. (28) is possible because of the simplicity assumed in the demand profile. A simple linear interpolation is used to compute arrival times for each mode in (27) in the case of constant demand, which yields an exact solution. Whereas in the case of non-constant demand, a higher order interpolation is desired to minimize the errors. However, the underlying time integration scheme introduces errors of its own to the numerical solution. In the current work, the first-order forward Euler scheme (second-order accurate within each time step) is used to discretize the ODEs. A linear interpolation introduces errors of order 2 as well. Since the errors introduced by both time integration and FIFO-entry rule are of the same order, the convergence of the entire method is unchanged. If higher order time integration schemes like Backward Differentiation Formulae (BDF), Runge–Kutta (RK) are used, higher order interpolation schemes must be used in the FIFO-based entry supply functions. Therefore, it is concluded from this work that linear interpolation is sufficient and gives satisfactory results.

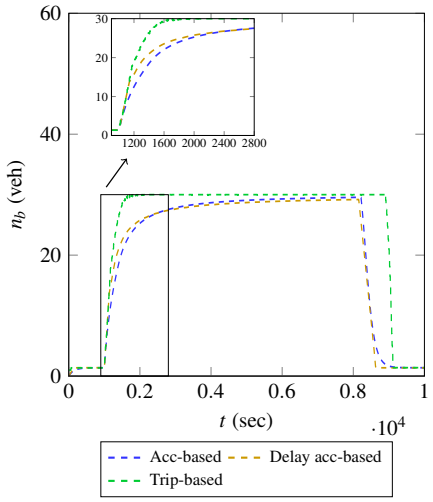
The proposed FIFO-based entry supply function uses eq. (10) to predict the inflow values before correcting them using the FIFO rule. Hence, an accurate and realistic entry flow function is necessary for accurate modeling of congestion spill-backs. The entry supply function

defined in eq. (10) is a straightforward extension from 2D-MFD, which is in turn based on CTM model. Mariotte and Leclercq (2019) discussed in-detail the calibration of the entry flow function in the context of 2D-MFD. The stated work concluded that the shape of the entry supply function is characterized by higher critical capacity and critical accumulation compared to the underlying MFD. This conclusion is verified using micro-simulations on a rectangular grid network in Mariotte *et al.* (2019). This study can be extended to the case of 3D-MFD as well, where the entry supply function can be calibrated based on the total capacity of links that are entering the assumed reservoir. The proposed FIFO-based entry rule can be applied to any generic entry flow function, irrespective of its shape. Another important aspect is to consider a reservoir with a perimeter separated in several borders (collection of entry points), which can arise in the multi-reservoir MFD-based simulations. In this case, each border capacity can be estimated using *pro rata* demand merge coefficient (Mariotte and Leclercq, 2019). Then, the FIFO-based entry flow function can be applied to every border and each mode. Even though the predicted inflows at different entry points for different models are different, the inflow at each entry is driven by the FIFO discipline, which results in unique inflow evolution for all models. As the bi-modal share for all separate borders evolves the same way in all the models, the steady state is unique irrespective of the model. Hence, the extension of this framework to multi-reservoir simulations is also straight-forward.

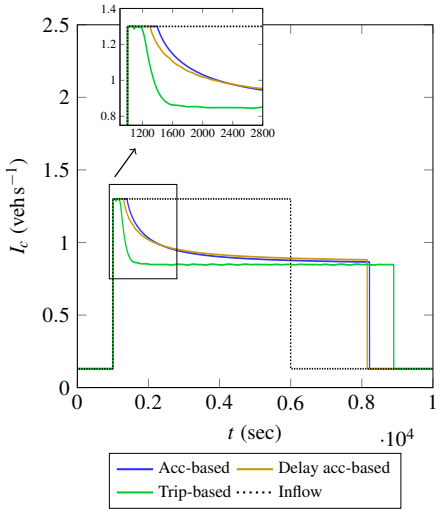
Figure 8 presents the evolution of accumulation, inflow and outflow with FIFO-based entry flow for MFD-based models. The first remark to be made here is that all models reach, as expected, the same steady state irrespective of their differences in evolution during transition periods. Observing the evolution of accumulation in Figs. 8a and 8b for cars and buses, respectively, it is clear that the steady-state value is approximately the same for all three models. The trip-based model has slight oscillations that come from smoothening of the solution. Both accumulation-based and delay accumulation-based models behave very similarly in the transition and the differences are less obvious compared to the free-flow case. This is due to averaging the outflow in the delay accumulation-based model to avoid unstable oscillations. This phenomenon is discussed in-detail in the following part of the section. Inflow evolution shown in Figs. 8c and 8d conclude that the inflow drop occurs in the trip-based model first followed by the delay accumulation-based and the accumulation-based models. This is due to the fact that the shock wave propagation is estimated more accurately in the trip-based model. Also, it can be observed that the drop in inflow is sharper in the trip-based and diffused in the accumulation-based models. This is also an expected conclusion based on theoretical frameworks of the models. Figs. 8e and 8f present the evolution of outflow and an important point worth making here is the peak outflow reached by the trip-based model. It is evident from the plot that both accumulation-based models experience a higher outflow than trip-based model during demand surge. This is the result of two different phenomena that are inherent to trip-based namely, faster propagation of information to the entry and causality effect. Even though the trip-based model allows the maximum inflow, *i.e.*, demand



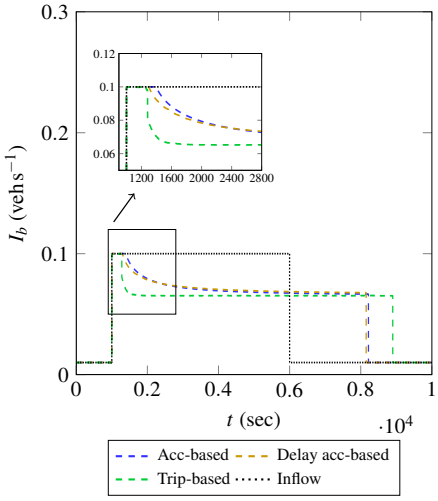
(a) Accumulation *vs.* time for cars.



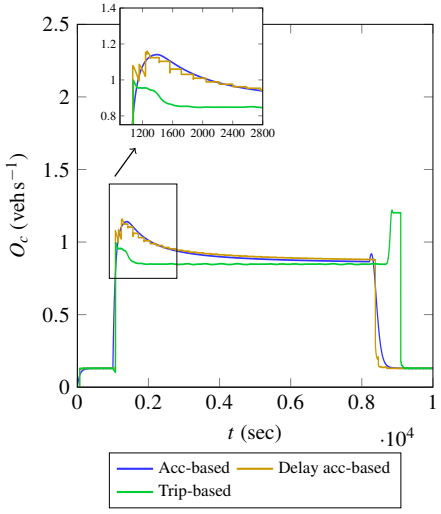
(b) Accumulation *vs.* time for buses.



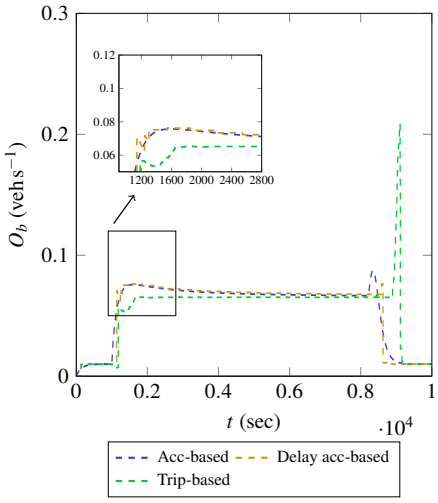
(c) Inflow *vs.* time for cars.



(d) Inflow *vs.* time for buses.

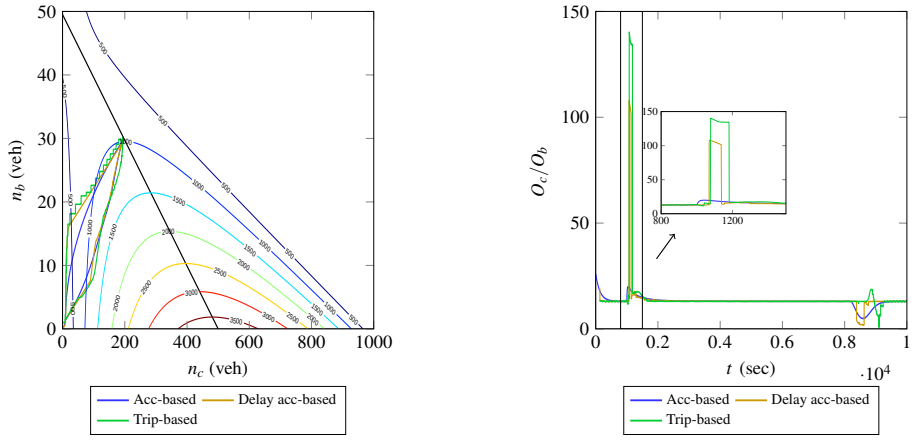


(e) Outflow *vs.* time for cars.



(f) Outflow *vs.* time for buses.

Figure 8: High demand congested-flow scenario: Evolution of accumulation, inflow and outflow with time for cars and buses with FIFO discipline at the entry for all MFD-based models considered.



(a) Evolution of accumulation on MFD surface.

(b) Outflow mode share ratio.

Figure 9: High demand congested-flow scenario: Evolution of accumulation on MFD surface and mode share in outflow.

for a short period of time before dropping, the mean speed of these vehicles is reduced by the time they reach the exit of the reservoir due to the reverse causality effect. Finally, Fig. 9a shows the evolution of traffic dynamics on the MFD surface. Unlike in the case of conventional entry flow presented in Fig. 6b, it is clear that all MFD-based models follow the same trend on the MFD surface. All MFD models reach the same equilibrium point on the surface and the differences between the models are noticed in the transition period. Similarly, the mode share in the outflow is presented in Fig. 9a and as concluded in the case of free-flow, the accumulation-based model is unable to capture transition dynamics accurately. On the other hand, delay accumulation-based model and trip-based give satisfactory results.

Weak internal FIFO discipline and outflow stabilization

There are few important remarks to be made in the case of the delay accumulation-based model. As presented earlier, this model is consistent and stable in the free-flow regime. However, it suffers from stability issues and internal FIFO violation in the case of congested scenarios. The stability in this context refers to numerical stability, where the underlying numerical scheme tends to diverge. The results presented in Figs. 8 and 9 correspond to the stabilized delay accumulation-based model. In the following, the stability issues are explained along with the stabilization techniques proposed in the present work.

Notice from eq. (19), outflow at time $t+\tau_m(t)$ is proportional to inflow at time t and proportionality constant is $\left(1+\frac{\mathrm{d}\tau_m}{\mathrm{d}t}\right)^{-1}$. As outflow cannot be negative, one of the conditions for a stable numerical scheme is $\frac{\mathrm{d}\tau_m}{\mathrm{d}t} > -1$ as discussed earlier. In the present case of two modes, this

condition can be expressed as,

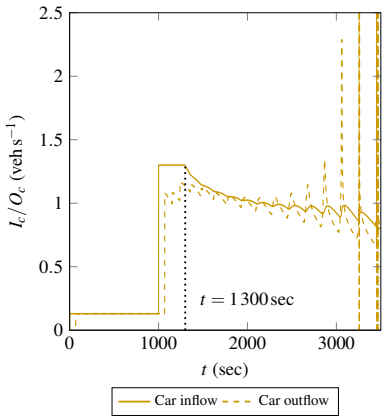
$$\frac{\mathrm{d}\tau_c}{\mathrm{d}t} = \frac{\partial \tau_c}{\partial n_c}\left(q_{c,\mathrm{in}}(t) - q_{c,\mathrm{out}}(t)\right) + \frac{\partial \tau_c}{\partial n_b}\left(q_{b,\mathrm{in}}(t) - q_{b,\mathrm{out}}(t)\right) > -1 \quad \text{for cars,} \tag{29a}$$

$$\frac{\mathrm{d}\tau_b}{\mathrm{d}t} = \frac{\partial \tau_b}{\partial n_c}\left(q_{c,\mathrm{in}}(t) - q_{c,\mathrm{out}}(t)\right) + \frac{\partial \tau_b}{\partial n_b}\left(q_{b,\mathrm{in}}(t) - q_{b,\mathrm{out}}(t)\right) > -1 \quad \text{for buses,} \tag{29b}$$

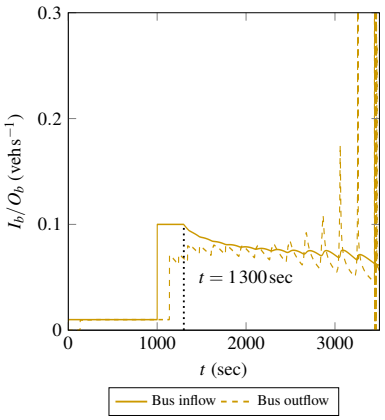
where subscripts c and b denote cars and buses, respectively. It is evident that both τ_c and τ_b are increasing functions with respect to n_c and n_b and hence, their derivatives are always positive. In the case of increasing inflow demand, the conditions (29) are always satisfied as $(q_{c,\mathrm{in}}(t) - q_{c,\mathrm{out}}(t)) \geq 0$. However, if $(q_{c,\mathrm{in}}(t) - q_{c,\mathrm{out}}(t)) < 0$ or $(q_{b,\mathrm{in}}(t) - q_{b,\mathrm{out}}(t)) < 0$, the resulting delay outflow might be negative, thereby yielding physically inadmissible solutions. There can be two situations, where the conditions $(q_{c,\mathrm{in}}(t) - q_{c,\mathrm{out}}(t)) < 0$ or $(q_{b,\mathrm{in}}(t) - q_{b,\mathrm{out}}(t)) < 0$ hold true. The first case being the unloading of the network, where there is a significant drop in the inflow demand with respect to the outflow and the second case arises in the event of congestion spill-backs reaching the entry, which limits the inflow. In the first case, the conditions (29) are violated, which in-turn violates the internal FIFO discipline. Whereas in the second case, the value $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$ tend to be *very close* to -1, which sets the outflow into unbounded oscillations.

Now consider the second case, where the congestion reaches the entry of the reservoir thereby limiting the inflow. As already discussed, outflow tends to the steady state value in a series of discontinuous jumps in the delay accumulation-based model. As the inflow demand increases suddenly, the travel times of the vehicles inside the reservoir increases rapidly resulting in a sharp gradient in the time derivative of the travel time function, $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$. As the outflow asymptotically reaches the inflow, the gradient $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$ decays slowly towards 0 in series of jumps as well. During this period, if inflow drops due to the congestion spill-backs before outflow reaching the steady state, the jumps in $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$ amplify slowly, instead of decaying. If neither inflow nor outflow reaches steady state quick enough, the value of $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$ tend to -1, which leads to unbounded oscillations in the outflow. This phenomenon can be observed in the case, where congestion spill-backs reduce the inflow at the entry. Using the previous example of high demand congested-flow scenario, this stability issue can be demonstrated.

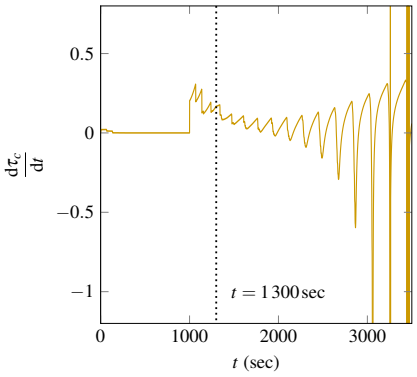
Figures 10a and 10b present the inflow and outflow for cars and buses, respectively. The spill-backs reach the entry of the reservoir around 1 300 sec and inflow starts to decrease. The presence of oscillations in the outflow with the decrease of the inflow can be noticed in the plots. Notice that the jumps in outflow prior to 1 300 sec are the result of the delay accumulation-based model framework, where outflow reaches the steady state in finite discontinuous jumps. This phenomenon is inherent to the delay accumulation-based model and it is stable in the absence of spill-backs. Since the spill-backs reduce the inflow, before the outflow reaches the steady-state, these discontinuous jumps amplify proportional to the reduction in the inflow. This is evident from Figs. 10c and 10d, which present the time derivative of travel time functions of cars and buses, respectively. After $t = 1\,300$ sec, the oscillations become unbounded and eventually violate



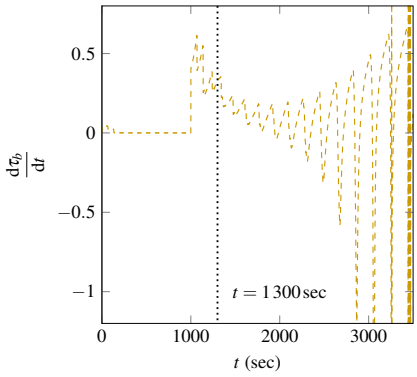
(a) Inflow and outflow of cars.



(b) Inflow and outflow of buses.

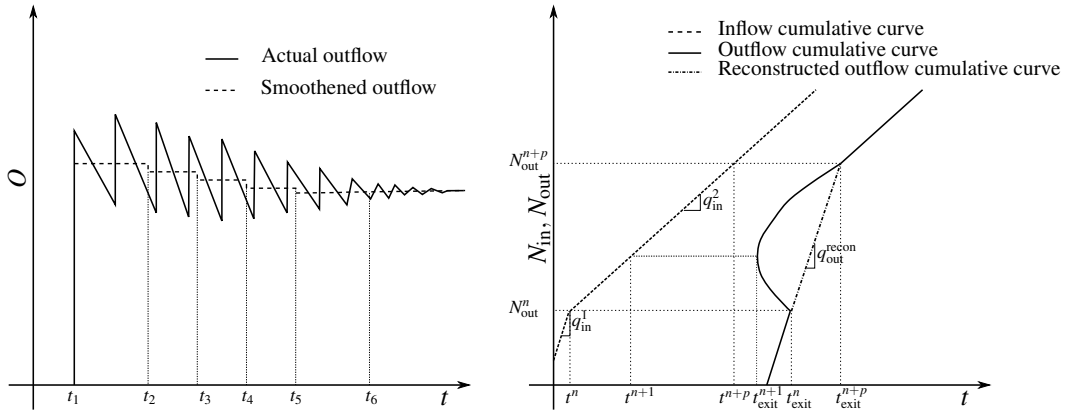


(c) Time derivative of travel time function for cars.



(d) Time derivative of travel time function for buses.

Figure 10: High demand congested-flow scenario: Instabilities in inflow and outflow for the delay accumulation-based model.



(a) Stabilization of oscillations in outflow during transition periods.

(b) Ilustration of outflow re-construction technique when there is internal FIFO violation locally.

Figure 11: Stabilization techniques for the delay accumulation-based model.

the stability condition around 3 000 sec. Previous works that used this model either simplified the model using a constant travel time (Haddad and Zheng, 2018) or illustrated the model using test scenarios (mostly in the free-flow regime) that do not pose stability issues (Zhong *et al.*, 2018). So, no stability problems were reported for the delay accumulation-based model in the context of the MFD-based framework. Hence, it is important to consider the possibility of occurrence of these spurious oscillations under spill-back conditions and stabilize the model.

In this work, oscillations in the outflow are decayed in order to stabilize the delay accumulation-based solution. As discussed earlier, one of the reasons for the amplification of oscillations is that the inflow starts to decrease before the outflow reaches a steady state. This results in a cascading effect in the oscillations that become unbounded. So, the idea behind the stabilization technique is to smoothen the outflow in order to damp the unbounded oscillations. Fig. 11a illustrates the stabilization idea. Consider there is a demand surge at time $t_1 - T$, where T is the travel time. As the delay accumulation-based model reacts to changes in inflow as finite discontinuous steps, it can be represented as shown in the plot as a solid line. In this technique, the outflow is averaged for a certain time period and the outflow value is replaced by the mean outflow in that time period. For example, in Fig. 11a the actual outflow between times t_1 and t_2 is replaced by its mean value (represented in dashed lines). The same is done between times t_2 and t_3 and so on. These time periods can be either fixed or variable based on the evolution of dynamics. For very small periods, the oscillations are not well damped and the model can become unstable. On the contrary, for very large time periods, errors arising from smoothening the solution might affect the traffic dynamics. Using this method, the outflow is in a pseudo steady state within each period and after sufficient time, the actual outflow converges to the smoothened outflow for the case of constant inflow demand. It can be regarded that this stabilization technique is based on a low-pass filter, where high frequency oscillations are filtered out. This can be clearly noticed in Figs. 8e and 8f, where piecewise constant outflows are present in both cars and buses. As all

the stabilization schemes in the numerical methods, this scheme also introduces a few errors of its own to the solution. A brief error analysis is presented in Appendix A. However, the proposed stabilization method is *requisite* in the present context, without which it is impossible to obtain a converged solution. It is also worth noting that this stabilization technique is only used when there are spurious oscillations in the outflow. In the absence of unbounded oscillations, the model reduces to the original delay accumulation-based model.

Now coming back to the first case, where there is a drop in the inflow demand. Depending on the magnitude of demand drop, from eqs. (29), $\frac{\mathrm{d}\tau_m}{\mathrm{d}t}$ can be less than -1. As stated earlier, $\frac{\mathrm{d}\tau_m}{\mathrm{d}t} < -1$ implies the violation of internal FIFO discipline. From eq. (21) the outflow becomes negative and exit times do not follow FIFO rule. As illustrated in Fig. 11b, the internal FIFO rule is violated when the demand decreases from q^1_{in} to q^2_{in}. As the inflow decreases, the number of vehicles inside the reservoir decrease and hence, the speed of the new vehicles that enter the reservoir is higher than the ones that are already inside the reservoir. If the rate at which the vehicles decrease, which depends on the differential between two demand values at the unloading, is above a certain threshold, the newly entered vehicles can overtake the existing ones and leave the reservoir before. Hence, unrealistic flow values are observed for a short period during the unloading phase. However, this is more of a local phenomenon and globally the system will reach a steady state that corresponds to the new demand. This can be illustrated using outflow cumulative curve of reservoir. Fig. 11b shows a case where internal FIFO discipline is violated during the demand drop period using inflow and outflow cumulative curves. It is obvious from the inflow curve that there is a demand drop, *i.e.*, $q^2_{\text{in}} < q^1_{\text{in}}$. Consider that the inflow demand drops at time t^n and the vehicles that enter the reservoir at that time, exit the reservoir at time t^n_{exit}. Since the internal FIFO discipline is violated, vehicles that enter after time t^n, say t^{n+1}, leave the reservoir before t^n_{exit}. This is evident from the plot that $t^n_{\text{exit}} > t^{n+1}_{\text{exit}}$. As stated earlier, this phenomenon happens only in transition period and eventually, the system reaches the steady state. After time t^{n+p}_{exit}, it is clear from the plot that the outflow is in equilibrium with new inflow demand. The period of internal FIFO violation is between t^n_{exit} and t^{n+p}_{exit}. The outflow cumulative curve is re-constructed between in time period, *i.e.*, $\left[t^n_{\text{exit}}, t^{n+p}_{\text{exit}}\right]$ in order to obtain admissible solutions. As times t^n_{exit} and t^{n+p}_{exit} and corresponding N^n_{out} and N^{n+p}_{out} are known *a priori*, the outflow is computed based on the mean outflow between t^n_{exit} and t^{n+p}_{exit}. The re-constructed outflow, $q^{\text{recon}}_{\text{out}}$, can be expressed as,

$$q^{\text{recon}}_{\text{out}} = \frac{N^{n+p}_{\text{out}} - N^n_{\text{out}}}{t^{n+p}_{\text{exit}} - t^n_{\text{exit}}}. \tag{30}$$

As the delay accumulation-based model projects the solution at time t onto time $t+T$, where T is travel time, this reconstruction process does not effect the solution at the current time t. Physically, this method can be interpreted as holding up the vehicles in the reservoir and discharging them in the order of internal FIFO discipline. Hence, it is clear that this method conserves the total number of vehicles during the simulation duration.

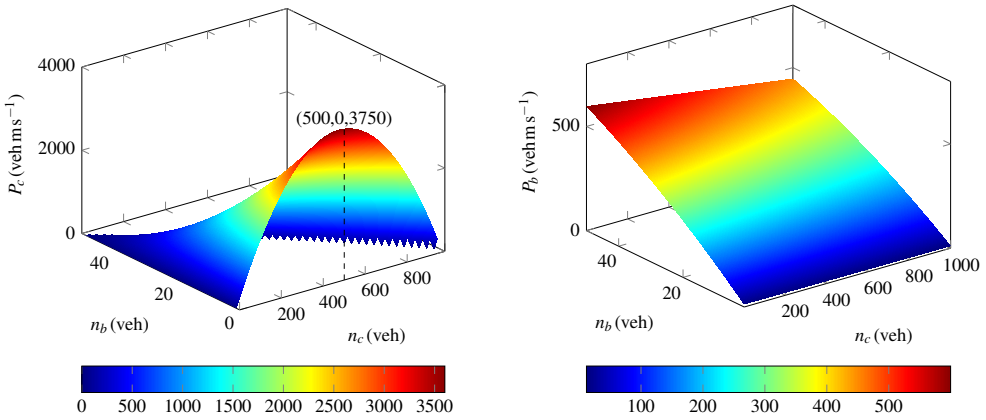
To conclude the present test case scenario, a non-constant time varying demand is also considered. It is noticed that the proposed FIFO-based entry flow function for all the MFD-based models, weak internal FIFO discipline and outflow stabilization in the delay accumulation-based model are working as expected. It is inferred that there are neither new insights nor new limitations with a non-constant demand. Therefore, to avoid redundancy, the results with non-constant demand are omitted. Moreover, the MFD framework is well suited to the slow-varying demand as the MFD is based on the steady state approximation.

3.4. Segregated 3D-MFDs

Until now, assumed 3D-MFD is the function of *total production* in the network with respect to the accumulation of cars and buses. This is the classical approach for 3D-MFD in the literature. For example, Loder *et al.* (2017) collected the data of mean speeds of cars and buses and approximated the mean speed of each mode as the linear combination of two modes. These mean speed fits are aggregated to obtain the 3D-MFD in total production in the network, which is the sum of car and bus partial productions. This kind of 3D-MFD in total production is resourceful for practitioners to understand the impact of each mode in the network. Moreover, for the modeling purpose, it is more appealing to keep the data of production per mode and segregate the 3D-MFD into two 3D-MFDs in *partial productions*. Perhaps using a single 3D-MFD for determining network performance is a viable solution, averaging the speed of all vehicles even though when the mean speed per mode is readily available, may lead to a significant approximation for dynamic simulations. Furthermore, dedicated bus lanes form a specific network with few interactions with other vehicles. Thus, it may be interesting to distinguish the MFD for these specific lanes and have a different equation for the dedicated networks.

The purpose of this study is to compare the different levels of aggregation for 3D-MFD dynamic models: a single 3D-MFD *vs.* two segregated partial 3D-MFDs per mode. This case compares the results between two approaches and attempts to explain the differences if there are any.

Figure 12 presents the segregated 3D-MFDs for the coefficients presented in Table 1. It is evident that both modes have a peak production when the other mode is absent in the network and the productivity of the network decreases gradually as the accumulation of other mode increases. The fundamental difference between using a single 3D-MFD and segregated 3D-MFDs is that the computation of partial outflow in the accumulation-based model or mean speed in the trip-based framework. Consider the partial outflow in the accumulation-based model given in eq. (14), when $n \leq n_{cr}$, which can be expressed as $\frac{n_m}{L_m} v$. When a single 3D-MFD is used, v is the mean speed of all modes, *i.e.*, $\frac{n_c v_c + n_b v_b}{n_c + n_b}$. On the other hand, when segregated 3D-MFDs are used, v becomes the mean speed of each mode, *i.e.*, $v_c = \frac{P_c}{n_c}$ and $v_b = \frac{P_b}{n_b}$. Since, the mean speed of cars is generally higher than the mean speed of buses, it is easy to verify that $\frac{n_c v_c + n_b v_b}{n_c + n_b} < v_c$. Consequently,



(a) Car production MFD surface.

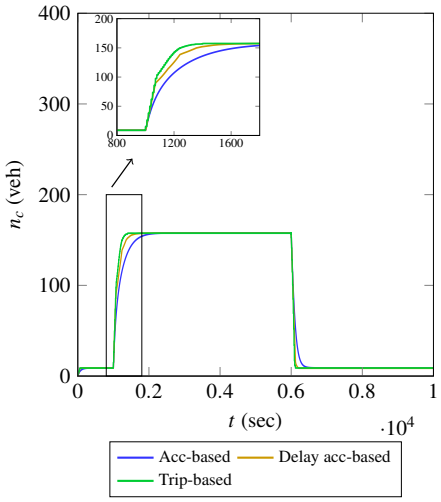
(b) Bus production MFD surface.

Figure 12: Segregated car and bus 3D-MFD surfaces.

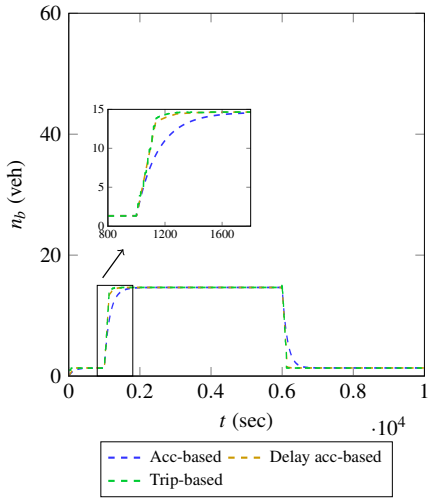
the outflow of cars using segregated 3D-MFDs is higher than in the case of single 3D-MFD. On the contrary, the outflow of buses is higher using a single 3D-MFD, as the mean speed with both modes combined is higher than the mean speed of buses. Hence, the differences in the network dynamics using a single 3D-MFD or segregated 3D-MFDs depend on the demands of each mode and the 3D-MFD *per se*.

In order to compare the results from a single 3D-MFD and two segregated 3D-MFDs, a high demand test case scenario (eq. (23)) is considered. Fig. 13 shows the evolution of accumulation, outflow with time for cars and buses using segregated 3D-MFDs. It is clear from Figs. 13c and 13d, that the reservoir is in the free-flow regime, where the outflow matches the inflow demand. This is a major difference between two approaches of using 3D-MFD using the same demand profile. Using a single 3D-MFD, congestion spill-backs appear that reach the entry in a very short time. Whereas in the second approach of using two segregated 3D-MFDs, the reservoir is in the free-flow regime. This is due to the relative magnitudes of peak demands of cars and buses. Since, the demand for cars is higher than buses (usual case in realistic networks), using the segregated 3D-MFD approach improves the outflow of cars in the accumulation-based models (mean speed of cars in the trip-based). At the same time, the outflow (or mean speed) of buses is reduced compared to the single 3D-MFD approach. However, the increase in the outflow of cars outweighs the decrease in the outflow of buses, which results in the increased productivity of the network. The peak accumulations of cars and buses in the present case are 158 veh and 15 veh, respectively are comparatively lower than the peak accumulations obtained with the single 3D-MFD approach, which are 200 veh and 30 veh, respectively. Figs. 13e and 13f show the evolution of accumulations on car and bus 3D-MFD surfaces, respectively. It can be noticed that the accumulations stay in the free-flow branch of the MFD surfaces.

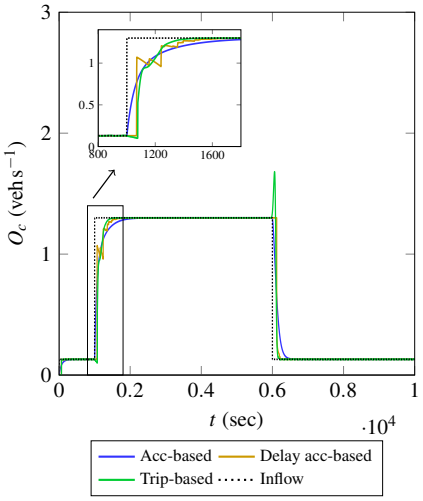
Hence, it is important to investigate which approach is more realistic for practical applications. Theoretically, two segregated 3D-MFDs are more accurate than one aggregated 3D-MFD



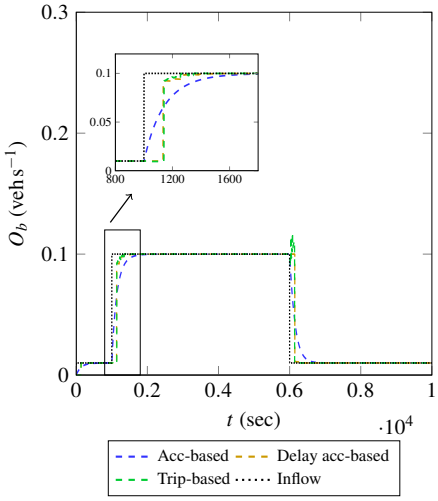
(a) Accumulation *vs.* time for cars.



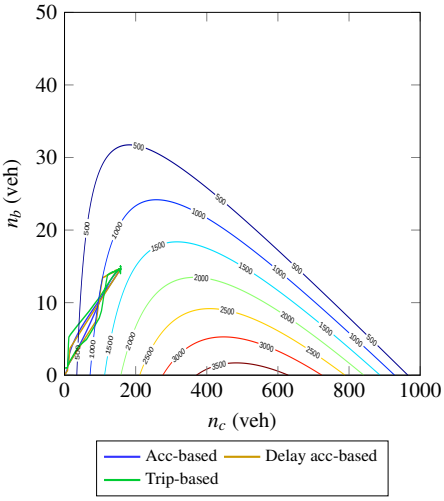
(b) Accumulation *vs.* time for buses.



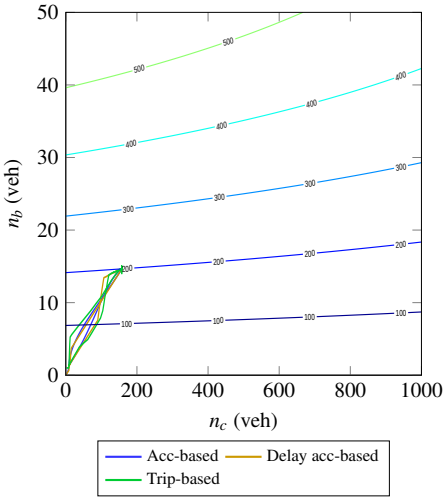
(c) Outflow *vs.* time for cars.



(d) Outflow *vs.* time for buses.



(e) Evolution of accumulations on car MFD surface.



(f) Evolution of accumulations on bus MFD surface.

Figure 13: High demand congested-flow scenario: Evolution of accumulation, outflow with time for cars and buses using segregated 3D-MFDs.

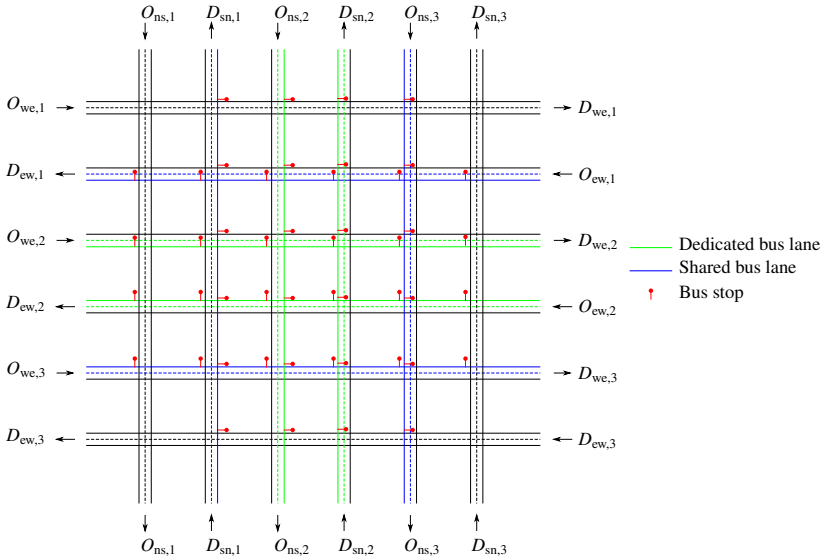


Figure 14: Idealized grid network used in micro-simulations for verification of MFD-based models. It is a 6×6 square grid network, where each block is 1 000 m. The links represented in *green* are dedicated bus lanes from west to east, north to south and *vice-versa*. Similarly, links that are in *blue* correspond to bus route shared with cars. Bus stops are located at the end of each block.

as aggregation introduces additional errors. Considering the mean speed of each vehicle category significantly changes the traffic dynamics and hence, should not be omitted in practice. In the following section, micro-simulation results of a simple grid network are used to compare the two approaches of using 3D-MFDs and to provide insights about which one to consider for the practical applications.

4. Comparison of MFD-based models solutions with bi-modal micro-simulation

4.1. Description of network and micro-simulation settings

An idealized 6×6 grid network is considered to verify the MFD-based models with micro-simulation results. All links are uni-directional with two lanes and have a length of 1 000 m. Fig. 14 presents the schematic of the network, where there are 12 entries and 12 exits in the whole network. All the intersections are controlled by traffic signals with the same green and red times of 120 sec. There is no offset in the signal timings. Origins and destinations are denoted by O and D, respectively. There are two different types of bus routes considered in this network namely, routes with dedicated bus lanes and routes that share the lane with cars. The dedicated bus lanes are marked in green and shared bus lanes are marked in blue in the network schematic. The markings also denote the proposed bus routes, *i.e.*, the sequence of links colored in green represents a bus route between $O_{we,2}$ and $D_{we,2}$ in the direction of west to east having a dedicated bus lane. Similarly, the sequence of links marked in blue shows a route between $O_{we,3}$ and $D_{we,3}$ in the direction of west to east as well, but sharing the road with cars. Therefore, it is clear that there are a total of 8 bus routes of which 4 have dedicated bus lanes and the rest share the lanes

Table 3: Normalized OD matrix for the grid network considered. O_{ew} corresponds to collection of origins on west side of network, *i.e.*, $\{O_{ew,1}, O_{ew,2}, O_{ew,3}\}$ and this notation is applicable to all origins and destinations.

	D_{ew}	D_{ns}	D_{we}	D_{sn}
O_{ew}	0.90	0.05	0.00	0.05
O_{ns}	0.05	0.90	0.05	0.00
O_{we}	0.00	0.05	0.90	0.05
O_{sn}	0.05	0.00	0.05	0.90

with cars. The bus stops are located at the end of each block. The duration of stop is assumed to be a constant of 30 sec for all bus stops.

For the sake of simplicity, only transfer trips are considered in this example. Table 3 presents the normalized OD matrix used in the test study. The OD matrix is chosen in a way that the network will have relatively homogeneous traffic conditions. The normalized OD matrix is time-invariant and taking the product of each row with respective time-dependent demand value gives time varying OD flow matrix. Only two classes of vehicles namely, cars and buses are considered in the simulation. Inflow demand with a demand surge and demand drop, as presented in earlier test cases, is considered for a total simulation period of 6 h. The bus schedules are also chosen based on the peak hour demand profile. The bus routes having the dedicated bus lanes have higher frequency during peak hour than the ones sharing the road with cars. The details of the demand profile and bus frequencies are provided while presenting the results.

Triangular fundamental diagram (FD) is assumed for both classes of vehicles. The parameters for cars are: free-flow speed, $u_c = 15\,\mathrm{m\,s^{-1}}$, wave speed, $w_c = -5.88\,\mathrm{m\,s^{-1}}$ and jam density, $\kappa_c = 0.17\,\mathrm{veh\,m^{-1}}$. In the case of buses, parameters of FD are: free-flow speed, $u_b = 12\,\mathrm{m\,s^{-1}}$, wave speed, $w_c = -5\,\mathrm{m\,s^{-1}}$ and jam density, $\kappa_c = 0.1\,\mathrm{veh\,m^{-1}}$. Micro-simulations are performed using *Symuvia* that is being developed within the *LICIT* laboratory. The simulator is based on car-following model of (Newell, 2002) and further extensions for node merge models based on the work of Leclercq *et al.* (2007). A Dynamic Traffic Assignment (DTA) component determine User Equilibrium (UE) conditions in the simulations. The routing assignment is performed for every 1 h. The optimization problem is solved iteratively using the Method of Successive Averages (MSA) (Lu *et al.*, 2009; Ameli *et al.*, 2017).

In order to calibrate the MFD for the chosen network, micro-simulations with various network configurations are performed. Simulations with different levels of bus demand, including the case of no bus demand are used to compute the 3D-MFD. The mean speeds of cars and buses are approximated based on the hypothesis of Loder *et al.* (2017).

Two different network configurations, with and without dedicated bus lanes, are considered in this study. The aim is to demonstrate the importance of segregating the dedicated bus lane network from 3D-MFD for modeling applications. In the first case, no dedicated lanes are considered, where all the bus lines share the network with cars. The calibrated coefficients for the mean

Table 4: Coefficient values for mean speed fits for cars and buses for the case with no dedicated bus lanes.

Cars		Buses	
Coefficient	Value	Coefficient	Value
$\beta_{c,0}$	10.045	$\beta_{b,0}$	7.231
$\beta_{c,c}$	-4.07×10^{-4}	$\beta_{c,b}$	-2.07×10^{-4}
$\beta_{b,c}$	-0.014	$\beta_{b,b}$	-0.0072

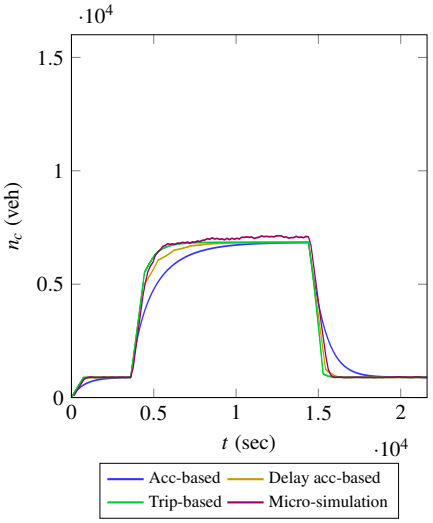
speed fits for cars and buses, in this case, are presented in Table 4. All the bus lanes use identical bus schedule and total demand for the network is given as follows:

$$[\lambda_c,\, \lambda_b]^T = \begin{cases} [1.2,\, 0.0089]^T & \text{if } 0 < t \leq 3600 \text{ or } 14400 < t \leq 21600 \\ [6.6,\, 0.0267]^T & \text{otherwise.} \end{cases} \tag{31}$$

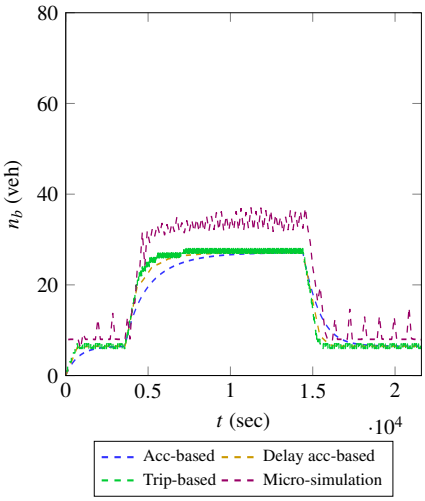
Following the demand profile (31), it is clear that there is a demand surge at $t = 3\,600$ sec and demand drop at $t = 14\,400$ sec. The trip lengths from micro-simulations are averaged per mode, which yields $L_c = 7\,129$ m and $L_b = 6\,991$ m. The results of micro-simulation are aggregated for a period of 150 sec to compare with the MFD-based ones. Production and accumulation are computed using Total Traveled Distance (TTD) and Total Traveled Time (TTT) within each aggregation period. Inflow and outflow values are constructed based on entry and exit times of each vehicle, respectively and instantaneous flow values are again aggregated within each period to damp the oscillations.

4.2. Test case with no dedicated bus lanes

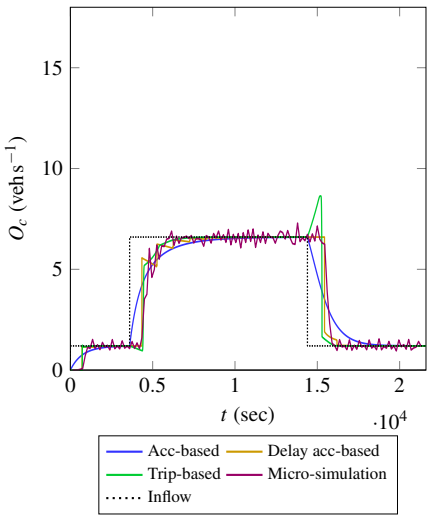
Figure 15 presents the evolution of accumulation and outflow for cars and buses using a single aggregated 3D-MFD. The demand corresponds to the free-flow regime of the MFD, where inflow follows the demand profile. Hence, plots of inflow evolution are omitted in this case. Firstly, it is clear from the plots that the steady state values of accumulations obtained from MFD-based models are different from micro-simulation. On the other hand, the outflow evolution of MFD-based models for both cars and buses follows the micro-simulation reaching the same steady state values. Since steady state inflow and outflow values for MFD-based models and micro-simulation are similar, the differences in steady state values for accumulation can be due to discrepancies in either trip lengths or 3D-MFD fit. The OD matrix is chosen in such a way that the variance in trip lengths is small, at the same time having sufficient interactions between different OD pairs. The trip length of buses is exact as the bus lines have a pre-determined route defined *a priori*. In the case of cars, the Coefficient of Variation (CV) of the trip length is estimated at 0.15, which means that the standard deviation is 15% of the mean trip length. Therefore, the reason for reaching a different steady state in the MFD-based modeling framework is due to the estimated 3D-MFD fit. The partial productions of cars and buses from the micro-simulations are aggregated to defined a single 3D-MFD. As discussed earlier, this total production is split into



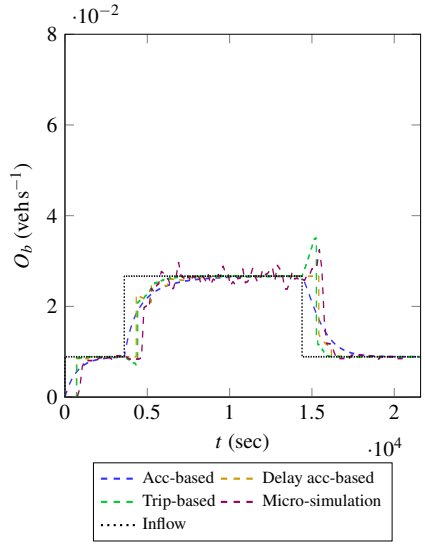
(a) Accumulation *vs.* time for cars.



(b) Accumulation *vs.* time for buses.



(c) Outflow *vs.* time for cars.



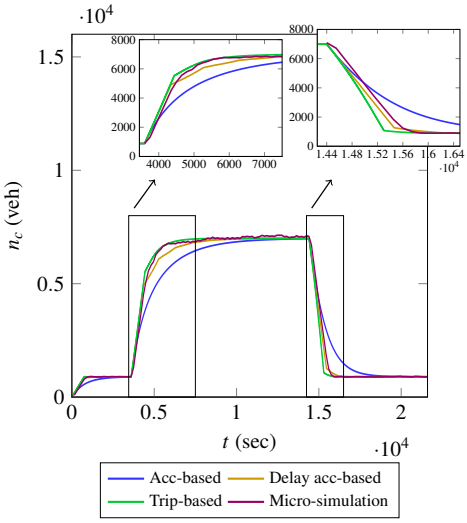
(d) Outflow *vs.* time for buses.

Figure 15: Verification with micro-simulations: Evolution of accumulation, outflow with time for cars and buses using single 3D-MFD.

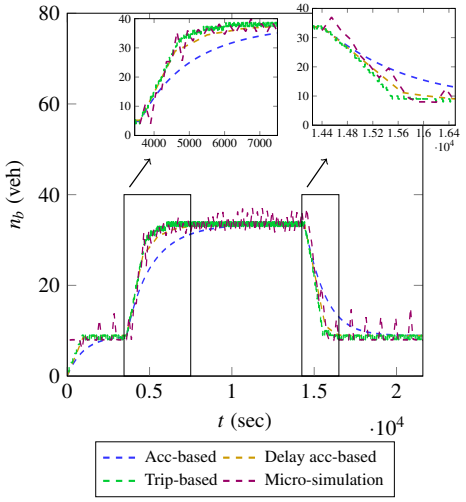
partial productions based on the partial accumulation values in the MFD-based modeling. In the present case, the demand for buses is very small compared to cars, which results in a smaller accumulation of buses compared to cars and hence, smaller partial production. This can drive the solution into a different equilibrium state on the 3D-MFD plane. Now, the same scenario is considered with two segregated 3D-MFDs instead of a single aggregated 3D-MFD in order to verify if separating the 3D-MFD can result in accurate solutions.

Figure 16 presents the evolution of accumulation, outflow and mode share for the same scenario considered before, but using two segregated 3D-MFDs. The notable observation compared to previous results is that all MFD-based models reach the same steady state as micro-simulation. It is clear from the evolution of accumulation plots in Figs. 16a and 16b. Analyzing the transition period in the accumulation evolution plots, it is clear that the conclusions made in previous test cases hold true. Similarly, the outflow evolutions of MFD-based models follow the micro-simulation ones. The delay in the outflow increase (or decrease) to the increase (or decrease) in demand can be clearly noticed in the delay accumulation-based and the trip-based models, where the results of stated MFD-based models are in very good agreement with the micro-simulation. The outflow decrease for cars during demand drop in the micro-simulation is little diffusive owing to the aggregation process. Finally, Figs. 16e and 16f present the mode share ratio evolution with time for MFD-based models and micro-simulation. As elaborated in earlier test cases, during demand surge period, the increase in mode share and then decreasing to the equilibrium value can be noticed in micro-simulation. However, the maximum ratio of mode share in micro-simulation is considerably low compared to MFD-based results. This is due to the aggregation of outflow for each aggregate interval, which adds diffusion to the solution. This is evident from top Fig. 16f, where the increase in mode share ratio of MFD-based models is sharp. On the other hand, the rise in the mode share ratio of micro simulation is more gradual and diffusive. The same argument holds during the demand drop period, where the mode share ratio of micro-simulation is diffusive compared to MFD-based ones. Hence, conclusions made from earlier test cases are verified using micro-simulation in this study.

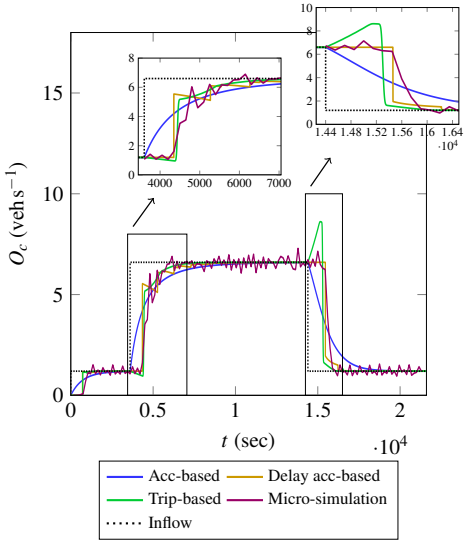
Another important conclusion to be made in this context is that using partial 3D-MFDs yields more accurate results compared to single 3D-MFD. In order to reinforce this conclusion, the relative errors in the accumulation and outflow are compared for the cases of using single 3D-MFD and two segregated 3D-MFDs. Since the results of the micro-simulations have oscillations, $\mathscr{L}_2$ norm over the simulation duration might not give a representative error for the MFD-based models. On the other hand, smoothening the micro-simulation results using techniques like moving average adds diffusion to the transition periods. Hence, the solutions of accumulation and outflow are averaged during the steady state period and the relative errors are computed using the time averaged mean values. The solutions are averaged from $t = 6\,000\,\text{s}$ to $t = 14\,400\,\text{s}$. Table 5 shows the errors in the accumulation and outflow using two approaches of 3D-MFD. Firstly, it can be inferred from the direct comparison of the errors between the single and segregated 3D-MFDs



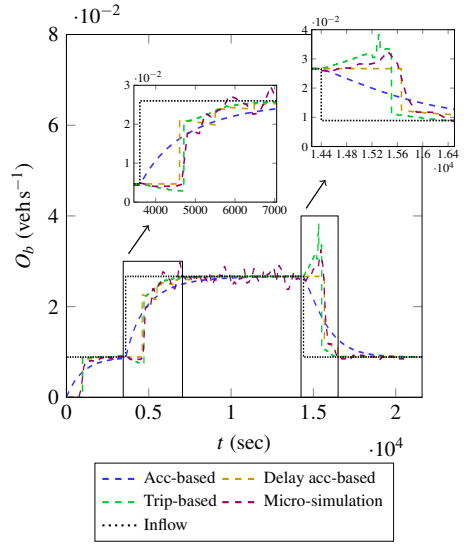
(a) Accumulation *vs.* time for cars.



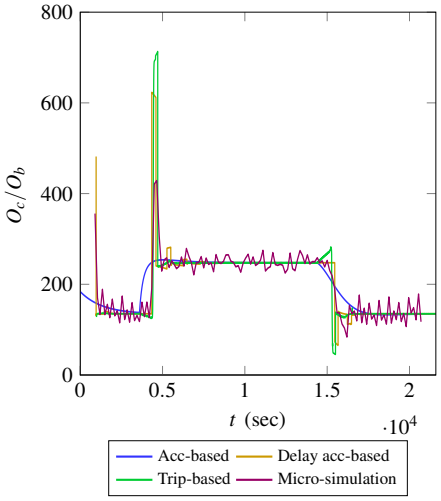
(b) Accumulation *vs.* time for buses.



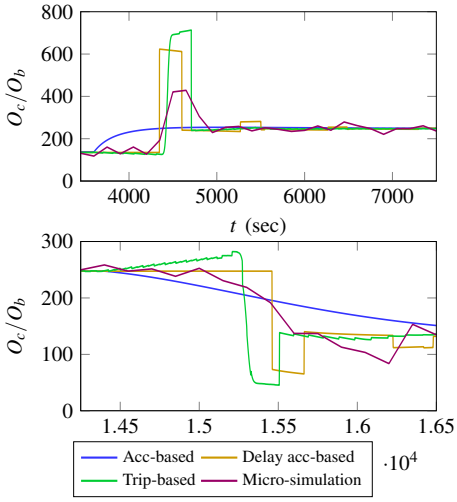
(c) Outflow *vs.* time for cars.



(d) Outflow *vs.* time for buses.



(e) Evolution of accumulations on car MFD surface.



(f) Evolution of accumulations on bus MFD surface.

Figure 16: Verification with micro-simulations: Evolution of accumulation, outflow and mode share ratio with time for cars and buses using two segregated 3D-MFDs.

Table 5: Relative errors for time averaged accumulation and outflow using single and segregated 3D-MFDs with respect to the solution of micro-simulation.

Model	Error in n_c		Error in O_c	
	Single 3D-MFD	Two 3D-MFDs	Single 3D-MFD	Two 3D-MFDs
Acc-based	0.0390	0.0198	0.0016	0.0017
Delay acc-based	0.0338	0.0124	0.0022	0.0020
Trip-based	0.0324	0.0110	0.0029	0.0029
Model	Error in n_b		Error in O_b	
	Single 3D-MFD	Two 3D-MFDs	Single 3D-MFD	Two 3D-MFDs
Acc-based	0.2103	0.0204	0.0058	0.0047
Delay acc-based	0.2063	0.0141	0.0088	0.0086
Trip-based	0.2048	0.0130	0.0094	0.0094

that using two 3D-MFDs indeed improves the accuracy of all the MFD-based models. The difference in the errors is more pronounced in the accumulation values as evident from Fig. 15, as the discrepancy in the accumulation, is visible. As concluded previously, a similar trend in the MFD-based models, *i.e.*, the trip-based model being most accurate and the accumulation-based model being least accurate, is observed in the present case too. It is clear from Figs. 16a and 16b that at time $t = 6000\,\text{s}$, the differences in the transition periods are still significant between different MFD-based models. Since the solutions are averaged starting from $t = 6000\,\text{s}$, this difference in the transition periods is quantified in the relative errors. There is no clear trend observed in the error values of the outflow, as they are comparatively quite low. This can be verified by comparing the errors in the outflow in Table 2, where the errors in the present case are one order less in the magnitude.

4.3. Test case with dedicated bus lanes included

Now, dedicated bus lanes are included in the scenario as shown in Fig. 14 to study the effect of aggregation of dedicated bus lanes with shared bus lanes in a network. In this case, the frequency of buses in dedicated bus lanes is chosen to be higher than that of shared bus lanes. The demand profile chosen for this scenario is as follows,

$$[\lambda_c,\, \lambda_b,\, \lambda_{db}]^T = \begin{cases} [1.2,\, 0.0044,\, 0.0044]^T & \text{if } 0 < t \leq 3600 \text{ or } 14400 < t \leq 21600 \\ [5.54,\, 0.0067,\, 0.067]^T & \text{otherwise,} \end{cases} \tag{32}$$

where λ_{db} corresponds to the demand of buses in the dedicated bus lane. The trip lengths for cars and buses are estimated to be 7121 m and 6990 m, respectively. As the previous test case concluded that using partial 3D-MFDs results in a more accurate resolution of traffic dynamics, only partial 3D-MFDs are used in the present case. Firstly, all the buses, irrespective of using shared or dedicated bus lanes, are aggregated together to compute 3D-MFD fits. Table 6 presents

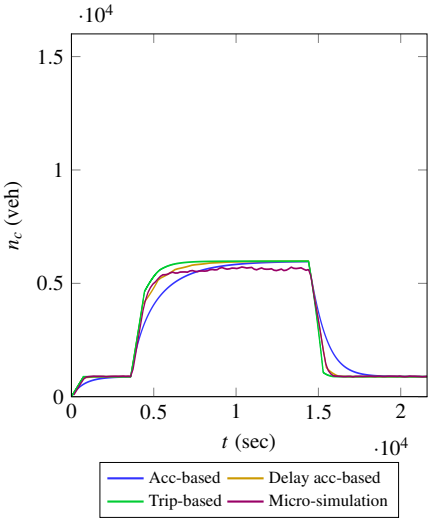
Table 6: Coefficient values for mean speed fits for cars and buses for the case with dedicated bus lanes. Both shared and dedicated bus lanes are aggregated together to fit 3D-MFDs in this case.

Cars		Buses	
Coefficient	Value	Coefficient	Value
$\beta_{c,0}$	10.172	$\beta_{b,0}$	7.872
$\beta_{c,c}$	-5.134×10^{-4}	$\beta_{c,b}$	-9.958×10^{-6}
$\beta_{b,c}$	-0.0062	$\beta_{b,b}$	-1.076×10^{-4}

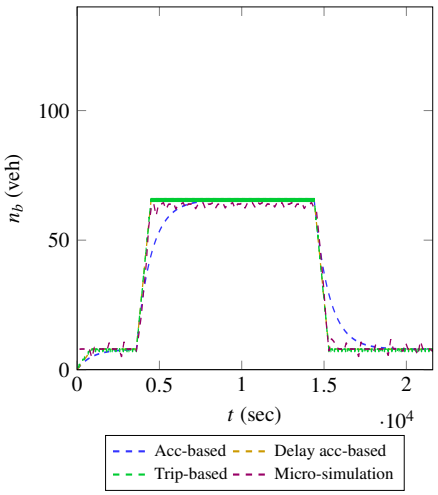
the values of coefficients for the speed MFD fits for cars and buses. It is evident that the coefficient, which characterizes the marginal effect of buses on car speeds ($\beta_{b,c}$) is at least two times lower in magnitude than the previous case presented in Table 4. Since the same network is used to compute MFD fits in both cases, albeit the difference of having dedicated bus lanes, this result is misleading. The demand for buses in dedicated bus lanes is ten times higher than the shared bus lanes during peak demand and hence, at any instance, there are more buses in dedicated lanes than shared ones. However, the buses in dedicated lanes do not have any marginal effect on the mean speed of cars as per the design of the network in this case. Hence, lower $\beta_{b,c}$ is due to the aggregation of *all* buses in the network, where a large fraction of buses have no marginal effect on the mean speed of cars. The result is an inaccurate 3D-MFD calibration, which can affect the resolution of traffic dynamics.

Figure 17 presents the evolution of accumulation and outflow for cars and buses using the segregated 3D-MFDs from coefficients presented in Table 6. It is clear from the accumulation plot of cars, that the MFD-based models reach a different steady state compared to micro-simulation. The steady state accumulation of the micro-simulation is around 5 600 veh, while MFD-based models reach up to 6 000 veh. The same is true for steady state accumulation of buses, where MFD-based models predict higher accumulation than the micro-simulation counterpart. As stated in an earlier case, this is due to inaccurate 3D-MFD fit calibrated from the micro-simulation data. In cases where dedicated bus lanes make up the majority of the bus network, it is clear from the present example, aggregating dedicated and shared bus lanes do not provide accurate 3D-MFD. Even though in the present case, both shared and dedicated bus lanes have an equal share of the total bus network, the demand in the dedicated bus network is chosen to be considerably higher than its counterpart to emulate the case of having a bigger dedicated bus network.

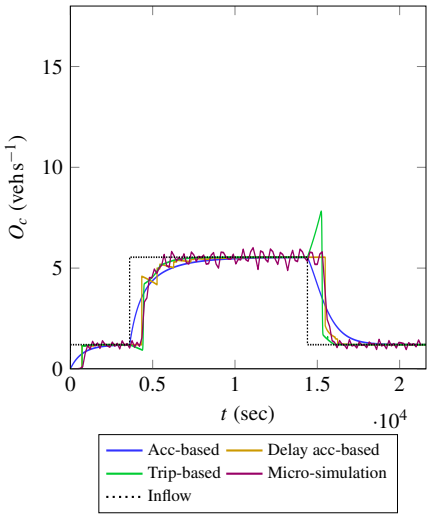
To address this shortcoming, this work proposes to segregate, yet again, the dedicated bus network from the shared one. Since buses in dedicated lanes have little or no marginal effect on other vehicles in the network, a third 2D-MFD can be proposed for let alone buses in the dedicated bus lanes. In the present case, this third 2D-MFD is estimated to be the free-flow speed of buses in dedicated bus lanes, *i.e.*, $v_{bd}(n_{bd}, n) = 8.32$, where v_{bd} and n_{bd} are the mean speed and accumulation of buses in dedicated bus lanes, respectively. Thus, new 3D-MFDs are calibrated



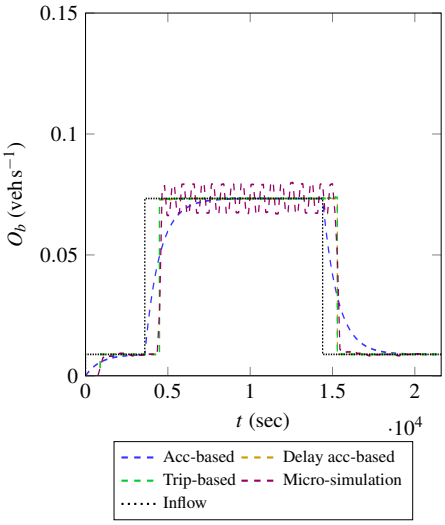
(a) Accumulation *vs.* time for cars.



(b) Accumulation *vs.* time for buses.



(c) Outflow *vs.* time for cars.



(d) Outflow *vs.* time for buses.

Figure 17: Verification with micro-simulations: Evolution of accumulation, outflow with time for cars and buses using segregated 3D-MFDs computed by treating shared and dedicated bus lanes alike.

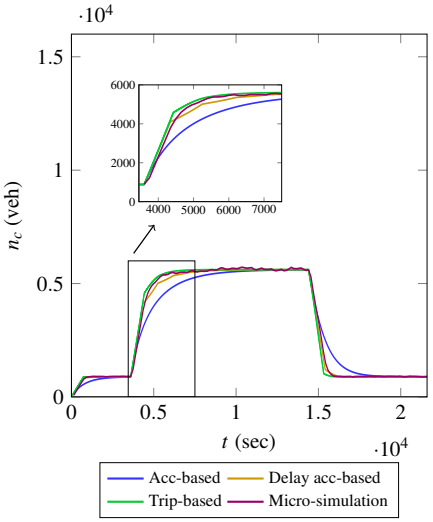
Table 7: Coefficient values for mean speed fits for cars and buses for the case with dedicated bus lanes included. In this case, only buses that share lanes with cars are included in estimation of coefficients.

Cars		Buses	
Coefficient	Value	Coefficient	Value
$\beta_{c,0}$	10.172	$\beta_{b,0}$	7.481
$\beta_{c,c}$	-5.134×10^{-4}	$\beta_{c,b}$	-2.987×10^{-4}
$\beta_{b,c}$	-0.0209	$\beta_{b,b}$	-0.011

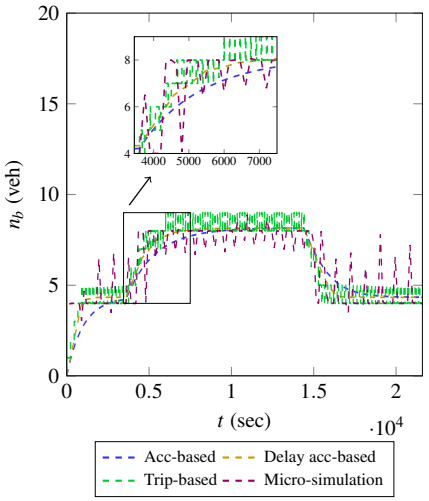
Table 8: Relative errors for time averaged accumulation and outflow using with and without segregating dedicated bus lanes with respect to the solution of micro-simulation.

Model	Error in n_c		Error in O_c	
	All bus lanes alike	Ded. lanes sep.	All bus lanes alike	Ded. lanes sep.
Acc-based	0.0494	0.0102	0.0053	0.0034
Delay acc-based	0.0578	0.0053	0.0013	0.0007
Trip-based	0.0596	0.0046	0.0008	0.0002
Model	Error in n_b and n_{db}		Error in O_b and O_{db}	
	All bus lanes alike	Ded. lanes sep.	All bus lanes alike	Ded. lanes sep.
Acc-based	0.0303	0.0056	0.0058	0.0060
Delay acc-based	0.0306	0.0062	0.0057	0.0057
Trip-based	0.0306	0.0060	0.0056	0.0056

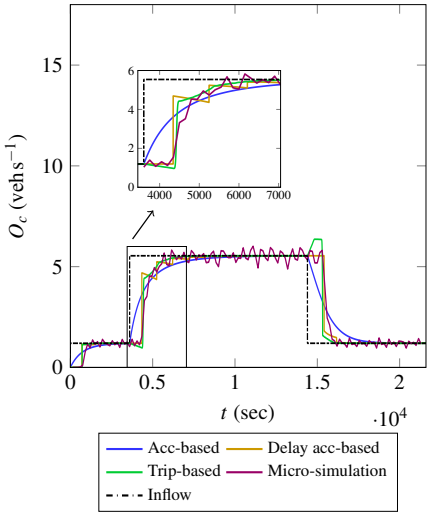
considering the buses only in shared bus lanes and coefficients are tabulated in Table 7. The absolute value of coefficient $\beta_{b,c}$ in this case is higher than the previous case (Table 6) and in the same order as the case with no dedicated lanes (Table 4). This infers that the average interactions of bus sharing the lanes with other cars are constant for a given network and proposed bus routes. Fig. 18 presents the evolution of accumulation and outflow for cars, buses and buses in dedicated bus lanes, respectively. As there are three classes of vehicles, there are three different trip lengths in the MFD-based simulation, where the trip length of dedicated bus lane is $L_{db} = 6989\,\text{m}$. It is obvious from the plots that the MFD-based models and the micro-simulation reach the same steady state accumulation for cars and buses. Figs. 18e and 18f show the accumulation and outflow for buses in the dedicated bus lanes and it is clear that the MFD-based models reproduce the micro-simulation results with very good accuracy. Table 8 presents the errors in the accumulation and outflow for the cases of with and without dedicated bus lanes segregation. A similar procedure briefed for estimating the errors in Table 5 is used here too. As concluded from the plots, the error values show that separating the dedicated bus lanes attributes to the improved accuracy of the MFD-based models. As in the previous case, the errors in the outflow are comparatively small to observe a trend. In addition, all the conclusions made earlier about the transition regime and



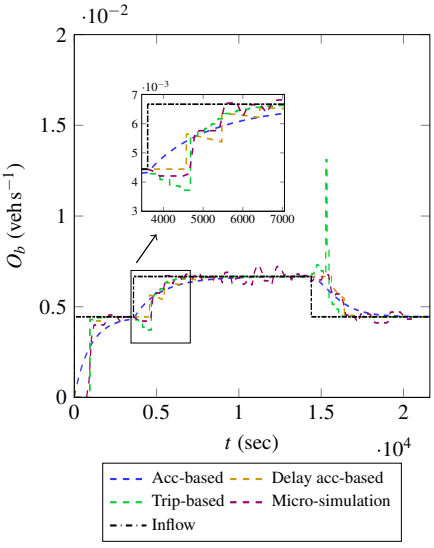
(a) Accumulation *vs.* time for cars.



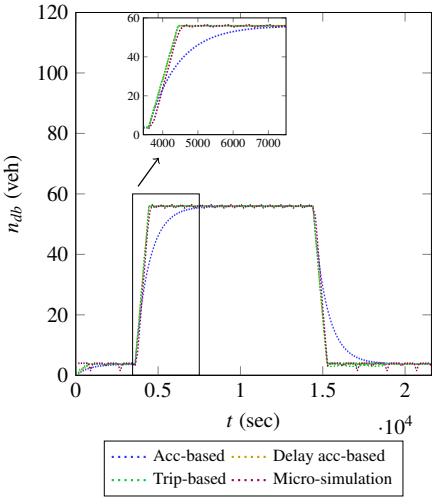
(b) Accumulation *vs.* time for buses.



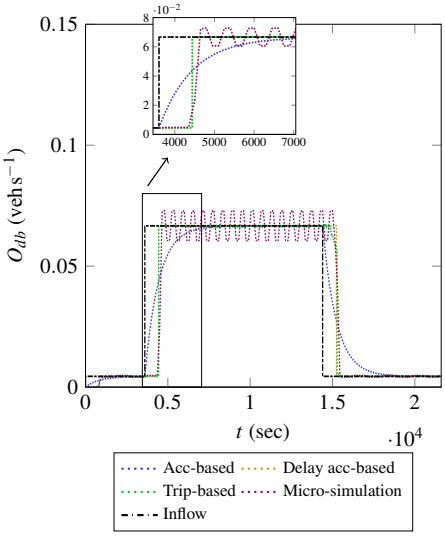
(c) Outflow *vs.* time for cars.



(d) Outflow *vs.* time for buses.



(e) Accumulation *vs.* time for buses in dedicated bus lanes.



(f) Outflow *vs.* time for buses in dedicated bus lanes.

Figure 18: Verification with micro-simulations: Evolution of accumulation, outflow and mode share ratio with time for cars and buses using two segregated 3D-MFDs by treating shared and dedicated bus lanes different.

delay in outflow are well captured in the present case as well for the delay accumulation-based and the trip-based models. Therefore, it is evident from the present results that segregating dedicated bus lanes improves the accuracy of MFD-based models. In a nutshell, it is advisable to use three MFDs, a 3D-MFD for cars, a 3D-MFD for buses using shared bus lanes and a 2D-MFD for buses using dedicated bus lanes for a multi-modal MFD simulation. Note that the two 3D-MFDs considered depend on the same variables n_c and n_b, whereas the 2D-MFD depends on only buses in dedicated bus lanes (n_{db}).

5. Conclusions

The primary objective of this work is to analyze the different MFD-based models during the transition and steady state conditions using a 3D-MFD. To this extent, accumulation-based, delay accumulation-based and trip-based models are considered. In order to be able to compare the accuracy of different MFD-based models, a reference continuum space-time model based on the linear hyperbolic conservation equation is also presented and two benchmark examples are proposed. Following this, an idealized grid network is considered to verify the proposed MFD-based models' results with micro-simulations.

Firstly, the entry flow and exit demand functions are extended to the case of 3D-MFD and the importance of non-constant critical accumulation is discussed. This work provides deeper investigation and presents the limitations of the delay accumulation-based model with a non-linear travel time MFD without stabilizations. The first test case corresponds to a demand profile with the demand surge and drops in the free-flow regime of assumed 3D-MFD. It is concluded that the trip-based model resolves the traffic dynamics most accurately followed by the delay accumulation-based model. It also highlights several limitations of the classical accumulation-based model and they are demonstrated using the reference continuum space-time model. These insights are already discussed in Mariotte *et al.* (2017); Mariotte and Leclercq (2019), however in the present work, the relative error norms reveal that the delay accumulation-based model has the best features of both accumulation-based and trip-based models. The second test case uses the same demand profile as the first one, except that the peak demand is high enough to create congestion internally inside the reservoir. This test case is designed to show the limitation of using only the conventional entry flow functions in the 3D-MFD framework. The differences in the evolution of mode share ratio for different models result in different inflow rates and consequently, different steady states. This phenomenon occurs only in the case of 3D-MFD and is absent in 2D-MFD. This is addressed by proposing a new FIFO-based entry function, where the internal FIFO discipline at the entry of queued vehicles guarantees that the share of inflows matches the share of demand. It is noticed that the FIFO-based entry function is essential for the MFD-based models founded on 3D-MFD to reach the same steady state. Using the same test case, the limitations of the delay accumulation-based model are also demonstrated. It is already well established in the literature in the context of link flow dynamics that the delay accumulation-based models

reach the steady state in finite discontinuous jumps. It is showed that these jumps will grow in amplitude for time varying inflow demand close to network saturation state and eventually results in a diverging solution. This can be addressed by smoothening the solution, which damps the oscillations in time and gives a stabilized solution. Another limitation in this framework arises during the demand drop, where the internal FIFO discipline can be locally violated depending on the magnitude of demand drop. A novel concept called weak internal FIFO discipline is proposed to circumvent this limitation by reconstruction of outflow cumulative curve, which conserves the total number of vehicles. Following this, the same test case is solved using a single 3D-MFD and two segregated 3D-MFDs, where cars and buses assume different mean speeds. It is concluded that two cases result in different solutions for the same scenario and a verification test case with micro-simulation is proposed in order to demonstrate the relative accuracy of both approaches.

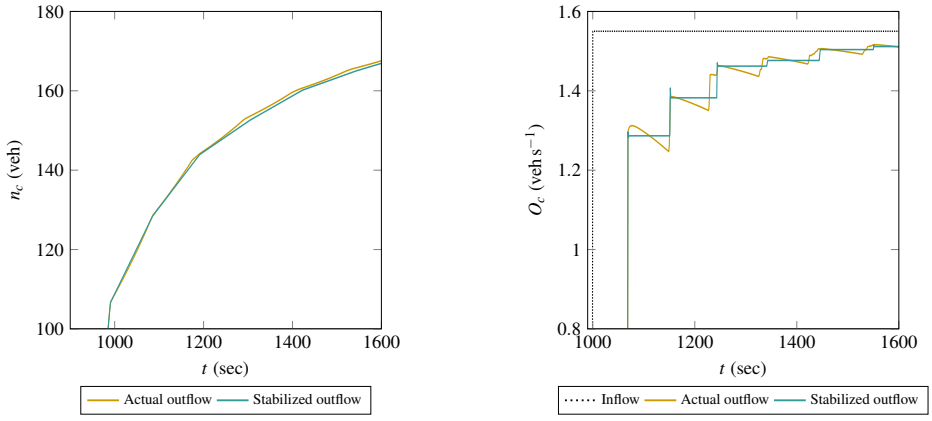
An idealized grid network comprising of cars, buses in shared bus lanes and buses in dedicated bus lanes is considered. As a first step, no dedicated bus lanes are considered in order to simplify the case further and coefficients of 3D-MFD are calibrated using several static simulations. Using the calibrated coefficients, MFD-based simulations with a single 3D-MFD and segregated 3D-MFDs are performed and it is concluded that traffic dynamics are resolved more accurately when segregated 3D-MFDs are used. In the second step, dedicated bus lanes are included in the network and coefficients of 3D-MFD are re-calibrated. This study reveals that separating the buses in dedicated bus lanes and representing them with a third 2D-MFD is necessary for MFD-based modeling. The buses in dedicated bus lanes have very few interactions with the rest of the vehicles in the network and thus, separating them from other buses resolve the dynamics more accurately in MFD-based modeling. In other words, the MFD is segregated based on vehicle categories that have different mean speeds at the same time in a given network. This results in three different MFDs namely, two partial 3D-MFDs that provide mean speed per mode for mixed network depending on partial accumulations, n_c and n_b and a 2D-MFD gives the mean speed of buses in the dedicated bus lane network.

Acknowledgment

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 646592 – MAGnUM project).

Appendix A. Error analysis for the outflow stabilization in the delay accumulation-based model

It is already stated that the stabilization of the outflow in the presence of spurious oscillations introduces some errors. This is due to replacing the actual outflow with the time averaged outflow for certain time horizon, which adds artificial diffusion to the solution. The error introduced by the



(a) Evolution of accumulation for cars

(b) Evolution of outflow for cars.

Figure A.19: Low demand free-flow scenario: Comparison of evolution of accumulation and outflow with and without outflow stabilization for delay accumulation-based model.

Table A.9: Low demand free-flow scenario: Relative errors in $\mathscr{L}_2$ norm for accumulation and outflow for delay accumulation-based model with outflow stabilization using the solution of original model as the reference.

	Cars	Buses
Accumulation	0.0044	0.0041
Outflow	0.0066	0.0073

outflow averaging is estimated using the free-flow case scenario (22). It is not possible to compute the error for the congested case scenario due to the unavailability of neither an analytical solution nor a converged numerical solution. Hence, the low demand free-flow scenario is considered to quantify the error. Note that this scenario is stable *without* stabilizing the outflow, as the inflow is constant and the peak demand corresponds to the free-flow regime. Fig. A.19 presents the evolution of accumulation and outflow during the loading phase for the models with and without outflow stabilization. As shown in Fig. A.19b, the series of jumps are replaced by the mean outflow for certain time horizon as illustrated in the stabilization technique in Fig. 11a. Hence, *two converged solutions* are obtained with and without outflow stabilization for the present scenario.

Even though the evolution of accumulation for both models seems to be close, there is a visual discrepancy between the results. This is already stated as averaging the outflow introduces certain errors in the solution. The relative $\mathscr{L}_2$ norm of errors are computed for the stabilized outflow solution using the unstabilized/actual solution as the reference. Table A.9 presents the relative errors for the accumulation and the outflow for cars and buses. It is clear that the errors are one order less than the errors presented in Table 2, which shows the errors of each MFD-based model with respect to the continuum space-time model in the manuscript. Hence, it can be concluded that the errors introduced by the outflow stabilization technique does not influence the global error convergence of the delay accumulation-based model.

References

Ameli, M., Lebacque, J.P., Leclercq, L., 2017. Multi-modal multi-class traffic network equilibrium model: a new heuristic optimization approach, in: The 12th Conference on Traffic and Granular Flow (TGF).

Ampountolas, K., Zheng, N., Geroliminis, N., 2017. Macroscopic modelling and robust control of bi-modal multi-region urban road networks. Transportation Research Part B: Methodological 104, 616 – 637. URL: http://www.sciencedirect.com/science/article/pii/S0191261515300370, doi:https://doi.org/10.1016/j.trb.2017.05.007.

Arnott, R., 2013. A bathtub model of downtown traffic congestion. Journal of Urban Economics 76, 110 – 121. URL: http://www.sciencedirect.com/science/article/pii/S0094119013000107, doi:https://doi.org/10.1016/j.jue.2013.01.001.

Astarita, V., 1996. A continuous time link model for dynamic network loading based on travel time function, in: The 13th International Symposium on Theory of Traffic flow, Lyon.

Benzoni-Gavage, S., Colombo, R.M., 2002. An n-Populations Model for Traffic Flow. Europ. J. Appl.Math 14, 587–612.

Boyac, B., Geroliminis, N., 2011. Estimation of the network capacity for multimodal urban systems. Procedia - Social and Behavioral Sciences 16, 803 – 813. URL: http://www.sciencedirect.com/science/article/pii/S1877042811010457, doi:https://doi.org/10.1016/j.sbspro.2011.04.499. 6th International Symposium on Highway Capacity and Quality of Service.

Cao, J., Menendez, M., 2015. System dynamics of urban traffic based on its parking-related-states. Transportation Research Part B: Methodological 81, 718 – 736. URL: http://www.sciencedirect.com/science/article/pii/S0191261515001654, doi:https://doi.org/10.1016/j.trb.2015.07.018. iSTTT 21 for the year 2015.

Carey, M., McCartney, M., 2002. Behaviour of a whole-link travel time model used in dynamic traffic assignment. Transportation Research Part B: Methodological 36, 83 – 95. URL: http://www.sciencedirect.com/science/article/pii/S0191261500000394, doi:https://doi.org/10.1016/S0191-2615(00)00039-4.

Chevallier, E., Leclercq, L., 2008. A macroscopic single-lane roundabout model to account for insertion delays and O–D patterns. Computer-Aided Civil and Infrastructure Engineering 23, 104–115. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8667.2007.00527.x, doi:10.1111/j.1467-8667.2007.00527.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8667.2007.00527.x.

Chiabaut, N., Xie, X., Leclercq, L., 2014. Performance analysis for different designs of a multimodal urban arterial. Transportmetrica B: Transport Dynamics 2, 229–245. URL: https://doi.org/10.1080/21680566.2014.939245, doi:10.1080/21680566.2014.939245, arXiv:https://doi.org/10.1080/21680566.2014.939245.

Daganzo, C.F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological 28, 269 – 287. URL: http://www.sciencedirect.com/science/article/pii/0191261594900027, doi:https://doi.org/10.1016/0191-2615(94)90002-7.

Daganzo, C.F., 1995. Properties of link travel times under dynamic load. Transportation Research Part B: Methodological 29, 95–98. doi:10.1016/0191-2615(94)00026-V.

Daganzo, C.F., 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. Transportation Research Part B: Methodological 41, 49 – 62. URL: http://www.sciencedirect.com/science/article/pii/S0191261506000282, doi:https://doi.org/10.1016/j.trb.2006.03.001.

Daganzo, C.F., Lehe, L.J., 2015. Distance-dependent congestion pricing for downtown zones. Transportation Research Part B: Methodological 75, 89 – 99. URL: http://www.sciencedirect.com/science/article/pii/S0191261515000387, doi:https://doi.org/10.1016/j.trb.2015.02.010.

Dakic, I., Yang, K., Menendez, M., 2019. Evaluating the effects of passenger occupancy dynamics on a bi-modal perimeter control, in: Transportation Research Board 98rd Annual Meeting Transportation Research Board, Washington.

Fan, S., Work, D., 2015. A heterogeneous multiclass traffic flow model with creeping. SIAM Journal on Applied Mathematics 75, 813–835. URL: https://doi.org/10.1137/140977977, doi:10.1137/140977977, arXiv:https://doi.org/10.1137/140977977.

Friesz, T.L., Luque, J., Tobin, R.L., Wie, B.W., 1989. Dynamic network traffic assignment considered as a continuous time optimal control problem. Operations Research 37, 893–901. URL: http://www.jstor.org/stable/171471.

Geroliminis, N., 2009. Dynamics of peak hour and effect of parking for congested cities, in: Transportation Research Board 88th Annual Meeting Transportation Research Board, Washington.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B: Methodological 42, 759 – 770. URL: http://www.sciencedirect.com/science/article/pii/S0191261508000180, doi:https://doi.org/10.1016/j.trb.2008.02.002.

Geroliminis, N., Zheng, N., Ampountolas, K., 2014. A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. Transportation Research Part C: Emerging Technologies 42, 168 – 181. URL: http://www.sciencedirect.com/science/article/pii/S0968090X14000709, doi:https://doi.org/10.1016/j.trc.2014.03.004.

Godfrey, J., 1969. The mechanism of a road network. Traffic Engineering and Control 11, 323–327.

Gu, Z., Shafiei, S., Liu, Z., Saberi, M., 2018. Optimal distance- and time-dependent area-based pricing with the Network Fundamental Diagram. Transportation Research Part C: Emerging Technologies 95, 1 – 28. URL: http://www.sciencedirect.com/science/article/pii/S0968090X18300573, doi:https://doi.org/10.1016/j.trc.2018.07.004.

Haddad, J., Mirkin, B., 2017. Coordinated distributed adaptive perimeter control for large-scale urban road networks. Transportation Research Part C: Emerging Technologies 77, 495 – 515. URL: http://www.sciencedirect.com/science/article/pii/S0968090X16302509, doi:https://doi.org/10.1016/j.trc.2016.12.002.

Haddad, J., Zheng, Z., 2018. Adaptive perimeter control for multi-region accumulation-based models with state delays. Transportation Research Part B: Methodological URL: http://www.sciencedirect.com/science/article/pii/S0191261518303096, doi:https://doi.org/10.1016/j.trb.2018.05.019.

Huang, C., Zheng, N., Zhang, J., 2019. Investigation on the bi-modal macroscopic fundamental diagrams in large-scale urban networks - an empirical study of Shenzhen city, in: Transportation Research Board 98rd Annual Meeting Transportation Research Board, Washington.

Kavianipour, M., Saedi, R., Zockaie, A., Saberi, M., 2019. Traffic state estimation in heterogeneous networks with stochastic demand and supply: Mixed lagrangian–eulerian approach. Transportation Research Record 0, 0361198119850472. URL: https://doi.org/10.1177/0361198119850472, doi:10.1177/0361198119850472, arXiv:https://doi.org/10.1177/0361198119850472. (In press).

Keyfitz, B.L., Kranzer, H.C., 1980. A system of non-strictly hyperbolic conservations laws arising in elasticity theory. Archives for Rational Mechanics and Analysis 72, 219 – 241.

Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. Transportation Research Part B: Methodological 46, 1393 – 1403. URL: http://www.sciencedirect.com/science/article/pii/S0191261512000926, doi:https://doi.org/10.1016/j.trb.2012.06.008.

Knoop, V.L., Hoogendoorn, S.P., 2014. Network transmission model: a dynamic traffic model at network level, in: Transportation Research Board 93rd Annual Meeting Transportation Research Board, Washington.

Lamotte, R., Geroliminis, N., 2016. The morning commute in urban areas: Insights from theory and simulation, in: Transportation Research Board 95th Annual Meeting Transportation Research Board, Washington.

Leclercq, L., Laval, J., Chevallier, E., 2007. The Lagrangian coordinates and what it means for first order traffic flow models, in: Proceedings of the 17th International Symposium on Transportation and Traffic Theory, Elsevier, London.

Leclercq, L., Paipuri, M., 2019. Macroscopic traffic dynamics under fast-varying demand. Transportation Science 53, 1526–1545. URL: https://doi.org/10.1287/trsc.2019.0908, doi:10.1287/trsc.2019.0908, arXiv:https://doi.org/10.1287/trsc.2019.0908.

Leclercq, L., Parzani, C., Knoop, V.L., Amourette, J., Hoogendoorn, S.P., 2015. Macroscopic traffic dynamics with heterogeneous route patterns. Transportation Research Part C: Emerging Technologies 59, 292 – 307. URL: http://www.sciencedirect.com/science/article/pii/S0968090X15001783, doi:https://doi.org/10.1016/j.trc.2015.05.006. special Issue on International Symposium on Transportation and Traffic Theory.

Leclercq, L., Sénécat, A., Mariotte, G., 2017. Dynamic macroscopic simulation of on-street parking search: A trip-based approach. Transportation Research Part B: Methodological 101, 268 – 282. URL: http://www.sciencedirect.com/science/article/pii/S0191261516309717, doi:https://doi.org/10.1016/j.trb.2017.04.004.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II. a theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 229, 317–345. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1955.0089, doi:10.1098/rspa.1955.0089, arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1955.0089.

Loder, A., Ambühl, L., Menendez, M., Axhausen, K.W., 2017. Empirics of multi-modal traffic networks–using the 3D macroscopic fundamental diagram. Transportation Research Part C: Emerging Technologies 82, 88 – 101. URL: http://www.sciencedirect.com/science/article/pii/S0968090X17301626, doi:https://doi.org/10.1016/j.trc.2017.06.009.

Loder, A., Bressan, L., Dakic, I., Ambühl, L., Bliemer, M., Menendez, M., Axhausen, K., 2019. Modeling multi-modal traffic in cities using the 3D macroscopic fundamental diagram, in: Transportation Research Board 98rd Annual Meeting Transportation Research Board, Washington.

Lu, C.C., Mahmassani, H.S., Zhou, X., 2009. Equivalent gap function-based reformulation and solution algorithm for the dynamic user equilibrium problem. Transportation Research Part B: Methodological 43, 345 – 364. URL: http://www.sciencedirect.com/science/article/pii/S0191261508000829, doi:https://doi.org/10.1016/j.trb.2008.07.005.

Mahmassani, H., Williams, J., Herman, R., 1984. Investigation of network-level traffic flow relationships: Some simulation results. Transportation Research Record , 121–130.

Mariotte, G., 2018. Dynamic Modeling of Large-Scale Urban Transportation Systems. Phd thesis.

Mariotte, G., Leclercq, L., 2019. Flow exchanges in multi-reservoir systems with spillbacks. Transportation Research Part B: Methodological 122, 327 – 349. URL: http://www.sciencedirect.com/science/article/pii/S019126151731175X, doi:https://doi.org/10.1016/j.trb.2019.02.014.

Mariotte, G., Leclercq, L., Laval, J.A., 2017. Macroscopic urban dynamics: Analytical and numerical comparisons of existing models. Transportation Research Part B: Methodological 101, 245 – 267. URL: http://www.sciencedirect.com/science/article/pii/S0191261516307846, doi:https://doi.org/10.1016/j.trb.2017.04.002.

Mariotte, G., Paipuri, M., Leclercq, L., 2019. Flow exchanges in multi-trip MFD-based systems: A validation study against microscopic simulation, in: Transportation Research Board 98rd Annual Meeting Transportation Research Board, Washington.

Mohajerpoor, R., Saberi, M., Vu, H.L., Garoni, T.M., Ramezani, M., 2019. H_∞ robust perimeter flow control in urban networks with partial information feedback. Transportation Research Part B: Methodological URL: http://www.sciencedirect.com/science/article/pii/S0191261518308609, doi:https://doi.org/10.1016/j.trb.2019.03.010. (In press).

Newell, G., 2002. A simplified car-following theory: a lower order model. Transportation Research Part B: Methodological 36, 195 – 205. URL: http://www.sciencedirect.com/science/article/pii/S0191261500000448, doi:https://doi.org/10.1016/S0191-2615(00)00044-8.

Ortigosa, J., Zheng, N., Menendez, M., Geroliminis, N., 2015. Analysis of the 3D-vMFDs of the urban networks of Zurich and San Francisco, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pp. 113–118. doi:10.1109/ITSC.2015.27.

Paipuri, M., Leclercq, L., Krug, J., 2019. Validation of Macroscopic Fundamental Diagrams-based models with microscopic simulations on real networks: Importance of production hysteresis and trip lengths estimation. Transportation Research Record 2673, 478–492. URL: https://doi.org/10.1177/0361198119839340, doi:10.1177/0361198119839340, arXiv:https://doi.org/10.1177/0361198119839340.

Ramezani, M., Haddad, J., Geroliminis, N., 2015. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. Transportation Research Part B: Methodological 74, 1 – 19. URL: http://www.sciencedirect.com/science/article/pii/S0191261515000028, doi:https://doi.org/10.1016/j.trb.2014.12.010.

Ran, B., Boyce, D.E., LeBlanc, L.J., 1993. A new class of instantaneous dynamic user-optimal traffic assignment models. Operations Research 41, 192–202. URL: https://doi.org/10.1287/opre.41.1.192, doi:10.1287/opre.41.1.192, arXiv:https://doi.org/10.1287/opre.41.1.192.

Richards, P.I., 1956. Shock waves on the highway. Operations Research 4, 42–51. URL: https://doi.org/10.1287/opre.4.1.42, doi:10.1287/opre.4.1.42, arXiv:https://doi.org/10.1287/opre.4.1.42.

Yildirimoglu, M., Geroliminis, N., 2014. Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams. Transportation Research Part B: Methodological 70, 186 – 200. URL: http://www.sciencedirect.com/science/article/pii/S0191261514001568, doi:https://doi.org/10.1016/j.trb.2014.09.002.

Yildirimoglu, M., Ramezani, M., Geroliminis, N., 2015. Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. Transportation Research Part C: Emerging Technologies 59, 404 – 420. URL: http://www.sciencedirect.com/science/article/pii/S0968090X15001813, doi:https://doi.org/10.1016/j.trc.2015.05.009. special Issue on International Symposium on Transportation and Traffic Theory.

Zhong, R., Huang, Y., Xiong, J., Zheng, N., Lam, W., Sumalee, A., 2018. An optimal control framework for multi-region macroscopic fundamental diagram systems with time delay, considering route choice and departure time choice, in: 21st IEEE International Conference on Intelligent Transportation Systems, pp. 1962–1967. doi:10.1109/ITSC.2018.8569905.