Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 38 (2019) 119–138

Transportation Research Procedia

www.elsevier.com/locate/procedia

23rd International Symposium on Transportation and Traffic Theory, ISTTT 23, 24-26 July 2019, Lausanne, Switzerland

Responsive bus dispatching strategy in a multi-modal and multi-directional transportation system: a doubly dynamical approach

Fangni Zhang[a], Wei Liu[b,c,*]

[a] *School of Aviation, University of New South Wales, Sydney NSW 2052, Australia*

[b] *School of Computer Science and Engineering, University of New South Wales, Sydney NSW 2052, Australia*

[c] *Research Centre for Integrated Transport Innovation, School of Civil and Environmental Engineering, University of New South Wales, Sydney NSW 2052, Australia*

Abstract

This paper examines the time-dependent bus dispatching problem in a multi-modal context. Traditional studies along this line often optimize the bus frequency or schedule. However, they may fail as the realized bus frequency or schedule is constrained by the time-varying traffic congestion on the road. Adding more buses to service does not necessarily increase the service frequency. Given this, we look into the time-dependent bus dispatching (number of buses in service on road) when taking into account complex multi-modal and multi-directional flow interactions on the road. In particular, the traffic dynamics over clock time is modeled through an aggregate traffic representation with flow interactions between cars and buses, and interactions between traffic in opposite moving directions. Instead of explicitly optimizing the size of dispatched bus fleet, we propose an adaptive fleet size adjustment mechanism where we have a target level of bus loading factor. This adaptive or responsive approach, by taking advantage of the doubly dynamical system proposed in Liu and Geroliminis (2017), adjusts the size of dispatched bus fleet over calendar time and accommodates day-to-day variations of mode choices and traffic patterns. Numerical studies show that the proposed approach can help bus operator to reduce operating cost and improve net benefit while maintaining comparable user costs for passengers. This study offers a new perspective for dynamic bus dispatching strategy and improves our understanding of multi-modal traffic dynamics.

Peer-review under responsibility of the scientific committee of the 23rd International Symposium on Transportation and Traffic Theory.

Keywords: Multi-modal; bus fleet; day-to-day; MFD; two-way traffic

* Corresponding author. Tel.: +61 2 9385 6926.
E-mail address: Wei.Liu@unsw.edu.au

2352-1465
Peer-review under responsibility of the scientific committee of the 23rd International Symposium on Transportation and Traffic Theory.
10.1016/j.trpro.2019.05.008

1. Introduction

Urban transit services are known for the scale economies, where additional users incur positive externalities by increasing service frequency, which is referred to as the "Mohring effect" (Mohring (1972)). It is implicitly assumed in this reasoning that the change in transit service frequency does not affect road traffic congestion. However, in reality, transportation systems are multi-modal in nature where cars and buses have complex and dynamic interactions. Increasing bus vehicles boosts traffic density which in turn exacerbates the traffic congestion in general. Despite the prevalence, the complex congestion interaction between cars and buses is rarely studied by most of the traditional 'two-mode' studies due to its limitation on analytical tractability. However, it is relevant and necessary to well incorporate the mentioned flow interactions when decide road and transit service policies. This paper addresses the bus dispatching problem with time-dependent multi-modal congestion interaction.

Most studies of 'bus service optimization' focused on the optimization of service frequency (or timetabling), see, e.g., Ibeas et al. (2010), Szeto and Wu (2011). However, the designed values for bus frequency might not be realized due to time-varying traffic conditions. If the assumed road traffic conditions are not appropriate and/or the multi-modal flow interactions on the road are not well captured, the design value for bus frequency might be in the wrong direction, which can be inefficient and misleading. Considering the multi-modal traffic dynamics, the number of buses actively deployed in service (for abbreviation, it is referred to as 'dispatched fleet size' or 'bus dispatching strategy' interchangeably) is a more straightforward decision variable, while the frequency is endogenously determined by traffic conditions given the dispatched fleet size. This motivates our study to examine the time-dependent dispatched fleet size optimization strategy. In practice, the frequency and/or schedule information can still be provided to travelers since the endogenously determined frequency can be estimated or observed based on real operation data.

Earlier literature on multi-modality primarily often focused on static or steady-state analysis. To account for the time-varying traffic conditions and congestion dynamics, much effort has been dedicated to the dynamic traffic assignment problem, such as Friesz et al. (1993), Lo and Szeto (2002), and Yildirimoglu and Geroliminis (2014). In particular, the within-day dynamic traffic patterns are governed by travelers' trip-timing choices, which has been modeled by the bottleneck model proposed by Vickrey (1969). Some studies have adopted the bottleneck model for time-dependent transit services, such as Amirgholy and Gonzales (2016) and Amirgholy et al. (2017). Recent works have also embedded the bottleneck model in the bi-modal context, such as Tabuchi (1993), Gonzales and Daganzo (2012), Yang et al. (2013), Wu and Huang (2014), Gonzales (2015), and Wang et al. (2017).

It is generally difficult to develop an analytically tractable model for large-scale mixed traffic situations where buses and cars are interacting with one another dynamically. The recently proposed aggregate traffic flow model, i.e., Macroscopic Fundamental Diagram (MFD), offers a way to model the traffic dynamics with multi-modal flow interactions. Along this line, Geroliminis et al. (2014), Chiabaut (2015), and Loder et al. (2017) studied the 3D-MFD for multi-modal flows. Zheng and Geroliminis (2013) and Zhang et al. (2018) analyzed the distribution of road space to different modes in the multi-modal system with MFD-based traffic dynamics. Zheng and Geroliminis (2016) and Liu and Geroliminis (2017) considered pricing strategy in a multi-modal system with a well-defined MFD curve for the road network. Thanks to its tractability and flexibility for large-scale networks, the MFD concept has attracted a significant amount of attentions in the literature for a number of modeling, planning, and operation issues, e.g., Gayah and Daganzo (2011), Keyvan-Ekbatani et al. (2012), Geroliminis et al. (2013), Gayah et al. (2014), Leclercq et al. (2014), Haddad (2015), Liu and Geroliminis (2016), Leclercq et al. (2017), Kouvelas et al. (2017), Zhong et al. (2018).

Nevertheless, how to optimally set the size of dispatched bus fleet and how it will affect system dynamics have rarely been modeled. Particularly important is that the congestion dynamics will affect the realized frequency or schedule of buses on the road, which might significantly deviate from the planned one due to congestion. In this context, we examine the time-dependent bus dispatching problem on account of complex multi-modal flow interactions on the road. Based on the MFD concept, the impacts of buses on traffic congestion are modeled by incorporating the Passenger Car Equivalent (PCE) concept. Moreover, traffic in different directions in the network might affect each other. Therefore, the speeds of traffic of different modes and different directions can be interdependent on each other.

Instead of explicitly optimizing the size of dispatched bus fleet, we propose a more practical adaptive feet size adjustment mechanism where we have a target level of regional bus loading. This adaptive or responsive approach follows the adaptive congestion pricing strategy proposed by Liu and Geroliminis (2017) with a doubly dynamical system. The day-to-day variations of mode choices and traffic patterns under a particular bus dispatching adjustment offers information for the adjustment of dispatched fleet size in the future on a day-to-day basis. While most studies modeled the day-to-day traffic evolution process looked at the standalone driving mode, e.g., Smith (1984),Watling (1999), Cantarella (2013), Guo et al. (2015), Smith and Watling (2016), and Guo and Huang (2016), only a few examined multi-modal systems, e.g., Cantarella et al. (2015), Li and Yang (2016), Liu and Geroliminis (2017), and Guo and Szeto (2018). It is found in previous studies that observations of day-to-day traffic patterns can be utilized to develop management or operation strategies. Li and Yang (2016) developed a responsive transit frequency adjustment strategy in the bi-modal context. However, the flow interaction between cars and buses is ignored in their study.

This study employs the doubly dynamical framework (as mentioned in the above), which has been developed by Ben-Akiva et al. (1984), Iryo (2008), Liu et al. (2017), and Guo et al. (2018). The doubly dynamical framework is able to accommodate both day-to-day evolution (over calendar time) and within-day dynamics (over clock time) of traffic patterns. In particular, the day-to-day dynamics describes the evolution of the within-day dynamic traffic patterns. Existing studies along this direction, while insightful, often rely on simplified network or demand settings or traffic flow models (e.g., single origin-destination, single-mode systems, and point-queue traffic models). The current study is the first to model the day-to-day traffic variation where multi-modal and multi-directional flows directly interact with each other dynamically. Moreover, we incorporate the effect of real-time traffic condition on travelers' choices following Liu and Geroliminis (2017). The impact of traffic information is examined by Iryo (2016), Bifulco et al. (2016), and Li et al. (2018) based on static models, and by Mahmassani and Liu (1999), Xiao and Lo (2016), and Liu et al. (2017) based on dynamic models.

This study shows that the proposed doubly dynamical model can reproduce the bi-modal traffic dynamics and complex flow interactions. Besides vehicle flows (including cars and buses), it can also characterize the passenger flows in the public transport systems. The numerical studies also demonstrate that the proposed adaptive strategy has the potential to help bus operator reduce operating cost and improve net benefit while maintaining comparable user costs for passengers.

This study contributes to the literature in several ways. Firstly, it provides a novel methodological framework for multi-modal traffic dynamics with an emphasis on bus dispatching, which is flexible for modeling large-scale systems. Secondly, it provides a novel and straightforward way to adjust bus dispatching in order to improve bus operation efficiency. Thirdly, it points out the often ignored issue regarding endogenously determined bus frequency based on real-time bi-modal traffic dynamics, and improves our understanding of multi-modal traffic dynamics.

The rest of the paper is organized as follows. Section 2 introduces the basic settings for the bi-modal commuting problem, where different modes and traffic in different directions are interacting with each other, and then formulates the day-to-day travel choice updating process and the within-day multi-modal traffic dynamics. Section 3 discusses the effect of bus dispatching on system performance and presents the adaptive bus dispatching adjustment given the day-to-day traffic variations. Numerical studies are presented in Section 4. Finally, Section 5 concludes the paper.

2. Methodological framework

We start with a brief description of the rush-hour commuting problem with two modes to serve the travel demand. We consider the city of one region in Fig. 1, where private cars and buses interact one another in the same road network. Fig. 1 is only illustrative with a simplified network structure. Travelers can either drive a private car (mode a) or take a bus (mode b) to reach their destinations.

As this study examines the single-region city with an aggregate approach, we do not specify origin-destination pairs of travelers. Instead, travelers are grouped based on their trip directions and trip lengths. Let d denote the trip direction. Typically, we have two main opposite trip directions in the peak duration, i.e., $d \in \{1, 2\}$, as indicated in Fig. 1. Note that the two opposite directions do not necessarily refer to the two directions for a road link. Instead, they can refer to the approximate moving directions for the regional traffic such as inbound and outbound. Let w denote the group index for the trip distance and W denote the set of all groups, thus $w \in W$. For different groups, the trip distance is different. A list of main notations is provided in the Nomenclature.

The clock time within a day is denoted by t, and the travel demand of a particular group on a particular direction at a particular time is denoted by $r_{d,w}(t)$. It is assumed that the demand pattern is exogenously given and does not vary over days. While the total demand intensity is given, the realized travel choices and the bi-modal traffic pattern of a

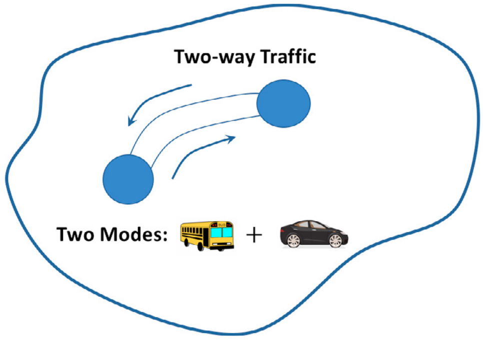

Fig. 1. City network.

single day can vary from day to day. Let q denote the day. On day q, among cohort $r_{d,w}(t)$, the number of travelers choosing mode m is represented by $r^{(q)}_{d,w,m}(t) \geq 0$, where $m \in \{a, b\}$, and $\sum_m r^{(q)}_{d,w,m}(t) = r_{d,w}(t)$. Every day, travelers choose between the two modes based on their perceptions of travel costs for the two options (often termed as the perceived cost or utility/dis-utility).

Nomenclature

m	travel mode: $m = a$ for private car mode and $m = b$ for bus mode
w	traveler group for a specific trip distance
W	set of all traveler groups with different trip distances where $w \in W$
d	travelers' traveling direction, where $d \in \{1, 2\}$
t, μ	time points
$[t_s, t_e]$	the modeling duration (for departures) where t_s is the start time and t_e is the ending time
Δt	interval length where $[t_s, t_e]$ has been equally discretized into multiple intervals with a length of Δt
q	day index or calendar time
$r_{d,w}(t)$	travel demand for travel direction d, trip distance group w, departure time t
$r^{(q)}_{d,w,m}(t)$	travel demand among $r_{d,w}(t)$ who choosing mode m on day q
α	value of time (VOT)
$c^{(q)}_{d,w,m}(t)$	mean perceived travel cost of mode m for travel demand $r_{d,w}(t)$ on day q
$C^{(q)}_{d,w,m}(t)$	experienced travel cost for travel demand among $r_{d,w}(t)$ who choosing mode m on day q
$p^{(q)}_{d,w,m}(t)$	estimated cost of mode m for travel demand $r_{d,w}(t)$ based on real-time traffic conditions on day q
η_c	learning parameter associated with yesterday's perceived cost
η_C	learning parameter associated with yesterday's experienced cost
η_p	learning parameter associated with today and yesterday's estimated cost based on real-time conditions
$Pr^{(q)}_{d,w,m}(t)$	proportion of travelers among demand $r_{d,w}(t)$ who choosing model m on day q
$\theta_{d,w}$	Logit-model parameter associated with travel direction d and trip distance group w
$v_{d,m}$	regional speed for mode m traffic with direct d
$v(n_x, n_y)$	regional speed function where n_x and n_y are traffic in different moving directions
ρ	parameter in the speed function $v(n_x, n_y)$
v_{cri}	maximum speed for the speed function $v(n_x, n_y)$
n_{cri}	critical accumulation for the speed function $v(n_x, n_y)$
$n_{d,w,a}$	car accumulation in direction d and group w
$n_{d,w,b}$	passenger accumulation in direction d and group w
$N_{d,b}$	bus accumulation in direction d

$v_{d,m}$	regional speed for car if $m = a$ or bus if $m = b$ in direction d
$\lambda_{d,b}$	bus loading factor for direction d
λ_{cri}	critical bus loading factor for calculate the discounted speed for buses
$V_{d,w,b}$	average speed for passengers (direction d and group w) after taking into account waiting time
$l_{d,w,m}$	trip distance for travelers (direction d, group w), and mode m
$L_{d,b}$	trip distance for the one-way of the bus route
$o_{d,w,m}$	outflow for cars when $m = a$ and for bus passengers when $m = b$ (direction d, group w)
$O_{d,b}$	outflow for buses (direction d)
$T^{(q)}_{d,w,m}(t)$	experienced travel time for travel demand among $r_{d,w}(t)$ who choosing mode m on day q
$\tau^{(q)}_{d,w,m}(t)$	monetary cost for travel demand among $r_{d,w}(t)$ who choosing mode m on day q
β	the operating cost per unit time for one bus
$TBF^{(q)}(t)$	total bus fee collected on day q
$TBC^{(q)}(t)$	total bus operating cost on day q
$TBN^{(q)}(t)$	total bus net revenue on day q
$TUC^{(q)}(t)$	total user cost on day q
Q	the number of days between two adjacent dispatched bus fleet adjustment
$F^{(q)}_{d}(t)$	planned size of dispatched bus fleet in direction d for time t on day q
$F_{d,0}$	parameter for the dispatched fleet size adjustment
M	number of time intervals with equal length for the modeling duration
δ	length of each of the M time intervals
ϵ	error term for the traffic flow pattern to check convergence

2.1. Day-to-day dynamics

We consider a discrete-time day-to-day evolution model, following Liu and Geroliminis (2017). For group w and departure time t, the mean perceived travel cost on day q for mode m is denoted by $c^{(q)}_{d,w,m}(t)$. Similarly, the experienced cost is $C^{(q)}_{d,w,m}(t)$, and the estimated cost based on perception of real-time traffic condition is $p^{(q)}_{d,w,m}(t)$ ("real-time instantaneous cost estimate" hereafter).

Following Liu and Geroliminis (2017), the mean perceived travel cost of departing at time t on day $q+1$ is updated based on the following (perception updating for the demand):

$$c^{(q+1)}_{d,w,m}(t) = \eta_c \cdot c^{(q)}_{d,w,m}(t) + \eta_C \cdot C^{(q)}_{d,w,m}(t) + \eta_p \cdot \left(p^{(q+1)}_{d,w,m}(t) - p^{(q)}_{d,w,m}(t)\right), \tag{1}$$

where $\eta_c > 0$, $\eta_C > 0$, and $\eta_p > 0$ are three learning parameters associated with the previous mean perceived cost, experienced cost, and estimated cost based on real-time conditions, respectively, and $\eta_c + \eta_C = 1$. Eq.(1) means that the travelers' mean perceived cost on day $q + 1$ is a linear combination of previous day's mean perceived cost and experienced cost, and then plus the difference between real-time estimated costs for days $q + 1$ and q. The first and second terms in the right-hand side of Eq.(1) have been modeled in many studies, e.g., Bie and Lo (2010). The third term is recently introduced by Liu and Geroliminis (2017). It means that travelers will compare the current traffic conditions with those from the previous day, and evaluate whether the situation on today (of different modes) is worse or better. In particular, $p^{(q+1)}_{d,w,m}(t) - p^{(q)}_{d,w,m}(t) > 0$ indicates that the travel cost estimate based on the current traffic condition on day $q + 1$ is larger than that on the previous day, which leads to an increase in the perceived cost since $\eta_p > 0$, and vice versa. With $\eta_p > 0$, Liu and Geroliminis (2017) shows the potential for information to help the system traffic to converge to an equilibrium state or fixed point. This is also verified in a recent study of Li et al. (2018) with static within-day traffic patterns.

Based on the perceived costs, the proportion choosing mode m is

$$Pr^{(q)}_{d,w,m}(t) = \frac{e^{-\theta_{d,w} \cdot c^{(q)}_{d,w,m}(t)}}{e^{-\theta_{d,w} \cdot c^{(q)}_{d,w,a}(t)} + e^{-\theta_{d,w} \cdot c^{(q)}_{d,w,b}(t)}}, \tag{2}$$

where $\theta_{d,w}$ is a trip direction and trip distance specific parameter for the Logit-model. It is obvious that $Pr^{(q)}_{d,w,m}(t) \geq 0$ and $\sum_m Pr^{(q)}_{d,w,m}(t) = 1$. The Logit-model in the above requires that the random terms associated with the mean utilities $-\theta_{d,w} \cdot c^{(q)}_{d,w,m}(t)$ for different modes are identically and independently distributed with a Gumbel probability distribution, where the mean is zero and the variance is $\frac{1}{6}\pi^2\theta_{d,w}$. The travel demand for mode m is then

$$r^{(q)}_{d,w,m}(t) = r_{d,w}(t) \cdot Pr^{(q)}_{d,w,m}(t). \tag{3}$$

Eqs.(1), (2), and (3) together establish a day-to-day dynamical system. At the fixed point of this dynamical system, we should have $c^{(q+1)}_{d,w,m}(t) = c^{(q)}_{d,w,m}(t) = C^{(q+1)}_{d,w,m}(t) = C^{(q)}_{d,w,m}(t)$ and $p^{(q+1)}_{d,w,m}(t) = p^{(q)}_{d,w,m}(t)$. As discussed in Li et al. (2018), this fixed point is equivalent to the multi-modal stochastic user equilibrium under a Logit form for mode choice. According to, e.g., Smith and Wisten (1995), the existence of a fixed point or equilibrium solution is expected under the following two conditions. Firstly, the feasible flow set is closed and convex. This holds by the formulation established in the current study. Secondly, the cost function and the mode choice function are continuous over the flows. Furthermore, the uniqueness of the fixed point or equilibrium solution requires the experienced cost to be strongly monotonic over the flows (refer to e.g., Nagurney (1993)), which does not necessarily hold in this paper under the aggregate traffic flow model with complex and dynamic multi-modal flow interactions. We leave these questions for future research while here we focus on examining the multi-modal traffic dynamics with an emphasis on the bus dispatching strategy and adjusting the dispatched fleet size over time to improve efficiency.

2.2. Within-day dynamics with bi-modal and bi-directional flow interactions

We now describe the road traffic flow model for the bi-modal system. For the roadway network in the region (we consider a single region city as indicated in Fig. 1), instead of modeling a detailed node-link network, we adopt an aggregate approach based on recently proposed Macroscopic Fundamental Diagram (MFD), see, e.g., Geroliminis and Daganzo (2008). Different from the single-region single-speed MFD framework, we consider two-directional traffic flow in the network. The modeling of direction is in an aggregate manner, i.e., we consider two aggregate approach for traffic movements, they are not necessarily on the same two-way street. Traffic in different directions can have different speeds (this better captures the heterogeneity in the network than single-region single-speed MFD framework). Particularly, the regional on-road speed has a relationship with regional traffic accumulation as follows:

$$v_{1,a} = v(n_1, n_2)\,; v_{2,a} = v(n_2, n_1)\,, \tag{4}$$

where $v_{d,a}$ is the regional space-mean speed for road traffic in direction d, n_d is the directional traffic accumulation in unit of Passenger Car Equivalent (PCE), and $v(\cdot)$ is the regional speed function. Note that we consider a typical two-directional commuting patterns, and thus we have both n_1 and n_2 (with different directions) in the speed function, where n_1 and n_2 might affect the speed $v_{1,a}$ differently (as well as for $v_{2,a}$). Generally, we expect that the traffic in the same direction has a larger unit impact on its speed than the traffic in the opposite direction. Eq.(4) covers a typical case that the cars and buses in opposite directions are interacting with each other in the network since they share the same network (e.g., they may conflict with each other at the same intersection). In case we consider more than two directions, we have to include relevant accumulation variables in Eq.(4). Since we adopt an aggregate traffic flow approach without detailed network setting, later we may frequently use 'region' to refer to the roadway system or network. The accumulation $n_d = \sum_w n_{d,w,a} + 2.5 \cdot N_{d,b}$, where $n_{d,w,a}$ is the car traffic in direction d, $N_{d,b}$ is the bus traffic in direction d, and 2.5 is the Passenger Car Equivalent for one bus.

Specifically, the following specific form for the regional speed is adopted for numerical analysis:

$$v_{x,a}\left(n_x, n_y\right) = v_{cri} \cdot \exp\left\{ -0.5 \cdot \left(\frac{n_x + \rho \cdot n_y}{n_{cri}}\right)^2 \right\}, \tag{5}$$

where $x, y \in \{1, 2\}$, $x \neq y$. Eq.(5) involves three parameters: v_{cri} represents the maximum speed for the region, n_{cri} is a critical value for accumulation, and ρ is a parameter to measure the impact of traffic in opposing direction and $\rho \in [0, 1]$. In practice, v_{cri}, n_{cri}, and ρ can be calibrated with real observations from specific network topology; in this paper, we assume that these parameters are exogenously given. There are two special cases associated with the two extreme values of ρ: $\rho = 0$ and $\rho = 1$. $\rho = 0$ means that while we have a single region, traffic moving in each direction has a separate and independent speed-accumulation function, i.e., an independent MFD for each sub-network in the region. The clustering of adjacent road links with directional flows into different sub-networks with a well-defined MFD is studied in Saeedmanesh and Geroliminis (2016). $\rho = 1$ means that the whole network including the two directions has one well-defined speed-accumulation function, i.e., a single well-defined MFD for the whole region. If we have more than two directions, we then have to specify a set of coefficients ρ for traffic in different directions.

Eq.(4) or (5) characterizes the speed of normal traffic comprises of cars and buses. It is assumed that buses can be operated at the same speed with cars if buses do not carry passengers and do not stop at bus stations. However, in practice, buses are delayed due to boarding and alighting processes and crowding effects, depending on passenger loading. Following Zheng and Geroliminis (2013), we introduce an average bus speed function as a discounted speed as opposed to private cars to account for these delays

$$v_{d,b} = \varphi(\lambda_{d,b}) \cdot v_{d,a}, \tag{6}$$

where φ represents the 'discount rate' which decreases with the regional loading factor $\lambda_{d,b}$ in direction d, i.e., $\varphi'(\lambda_{d,b}) < 0$. While Eq.(6) follows the notion of Zheng and Geroliminis (2013), it is introduced here in a more general form. Instead of explicitly aggregating dwelling delays from individual boarding and alighting times as in Zheng and Geroliminis (2013), the effect of passenger loading is designated by the term $\varphi(\lambda_{d,b})$. The general property that more passengers lead to more large delays holds in this form and it meanwhile allows for flexibility on the specific function form.

The analytical analysis in Section 3, by treating $\varphi(\lambda_{d,b})$ as a general function, does not assume any specific function form for it. To implement the numerical analysis in Section 4, the exponential function $\varphi(\lambda_{d,b}) = \exp\left\{-0.5 \cdot \left(\frac{\lambda_{d,b}}{\lambda_{cri}}\right)^2\right\}$ is adopted, where λ_{cri} represents a critical loading factor that is exogenously given. This function form reflects a specific characteristics of how the 'discount rate φ' changes with 'loading factor $\lambda_{d,b}$': it decreases at an increasing rate when $\lambda_{d,b}$ is smaller than the critical value λ_{cri}; when $\lambda_{d,b}$ exceeds λ_{cri}, it still decreases with $\lambda_{d,b}$ but the increasing rate decreases with $\lambda_{d,b}$. In general, $v_{d,b}$ governs a lower bus speed under a larger bus loading factor. In practice, the selection of $\varphi(\lambda_{d,b})$ entails dedicated sensitivity and calibration analysis.

Throughout this paper, the bus 'loading factor' ($\lambda_{d,b}$), refers to the average number of passengers on a bus in a particular direction d. By definition, the loading factor is calculated by $\lambda_{d,b} = \sum_w n_{d,w,b}/N_{d,b}$ (passenger-minute per bus).

From the bus passengers' point of view, the journey time includes not only in-vehicle time $l_{d,w,b}/v_{d,b}$ but also waiting time at bus station, where $l_{d,w,b}$ is the trip length for direction d, group w, and mode b. Denote $L_{d,b}$ as the trip distance for the one-way of the bus route on direction d, the average one-way travel time of a bus is $L_{d,b}/v_{d,b}$. When the average proportion of buses serving overlapped routes is $\gamma_{d,b}$ (the average proportion of buses that is suitable for a particular traveler trip among all buses operating in the same direction), the mean bus headway is $h_{d,b} = L_{d,b}/(\gamma_{d,b} \cdot N_{d,b} \cdot v_{d,b})$. The waiting time is approximated by the half of the mean bus headway, which is $0.5 \cdot h_{d,b} = 0.5 \cdot L_{d,b}/(\gamma_{d,b} \cdot N_{d,b} \cdot v_{d,b})$. Thus, the average traveling speed of bus passengers is

$$V_{d,w,b} = \frac{l_{d,w,b}}{\frac{l_{d,w,b}}{v_{d,b}} + 0.5 \cdot \gamma_{d,b} \cdot \frac{L_{d,b}}{\gamma_{d,b} \cdot N_{d,b} \cdot v_{d,b}}} = \frac{l_{d,w,b}}{l_{d,w,b} + 0.5 \cdot L_{d,b}/(\gamma_{d,b} \cdot N_{d,b})} \cdot v_{d,b}. \tag{7}$$

Based on speeds and trip lengths, we can then determine the outflows of car and bus traffic, as well as passenger flows. Following many studies based on the MFD concept, we have the outflow for cars (or flow rate leaving the network) as follows:

$$o_{d,w,a} = \frac{n_{d,w,a} \cdot v_{d,a}}{l_{d,w,a}}. \tag{8}$$

The outflow for buses is

$$O_{d,b} = \frac{N_{d,b} \cdot v_{d,b}}{L_{d,b}}. \tag{9}$$

The outflow for bus passengers is

$$o_{d,w,b} = \frac{n_{d,w,b} \cdot V_{d,w,b}}{l_{d,w,b}}. \tag{10}$$

2.3. Travel cost formulation

Now we discuss how the experienced cost and real-time instantaneous cost estimates (which simulate the perception of real-time traffic conditions) mentioned in Section 2.1 can be determined.

Experienced conditions and costs. For all travelers from group w and in direction d, the trip distance is $l_{d,w,m}$ for mode m. The experienced car travel time (driving time) $T^{(q)}_{d,w,a}(t)$ on day q can be determined by solving the following equation numerically (the driving time is in the upper limit of the integral):

$$\int_{t}^{t+T^{(q)}_{d,w,a}(t)} v^{(q)}_{d,a}(\mu)\, d\mu = l_{d,w,a}, \tag{11}$$

where $v^{(q)}_{d,a}(\mu)$ is the realized regional speed at time μ on day q for car traffic in direction d, and $l_{d,w,a}$ is the driving distance, which is time-invariant. Similarly, we can determine the experienced journey time for bus trips, which is obtained by solving

$$\int_{t}^{t+T^{(q)}_{d,w,b}(t)} V^{(q)}_{d,w,b}(\mu)\, d\mu = l_{d,w,b}, \tag{12}$$

where $V^{(q)}_{d,w,b}(\mu)$ is the realized speed at time μ for bus trips, and $l_{d,w,b}$ is the bus trip distance, which is also time-invariant.

We are now ready to formulate the experienced travel costs. For group w travelers departing at time t, the experienced travel cost of mode m on day q is the sum of travel time cost and monetary cost, which is given by

$$C^{(q)}_{d,w,m}(t) = \alpha \cdot T^{(q)}_{d,w,m}(t) + \tau^{(q)}_{d,w,m}(t), \tag{13}$$

where α is value of time (VOT), $T^{(q)}_{d,w,m}(t)$ is the experienced travel time of mode m, and $\tau^{(q)}_{d,w,a}(t)$ is the monetary cost for private car mode including private car running cost, parking fee, and $\tau^{(q)}_{d,w,b}(t)$ is the monetary cost for bus mode, i.e., the distance-dependent bus fee in this study.

Perception of real-time traffic condition. We now discuss the instantaneous cost estimate based on real-time traffic conditions. The instantaneous cost estimate for private cars can be determined in a similar way as the experienced cost. We simply need to replace the experienced speed $v^{(q)}_{d,w,a}(\mu)$ in Eq.(13) with the speed $v^{(q)}_{d,w,a}(t)$ at departure time t to have an instantaneous estimation of travel time. It follows then

$$p^{(q)}_{d,w,a}(t) = \alpha \cdot \left[\frac{l_{d,w,a}}{v^{(q)}_{d,w,a}(t)}\right] + \tau^{(q)}_{d,w,a}(t). \tag{14}$$

Similarly, we can determined that for buses, which is

$$p^{(q)}_{d,w,b}(t) = \alpha \cdot \left[\frac{l_{d,w,b}}{V^{(q)}_{d,w,b}(t)}\right] + \tau^{(q)}_{d,w,b}(t). \tag{15}$$

2.4. System efficiency measures

The major objective in this study is to improve bus utilization and increase net revenue for the bus operator. Thus, the main efficiency measure of interest in this paper is the bus operator's revenue. For alternative objective functions, one can found in, e.g., Zhang et al. (2016). We now list a few relevant efficiency measures.

Firstly, the total bus fare-box revenue (collected from the travelers) is calculated as follows:

$$TBF^{(q)} = \sum_{d}\sum_{w}\int_{t} r^{(q)}_{d,w,b}(t) \cdot \tau^{(q)}_{d,w,b}(t)dt, \tag{16}$$

Secondly, the total operating cost for buses is given by

$$TBC^{(q)} = \beta \cdot \sum_{d}\int_{t} N^{(q)}_{d,b}(t)dt, \tag{17}$$

where β is the operating cost per unit time for one bus, and $\sum_d \int_t N_{d,b}^{(q)}(t)dt$ is the total bus operating time. The total net revenue (profit) is then

$$TBN^{(q)} = TBF^{(q)} - TBC^{(q)}. \tag{18}$$

From the traveler's perspective, the total user cost is

$$TUC^{(q)} = \sum_d \sum_w \sum_m \int_t r_{d,w,m}^{(q)}(t) \cdot C_{d,w,m}^{(q)}(t)dt, \tag{19}$$

and the total system cost is

$$TSC^{(q)} = TUC^{(q)} + TBC^{(q)} - TBF^{(q)}. \tag{20}$$

In the following analysis of bus dispatching strategy, we are interested in not only bus operator's revenue but also user cost. We will examine how much the proposed adaptive bus operation to improve revenue could harm the travelers in the bi-modal system as a whole, or how much it will help to reduce user cost.

3. Responsive bus dispatching strategy and effect of dispatched fleet size

In last section, we discuss both the day-to-day traffic variation and the within-day multi-modal traffic interactions. Given these, we now develop a responsive (adaptive) bus dispatching strategy for determining time-dependent dispatched fleet size, which is to improve bus utilization efficiency and net revenue while maintaining certain level of services for passengers. Responsive bus operating strategies have been considered in Zhang et al. (2014), Zhang et al. (2016), and Li and Yang (2016). However, dynamic dispatching strategy considered here is much more complex and realistic than those in the literature. The main reason is that providing a larger dispatched fleet size does not guarantee a higher frequency due to potential larger congestion delays resulting from the increased number of bus in service. In short, this study takes into account the endogenously multi-modal and multi-directional flow interaction when adjusting the dispatched fleet size. Those in the literature often ignore this issue and implicitly assume that the planned bus frequency will not be affected by endogenous congestion and can be achieved as planned. Throughout this paper, the variable of interest is the size of dispatched fleet rather than whole available fleet (including both operating and idling vehicles) with the assumption that the size of whole fleet is large enough to cover the studied period of time.

3.1. Objective loading factor

The bus dispatching strategy discussed in this paper aims at an exogenously proposed target loading factor λ_{obj}. In practice, the target loading factor can be determined by comprehensively calibrating the relationship between dispatched fleet size and net revenue. Alternatively, the optimal loading factor can be estimated by the aggregate demand elasticity with respect to the dispatched fleet size. We provide a brief discussion in the below. For simplicity, the trip direction, trip distance, and time-dimension are ignored, and indices d, w, q, and t are thus omitted (this also eases the notation burden). We introduce the elasticity of the travel demand for bus with respect to the dispatched fleet size, which is given by

$$\xi_{N_b}^{r_b} = \frac{dr_b}{dN_b}\frac{N_b}{r_b}, \tag{21}$$

where r_b represents the aggregate travel demand and N_b represents the dispatched fleet size. It is generally expected that employing more buses will attract more demand, namely the demand elasticity is normally positive. With the abbreviated notation, the bus operator's fare-box revenue is $BF = r_b \cdot \tau_b$, the operating cost is $BC = \beta \cdot N_b$, and the net revenue (profit) of the bus operator is $BN = BF - BC$. When the dispatched fleet size changes, the marginal revenue is

$$\frac{dBF}{dN_b} = \frac{dr_b}{dN_b} \cdot \tau_b, \tag{22}$$

and the marginal operating cost is

$$\frac{dBC}{dN_b} = \beta. \tag{23}$$

When selecting the dispatched fleet size to maximize net revenue, the optimal dispatched fleet size should strike a balance between the marginal fare-box revenue and the marginal operating cost, i.e, $dBF/dN_b = dBC/dN_b$, or equivalently,

$$\frac{dr_b}{dN_b} \cdot \tau_b = \beta. \tag{24}$$

Substituting Eq.(21) into Eq.(24) and with some manipulations, we obtain the following proposition regarding the objective loading factor.

Proposition 1. Given the unit operating cost, ticket price and the demand elasticity with respect to dispatched fleet size, the optimal loading factor is

$$\lambda_{obj} = \frac{\beta}{\tau_b \xi^{r_b}_{N_b}}. \qquad \square \tag{25}$$

Proposition 1 indicates that the objective loading factor should be determined based on unit operating cost, ticket price and demand elasticity. Given that β represents the unit operating cost per unit time per bus and τ the unit ticket price, β/τ signifies the minimum patronage that is required to cover the operating cost per unit time. Eq.(25) prescribes that the optimal loading factor is equal to the minimum patronage multiplied by the inverse demand elasticity. This implies that the more sensitive the demand is to the density of bus service, the lower the optimal loading factor should be.

Proposition 1 provides the analytical form of the objective loading factor, which entails precedent information of demand elasticity. In practice, the selection of objective loading factor requires dedicated sensitivity analysis and parameter calibration.

3.2. Effect of dispatched fleet size

This section presents a comparative analysis of the effect of dispatched fleet size on the level of service, in particular reflected by the mean bus headway. Following Section 3.1, this section adopts the steady-state analysis with abbreviated notions. As shown in Section 2.2, the mean bus headway is given by $h_b = L_b/(\gamma_b \cdot N_b \cdot v_b)$. When there is a marginal increase in dispatched fleet size, the change in the mean bus headway is

$$\frac{dh_b}{dN_b} = -\frac{L}{\gamma_b v_b^2 N_b^2} \cdot (v_b + N_b \frac{dv_b}{dN_b}). \tag{26}$$

The first term on the RHS of Eq.(26) is negative, and thus h_b decreases with N_b if and only if the term in the brackets is positive as a whole, i.e., $\frac{dv_b}{dN_b} > -v_b/N_b$. Eq.(26) suggests that the change direction of bus headway depends on how the bus speed changes with the fleet size $\frac{dv_b}{dN_b}$. Since $v_b = \varphi(\lambda_b) \cdot v_a$ and $\lambda_b = n_b/N_b$, we can further derive

$$\frac{dv_b}{dN_b} = \varphi' v_a \cdot \frac{d\lambda_b}{dN_b} + \varphi \cdot \frac{dv_a}{dN_b}. \tag{27}$$

$$\frac{d\lambda_b}{dN_b} = \frac{1}{N_b}\frac{dn_b}{dN_b} - \frac{n_b}{N_b^2} \tag{28}$$

where $\varphi' = d\varphi/d\lambda_b < 0$. Substituting Eqs.(27) and (28) into Eq.(26), we have

$$\frac{dh_b}{dN_b} < 0 \Leftrightarrow \varphi' v_a \frac{d\lambda_b}{dN_b} + \varphi \frac{dv_a}{dN_b} > -\frac{v_b}{N_b} \Leftrightarrow \frac{d\lambda_b}{dN_b} < -\frac{1}{N_b}\frac{\varphi}{\varphi'}(1 + \frac{dv_a}{dN_b}\frac{N_b}{v_a}) \Leftrightarrow \frac{d\lambda_b}{dN_b}\frac{N_b^2}{n_b} < -\frac{\varphi}{\lambda_b \varphi'}(1 + \frac{dv_a}{dN_b}\frac{N_b}{v_a}). \tag{29}$$

We thus obtain the following proposition regarding the effect of dispatched fleet size.

Proposition 2. An increase in the dispatched fleet size leads to the increase in the level of service of bus service (reduced mean headway) if and only if

$$\xi_{N_b}^{\lambda_b} < -\frac{\varphi}{\lambda_b \varphi'}(1 + \xi_{N_b}^{v_a}), \tag{30}$$

where $\xi_{N_b}^{\lambda_b}$ and $\xi_{N_b}^{v_a}$ represent the elasticities of loading factor λ_b and normal traffic speed v_a with respect to the dispatched fleet size N_b respectively. □

Proposition 2 establishes the sufficient and necessary condition for the dispatched fleet size positively improving the level of service. It implies that how the level of service changes with the dispatched fleet size is driven by how loading factor and traffic congestion change with it. With endogenous mode choice, the effects of dispatched fleet size on these key performance indicators (level of service, loading factor, traffic speed) are complex. This is because the change of dispatched fleet size has not only direct impacts on the bus vehicle density, but also indirect impacts through modal split and demand redistribution. The traffic speed, for example, is determined by the total vehicle accumulation which is the sum of car accumulation and bus accumulation. Increasing dispatched fleet size directly enlarges the bus accumulation which may reduces the traffic speed, but its impact on car accumulation is not straightforward. Given denser bus service, there will a modal shift from driving mode to bus mode leading to a decreased driving demand which in turn reduces the car accumulation. As the consequence, the actual change of traffic speed depends on the relative magnitude of the direct and indirect impacts. When the mode shift effect is large enough, enlarging dispatched fleet size may improve the overall traffic speed by substantially reducing driving demand. This is to say, the elasticity of the traffic speed with respect to the dispatched fleet size is not necessarily negative (direct impact dominating) nor positive (indirect impact dominating); instead, it can take any value. The similar argument is applicable to the effect of dispatched fleet size on loading factor – the direct impact reduces the loading factor and the indirect impact increases the demand for bus travel which potentially intensifies the loading factor.

Therefore, the sufficient and necessary condition of Eq.(30) does not always stand. When the direct impact of dispatched fleet size on traffic speed substantially overwhelms, $\xi_{N_b}^{v_a}$ can be less than -1, and the RHS of Eq.(30) is negative. If at the same time, the direct impact on loading factor is dominated by the indirect impact, the LHS of Eq.(30) is positive, which contradicts Eq.(30). We thus have the following Corollary.

Corollary. When $\xi_{N_b}^{v_a} < -1$ and $\xi_{N_b}^{\lambda_b} > 0$, increasing the dispatched fleet size reduces the service headway.

This Corollary provides a sufficient condition for the counterproductive outcome to occur when the dispatched fleet size is increased. Intuitively, this Corollary states that if adding buses in service attracts more bus demand and slows the traffic speed substantially and simultaneously, the level of service will be reduced.

This is likely to happen when the traffic congestion is highly intensive to the extent that both traffic speed and modal split are very sensitive to the change of dispatched fleet size. In practice, in the downtown areas of many mega-cities with limited road space, such as London, Sydney, Hong Kong, and Beijing, peak-hour traffic congestion is significant. It is frequently observed that the density of transit vehicles (especially long bus vehicles) stuck in the traffic congestion is substantial. In such systems, it is expected that adding more bus vehicles will further exacerbate traffic congestion and reduce the level of service.

3.3. Responsive bus dispatching strategy

Now we proceed to discuss how the size of dispatched bus fleet should be adjusted over calendar time given day-to-day traffic variations for a given target loading factor λ_{obj}. The strategy we introduce here is to update the dispatched fleet size adaptively to drive the actual average loading factor $\lambda_{d,b} = \frac{\sum_w n_{d,w,b}}{N_{d,b}}$ to the objective λ_{obj}. Note that the dynamic dispatched fleet size is not updated from day to day; instead, it is updated once over a certain period of time, i.e., a certain number of days (denoted by Q), e.g., a month, or a quarter.

The modeling period is $[t_s, t_e]$, and $t \in [t_s, t_e]$. Suppose we divide $[t_s, t_e]$ into M intervals with an identical length of $\delta = \frac{t_e - t_s}{M}$, and the i-th interval is $[(i-1)\cdot\delta, i\cdot\delta]$. For each time interval, the dispatched fleet size is identical.

If the time-dependent dispatched fleet size is lastly updated on day $q - Q$, then the next update round will take place on day $q + 1$ as follows:

$$F_d^{(q+1)}(t) = F_d^{(q)}(t) + F_{d,0} \cdot \left(\frac{\int_{\mu \in [(i-1)\cdot\delta, i\cdot\delta]} \lambda_{d,b}^{(q)}(\mu) d\mu}{\delta} - \lambda_{obj} \right) \tag{31}$$

where $t \in [(i-1)\cdot\delta, i\cdot\delta]$. If we increase M and decrease δ, for every t, we can find $(i-1)\cdot\delta \to t$ and $i\cdot\delta \to t$. This means, the dispatched fleet size at time t is updated based on system conditions at time t. Later we will test different values for δ and explore how it will affect the efficiency of the adaptive fleet management strategy.

The planned dispatched fleet size $F_d^{(q)}(t)$ is realized by adding buses into the roadway when we have to increase the dispatched fleet over clock time, and by stopping buses completing their trips and arriving at the end of the bus route when we want to decrease real-time dispatched fleet size. Note that the rate/intensity for stopping bus services is bounded by the rate of or the number of arriving buses at the final stops.

4. Numerical studies

This section presents numerical examples to illustrate the model developed and analysis presented. Firstly, we present and elaborate the common numerical setting. Secondly, we examine day-to-day traffic variations, under either asymmetric or symmetric demand profile for the two trip directions, and how bus frequency (the number of running or dispatched buses) evolve with clock time governed by congestion dynamics. Thirdly, we examine the adaptive strategy for the time-dependent bus dispatching. In particular, we look into the impact of the value of δ, which is the bus operation interval where the dispatched fleet size remains constant. We test and compare the model performance under different demand and congestion conditions. We also compare our adaptive approach with a direct optimization approach for the time-dependent bus dispatching strategy.

4.1. Numerical setting

We start with describing the common numerical setting, a significant part of which is summarized in Table 1. The modeling duration is $[t_s, t_e] = [0, 3]$(h), which has been discretized into time intervals with a length of $\Delta t = 1$ (min) for computing the traffic dynamics.

Table 1. Basic numerical setting.

Variables or parameters	Specification
Modeling duration	3 hours: $[t_s, t_e] = [0, 3]$ (h)
VOT	$\alpha = 48.45$ (AUD/h)
Monetary cost for a car trip	$\tau_{d,w,a} = 0.703 \cdot l_{d,w,a} + 15$ (AUD per trip): $0.703 \cdot l_{d,w,a}$: running cost; 15 (AUD): parking fee
Speed function parameters	$v_{cri} = 40$ (km/h); $n_{cri} = 1 \times 10^4$ (PCE); $\rho = 0.2$; $\lambda_{cri} = 25$ (passenger-minute per bus)
Day-to-day traveler learning parameters	$\eta_c = 0.5$, $\eta_C = 0.5$, $\eta_p = 0.8$
Logit model parameter	$\theta_{d,w} = 0.15$
Bus route	one-way length $L_{d,b} = 30$ (km), proportion of overlapping routes $\gamma_{d,b} = 5\%$
Operating cost for unit time of one running bus	$\beta = 98.10$ (AUD per one hour operation of one bus)
Bus target loading factor	$\lambda_{obj} = 15$ (passenger-minute per bus)
Bus dispatching adjustment factor	$F_{d,0} = 1.75$ (bus per loading factor unit)
Number of days between two fleet adjustments	$Q = 30$ (day)

The travel demand intensity over time is presented in Fig. 2(a), where demands in direction $d = 1$ and direction $d = 2$ are different (asymmetric demand). While this is the benchmark demand, we have also tested the case with symmetric demand (however, to be comparable, the total demand density for both directions is still the same). This reflects typical patterns in morning or evening peak durations. Moreover, there are five groups of trip distances among the demand, i.e., for $w = 1, 2, 3, 4, 5$, the trip lengths are 2, 4, 6, 8,10 (km) for car mode and 2.15, 4.30, 6.45, 8.60, 10.75 for bus mode, respectively. The proportions of these groups among the total demand are 56.0%, 25.2%, 11.4%, 5.1%, 2.3%. Note that the length of the bus route (30km for both directions) is larger than the trip distance of all passengers.

To examine the adaptive approach for adjusting the running bus fleet size (the number of buses dispatched), we look at 10 periods and each period has 30 days, i.e., $Q = 30$. Time-dependent bus dispatching is adjusted from period to period. The whole modeling duration (3 hours) is equally divided into 36 intervals with a length of $\delta = 5$ (min). Within each interval, the dispatched fleet size is the same, and is adjusted identically over periods. Note that δ is the

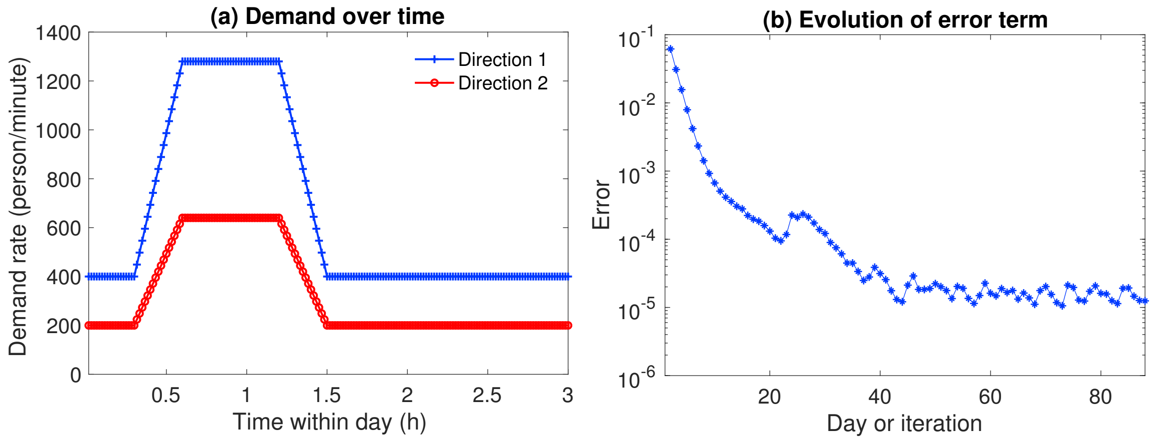

Fig. 2. (a) Total demand rate over time; (b) Evolution of error term ϵ in Eq.(32) from day to day.

length of a bus operation time interval with the same bus dispatching frequency, which is different from Δt. Different values of δ (2, 5, 15, 30, 45, and 60) are tested for sensitivity analysis.

The dispatched bus fleet size for the whole region is constantly equal to 300 as a benchmark case. The whole dispatched fleet is proportional to the demand level for both travel directions at the beginning of the modeling duration (2:1 in the benchmark case with asymmetric demand for both directions). The objective loading factor is set to be $\lambda_{obj} = 15$ (passenger-minute per bus). In practice, determination of λ_{obj} requires calibration and sensitivity analysis following the discussion in Section 3.1 and Section 3.2. In this paper, we can approximate its value by our simulation model. In particular, we test the performance (in terms of revenue) under different values of λ_{obj} and choose the λ_{obj} with the best performance.

4.2. Day-to-day dynamics

Firstly, we examine the convergence of traffic patterns to an equilibrium state after day-to-day evolution under the benchmark numerical setting. We define the error term (for a specific day q) as follows:

$$\epsilon^{(q)} = \frac{\sum_d \sum_w \sum_m \int_t r^{(q)}_{d,w,m}(t) \cdot \left| c^{(q)}_{d,w,m}(t) - C^{(q)}_{d,w,m}(t) \right| dt}{\sum_d \sum_w \int_t r_{d,w}(t) dt} \tag{32}$$

The error term is the absolute error between mean perceived cost and experienced cost, which is averaged over the direction, group, and departure time. Fig. 2(b) shows the error term for 90 days of traffic evolution, which indicate convergence of the traffic pattern. Note that here the ”90 days” are used for illustration purpose (for convergence). Later on we adopt ”30 days” for each period, as mentioned earlier.

We now compare the car accumulation profile, bus passenger accumulation, bus accumulation profile, and the PCE-based accumulation profile over clock time at the equilibrium solution under asymmetric demand and symmetric demand (for two directions) cases. The total demand intensity over time is identical. In Fig. 3, ‘AS’ and ‘S‘ refer to cases with asymmetric and symmetric directional demand respectively. ‘AS1’ and ‘AS2’ refer to two initial dispatched bus fleet allocations for two directions: ‘AS1’ means the initial allocation is proportional to the total travel demand, and ‘AS2’ is for equal allocation for both directions. Note that for the ‘S’ case, the two different allocations for ‘AS’ are identical since the demand is symmetric. $d = 1$ and $d = 2$ refer to moving directions of traffic. We have several key observations from Fig. 3. Firstly, traffic conditions under asymmetric and symmetric demand cases are very different, although the total demand intensity over time is the same. Therefore, it is important to incorporate asymmetrical demand in the multi-directional system. Secondly, under two different initial bus allocations for two directions in the asymmetric demand case, the bus accumulation over clock time (also the bus density or frequency)

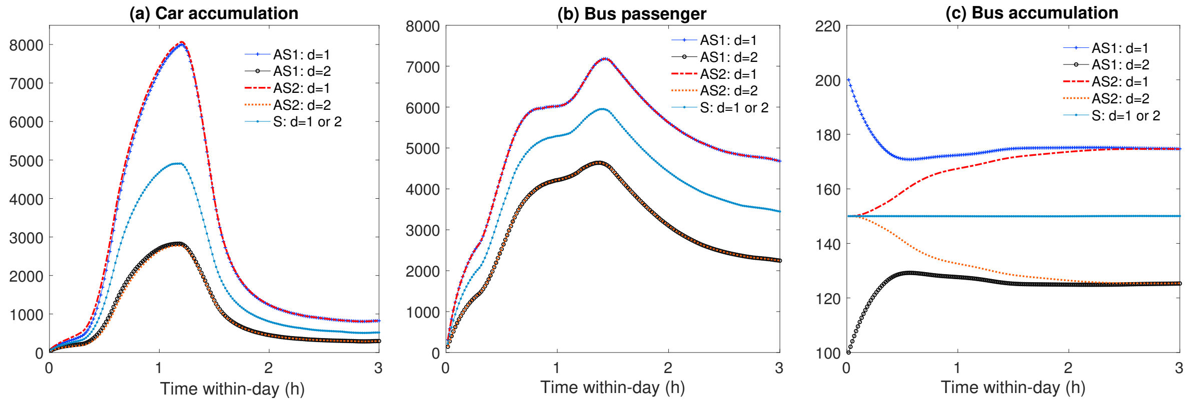

Fig. 3. Asymmetric and symmetric demand cases.

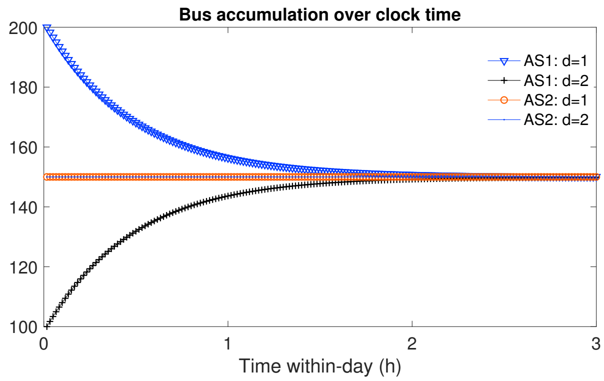

Fig. 4. Asymmetric demand case with constant bus speed.

endogenously evolve to the same pattern (refer to Fig. 3(c)). Thirdly, in the asymmetric case, due to spatially and temporally heterogeneous congestion, the real time bus densities for the two directions are different. This means that, traditional ways for setting bus schedules or service frequency might not necessarily provide the planned level of service due to congestion. Fourthly, the traffic condition in the symmetric case is generally in between of those in directions $d = 1$ and $d = 2$ since the demand level is also in between.

We now further illustrate the importance to incorporate traffic dynamics, where bus accumulation in the network as shown in Fig. 3(b) is endogenously determined. Fig. 4 shows the bus accumulations (for both directions) when we adopt constant speed for buses (i.e., $0.8 \times v_{cri} = 32km/h$ here), while all other demand setting and parameters are identical to those for Fig. 3(b). It can be seen that for case 'AS1' (the initial allocation is proportional to the total travel demand), the bus accumulations for both directions will evolve to the same value, i.e., steady state bus operation for two directions with equal bus frequency. In line with this, for the case 'AS2' (i.e., equal bus allocation for both directions initially), the bus accumulation remains constant over clock time (steady state again). These only occur where we adopt constant speed (without consideration of congestion). The performance is very different from those in Fig. 3(b). The observation in Fig. (4) might occur when there are dedicated bus lanes. If this is not the case, these results indicate that without taking into traffic dynamics, modeling of bus operation can be largely inaccurate.

4.3. Adaptive bus dispatching fleet size

We now illustrate the adaptive time-dependent bus dispatching shown in Fig. 5(a). For $\delta = 5$ (min), we adjust the planned bus dispatching for each discretized time interval with a length of $\delta = 5$ (min). In Fig. 5(a), '2P6', for example,

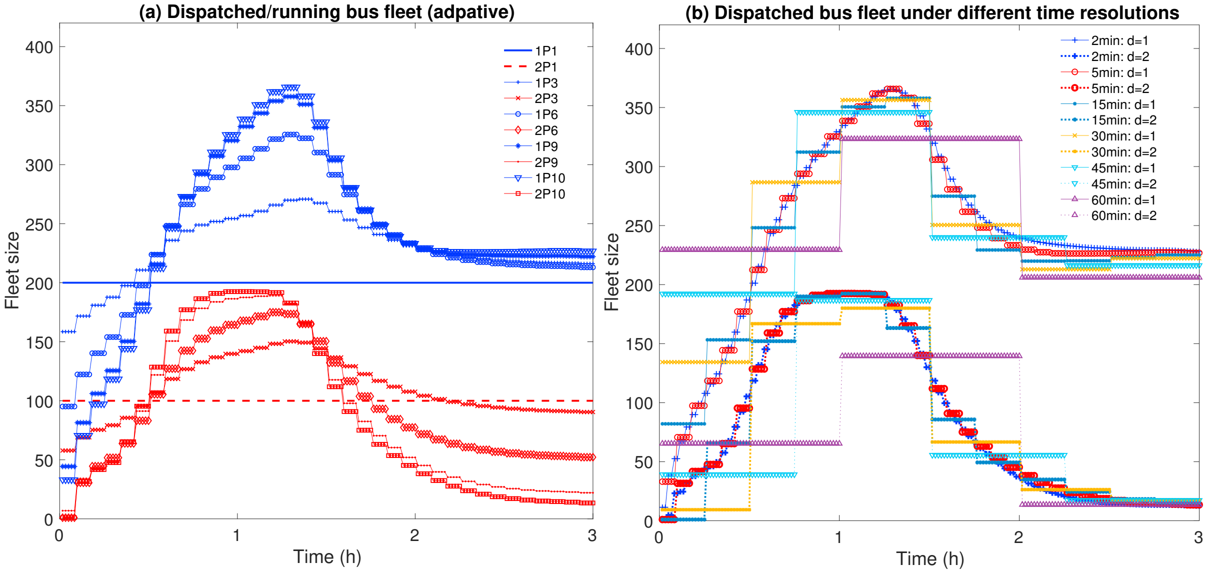

Fig. 5. (a) Evolution of bus dispatching (adaptive bus dispatching) with $\delta = 5$; (b) Evolution of the bus dispatching (adaptive bus dispatching) under different values for δ.

refers to the dispatched fleet size plan for direction $d = 2$ traffic in period 6 (after 5 adjustments, i.e., $5 \times Q = 120$ days). The same interpretation applies for others in Fig. 5(a), where we pick up those in periods 1, 3, 6, 9, and 10 for illustration purpose. As can be seen, the planned bus dispatching approaches a time-dependent pattern, which is compatible with the travel demand intensity over time. In particular, it is noted that there is a time lag between the peaked bus service and the peak demand duration. This is caused by the time lag between departure time and arrival time of travelers.

We also examine the cases with different values of δ, representing how frequent the dispatched fleet size is updated. In particular, we test six values, i.e., 2, 5, 15, 30, 45, and 60. Fig. 5(b) displays the planned time-dependent bus dispatching after 8 periods (with 7 adjustments). Fig. 5(b) is comparable to Fig. 5(a) with only the case of $\delta = 5$. It can be seen that, with a smaller δ, the dispatched fleet size profile over time becomes smoother, which can better accommodate temporal variations of travel demand. This is reflected in Table 2, where resultant net revenues and total system costs under different values of δ are summarized. We have the following observations. Firstly, with the increase of δ, the net revenue generally decreases and total system cost generally increases. Secondly, the changing rates of the net revenue and system cost are sharper for $\delta > 15$, and is relatively flat for $\delta < 15$. This means that, in practice, a smaller time interval of updating the dispatched fleet size (or frequency) does not necessarily bring too much additional benefit.

Table 2. Efficiency comparison under different values for δ.

δ (minute)	2	5	15	30	45	60
Total net revenue (10^5 AUD)	1.434	1.435	1.422	1.401	1.358	1.204
Total system cost (10^6 AUD)	1.480	1.478	1.480	1.493	1.516	1.568

Under the period-dependent bus dispatching adjustment in Fig. 5(a), Fig. 6 further displays evolutions of the total bus fee collected TBF, the total bus operating cost TBC, the total net revenue for operating bus services TBN, the modal-split, total user cost TUC and total system cost TSC over calendar time. It is evident that these measures have relatively sharp changes when we update the planned dispatched fleet size. As shown in Fig. 6, both fare-box revenue and operating cost increase, and the net revenue increases (i.e., fare-box revenue increases at a sharper rate). At the

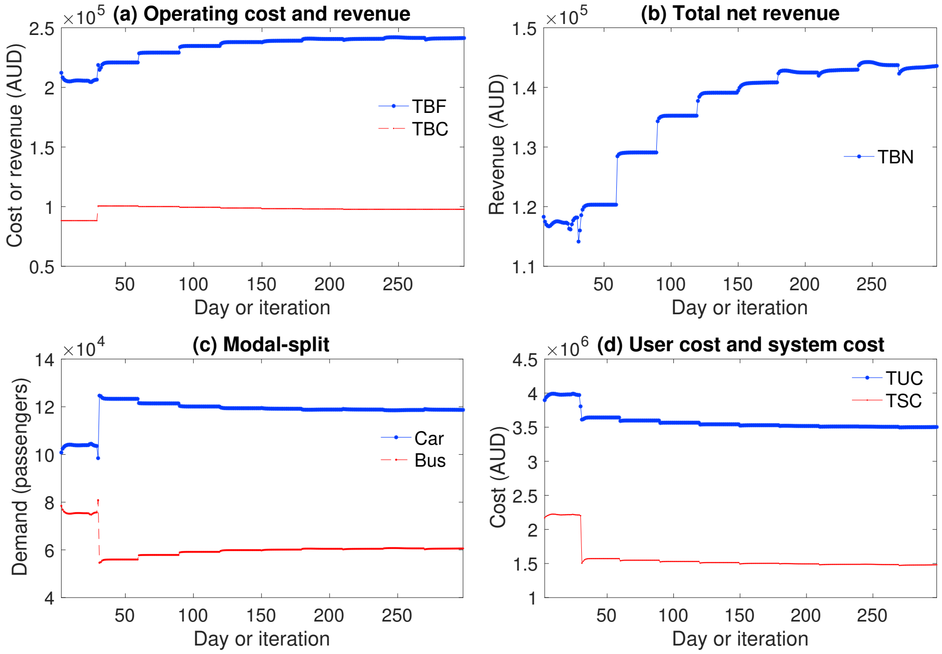

Fig. 6. Evolution of system efficiency measure against the bus dispatching.

same time, bus mode share increases (refer to Fig. 6(c)), and the total user cost and system cost almost remain constant. This is driven by the time-varying services. In response to demand, the general strategy is to dispatch a larger fleet size during the demand peak to increase competitiveness of bus services, a smaller fleet before the beginning of the peak to save operating cost, and a moderate size after the peak demand duration to transport the passenger accumulation to their destinations. as mentioned earlier, there is a time lag between departure time and arrival time).

There are several key points to be noted. Firstly, the bus operating cost is proportional to the total travel time of all operating buses. While this has taken into account the larger cost of a larger dispatched fleet, it does not consider that cost associated with the idling bus fleet during non-peak durations. This is reasonable in the sense that the life time of a bus is proportional to the time duration that it has been used, and the cost can be considered proportional to that time duration. However, this consideration can be readily relaxed by adding a cost for the un-utilized fleet during the non-peak. Secondly, the initial net revenue for bus operator is positive. This is dependent on the unit cost for operating bus and bus fares. If the unit operating cost is increased, or a lower bus fare is adopted, the initial net revenue can be negative (in some particular cases, the net revenue might never be positive). However, in this example, the revenue is positive and can be raised by the adaptive strategy by reducing the total operating cost. Thirdly, while the adaptive bus dispatching strategy aims at the net revenue of the bus operator, the user cost and total system cost is reduced at the same time. This implies that the two objectives do not necessarily conflict one another.

We now test our model under different demand conditions. Fig. 7 displays bus dispatching adjustment, the total net revenue for operating bus services, the total user cost, and total system cost over calendar time, where demand is only 5% of the benchmark example in Fig. 5(a) and Fig. 6 and other setting is identical. As can be seen in Fig 7(b), due to very low demand, initially the bus net revenue is significantly negative. With our adaptive bus dispatching adjustment, the dispatched bus fleet is decreasing over periods while the pattern over clock time still accounts for temporal variation for demand, which are shown in Fig. 7(a). This helps to reduce the loss of the bus operator, i.e., the net revenue becomes less negative. It should noted that reducing bus fleet under very low demand conditions, while reducing loss of the operator, may sacrifice user's benefit, as shown in Fig. 7(c). This means that the government might have to subsidize the bus operator to dispatch more buses for social welfare.

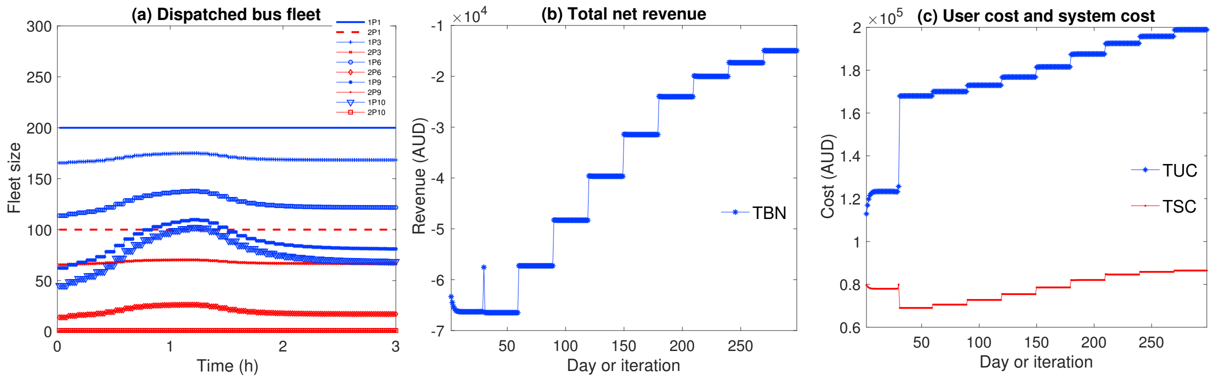

Fig. 7. Evolution of bus dispatching and system efficiency measure for the low demand case.

4.4. *Comparing the adaptive dispatching strategy with direct optimization*

We proceed to compare the adaptive strategy with the direct optimization of the time-dependent bus dispatching fleet size (to maximize net revenue). In particular, the numbers of dispatched buses in each time interval ($\delta = 5$(min)) are decision variables. One may find that due to the complexity of the traffic dynamics, this optimization problem is highly nonlinear and non-convex. How to obtain the global optimal solution is not the focus of this paper. Instead, we take the time-dependent bus dispatching strategy after implementing the proposed period to period adjustment (for 10 periods) as an initial solution and search locally to find a feasible and improved solution. We then implement the improved bus dispatching solution, which can lead to a net revenue of 1.451×10^5 AUD, as shown in Fig. 8(b). This is slightly higher than our adaptive approach. However, there are several advantages of our proposed that we would like to highlight in the following.

Firstly, implementing the directly optimized bus dispatching strategy (even if the numerical solution is sub-optimal), as can be seen in Fig. 8(a) for modal-split and in Fig. 8(b) for net revenue, can lead to larger fluctuations in the system. In practice, the fluctuation might be even larger, last for longer, and yield instability and unreliability for transport systems. Instead, our period-to-period adjustment could guide the system to an improved system gradually with less fluctuations. Secondly, the underlying assumption for the direct optimization is that we have all the parameters (including characteristics of the population and the system) calibrated and the solving algorithm to provide a good-quality optimization solution. If, for example, the value of time of the population is not well estimated, the direct optimization will not give us the ”correct” solution for bus dispatching. The proposed approach in this paper does not require such detailed information of the population and the system. As can be seen in Eq.(31), our approach mainly requires observation of passenger volumes or loading. These are usually readily observable with current smart transit card technologies. Thirdly, if the demand conditions or other parameters in the system change, with the direct optimization approach, the system needs to be re-calibrated and re-optimized. With our proposed adaptive approach, these changes will be automatically incorporated since we take into account day-to-day changes in the system and update the bus dispatching strategy from period-to-period.

5. Conclusions

This paper, by adopting the doubly dynamical approach proposed in Liu and Geroliminis (2017), examines the time-dependent bus dispatching problem in a multi-modal context. Existing explorations of the cross-modal congestion interaction often have limitations on modeling the time-dependent multi-modal and multi-directional flow interactions and the impact of bus dispatching on traffic conditions. Therefore, traditional studies optimizing the bus frequency or schedule may fail as the realized bus frequency or schedule is constrained by the dynamic traffic congestion on the road. In this context, this paper provides the methodology to model and optimize the time-dependent dispatched fleet size with complex multi-modal and multi-modal flow interactions taken into account. The traffic dynamics is modeled through the MFD framework. The proposed adaptive feet size adjustment mechanism utilizes

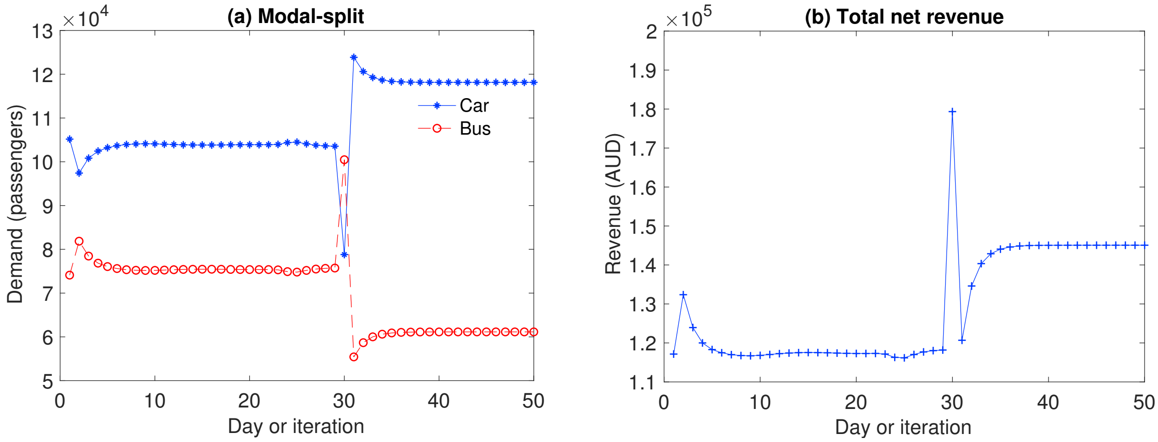

Fig. 8. Evolution of modal-split and total net revenue with an optimized bus dispatching scheme.

system conditions observed from day to day and drives the bus system to operate at a target level of bus loading. This study offers a new perspective for determining bus dispatching strategy and improves our understanding of multi-modal traffic dynamics.

This study is based on a single region city. It is of our interest to extend the current models to the multiple-region case, where a network can be partitioned into multiple regions with a well-defined Macroscopic Fundamental Diagram (MFD) by approaches studied in , e.g., Saeedmanesh and Geroliminis (2016). In this case, the modeling of cross-region buses becomes very relevant. Traffic conditions in one region might affect future bus supplies in other regions. Multiple-region pricing strategies or controllers and bus fleet management might have to be considered for improving traffic efficiency, such as those in Ramezani et al. (2015) and Kouvelas et al. (2017). While focusing on the mode choice, this study treats the departure time as exogenously given. Future study could incorporate the trip-timing choice into the doubly dynamic framework and examine the efficiency of the adaptive bus dispatching strategy. It is expected that the real-time traffic information will play a major role in the stability of such a system following the findings of Liu et al. (2017) and Guo et al. (2018).

Acknowledgements

We are grateful to three anonymous reviewers for their thoughtful and constructive comments, which helped improve both the exposition and technical quality of the paper.

References

Amirgholy, M., Gonzales, E.J., 2016. Demand responsive transit systems with time-dependent demand: user equilibrium, system optimum, and management strategy. Transportation Research Part B: Methodological 92, 234–252.
Amirgholy, M., Shahabi, M., Gao, H.O., 2017. Optimal design of sustainable transit systems in congested urban networks: A macroscopic approach. Transportation Research Part E: Logistics and Transportation Review 103, 261–285.
Ben-Akiva, M., Cyna, M., De Palma, A., 1984. Dynamic model of peak period congestion. Transportation Research Part B: Methodological 18, 339–355.
Bie, J., Lo, H.K., 2010. Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. Transportation Research Part B: Methodological 44, 90–107.
Bifulco, G.N., Cantarella, G.E., Simonelli, F., Velonà, P., 2016. Advanced traveller information systems under recurrent traffic conditions: Network equilibrium and stability. Transportation Research Part B: Methodological 92, 73–87.
Cantarella, G.E., 2013. Day-to-day dynamic models for intelligent transportation systems design and appraisal. Transportation Research Part C: Emerging Technologies 29, 117–130.
Cantarella, G.E., Velonà, P., Watling, D.P., 2015. Day-to-day dynamics and equilibrium stability in a two-mode transport system with responsive bus operator strategies. Networks and Spatial Economics 15, 485–506.

Chiabaut, N., 2015. Evaluation of a multimodal urban arterial: The passenger macroscopic fundamental diagram. Transportation Research Part B: Methodological 81, 410–420.
Friesz, T.L., Bernstein, D., Smith, T.E., Tobin, R.L., Wie, B.W., 1993. A variational inequality formulation of the dynamic network user equilibrium problem. Operations Research 41, 179–191.
Gayah, V.V., Daganzo, C.F., 2011. Clockwise hysteresis loops in the macroscopic fundamental diagram: an effect of network instability. Transportation Research Part B: Methodological 45, 643–655.
Gayah, V.V., Gao, X.S., Nagle, A.S., 2014. On the impacts of locally adaptive signal control on urban network stability and the macroscopic fundamental diagram. Transportation Research Part B: Methodological 70, 255–268.
Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B: Methodological 42, 759–770.
Geroliminis, N., Haddad, J., Ramezani, M., 2013. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. IEEE Transactions on Intelligent Transportation Systems 14, 348–359.
Geroliminis, N., Zheng, N., Ampountolas, K., 2014. A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. Transportation Research Part C: Emerging Technology 42, 168–181.
Gonzales, E.J., 2015. Coordinated pricing for cars and transit in cities with hypercongestion. Economics of Transportation 4, 64–81.
Gonzales, E.J., Daganzo, C.F., 2012. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. Transportation Research Part B: Methodological 46, 1519–1534.
Guo, R.Y., Huang, H.J., 2016. A discrete dynamical system of formulating traffic assignment: Revisiting Smith's model. Transportation Research Part C: Emerging Technologies 71, 122–142.
Guo, R.Y., Szeto, W., 2018. Day-to-day modal choice with a Pareto improvement or zero-sum revenue scheme. Transportation Research Part B: Methodological 110, 1–25.
Guo, R.Y., Yang, H., Huang, H.J., 2018. Are we really solving the dynamic traffic equilibrium problem with a departure time choice? Transportation Science 52, 603–620.
Guo, R.Y., Yang, H., Huang, H.J., Tan, Z., 2015. Day-to-day flow dynamics and congestion control. Transportation Science 50, 982–997.
Haddad, J., 2015. Robust constrained control of uncertain macroscopic fundamental diagram networks. Transportation Research Part C: Emerging Technologies 59, 323–339.
Ibeas, Á., dellOlio, L., Alonso, B., Sainz, O., 2010. Optimizing bus stop spacing in urban areas. Transportation Research Part E: logistics and transportation review 46, 446–458.
Iryo, T., 2008. An analysis of instability in a departure time choice problem. Journal of Advanced Transportation 42, 333–358.
Iryo, T., 2016. Day-to-day dynamical model incorporating an explicit description of individuals' information collection behaviour. Transportation Research Part B: Methodological 92, 88–103.
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.
Kouvelas, A., Saeedmanesh, M., Geroliminis, N., 2017. Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. Transportation Research Part B: Methodological 96, 26–45.
Leclercq, L., Chiabaut, N., Trinquier, B., 2014. Macroscopic fundamental diagrams: A cross-comparison of estimation methods. Transportation Research Part B: Methodological 62, 1–12.
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.
Li, X., Liu, W., Yang, H., 2018. Traffic dynamics in a bi-modal transportation network with information provision and adaptive transit services. Transportation Research Part C: Emerging Technologies 91, 77–89.
Li, X., Yang, H., 2016. Dynamics of modal choice of heterogeneous travelers with responsive transit services. Transportation Research Part C: Emerging Technologies 68, 333–349.
Liu, W., Geroliminis, N., 2016. Modeling the morning commute for urban networks with cruising-for-parking: An MFD approach. Transportation Research Part B: Methodological 93, 470–494.
Liu, W., Geroliminis, N., 2017. Doubly dynamics for multi-modal networks with park-and-ride and adaptive pricing. Transportation Research Part B: Methodological 102, 162–179.
Liu, W., Li, X., Zhang, F., Yang, H., 2017. Interactive travel choices and traffic forecast in a doubly dynamical system with user inertia and information provision. Transportation Research Part C: Emerging Technologies 85, 711–731.
Lo, H.K., Szeto, W.Y., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. Transportation Research Part B: Methodological 36, 421–443.
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.
Mahmassani, H.S., Liu, Y.H., 1999. Dynamics of commuting decision behaviour under advanced traveller information systems. Transportation Research Part C: Emerging Technologies 7, 91–107.
Mohring, H., 1972. Optimization and scale economies in urban bus transportation. The American Economic Review 62, 591–604.
Nagurney, A., 1993. Network economics: A variational inequality approach. volume 10. Springer Science & Business Media.
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.
Saeedmanesh, M., Geroliminis, N., 2016. Clustering of heterogeneous networks with directional flows based on "Snake" similarities. Transportation Research Part B: Methodological 91, 250–269.
Smith, M., Wisten, M., 1995. A continuous day-to-day traffic assignment model and the existence of a continuous dynamic user equilibrium.

Annals of Operations Research 60, 59–79.
Smith, M.J., 1984. The stability of a dynamic model of traffic assignment–An application of a method of Lyapunov. Transportation Science 18, 245–252.
Smith, M.J., Watling, D.P., 2016. A route-swapping dynamical system and Lyapunov function for stochastic user equilibrium. Transportation Research Part B: Methodological 85, 132–141.
Szeto, W.Y., Wu, Y., 2011. A simultaneous bus route design and frequency setting problem for Tin Shui Wai, Hong Kong. European Journal of Operational Research 209, 141–155.
Tabuchi, T., 1993. Bottleneck congestion and modal split. Journal of Urban Economics 34, 414–431.
Vickrey, W.S., 1969. Congestion theory and transport investment. The American Economic Review , 251–260.
Wang, W.W., Wang, D.Z., Zhang, F., Sun, H., Zhang, W., Wu, J., 2017. Overcoming the Downs-Thomson Paradox by transit subsidy policies. Transportation Research Part A: Policy and Practice 95, 126–147.
Watling, D., 1999. Stability of the stochastic equilibrium assignment problem: A dynamical systems approach. Transportation Research Part B: Methodological 33, 281–312.
Wu, W.X., Huang, H.J., 2014. Equilibrium and modal split in a competitive highway/transit system under different road-use pricing strategies. Journal of Transport Economics and Policy (JTEP) 48, 153–169.
Xiao, Y., Lo, H.K., 2016. Day-to-day departure time modeling under social network influence. Transportation Research Part B: Methodological 92, 54–72.
Yang, H., Liu, W., Wang, X., Zhang, X., 2013. On the morning commute problem with bottleneck congestion and parking space constraints. Transportation Research Part B: Methodological 58, 106–118.
Yildirimoglu, M., Geroliminis, N., 2014. Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams. Transportation Research Part B: Methodological 70, 186–200.
Zhang, F., Lindsey, R., Yang, H., 2016. The Downs–Thomson paradox with imperfect mode substitutes and alternative transit administration regimes. Transportation Research Part B: Methodological 86, 104–127.
Zhang, F., Yang, H., Liu, W., 2014. The Downs–Thomson Paradox with responsive transit service. Transportation Research Part A: Policy and Practice 70, 244–263.
Zhang, F., Zheng, N., Yang, H., Geroliminis, N., 2018. A systematic analysis of multimodal transport systems with road space distribution and responsive bus service. Transportation Research Part C: Emerging Technologies 96, 208–230.
Zheng, N., Geroliminis, N., 2013. On the distribution of urban road space for multimodal congested networks. Transportation Research Part B: Methodological 57, 326–341.
Zheng, N., Geroliminis, N., 2016. Modeling and optimization of multimodal urban networks with limited parking and dynamic pricing. Transportation Research Part B: Methodological 83, 36–58.
Zhong, R., Huang, Y., Chen, C., Lam, W., Xu, D., Sumalee, A., 2018. Boundary conditions and behavior of the macroscopic fundamental diagram based network traffic dynamics: A control systems perspective. Transportation Research Part B: Methodological 111, 327–355.