Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Schedule-Constrained Demand Management in Two-Region Urban Networks

Sakitha Kumarage , Mehmet Yildirimoglu , Mohsen Ramezani , Zuduo Zheng

http://pubsonline.informs.org/journal/trsc

TRANSPORTATION SCIENCE
Articles in Advance, pp. 1–26
ISSN 0041-1655 (print), ISSN 1526-5447 (online)

Schedule-Constrained Demand Management in Two-Region Urban Networks

Sakitha Kumarage,[a] Mehmet Yildirimoglu,[a] Mohsen Ramezani,[b] Zuduo Zheng[a]

[a] School of Civil Engineering, The University of Queensland, St Lucia, Queensland 4072, Australia; [b] School of Civil Engineering, The University of Sydney, Sydney, New South Wales 2006, Australia

Contact: s.kumarage@uq.edu.au, https://orcid.org/0000-0003-2422-4693 (SK); m.yildirimoglu@uq.edu.au, https://orcid.org/0000-0002-6534-4053 (MY); mohsen.ramezani@sydney.edu.au, https://orcid.org/0000-0001-6839-6839 (MR); zuduo.zheng@uq.edu.au, https://orcid.org/0000-0002-5289-2106 (ZZ)

Received: June 4, 2020
Revised: December 8, 2020
Accepted: February 14, 2021
Published Online in Articles in Advance: June 30, 2021

https://doi.org/10.1287/trsc.2021.1052

Abstract. Demand management aiming to optimize system cost while ensuring user compliance in an urban traffic network is a challenging task. This paper introduces a cooperative demand redistribution strategy to optimize network performance through the retiming of departure times within a limited time window. The proposed model minimizes the total time spent in a two-region urban network by incurring minimal disruption to travelers' departure schedules. Two traffic models based on the macroscopic fundamental diagram (MFD) are jointly implemented to redistribute demand and analyze travelers' reaction. First, we establish equilibrium conditions via a day-to-day assignment process, which allows travelers to find their preferred departure times. The trip-based MFD model that incorporates individual traveler attributes is implemented in the day-to-day assignment, and it is conjugated with a network-level detour ratio model to incorporate the effect of congestion in individual traveler route choice. This allows us to consider travelers with individual preferences on departure times influenced by desired arrival times, trip lengths, and earliness and lateness costs. Second, we develop a nonlinear optimization problem to minimize the total time spent considering both observed and unobserved demand—that is, travelers opting in and out of the demand management platform. The accumulation-based MFD model that builds on aggregated system representation is implemented as part of the constraints in the nonlinear optimization problem. The results confirm the resourcefulness of the model to address complex two-region traffic dynamics and to increase overall performance by reaching a constrained system optimum scenario while ensuring the applicability at both full and partial user compliance conditions.

Funding: This research was partly funded by iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

Keywords: demand management • departure times • network macroscopic fundamental diagram • day-to-day assignment • detour ratio

1. Introduction

Traffic congestion is a spatiotemporal process; it is the result of traffic demand exceeding capacity for a given roadway segment in a given period of time. Thus, curbing the imbalance of demand and supply (capacity) in time and space helps to alleviate traffic congestion. Travel demand management (TDM) strategies present an attractive direction in this regard by the application of techniques and policies aiming to reduce the demand or to redistribute it spatially and temporally.

Whereas improved public transport, ride-sharing incentive, and parking restriction strategies, among others, encourage travelers to shift to more sustainable travel modes, strategies such as route guidance, congestion pricing, and peak-hour pricing aim to redistribute the (private car) demand over space and time. These demand redistribution strategies become appealing when travelers have a strong attachment to their private cars due to diverse activity schedules and sparse distribution of attraction points.

We present a short- to medium-term (assuming inelastic demand) cooperative scheduling strategy that aims for effective temporal coordination of demand via the retiming of departure times within a limited time window in a large-scale urban network with two regions. Our goal is to minimize total time spent (TTS) in a two-region urban network by introducing limited schedule changes to travelers' departure times, anticipating that travelers have a tendency to comply if they are asked to make minor schedule changes. In fact, the resulting conditions represent "constrained" system optimum conditions (similar to

Jahn et al. 2005), where the optimal solution is constrained by limited schedule changes. Unlike parking restrictions or congestion pricing, the proposed method attempts to maximize social welfare by allowing all travelers to use public infrastructure by compensating their schedule rather than paying a price. An advanced traveler information systems (ATIS) platform in the form of a mobile app or a website, which has the capability of communicating with travelers and recommending a new departure time, is required for the success of the proposed TDM strategy.

Yildirimoglu and Ramezani (2020) introduce a demand management model that relies on limited changes in travelers' schedules and show its potential to increase mobility in congested homogeneous networks. Yildirimoglu and Ramezani (2020) focus on a single region network and assume that the demand management strategy could monitor the decision of all travelers in the network via an ATIS. This study (i) extends to a two-region network, (ii) enables the trip lengths to be updated with respect to time-varying traffic conditions using a detour ratio model, (iii) accounts for observed and unobserved demand within the optimization problem, and (iv) introduces a secondary control mechanism to increase the robustness.

The remainder of this paper is structured as follows. Section 2 evaluates the current methods and relatedness on demand management, MFD-based traffic control methods, and reaching equilibrium conditions with departure time assignments. Section 3 elaborates on the modeling methods and assumptions on two-region MFD-based traffic dynamics, day-to-day assignments, detour ratio modeling, and optimization frameworks. In Section 4, we present the results from numerical experiments and discuss the benefits of the proposed TDM strategy. Section 5 concludes this paper with findings and future research directives.

2. Related Works

Our focus in this work is to develop a demand management strategy for large-scale traffic networks to achieve a better network performance. We use two MFD-based traffic models to describe traffic dynamics; one representing the prediction (or optimization) model and the other representing the plant or the reality. As the proposed model builds on limited changes in departure times, it is crucial that the initial conditions realistically reflect the preferred departure times. Further, the stability of the proposed TDM strategy must be evaluated with respect to travelers' reactions. Thus, we implement a day-to-day assignment model both to capture preferred departure times and to explore the medium-term impact of the proposed TDM strategy. The following sections will provide the background information on three key pieces of the proposed framework: demand management, MFD modeling, and day-to-day assignment methods.

2.1. Demand Management

The spatial-temporal spread of vehicle volumes and the reduction of spent vehicle hours are the main objectives of any demand management strategy. However, the management of travel demand has been looked upon as an outcome of transport policy in the literature. Strategies such as road pricing (Yang 2005), congestion pricing (Lindsney and Verhoef 2001), parking restriction (Shoup 1997), ride sharing (Agatz et al. 2012), improving nonmotorized transport systems (Yu 2008), capacity increment of public transport (Wardman 2004), and implementing flexible work hours (Huang and Li 2011) could be identified as such methods to reduce or spread the traffic demand. The success of a TDM strategy depends on three pillars: (i) diversifying travel modes, (ii) providing acceptable incentives, and (iii) maintaining public participation and acceptance (Orski 1990, O'Flaherty 1997). The incapability of any TDM strategy to comply with these three pillars may result in inefficiency and lead to a stagnant share of private vehicles remaining on roads at peak hours. Our proposed strategy builds upon the aforementioned principles, where travelers are allocated minor schedule changes and are given the flexibility to comply. Further, it can operate at any level of public participation and maintain user retention. However, we assume that all travelers will keep using private cars due to the inherent restricted flexibility in schedules.

We harness the benefits of a connected environment or an ATIS with a cross-communication feature in the proposed TDM strategy. Although many traffic control strategies using ATIS are available in the literature (e.g., route guidance), less focus has been given to TDM strategies in large-scale traffic networks incorporating ATIS (Papageorgiou 2004). It should be noted that the envisioned ATIS infrastructure has more capability to influence traveler decision making (through cross-communication) than the conventional ATIS encountered in the literature. The available smart mobility apps in the market have the potential to facilitate requirements of the proposed TDM strategy. For instance, the ATIS infrastructure proposed by the U.S. patent of Chiu (2014) develops an incentive-based mobile application allowing people to choose departure times and routes. Travelers who comply with the recommended departure times and routes are given an incentive. The author does not develop a TDM strategy per se, but simply presents the ATIS concept. The proposed TDM strategy requires a similar ATIS infrastructure with the capability of collecting "preferred" departure times from travelers and recommending tailor-made "optimal" departure times subject to limited schedule change constraints; that is, recommended departure times cannot be far from preferred departure times.

It is obvious that existing ATIS infrastructures face challenges in maintaining 100% user penetration rates and user compliance. Hence, the observed demand obtained from an ATIS often underestimates the actual demand. ATIS infrastructure-based factors such as quality, reliability, and consistency of information, as well as individual-based factors such as awareness, willingness, and ability to use ATIS, cause partial compliance and lower penetration rates (Fox and Boehm-Davis 1998; Chen, Srinivasan, and Mahmassani 1999). Therefore, a robust and efficient TDM strategy should account for three different types of drivers: (i) travelers without ATIS, (ii) travelers with ATIS but in noncompliance with ATIS advice, and (iii) travelers with ATIS and in compliance with ATIS advice (Yin and Yang 2003). Moreover, the implementation of any TDM strategy will face an implementation period, where travelers adapt to the provided guidance and changing traffic conditions, and stages with different market penetration levels, which the system must cope with (Srinivasan and Mahmassani 2000). Hence, in this paper, we propose a TDM strategy with the capability of handling uncertain demand conditions and various market (user) penetration levels.

2.2. Macroscopic Fundamental Diagram and Network-Level Congestion Control Strategies

The demand redistribution strategy that we develop in this study requires traffic models that are capable of representing large-scale regional traffic dynamics. Traffic modeling in urban networks has been mostly based on microscopic or mesoscopic models, where traffic dynamics are defined at the link level. However, these models require a large amount of data and are computationally expensive, which limits their use for management and control purposes. The early works on macrolevel traffic control seen in Smeed (1966) and Wardrop (1968) highlighted control policies based on network operating capacity and road carriageway occupancy. The concept of network speed and network flow are explored in the works of Thomson (1967), Godfrey (1969), and Zahavi (1972). They found the existence of a linear decreasing relationship between space mean speed and average flow and concluded that maximum network speed is a property of the network. The concept of a network-level flow function with an optimal accumulation was initiated by Godfrey (1969). Similar ideas were later introduced by Herman and Prigogine (1979), Mahmassani, Williams, and Herman (1984), and Daganzo (2007), whereas the empirical existence of such a relationship was shown by Geroliminis and Daganzo (2008), who define it as a macroscopic fundamental diagram (MFD). An MFD is essentially a unimodal, low-scatter, and demand-insensitive relationship between average network flow (or production) and density (or accumulation).

MFD modeling assumes that outflow from a region depends on accumulation (n) rather than the history of inflows, and requires steady states or conditions with smooth demand changes over time. Geroliminis and Daganzo (2008) show that the outflow-to-production ratio (O/P) in a network is equal to the reciprocal of the average trip length (L). This relationship in fact refers to Little's formula (see Little 1961) being extended to the network level for steady-state systems. Thus, revealing a vital relationship between network accumulation, production, and outflow, $O(t) = P(n(t))/L$. The relationship between network speed and production could be obtained in a similar way by referring to the speed, density, and flow relationships. The so-called speed MFD can be calculated as $V(n) = P(n(t))/n$. Finally, the traffic dynamics of a neighborhood with homogeneous traffic conditions can be explained by $\frac{d}{dt}n(t) = f(t) - \frac{P(n(t))}{L}$, where $f(t)$ is the inflow to the neighborhood. Other empirical observations for real-world cities can be found in Buisson and Ladier (2009), Cassidy, Jang, and Daganzo (2011), and Geroliminis and Sun (2011). Modeling of urban network traffic via MFD paves the way for numerous traffic management applications. Table 1 presents some example applications from the literature, building on MFD modeling. Whereas there is a wide range of applications arising from MFD modeling, less research has been focused on demand management at the network level using MFD dynamics.

Despite many advantages, MFD does have several limitations. Buisson and Ladier (2009) show that the scatter and the shape of the MFD is highly influenced by the heterogeneity in traffic densities over the urban network. According to Leclercq et al. (2015), heterogeneity in traffic densities caused by network topology, trip length distribution, route choice in the network, and regional route patterns may eventually result in high scatter and hysteresis loops in MFD. Another issue is around the estimation of the MFD using empirical data; this is challenging, since collecting, storing, and processing a sufficient amount and quality of data in large-scale networks is costly and complex. Saffari, Yildirimoglu, and Hickman (2020) proposed a methodology to identify a small number of critical links in the network to monitor and estimate the MFD considering the detector measurements from those links.

Further, Batista, Leclercq, and Geroliminis (2019) discuss the impact of modeling with constant and dynamic trip lengths in regional and subregional networks. Mariotte, Leclercq, and Laval (2017) and Lamotte and Geroliminis (2018) discuss the trip length distribution and propose trip-based MFD models. This study implements a day-to-day assignment procedure, where the trip-based MFD model is essential to assess the costs that individuals experience from day to day.

Whereas the trip-based MFD model is a strong tool to model travelers' preferences, it does not account for changes in route choices that may happen as a result of changing traffic conditions. This study will explore the aggregate impact of individual route choices on a day-to-day assignment framework using detour ratio modeling. Detour ratio modeling accounts for the changes in the trip length with respect to changing traffic conditions. Yang, Ke, and Ye (2018) explore the relation of the detour ratio (actual distance/Euclidean distance) with respect to average trip length using taxi trajectory data, and Paipuri et al. (2020) reveal the behavior of the detour ratio with respect to average speed using mobile phone data. In this study, we propose to incorporate an abstract route choice model or a detour ratio model into the trip-based MFD model to manipulate trip lengths with respect to changing traffic conditions.

2.3. Traffic Equilibrium with Departure Time Choice

The equilibrium state of a transport system is a long-discussed topic with a myriad of research works. The majority of traffic equilibrium studies focus on the route choice aspect and the resulting fixed-point solution (Peeta and Ziliaskopoulos 2001), the empirical existence of which has been investigated by various

Table 1. Applications and Control Mechanisms of MFD

Purpose	Study	Characteristics
Perimeter control	Yang, Zheng, and Menendez (2018)	Hierarchical model predictive control strategy (MPC) with data from connected vehicles
	Aalipour, Kebriaei, and Ramezani (2019)	Prove optimal perimeter control is in form of bang-bang control
	Mohajerpoor et al. (2019)	Perimeter control with partial information feedback from the network
	Ingole, Mariotte, and Leclercq (2020)	Perimeter control considering user equilibrium conditions
Dynamic traffic assignment	Yildirimoglu and Geroliminis (2014)	Regional traffic assignment in multiregion traffic networks
	Batista and Leclercq (2019)	Dynamic traffic assignment in large-scale traffic networks
Route guidance	Leclercq et al. (2015)	Regional route assignment with partial accumulation values
	Yildirimoglu, Ramezani, and Geroliminis (2015)	Route guidance with an iterative assignment scheme
	Sirmatel and Geroliminis (2018)	Route guidance with regional paths and perimeter control with MPC
	Yildirimoglu, Sirmatel, and Geroliminis (2018)	Route guidance with regional paths and lower-level path assignment
Equilibrium with departure times	Liu and Geroliminis (2016)	Demand management under system optimum conditions with schedule changes
	Amirgholy and Gao (2017)	Formulate the user equilibrium over the peak as an ordinary differential equation
	Lamotte and Geroliminis (2018)	Modeling morning commute with departure times
Road pricing	Zheng and Geroliminis (2016)	MFD combined with an agent-based simulator to study road pricing
	Gu et al. (2018)	Distance-dependent area-based pricing models with MFD traffic dynamics
	Yang, Menendez, and Zheng (2019)	Congestion pricing integrated with perimeter control
Parking	Cao, Menendez, and Waraich (2019) Zheng and Geroliminis (2016)	MFD-based model to study time of cruising for parking under congested conditions and to derive dynamic feedback-based parking pricing schemes
Multimodal transport	Geroliminis, Zheng, and Ampountolas (2014) Loder et al. (2017)	Three-dimensional MFD modeling private and public transport traffic
	Ramezani and Nourinejad (2018)	MFD model with normal traffic flows and taxi dynamics

studies (Yildirimoglu and Kahraman 2018, González Ramírez et al. 2019). The initial works of Vickrey (1969) and Hendrickson and Kocur (1981) on revealing the deterministic user equilibrium with $\alpha-\beta-\gamma$ scheduling preferences, and subsequent works characterizing earliness and lateness costs (Newell 1987), equilibrium with first-in-first-out (FIFO) conditions (Daganzo 1985), stochastic dynamic equilibrium associated with disutility in travel time and schedule delay (De Palma et al. 1983), and deterministic dynamic equilibrium with experienced and perceived travel costs being equal and minimal (Ran, Hall, and Boyce 1996), established the theoretical base on which we frame our work (see Alfa 1986 for a summary). In our study, we consider an equilibrium-based approach for the choice/allocation of departure times (for travelers with a fixed desired arrival time) under constrained system optimum conditions and compare results with system optimum and user equilibrium approaches. Earlier studies on departure time choice focused on equilibrium concerning a single bottleneck. The expansion of a single bottleneck to a network level started with Ran, Hall, and Boyce (1996), who discussed equilibrium in a simple network with several origin-destination (OD) pairs, and Small and Chu (2003), who revealed equilibrium after a transient demand surge (hypercongestion) in a traffic network. Later, a more systematic approach was taken by Geroliminis and Levinson (2009) that incorporates Vickrey's equilibrium into MFD modeling and develops an area-based congestion pricing strategy. Subsequently, Arnott (2013), Fosgerau and Small (2013), Daganzo and Lehe (2015), Fosgerau (2015), and Lamotte and Geroliminis (2018) presented congestion pricing strategies based on Vickrey's equilibrium and MFD dynamics. In contrast to earlier studies, in this work we consider a scheduling strategy rather than a pricing strategy.

Without limiting to within-day equilibrium, we extend the TDM strategy to evaluate day-to-day equilibrium with a learning mechanism. According to Mahmassani, Williams, and Herman (1984), travelers make a departure time choice and a route choice on a day-to-day basis and reach an equilibrium point where no one has an incentive to change choice. Related work on convergence properties of the equilibrium in a day-to-day framework (Horowitz 1984), characteristics of stochastic and deterministic day-to-day learning models (Cantarella and Cascetta 1995), impact on leaning processes with asymmetric information (unavailability of an ATIS) on travel time (Nakayama, Kitamura, and Fujii 1999), and the day-to-day learning process for network-level applications (Chen and Mahmassani 2004, Liu and Geroliminis 2016, Guo et al. 2018, Yildirimoglu and Ramezani 2020) inspired us to formulate the day-to-day update model presented in this study.

3. Methods

In this study, we develop a demand management strategy to minimize TTS in the network by manipulating travelers' departure time within a limited time window. We adopt a model-plant approach, where the MFD-based traffic models represent the prediction model and the plant (reality). The trip-based MFD model, which builds on individual traveler attributes, serves as the plant. And the accumulation-based MFD model, which projects future aggregated traffic states, serves as the prediction model used in the optimization problem. As the proposed demand management strategy builds on departure time changes, it is imperative to start from an equilibrium scenario, where travelers' preferences regarding departure times are realistically captured. This will be done via the day-to-day assignment model, which updates travelers' departure time choices in response to changing traffic, and the trip-based MFD model, which evaluates the individual travel costs that travelers experience. Once the equilibrium scenario is established, the demand management strategy will be tested in a day-to-day manner. This allows us to explore travelers' reactions to the disruptive management scheme, which significantly changes traffic conditions and intervenes with travel decisions. In this section, we present (1) the two MFD-based traffic models deployed in developing the TDM strategy, (2) the proposed detour ratio model, which updates trip distances with respect to changing traffic conditions, (3) the day-to-day assignment model, and (4) the optimization problem.

3.1. Traffic Models

This study deploys two MFD-based traffic models: the accumulation-based MFD model and the trip-based MFD model. The accumulation-based MFD model found in the literature (see, e.g., Leclercq et al. 2015, Yildirimoglu and Ramezani 2020) is the conventional MFD model, which builds on the relationship between network production and the accumulation of vehicles. This is a parsimonious model; it does not consider individual vehicle attributes or driver characteristics but rather defines traffic dynamics through network production, average trip length, and vehicle accumulation.

The production MFD can be approximated by a right-skewed third-order polynomial function of accumulation (see Yildirimoglu and Ramezani 2020). MFD for each region r, $(r \in \{1,2\})$, is defined by $P_r(n_r(t)) = a_r n_r(t)^3 + b_r n_r(t)^2 + c_r n_r(t)$, where a_r, b_r, and c_r are estimated parameters for each region. Note that the average speed in region r is then $V_r(n_r(t))$

$= P_r(n_r(t))/n_r(t)$. Note that the production MFD may be sensitive to significant changes in the demand pattern and the loading profile (Leclercq and Paipuri 2019); however, the modeling of such changes on the MFD shape is outside the scope of this paper.

There are four different demand types in a two-region model, which are defined based on their origin and destination regions (see Figure 1(a)). These demand flows are denoted as $q_{rr}(t)$, $r \in \{1,2\}$ for intra-regional demand in each region, and $q_{rs}(t)$, $r \in \{1,2\}$ and $r \neq s$ for interregional demand from region r to region s. Similarly, four accumulation states are described as n_{rs}; $r,s \in \{1,2\}$, in which $n_{rs}(t)$ is the number of vehicles in region r with destination region s at time t. Therefore, the current accumulation of any region at a given time is $n_r(t) = \sum_{s=1}^{2} n_{rs}(t)$.

The accumulation-based MFD model builds on the relationship between network outflow and accumulation. The outflow of each region is calculated by $O_r(n_r(t)) = P_r(n_r(t))/L_r$, $r \in \{1,2\}$, considering that the demands are slow-varying and that the network average trip lengths L_r are constant. The transfer flow from region r to region s is given by $M_{rs} = \frac{n_{rs}(t)}{n_r(t)} \cdot \frac{P_r(n_r(t))}{L_r}$, where $r,s \in \{1,2\}$, $r \neq s$, and internal trip completion in each region is defined by $M_{rr} = \frac{n_{rr}(t)}{n_r(t)} \cdot \frac{P_r(n_r(t))}{L_r}$, where $r \in \{1,2\}$. Accordingly, the model state dynamics are

$$\begin{aligned} \frac{dn_{11}(t)}{dt} &= q_{11}(t) + M_{21}(t) - M_{11}(t) \\ \frac{dn_{12}(t)}{dt} &= q_{12}(t) - M_{12}(t) \\ \frac{dn_{21}(t)}{dt} &= q_{21}(t) - M_{21}(t) \\ \frac{dn_{22}(t)}{dt} &= q_{22}(t) + M_{12}(t) - M_{22}(t). \end{aligned} \tag{1}$$

Recently, a few studies (e.g., Arnott 2013, Lamotte and Geroliminis 2018, Mariotte and Leclercq 2019) explored the trip-based MFD model formulation, which provides a more detailed representation of traffic dynamics, as it represents each user in the network individually. Whereas the model accounts for each traveler's entrance and exit/transfer throughout the simulation, traffic performance in the network is based on the aggregate speed MFD relationship. The model executes a sequence of events, including departure, transfer, and arrival (completion) of trips in the network. At each event, the model calculates the region speed $V_r(n_r(t))$; $r \in \{1,2\}$ by referring to the current accumulation $n_r(t)$; $r \in \{1,2\}$ of the network using the speed MFD, and updates the distance traveled by all the vehicles in the network between previous and current events. These granular calculations repeated at every event makes the trip-based MFD model more detailed but computationally demanding. Whereas the trip-based model proposed by Mariotte and Leclercq (2019) extends to multiple reservoirs with multiple trip lengths and entry control restrictions, in this study we adopt a similar but more concise framework for a two-region network without entry control restrictions at boundaries between regions. Avoidance of entry control may lead to unrealistic gridlock conditions, particularly as demand surges. Nevertheless, the departure time choice mechanism, introduced in the paper is expected to curb such unstable conditions. However, evaluation of proposed TDM strategy with a multiregional trip-based model having such entry control restrictions could be a future research priority. In our two-region scenario (see Figure 1(b)), four trip types, namely, $R11, R12, R21,$ and $R22$, exist based on the origin-destination regions. The trip-based MFD model for the two-region network could be characterized by the following equations (see (2)), which relate departure time (t_i^{dep}), transfer time (t_i^{trs}), arrival time (t_i^{arv}), and network speed of the region ($V_r(.), r \in \{1,2\}$) to trip length ($l_i^{Rrs} r,s \in \{1,2\}$) of individual i depending on trip type:

$$\begin{aligned} l_i^{R11} &= \int_{t_i^{dep}}^{t_i^{arv}} V_1(n_1(t))dt \\ l_i^{R12} &= \int_{t_i^{dep}}^{t_i^{trs}} V_1(n_1(t))dt + \int_{t_i^{trs}}^{t_i^{arv}} V_2(n_2(t))dt \\ l_i^{R21} &= \int_{t_i^{dep}}^{t_i^{trs}} V_2(n_2(t))dt + \int_{t_i^{trs}}^{t_i^{arv}} V_1(n_1(t))dt \\ l_i^{R22} &= \int_{t_i^{dep}}^{t_i^{arv}} V_2(n_2(t))dt \\ t_i^{arv} &= t_i^{dep} + T_i^{exp}(t_i^{dep}), \end{aligned} \tag{2}$$

where $T_i^{exp}(t)$ is the experienced travel time for a traveler departing at time t. The iterative process of event-based updating is carried out until each vehicle completes its trip. The trip-based MFD model enables us to account for individual departure time choices and route choices, as it treats each traveler separately. The temporal variation of speed in the network tracked by the trip-based MFD model will be used in the detour ratio model to calculate the excess distance traveled by individuals in congested conditions. The excess distance is calculated using the detour ratio model when a traveler enters the network. Note that the regular trip-based MFD model does not account for the impact of trip rerouting or route choice in congested traffic conditions. We incorporate the detour ratio model into the trip-based MFD model and capture (approximately) the additional trip length due to congestion.

3.2. Detour Ratio Modeling

The detour ratio model presented in this section is implemented in the trip-based MFD model to account for the excess distance to be traveled by individuals in the congested traffic conditions. Travelers in the trip-based MFD model have identical origin-destination points across days (representing regular commuters), and the route they take or the distance they travel may change every day in response to changing traffic conditions in the network. Thus, we use detour ratio modeling to estimate the excess distance traveled with changing traffic conditions.

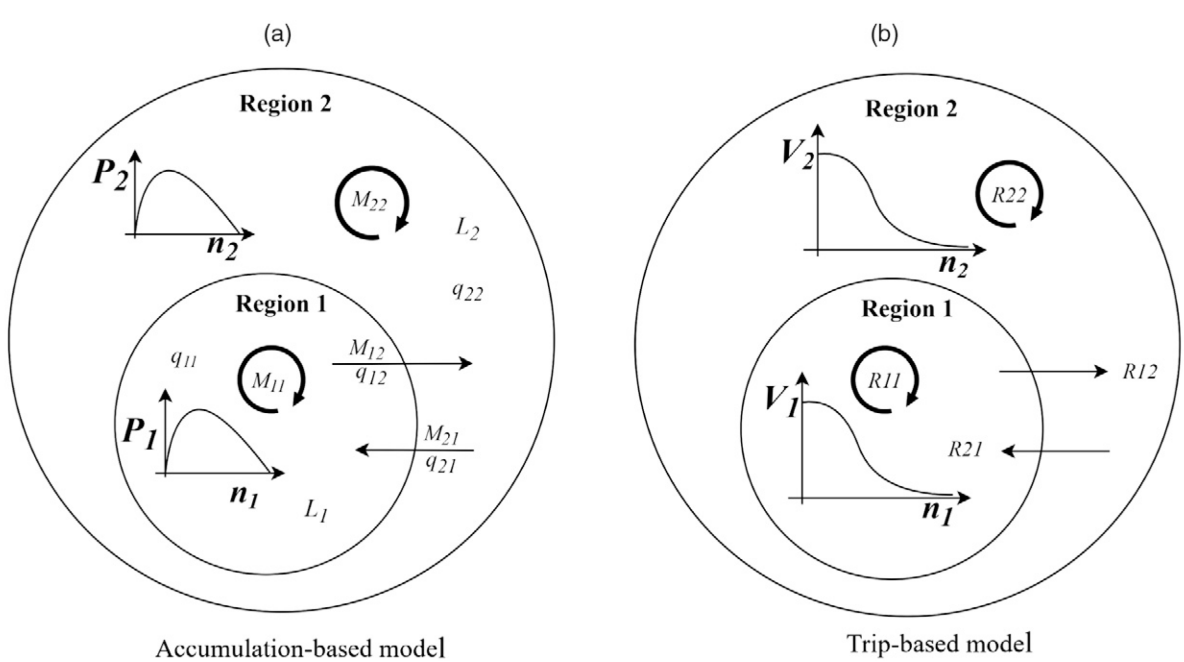
The actual traveled distance may change due to many factors, such as traffic conditions, network topology, and route choice. In this study, we are not interested in a traditional route choice model that estimates the paths chosen by travelers. We instead develop an abstract route choice model using detour ratios that represents the effect of average traffic conditions in a region on the traveled distance. Detour ratio was first introduced by Cole and King (1968), and the concept has since been used in other fields (Cardillo et al. 2006; Bebber et al. 2007; Boscoe, Henry, and Zdeb 2012; Zhang et al. 2015) to explain network topological characteristics and network accessibility and efficiency. Of particular interest to our study, it was found that the detour ratio is inversely proportional to the Euclidean distance with an intercept using taxi trajectory data (Yang, Ke, and Ye 2018) and that higher detour ratios at lower speeds and negligible detouring at highly congested conditions were observed using mobile phone data (Paipuri et al. 2020). However, none of these studies provide a multivariable analysis on detour ratios that incorporates both average network speed and average trip lengths.

Following these works, this section explores the combined relation between detour ratio, average network speed, and Euclidean distance. Detour ratio, in the context of this study, is the ratio of traveled distance to Euclidean distance between the origin-destination points.

We use an empirical data set to develop the detour ratio model; we use taxi trip data from Manhattan (Lower and Midtown Manhattan) on 28 days, extracted from New York City's open data portal (NYC OpenData 2018). The data set includes the origin-destination coordinates, the actual distance traveled by taxi, the time of the trip, and so on. We characterize the relation of detour ratio with two variables: network average speed and Euclidean distance. We calculate the average network speed at each time step by averaging the mean travel speed of taxis that start a journey at that time step. The time steps are discretized to five-minute intervals for network speed calculation. Finally, trips are grouped with respect to the average network speed (at the time of the departure) into intervals of 1.8 (km/h) (0.5 [m/s]) and with respect to Euclidean trip distances into intervals of 0.5 km.

In each group associated to a given network speed and a Euclidean trip distance, we observe a significant variation of detour ratios, which means that it is not

Figure 1. Two-Region Traffic Models: (a) Accumulation-Based MFD Model and (b) Trip-Based MFD Model



justifiable to consider a single average detour ratio for a given network speed and Euclidean trip length. Hence, detour ratios (D) in each group are fitted into a log-normal distribution with scale parameter μ and shape parameter σ. Panels (a) and (b) of Figure 2 show the fitted values of scale parameter μ and shape parameter σ. Note that each cell in the figure represents a set of trips grouped together with respect to the average network speed that they experience and their Euclidean distance. The Shapiro-Francia test (SF test), which measures the closeness to linearity in a quantile plot, was conducted for $\log(D)$ to identify the significance of goodness of fit and the agreement between observed data and log-normal distribution. The results show that there exists a significant fit with an SF test value W' greater than 0.9 in every single group, verifying the validity of the log-normal distribution to characterize the detour ratio distribution.

Panels (a) and (b) of Figure 2 indicate that μ and σ depend on both Euclidean distance and average network speed. Thus, we fit a two-variable two-degree polynomial function for μ and σ, where the independent variables are average network speed (x), Euclidean distance (y), and their two-degree combinations. The final formula for μ and σ after removing the insignificant variables is given by (3), and the parameters of the polynomial functions are in Table 2. Panels (c) and (d) of Figure 2 illustrate the estimated polynomials for $\hat{\mu}$ and $\hat{\sigma}$:

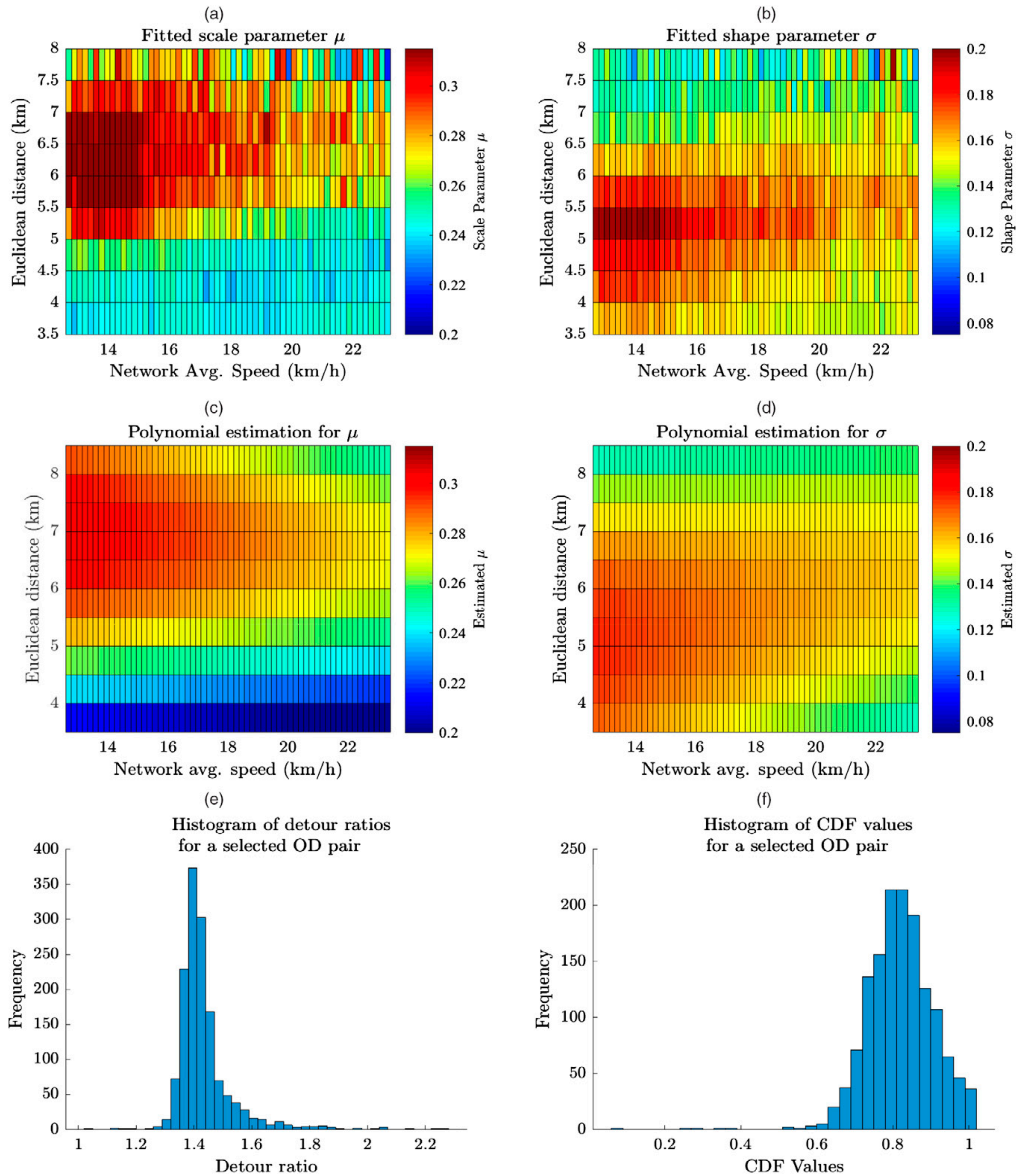
$$\begin{aligned} \hat{\mu} &= \alpha_0 + \alpha_1.y + \alpha_2.y^2 + \alpha_3.xy \\ \hat{\sigma} &= \beta_0 + \beta_1.x + \beta_2.y + \beta_3.y^2 + \beta_4.xy \end{aligned}. \quad (3)$$

Whereas the distribution parameters, μ and σ, are accurately modeled with a polynomial function, as shown in panels (c) and (d) of Figure 2, this does not allow the estimation of detour ratios for individual trips. Note that our final purpose is to integrate the detour ratios with the trip-based model and account for additional distance that individual travelers experience. In other words, there is still significant variation within each cell (or group of trips) that corresponds to a particular μ and σ, and it is not trivial to estimate the additional distance to be traveled by individuals. One way to estimate individual detour ratios is to draw random samples from the log-normal distributions presented in panels (c) and (d) of Figure 2. Nevertheless, since this procedure will be applied in a day-to-day assignment framework, it may result in inconsistent values across days. Our hypothesis is that there is a particular probability value for each trip, and its detour ratio can be defined as the inverse of the log-normal cumulative distribution function (CDF) (with μ and σ in the corresponding cell), evaluated at its probability value.

To verify our hypothesis, we choose an origin and a destination zone (O-D pair) and investigate the trips between them. It should be noted that each zone comprises 1 km^2 and that the zones are a 4.5-km Euclidean distance apart. Ideally, we would like to observe the same people repeating the same trips; however, this is not possible with the available taxi data set. Therefore, we assume that the trips between these two zones have the same characteristics and react similarly to changing traffic conditions. Figure 2(e) illustrates the histogram of detour ratios for taxi trips between the selected O-D pair. Whereas this represents a narrow distribution with more data closer to the mean, this analysis does not account for the varying traffic speed through time and changing distribution parameters. Note that we use 28 days of data and that the trips between these two zones are subject to changing average speeds in different time periods, which means that they are associated with different cells in panels (c) and (d) of Figure 2 and are subject to different distribution parameters. For each trip that we analyze, we compute the cumulative distribution function (CDF) value with respect to the distribution parameters, μ and σ, defined in the corresponding cell. Figure 2(f) shows the histogram of the resulting CDF values, which are distributed over a narrow range (80% of the trips that we analyze are distributed between CDF values of 0.75 and 0.9). The lesser variation in CDF values indicates that, despite changing distribution parameters, the detour ratios of the trips between two zones remain approximately within the same range in the associated distributions. Building on this analysis, we assume that there exists a particular probability or a CDF value for each trip. Traffic conditions, and therefore the average network speed, may significantly change in a day or across days. To determine the individual detour ratios, we will consider the associated log-normal distributions (corresponding to the Euclidean distance and the average network speed at the departure time of the trip) and calculate the inverse of the log-normal CDF at its probability value, which is defined prior to the day-to-day assignment procedure. This allows us to produce consistent detour ratios for the same trips across different days and eliminates the risk of fluctuating trip lengths in the day-to-day assignment framework.

The polynomials of μ and σ can be estimated for any other city, given that there are available data from probe vehicles. Data used to develop the detour ratio model had network speed ranging from 12.6 to 23.4 km/h (3.5 to 6.5 m/s) and Euclidean distance within the range 3.5–8.5 km. Hence, we normalized the polynomials developed for the New York taxi trip data set to match the network speeds (7.2–25.2 km/h) and Euclidean distances (2–7 km) observed in the numerical examples described in Section 4. Mathematically speaking,

Figure 2. (Color online) Estimation of Detour Ratio Distribution Parameters: (a) Fitted Scale Parameter μ, (b) Fitted Shape Parameter σ, (c) Polynomial Estimation of μ (See (3)), (d) Polynomial Estimation of σ (See (3)), (e) Histogram of Detour Ratios for a Selected OD Pair, and (f) Histogram of CDF Values for a Selected OD Pair



the detoured trip length (l_i^r) is given by $l_i^r = l_{i,0}^r \times D_i; r \in \{1,2\}$, where D_i is the detour ratio for the individual at given network conditions and $l_{i,0}^r$ indicates the Euclidean distance for traveler I (see Table 2 for details).

3.3. Day-to-Day Assignment

This section presents a short description of the day-to-day assignment model. In this study, a learning process is established via a day-to-day framework by allowing travelers to make a departure time choice on each day based on their perceived travel costs for several departure time alternatives. In the modeling framework, each traveler i has a desired arrival time (T_i^w) and a trip length ($l_i = l_i^1 + l_i^2$) with trip length components in region 1 (l_i^1) and region 2 (l_i^2). Every day, travelers make a departure time choice based on the perceived travel cost. The perceived cost for the next day is updated based on the historically perceived cost and the experienced or estimated travel cost on the current day. A traveler could experience the travel cost at the chosen departure time, but the cost for other departure times on a selected day has to be estimated using other sources of information.

The perceived generalized travel cost ($C_{i,d+1}^{\mathrm{p}}(t)$) of traveler i on the next day ($d+1$) is defined as a function of the perceived travel cost ($C_{i,d}^{\mathrm{p}}(t)$) on the current day (d), the experienced (or estimated) generalized travel cost ($C_{i,d}(t)$) on the current day (d), and a learning factor (ω_i), which characterizes the weight that individual i allocates to past and current travel costs. Mathematically, $C_{i,d+1}^{\mathrm{p}}(t) = \omega_i . C_{i,d}^{\mathrm{p}}(t) + (1-\omega_i)\cdot C_{i,d}(t)$, where $0 < \omega_i < 1$, $\forall i$. The higher learning factor indicates that the travelers rely more on historically

Table 2. Parameters Used in Numerical Simulations

Panel A: Accumulation-based model		
Production MFD function	$P_r(n_r(t)) = a_r.n_r(t)^3 + b_r.n_r(t)^2 + c_r.n_r(t)$, where $r \in \{1,2\}$ $a_r = 9.98.10^{-8}$, $b_r = -0.002$, $c_r = 9.78$, $r \in \{1,2\}$	
Jam accumulation	$n_r^{\mathrm{jam}} = 10^4$, $r \in \{1,2\}$	
Critical accumulation	$n_r^{\mathrm{cr}} = \frac{1}{3} n_r^{\mathrm{jam}}$, $r \in \{1,2\}$	
Trip-based model		
Initial average trip length	$\bar{L}_r = 4600$, $r \in \{1,2\}$	
Speed function	$V_r(n_r(t)) = P_r(n_r(t))/n_r(t)$ for $r \in \{1,2\}$	
Trip length at each region	$\begin{bmatrix} l_{i,0}^1 \\ l_{i,0}^2 \end{bmatrix} = \begin{bmatrix} \bar{L}_1 + \mathrm{N}(0,(0.2\bar{L}_1)^2) \\ \bar{L}_2 + \mathrm{N}(0,(0.2\bar{L}_2)^2) \end{bmatrix}$	
Trip length for each trip type	$\begin{bmatrix} l_i^{R11} \\ l_i^{R12} \\ l_i^{R21} \\ l_i^{R22} \end{bmatrix} = \begin{bmatrix} l_{i,0}^1 \\ l_{i,0}^1 \\ l_{i,0}^2 \\ l_{i,0}^2 \end{bmatrix} + \begin{bmatrix} 0 \\ l_{i,0}^2 \\ l_{i,0}^1 \\ 0 \end{bmatrix}$	
Updated distance from detour ratio model	$l_i^r = l_{i,0}^r .D_i \sim \log\!-\!\mathrm{normal}\,(\hat{\mu}, \hat{\sigma})$ subject to (s.t.) $D_i \in [1,2]$ for $r \in \{1,2\}$	
Earliness and lateness	$\begin{bmatrix} \Psi_i \\ \Lambda_i \end{bmatrix} = \begin{bmatrix} 0.5 \\ 4 \end{bmatrix} + \mathrm{N}\left(\begin{bmatrix} 0.05^2 & 0.1^2 \\ 0.1^2 & 0.4^2 \end{bmatrix}\right)$ s.t. $\Psi_i \in [0.3, 0.7]$ $\Lambda_i \in [2.5, 5.5]$	
Panel B: Day-to-day update		
Scenario	Before TDM strategy	During TDM strategy
Flexible time window	τ = 10 minutes	τ = 15 minutes, Ω = 10 minutes
Scale factor (θ)	$\theta = 0.01$	$\theta = 0.1$
Learning parameter (ω)	$\omega \in [0,1]$	$\omega \in [0,1]$
Detour ratio model		
Parameter estimate for μ, where x = Avg.networkspeed [m/s] y = Euclideandistance [m]	$\hat{\mu} = \alpha_0 + \alpha_1 y + \alpha_2 y^2 + \alpha_3 xy$ $R^2 = 0.583$ $SSE = 0.1742$ $\alpha_0 = 0.06240$, $\alpha_1 = 0.09442$, $\alpha_2 = -0.00869$, $\alpha_3 = -0.00131$	
Parameter estimate for σ, where x = Averagenetworkspeed y = Euclideandistance	$\hat{\sigma} = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 y^2 + \beta_4 xy$ $R^2 = 0.593$ $SSE = 0.1752$ $\beta_0 = 0.15730$, $\beta_1 = -0.01376$, $\beta_2 = 0.02558$, $\beta_3 = 0.00241$, $\beta_4 = 0.00241$	

perceived travel cost than the experienced travel cost on the current day. For simplicity, the learning factor (ω_i) is assumed to be identical for all the travelers ($\omega_i = \omega, \ \forall i$) in the numerical experiments.

The experienced (or estimated) generalized travel cost ($C_{i,d}(t)$) for traveler i, who departs at time t, is identified as the sum of the schedule cost and the travel time ($T_{i,d}(t)$). The schedule cost occurs to a traveler if the actual (estimated) arrival time ($t + T_{i,d}(t)$) is different from the desired arrival time (T_i^w). Hence, a traveler will experience an earliness cost (Ψ_i) or a lateness cost (Λ_i), depending on the arrival time. The experienced (or estimated) generalized travel cost for traveler i on day d departing at time t will be

$$C_{i,d}(t) = \begin{cases} T_{i,d}(t) + \Psi_i.(\mathrm{T}_i^w - t - T_{i,d}(t)), \\ \quad \text{if } (t + T_{i,d}(t)) < \mathrm{T}_i^w \\ T_{i,d}(t) + \Lambda_i.(T_{i,d}(t) + t - \mathrm{T}_i^w), \quad \text{otherwise} \end{cases}. \tag{4}$$

In this framework, we assume that travelers have access to instantaneous travel times for the alternative departure times through an ATIS ($T_{i,d}^{\text{ins}}(t) \,|\, t_{i,d}^{\text{dep}} - \tau < t < t_{i,d}^{\text{dep}} + \tau$), while they are fully aware of the experienced travel time for the chosen departure time($T_{i,d}^{\text{exp}}(t_{i,d}^{\text{dep}})$). As the discrepancy between instantaneous and experienced travel time may be significant in congested conditions, we assume that travelers correct the instantaneous travel time information and estimate the experienced travel times. This is achieved using the relative ratio between the experienced travel time($T_{i,d}^{\text{exp}}$) at departure time ($t_{i,d}^{\text{dep}}$) and the instantaneous travel time($T_{i,d}^{\text{ins}}$) at the departure time. See Yildirimoglu and Ramezani (2020) for further details.

We assume that travelers tend to make a departure time choice within a limited time window, which is centered around their previous day's departure time ($t_{i,d-1}^{\text{dep}}$), which means that a traveler has the flexibility to choose a departure time (t) within a window of ($t_{i,d-1}^{\text{dep}} - \tau < t < t_{i,d-1}^{\text{dep}} + \tau$), in which τ is half the size of the time window. The time window is discretized with (Δt) time steps: $t_{i,d}^{\text{dep}} \in \{t_{i,d-1}^{\text{dep}} + m \cdot \Delta t$ s.t. $m \in Z$ and $(-\tau/\Delta t) \le m \le (\tau/\Delta t)\}$.

The next-day departure times of individual travelers are selected by maximizing perceived utility. We use a concise logit model in this work and assume that the choice of departure time is independent of irrelevant alternatives. The incorporation of an extended random utility logit model can relax this assumption, but that is beyond the scope of this work. Therefore, the perceived utility of traveler i on day d and departure time t is defined as $U_{i,d}(t) = C_{i,d}^p(t) + \epsilon_i$, where the random error term (ϵ_i) is identically and independently distributed with a Gumbel distribution. Hence, the probability of choosing a departure time t by traveler i is given by

$$\Pr_i(t) = \frac{\exp(-\theta.U_{i,d}(t))}{\sum_{m=t_{i,d-1}^{\text{dep}}-\tau}^{t_{i,d-1}^{\text{dep}}+\tau} \exp(-\theta.U_{i,d}(m))}, \tag{5}$$

where θ is the scale factor.

The day-to-day model is considered to reach an equilibrium solution when the departure time choices of travelers remain relatively constant and they experience similar travel costs over subsequent days. Hence, at an equilibrium state, perceived costs and experienced costs should satisfy the following:

$$C_{i,*}^{\text{p}}(t) = C_{i,*}^{\text{p}}(t) + (1 - \omega_i).C_{i,*}(t) \quad \forall i. \tag{6}$$

This equation calls for a fixed-point solution, where perceived costs are equal to the experienced cost, $C_{i,*}^{\text{p}}(t) = C_{i,*}(t)$. The convergence of the two-region network to the equilibrium state is discussed in Section 4 with numerical experiments.

3.4. Network Optimization Problem

We develop an optimization problem that minimizes total time spent (TTS) in the network by introducing limited changes in travelers' schedule. The departure time choice that travelers make for a morning commute (or similar) depends upon the traffic conditions and schedule costs due to earliness or lateness. The travelers have less flexibility on desired arrival times due to externalities induced at work. Therefore, the study assumes that an individual traveler retains the same desired arrival time across days.

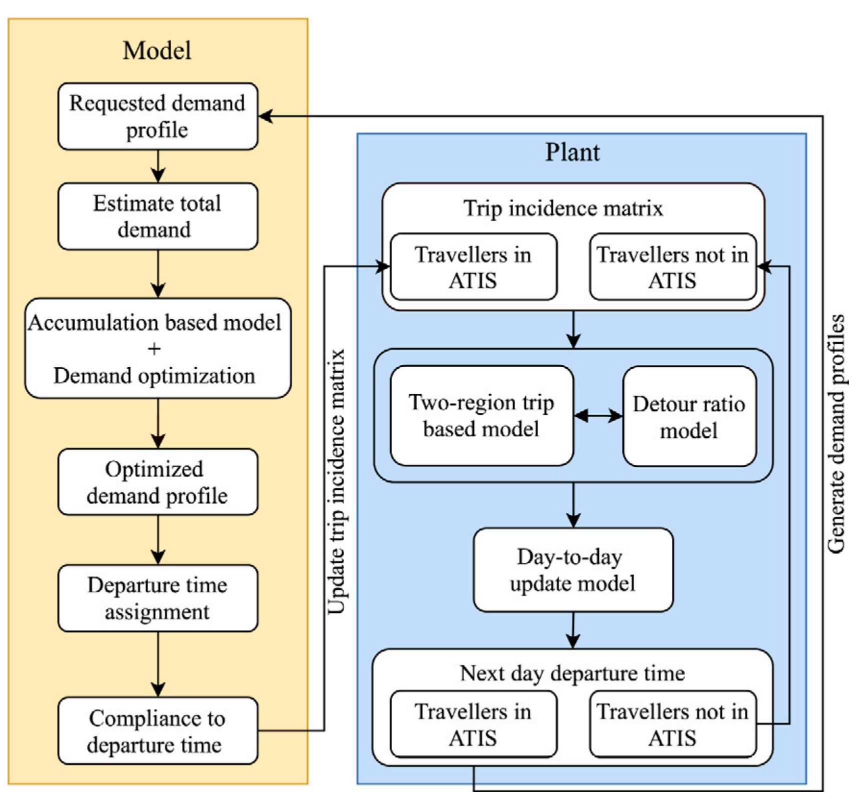
The success and robustness of the proposed scheme depends on its ability to maintain an active communication with travelers and to persuade them to comply with the given guidance. We refer to those who actively communicate with the demand management system (or those who actively use ATIS) as system users. However, the total demand consists of both system users (i.e., observed or requested demand) and nonsystem users (i.e., unobserved or estimated demand). To distinguish between these two types of users and to estimate the unobserved demand from the observed demand, we assume that the technology (user) penetration of the ATIS is known, since we can estimate the total demand (total number of users) in the network. Although we have no information about temporal variation of the unobserved demand, we assume that the system users are a representative sample of all network users, so that the temporal variation of total demand could be inferred from the representative sample. The system users request departure times from the ATIS and generate the requested demand profile $Q_{req}(t)$. The unobserved demand profile is estimated by scaling the requested demand profile using a time-independent scale factor or the penetration rate λ; that is,

$$Q_{est}(t) = (1/\lambda - 1) \times Q_{req}(t), \ \forall t.$$

Figure 3 demonstrates the modeling framework depicting the model-plant interaction. The plant includes both system users and nonsystem users. Given their departure time choice on the current day, we apply the two-region trip-based MFD model incorporating the detour ratio model and evaluate the experienced travel costs for both system and nonsystem users. With the newly experienced costs, we apply the day-to-day assignment model, which allows travelers to update their departure time choice while taking into account their perceived costs. The controller or the demand management system collects the departure time choices from the system users through the ATIS and builds the requested demand profile (i.e., time-dependent OD demand for system users). As this represents only a portion of the demand, we estimate the total demand in the system by up-scaling the requested demand with a known penetration rate. We then implement the optimization problem and minimize TTS in the system incorporating the accumulation-based MFD model and distinguishing between system and nonsystem users. Finally, the resulting "optimal" departure times are allocated to system users and are tested in the plant subject to compliance rules, which will be introduced in the following section.

The optimization algorithm assumes that the system users are willing (or incentivized) to make a shift in departure times within a limited time window $(-\Omega, \Omega)$, which means that travelers are willing to depart a few minutes earlier or later (e.g., 5 to 10 minutes) than their preferred or requested departure time. Further, we assume that the travelers request the departure time before the start of optimization, and thus the problem is formulated as a one-shot open-loop optimization problem iterated each day. It should be noted that this algorithm does not aim for socially optimal conditions by manipulating departure times but instead focuses on reaching a suboptimal (constrained optimal) equilibrium by constraining the problem to limited time windows. The optimization problem is formulated as a nonlinear nonconvex

Figure 3. (Color online) Modeling Framework of Network Optimization with Model and Plant Interactions



optimization problem considering all the aforementioned factors, as follows:

$$\text{minimize}_{Q_{\text{req}}} \quad \Delta k \sum_{k=0}^{k_f-1} \sum_{r=1}^{2} \sum_{s=1}^{2} n_{rs}(k) \tag{7a}$$

$$\text{subject to} \quad \text{for } k = 0, \dots, k_f - 1:,$$

$$\boldsymbol{N}(k+1) = g(\boldsymbol{N}(k), \boldsymbol{I}(k)), \tag{7b}$$

$$\boldsymbol{I}(k) = \left\{ \sum_{m=-\Omega}^{\Omega} \boldsymbol{Q}_{\text{req}}(k,m) \right\} + \boldsymbol{Q}_{\text{est}}(k), \tag{7c}$$

$$\boldsymbol{R}(k) = \sum_{m=-\Omega}^{\Omega} \boldsymbol{Q}_{\text{req}}(k+m, m), \tag{7d}$$

$$\boldsymbol{Q}_{\text{req}}(k,m), \boldsymbol{Q}_{\text{est}}(k) \geq \boldsymbol{O}, \tag{7e}$$

$$0 \leq n_{11}(k) + n_{12}(k) < n_{1,\text{jam}}, \tag{7f}$$

$$0 \leq n_{21}(k) + n_{22}(k) < n_{1,\text{jam}}, \tag{7g}$$

$$\boldsymbol{N}(k=0) = \boldsymbol{N}_0, \tag{7h}$$

where

$$\boldsymbol{N}(k) = \begin{pmatrix} n_{11}(k) & n_{12}(k) \\ n_{21}(k) & n_{22}(k) \end{pmatrix},$$

$$\boldsymbol{Q}_{\text{req}}(k,m) = \begin{pmatrix} q_{\text{req}11}(k,m) & q_{\text{req}12}(k,m) \\ q_{\text{req}21}(k,m) & q_{\text{req}22}(k,m) \end{pmatrix},$$

$$\boldsymbol{Q}_{\text{est}}(k) = \begin{pmatrix} q_{\text{est}11}(k) & q_{\text{est}12}(k) \\ q_{\text{est}21}(k) & q_{\text{est}22}(k) \end{pmatrix}, \boldsymbol{I}(k) = \begin{pmatrix} i_{11}(k) & i_{12}(k) \\ i_{21}(k) & i_{22}(k) \end{pmatrix},$$

$$\boldsymbol{R}(k) = \begin{pmatrix} r_{11}(k) & r_{12}(k) \\ r_{21}(k) & r_{22}(k) \end{pmatrix}, \quad \boldsymbol{N}_0 = \begin{pmatrix} n_{11}(0) & n_{12}(0) \\ n_{21}(0) & n_{22}(0) \end{pmatrix}.$$

The objective function in (7a) minimizes total time spent in the two-region network by the numerical integration of discrete accumulation values. The constant time step throughout the simulation, Δk, can be discarded from the objective function; t_f is the final time step. Equation (7b) defines the dynamics of the accumulation-based MFD model, where accumulation in the next time step $\boldsymbol{N}(k+1)$ is estimated by accumulation $N(k)$ and inflow $I(k)$ in the current time step. The function $g(.)$ predicts the accumulation in the next time step by solving the set of ordinary differential equations given in (1). The function is solved using the Runge-Kutta method (RK4) by temporal discretization of the prediction interval into a step size significantly smaller than Δk.

Equation (7c) describes the optimized demand profile or the resulting inflow $I(k)$ to the accumulation based model, which has both requested demand ($\boldsymbol{Q}_{\text{req}}$) and estimated demand ($\boldsymbol{Q}_{\text{est}}$) components. Note that $\boldsymbol{Q}_{\text{req}}(k,m)$ depends on two time indices, and it represents the number of travelers requesting to depart at time $k-m$ and being allocated to time k with a lag of $m \in [-\Omega, \Omega]$ time steps from their requested departure time. The requested demand component $\sum_{m=-\Omega}^{\Omega} Q_{\text{req}}(k,m)$ is the sum of all system users allocated to time step k, regardless of whether their requested departure time is early or late. Note that $[-\Omega, \Omega]$ is the flexible time window in which the departure times can be reallocated and that $\boldsymbol{I}(k)$ represents all the travelers who are allocated to depart at time k, irrespective of whether they are being shifted by m steps or whether they are system or nonsystem users.

Equation (7d) ensures that the requested demand is bounded within the flexible time window $[-\Omega, \Omega]$. Moreover, $\boldsymbol{R}(k)$ has only a $\boldsymbol{Q}_{\text{req}}$ component, because we manipulate the demand of system users only. We know that $\boldsymbol{Q}_{\text{req}}(k,m)$ indicates the travelers who initially request to depart at time $k-m$ and are allocated to the departure time of k with a time gap of m. Hence, $\boldsymbol{Q}(k+m,m)$ represents the system users that initially request to depart at time step k ($k = k+m-m$) and are allocated to time $k+m$. Inequality (7e) holds the nonnegative demand conditions. Inequalities (7f) and (7g) must hold to ensure that the region accumulation does not exceed the jam accumulation and remains positive. The $\boldsymbol{N}(k=0)$ in Equation (7h) are the initial accumulation states.

The optimization is based on the accumulation-based MFD model (model), which operates with system-level parameters (e.g., region accumulation, transfer flows). The trip-based MFD model (plant), however, operates with individual-level characteristics (e.g., individual trip length, departure time). Hence, certain assumptions are made when the outputs from the model are transferred to the plant due to the difference in granularity in the two MFD models. We follow a random sampling process of allocating departure time for particular travelers, irrespective of their trip length and schedule cost coefficients. For example, let the optimal allocated demand $Q_{\text{req}}(5,2) = 100$ [veh] (i.e., the number of vehicles allocated to the fifth step and departing two steps later than their request, which is the third step) and requested demand $R(3) = 400$ [veh]. We randomly sample 100 [veh] out of 400 [veh], independent of their attributes, and assign them to the fifth departure time period. Further, the accumulation-based MFD model (in the optimization formulation) is discretized with time steps Δk, whereas the trip-based MFD model (plant) can accommodate finer departure time choices (with steps Δt, where $\Delta t < \Delta k$). Hence, each allocated traveler needs to make a finer departure time choice within the time step k. The choice set for the traveler i allocated to time step k will be $t_i^{\text{dep}} + m.\Delta t$, where $\{t_i^{\text{dep}} + m.\Delta t \in \left[(k-1).\Delta k, k.\Delta k\right]\}$. The final departure time choice within this range is selected from this choice set using the logit formula presented in Equation (5).

Some travelers may experience relatively high costs as a result of the proposed optimization framework, which does not account for individual attributes. Hence, we introduce a secondary control mechanism to bound the optimization to what individual users are ready to accept. We propose to use perceived travel cost as the control criteria, and we assume that system users would not comply with the allocated departure time if the perceived cost of travel at the allocated departure time is more than $(1+r)\%$ of the travel cost experienced at equilibrium conditions (selfish Wardrop equilibrium). The rate of compliance r is subjective for decision, and we conduct a sensitivity analysis to understand the impacts of changing r for both individual benefits and network performance. In this approach, we ensure that system users are not maltreated by using the ATIS and for being willing to shift departure times.

Developing a demand management strategy that assumes that all users are system users is not realistic under operational conditions. It is not possible to assume that every traveler has access to and is willing to use the same level of technology, due to many demographic and behavioral reasons. Further, it is not practical to assume the consistent autonomous compliance of system users to allocated departure times in day-to-day operating conditions. Therefore, the proposed TDM strategy is tested with respect to varying penetration rates and compliance rules. User penetration rates are incorporated in the aforementioned optimization framework, and compliance rules are introduced as a secondary control mechanism. The robustness of the TDM strategy is tested while system users contribute to constrained system optimization by shifting their departure times and others (nonsystem users) travel according to user-optimal departure times. Evaluating the performance of overall network performance for different penetration rates helps us explore the feasibility of implementing a TDM strategy in realistic settings and the expected behavior at different phases where there may be different user penetration rates. Further, we can identify a minimum required user penetration rate for traffic congestion to be mitigated in the network. Section 4 will elaborate on this further and illustrate the benefits and the performance of the proposed TDM strategy.

4. Numerical Experiments

This section presents a numerical case study evaluating the impact of the proposed TDM strategy in the medium-term using a day-to-day assignment model. The demand profiles in the two-region network are heterogeneously loaded, as the inner region attracts 80% of the total demand while the rest are attracted to the outer region, demonstrating a monocentric city scenario with the central business district and the peripheral residential area. The numerical results are structured to elaborate on the simulation parameters, the performance of the TDM strategy with full compliance conditions, and the TDM strategy with partial compliance (secondary control) conditions, as well as to compare the TDM strategy with a pricing strategy.

4.1. Simulation Parameters

Before we test the TDM strategy, we create an equilibrium scenario (or no-TDM scenario). This is important not only for comparison purposes but also for producing an initial state in which the preference of travelers with respect to departure times is realistically captured. To build the equilibrium scenario, we let travelers choose a departure time based on the utility maximization principle and update their decisions every day. The trip-based model and day-to-day assignment model (components of the plant as presented in Figure 3) are run for several days (iterations) without incorporating the TDM model. This allows travelers to shift from their initial allocated departure time and converge to an equilibrium departure time choice. In this paper, we do not elaborate on the convergence properties and equilibrium conditions, as they have been thoroughly studied (Yildirimoglu and Ramezani 2020). Once the equilibrium scenario is created, we implement the TDM strategy via the modeling framework illustrated in Figure 3, using the *model* for demand optimization and the *plant* to evaluate the results. The iterative process is implemented for several days to explore the stability of the TDM strategy and its medium-term impacts. The optimization problem presented is a nonconvex nonlinear program (NLP) and was solved using the interior-point solver (IPOPT) (see Wächter and Biegler 2006) included in the optimization library of CasADi (see Andersson 2013) of MATLAB 9.4.0 (R2018a).

The presented simulation portrays a high congestion scenario, as the performance of a TDM strategy is more critical under such circumstances. The parameters used in the numerical simulation are presented in Table 2. The functional form of MFD, its coefficient values, jam accumulation, and critical accumulation for each region are the parameters used in the accumulation-based model. Note that, although we keep the average trip length used in the accumulation-based model (i.e., L_r) constant within the day, it is revised daily in the day-to day dynamics. We assume that we can obtain a representative sample of trip lengths via ATIS from 10% of the vehicles, and we update the average trip length in each region r with these observations for the upcoming day. The details of the trip-based MFD model are given with the functional form of speed MFD, trip length distributions, and earliness and lateness of cost distributions. In Table 2, the parameters used for the day-to-day update mechanism, such as a flexible time window, scale factor, and learning parameter, are given for both before and during the implementation of the TDM strategy. Two flexible time windows are in use during the implementation of TDM: one for the optimization problem (Ω) and one for the day-to-day assignment (τ), in which we set $\Omega < \tau$ to offer more flexibility to system users by increasing their choice set of departure times. The increase in scale parameter from $\theta = 0.01$ (before the TDM strategy) to $\theta = 0.1$ (during the TDM strategy) allows travelers to be more sensitive on departure time choice, as people are expected to be more alert when a new disruptive TDM strategy is introduced. Table 2 presents the parameters associated with the detour ratio model: fitted functions and coefficient estimates for μ and σ, threshold speed where the detour ratio model is activated, and the resulting trip length distribution after detour adjustment.

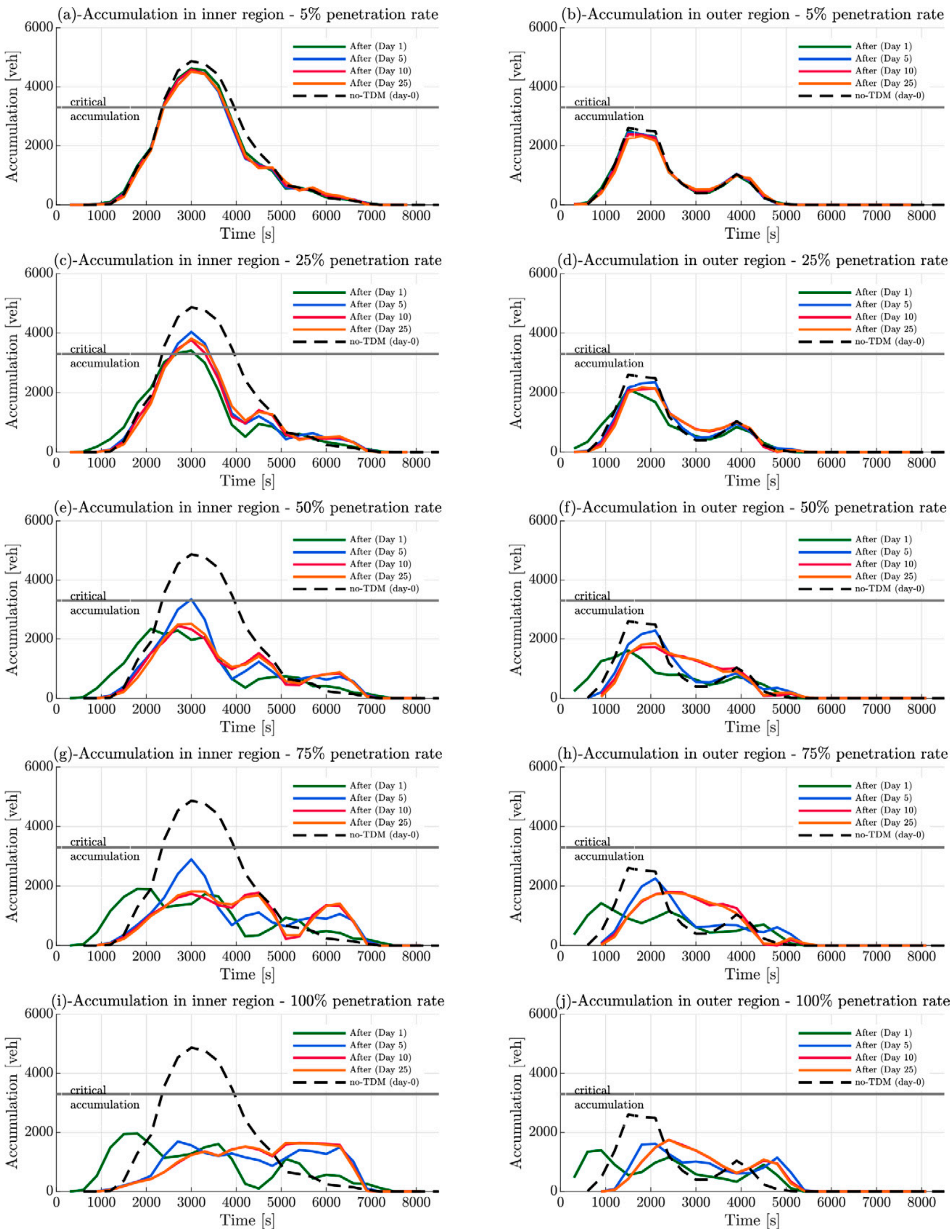
4.2. TDM with Full Compliance Conditions

The resulting impact of the TDM strategy could be observed by comparing regional accumulation levels before and after implementation. Figure 4 shows the variation of accumulation across days in the inner and outer regions when the TDM strategy is implemented with different user penetration rates, assuming full compliance of system users. The equilibrium accumulation pattern (black dashed lines) shows the existing traffic conditions in the no-TDM scenario (i.e., day 0—equilibrium conditions before the TDM strategy is introduced). The inner region accumulation on day 0 goes above the critical accumulation level, indicating highly congested traffic conditions. The day-to-day evolution of accumulation patterns after implementing the TDM strategy is shown in colored solid lines (i.e., accumulation profiles for days 1, 5, 10, and 25). These accumulation profiles show varying patterns across the penetration rates; nevertheless, in all scenarios, we observe that days 15 and 25 are almost identical. Although travelers are given an opportunity to change their schedules on a daily basis, a steady pattern is observed in day-to-day evolution toward the end of the 25-day period. This indicates the convergence of the system into a steady state, where the departure time choices and allocations become consistent throughout the days. Note that the choices and allocations are different from each other on day-25, which means that the TDM strategy is actively making schedule changes; nevertheless, the choices and allocations become individually consistent across days. Although travelers have the possibility to shift day-to-day, the traffic conditions are assembling to an equilibrium state, irrespective of the schedule changes caused by the TDM strategy. Overall, despite a disruptive demand management strategy that reshapes the demand profiles, the system reaches a stable state at the end of the medium-term period.

User penetration rate plays a vital role in reducing traffic congestion observed in the network. Panels (a) and (b) of Figure 4 present the traffic conditions when 5% of the total network users (10,000 network users) are controlled by the TDM strategy—ignoring the compliance issues that will be discussed later. The traffic conditions here do not exhibit a significant change from the equilibrium scenario, as 95% of the travelers make departure time choices based on their individual preference. On the other hand, the 25% penetration rate scenario presented in panels (c) and (d) of Figure 4 reveals a higher improvement, where the inner region accumulation could be lowered closer to the critical accumulation level. The accumulation profiles of inner and outer regions for 50%, 75%, and 100% penetration rates are depicted in panels (e)–(f), (g)–(h), and (i)–(j) of Figure 4, respectively. A diminishing marginal benefit is observed in accumulation levels with the increase in user levels beyond 25%. This implies that the proposed TDM strategy can curb traffic congestion in the network by controlling only a small percentage of travelers. Although the full control scenario (100% penetration rate) has a lower accumulation profile than that observed in the 25% user rate, both scenarios lower the peak accumulation below the critical accumulation level and demonstrate uncongested traffic conditions. Having an accumulation level closer to critical accumulation is in fact a sign of optimal network use without causing underutilization of the available infrastructure. An accumulation profile with a 25%–30% user rate enables an uncongested scenario with proper utilization of resources while demonstrating practical applicability, as it requires only 25%–30% of the travelers to use the TDM strategy.

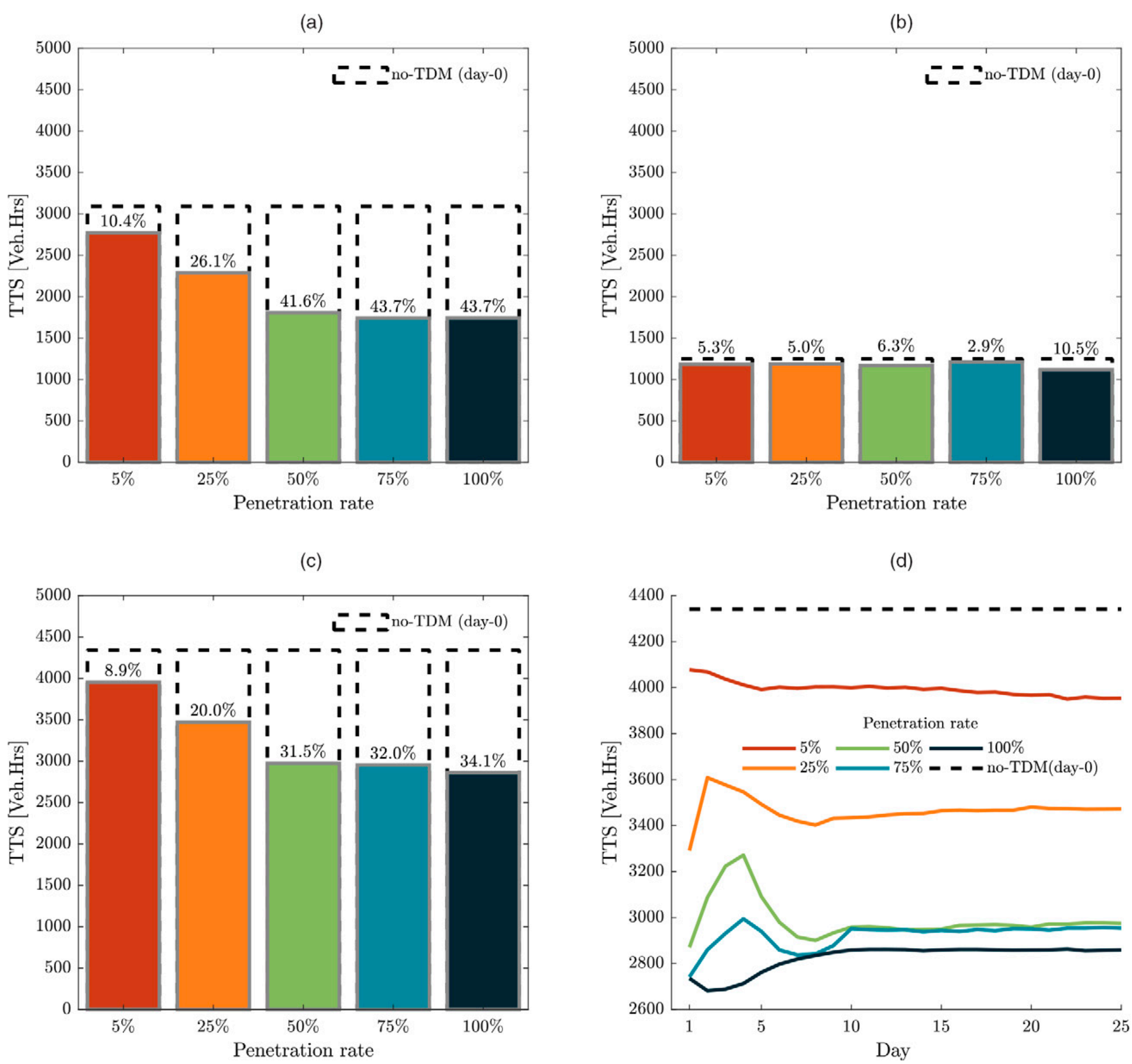
Figure 5 presents the network performance with respect to different (system) user penetration rates at the end of the 25-day period. The TTS before implementing the TDM strategy is indicated as the black dashed lines in the graphs. According to Figure 5(a), a significant reduction in TTS in the inner region is observed for all penetration rates, which confirms highly congested traffic conditions before the implementation of the TDM strategy. On the other hand, Figure 5(b) indicates a minor improvement for the outer region compared with the inner region. Because traffic conditions are not congested in the outer region in the no-TDM scenario, the TDM strategy brings only limited benefits. Further, it confirms that the outer region is not necessarily penalized for overall improvement, as it is in the perimeter control applications. The overall reduction in TTS for the whole network is shown in Figure 5(c). The results show that the TDM strategy is beneficial, even at low user penetration rates; for example, controlling 25% of travelers brings about a 20% reduction in TTS (5.4 minutes

Figure 4. (Color online) Accumulation in the Inner Region: (a) 5% Penetration Rate, (c) 25% Penetration Rate, (e) 50% Penetration Rate, (g) 75% Penetration Rate, (i) 100% Penetration Rate; Accumulation in the Outer Region: (b) 5% Penetration Rate, (d) 25% Penetration Rate, (f) 50% Penetration Rate, (h) 75% Penetration Rate, (j) 100% Penetration Rate



saved in trip time per traveler on average). We observe that the marginal benefit in TTS significantly decreases with increasing penetration rates; there is almost no difference between the 50%, 75%, and 100% scenarios. Figure 5(d) illustrates the day-to-day variation of TTS for different penetration rates. On the first day, where the TDM is introduced, 75% and 100% scenarios produce almost the same TTS, and over subsequent days, the 100% scenario converges to a slightly better TTS, which implies that the optimization problem produces equal or better results with more system users, as expected. We also see an increasing trend in TTS for the 25%, 50%, and 75% penetration levels in the first few days of the TDM strategy. In the initial days of implementing the TDM strategy, nonsystem users may experience a higher variation between the perceived costs and experienced costs (nonsystem users are not aware of the traffic conditions caused by the departure time allocations of the TDM strategy). But, over time, nonsystem users adapt to the new conditions and propagate stable conditions with the day-to-day update model, resulting in TTS profiles with sudden upsurges on initial days that smoothly converge to stable levels. However, this trend is not observed at the 100% penetration level (as all travelers are system users) and the 5% penetration level (as the penetration level is not significant enough), but it is evident that the learning process over subsequent

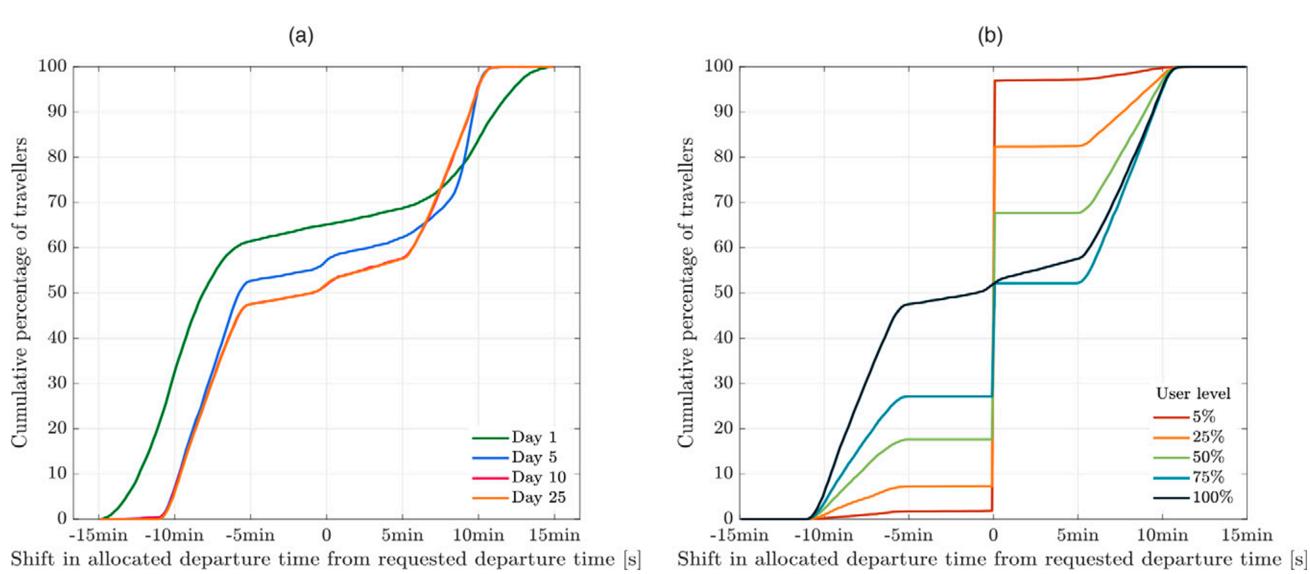
Figure 5. (Color online) TTS for Different Penetration Rates: (a) in the Inner Region; (b) in the Outer Region; (c) in the Whole Network; and (d) Across Days



days leads the traffic conditions to stabilize at the end of the 25-day period for all penetration levels, which implies the convergence of the system to an equilibrium at all scenarios. Note that system users follow the departure time guidance resulting from the TDM strategy (i.e., the optimization problem based on the accumulation-based model), whereas the nonsystem users simply rely on their experience in the trip-based MFD model. Although the TDM strategy pushes the system toward system optimum conditions by making changes in the schedule of system users, nonsystem users act in the principle of selfish equilibrium. Note that these two equilibrium states are not different in the uncongested traffic scenarios (Sheffi 1984). Moreover, 50%, 75%, and 100% scenarios are all uncongested; the peak accumulation is nowhere near the critical accumulation, which means that decisions based on user equilibrium and system optimality are practically the same in this range for system user levels.

The demand redistribution method adopted in the TDM strategy aims to achieve the desired outcome by assigning departure time shifts within a limited time window. Although the TDM significantly improves network traffic conditions, it is essential to examine the circumstances faced by individuals while the TDM strategy is being implemented. Figure 6(a) presents the day-to-day variation in the cumulative distribution of departure time shifts experienced by travelers in the 100% penetration rate scenario, that is, the difference between requested and allocated departure times on a given day. The shift becomes consistent over subsequent days and converges to an identical distribution; cumulative distributions of day 10 and day 25 are approximately identical. A similar pattern is observed with other user penetration rates. The flexible time window (Ω) in the optimization algorithm is set to two time steps (corresponding to 10 minutes), but we can see some system users are being shifted by more than 10 minutes on day 1 (see the green line). This occurs due to different time steps implemented in the optimization algorithm (model) and the day-to-day assignment (plant). Note that, whereas the time step in the optimization problem (or the accumulation-based MFD model) is Δk = 5min, the day-to-day assignment considers discrete intervals of Δt =1min. As the flexible time window in the optimization problem is Ω=2, it corresponds to a flexible period with five Δk time steps (i.e., −2, −1, 0, 1, 2), which, in turn, is equivalent to smaller Δt time steps (i.e., [−10, −5), [−5, 0), [0, 5), [5, 10), [10, 15)). Let us take, for example, a traveler requesting to depart at time step $k = 0$ and being allocated to time step $k = 2$. We also assume that travelers make a finer departure time choice t within the time step k. Therefore, this traveler can experience at most a shift of 14 minutes (from $t = 0$ to $t = 14$). Figure 6(b) depicts the cumulative distribution curves of the shift in allocated departure time from requested departure time for different penetration rates on day 25. In the full control scenario (100% penetration rate), all travelers may experience a shift in departure time, but, in the partial control scenario, only system users experience a shift in departure time. Nevertheless, the percentage of such travelers increases with a decreasing penetration rate, as expected. Irrespective of whether being a system user

Figure 6. (Color online) Cumulative Distribution of Departure Time Shift: (a) Across Days in the 100% Penetration Rate Scenario and (b) Across Penetration Rate Scenarios on the Final Day



or not, all travelers experience uncongested traffic conditions at partial control scenarios, as seen in Figure 4. Hence, the nonsystem users tend to experience uncongested traffic conditions at the expense of departure time shifts absorbed by the system users.

The two-region urban network has four distinct demand cases based on the origin-destination region as $R11$, $R12$, $R21$, and $R22$, where the total demand is distributed at the 40%, 10%, 40%, and 10% ratios, respectively, allowing the inner region to attract 80% of the total demand. For each demand case, there is a requested demand (Q_{req}) profile and an optimized demand profile (Q_{opt}) to compare. Figure 7 shows the variation of Q_{req} (dashed lines) and Q_{opt} (solid lines) for different demand types (in the 100% penetration rate scenario). The Q_{req} and Q_{opt} demand profiles of $R11$ and $R12$ are shown in Figure 7(a), as they originate in the inner region (region 1). Similarly, Figure 7(b) shows that the demand profiles of $R21$ and $R22$ originated in the outer region (region 2). The optimized demand profiles, in general, stretch the requested demand profiles along the two directions (to earlier and later periods). The stretch in the demand profiles should happen in harmony or in coordination so that the overall system performance is maximized. In Figure 7(b), we see a stretching effect occurring in both demand types and note that the optimized demand alternately prioritizes each demand type, such that when Q^{R21}_{opt} has a higher demand, Q^{R22}_{opt} exhibits a lower demand and vice versa. Further, we note that travelers in different demand categories naturally have different trip length distributions. The intraregional travelers ($R11$ and $R22$) have shorter trip lengths compared with interregional ($R12$ and $R21$) travelers. Whereas the two-region urban network allows us to represent two regions with distinct traffic conditions, it also captures travelers with short and long trip lengths and treats them separately.

4.3. TDM with Partial Compliance Conditions

Although the system users actively use the ATIS platform and exchange requested/allocated departure times, they may not comply with the allocated time if they sense a significant increment in their travel cost (experienced cost). Hence, we use the secondary control method where system users would not comply with the allocated departure time if the perceived cost of travel at the allocated departure time is more than $(1 + r)\%$ of the travel cost experienced at equilibrium conditions (no-TDM). Figure 8(a) presents the cumulative change in experienced cost for travelers under different compliance conditions varying $r\%$ from −15% to +50% ($(1 + r)\% = 85\%$–150%). The cumulative plots indicate the cumulative percentage of travelers with increased experienced cost compared with the no-TDM scenario. We see that more than 80% of travelers experience lower costs than in the no-TDM scenario, in all compliance conditions. Figure 8(b) plots the day-to-day variation on total time spent (TTS) for different compliance levels, and we observe higher TTS when the compliance levels are lower. However, the variation in TTS is insignificant and indicates that allowing a few

Figure 7. (Color online) Requested and Optimized Demand Profiles for (a) the Inner Region and (b) the Outer Region

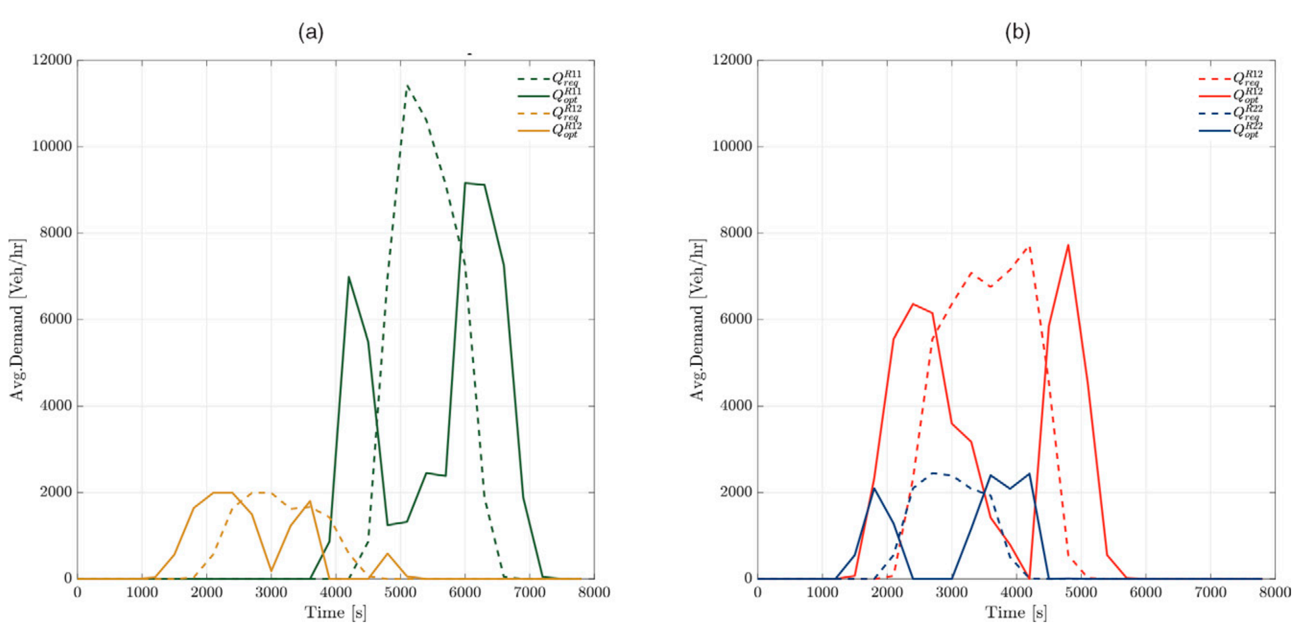
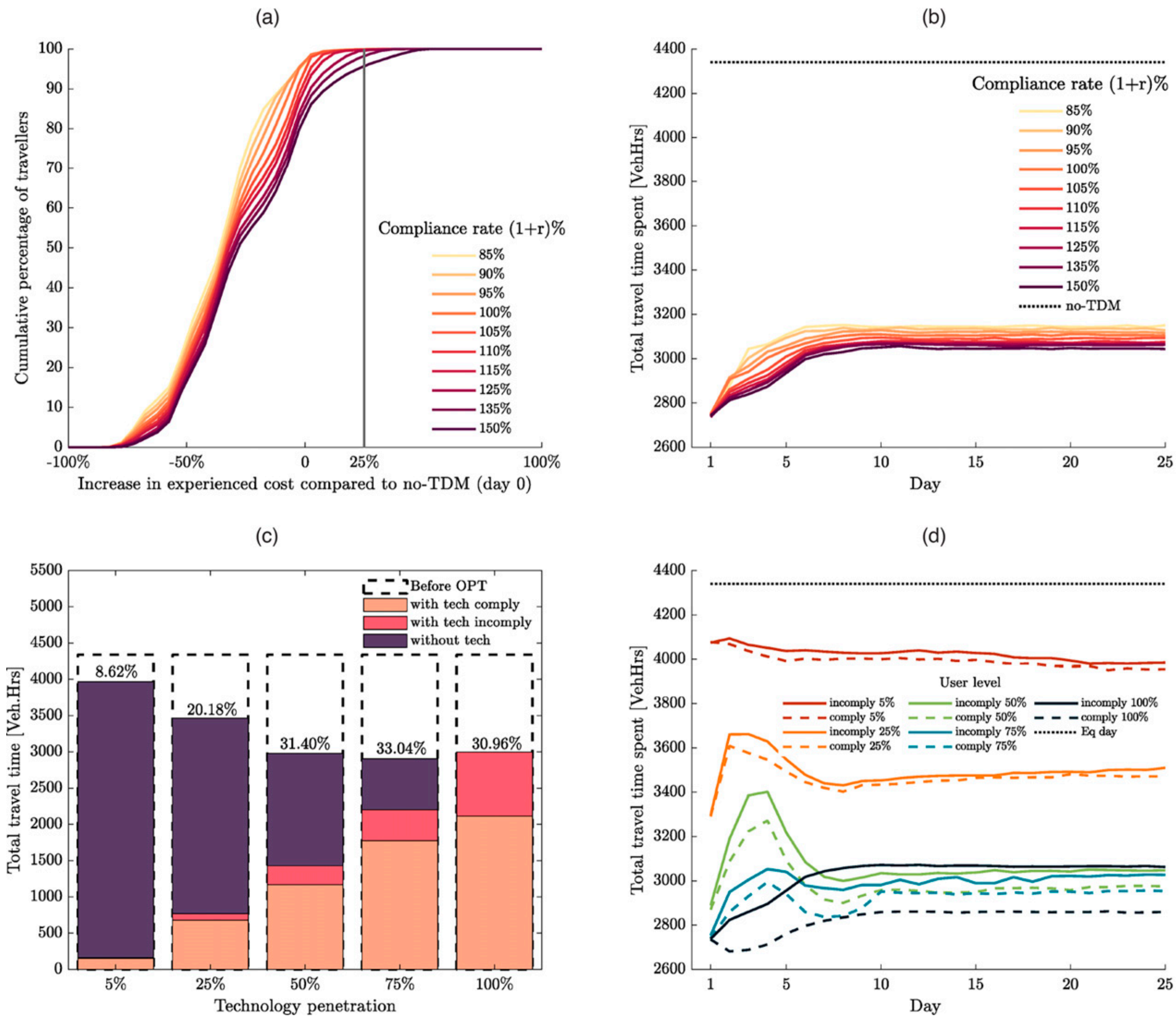


Figure 8. (Color online) (a) Cumulative Change in Experienced Cost with Compliance Rate; (b) Day-to-Day Variation of TTS with Compliance Rate; (c) TTS for Different Penetration Rates with Partial Compliance Conditions; (d) Comparison of Day-to-Day TTS for Different Penetration Rates with Partial and Full Compliance Conditions



travelers to not comply does not bring adverse effects to the overall performance of the network.

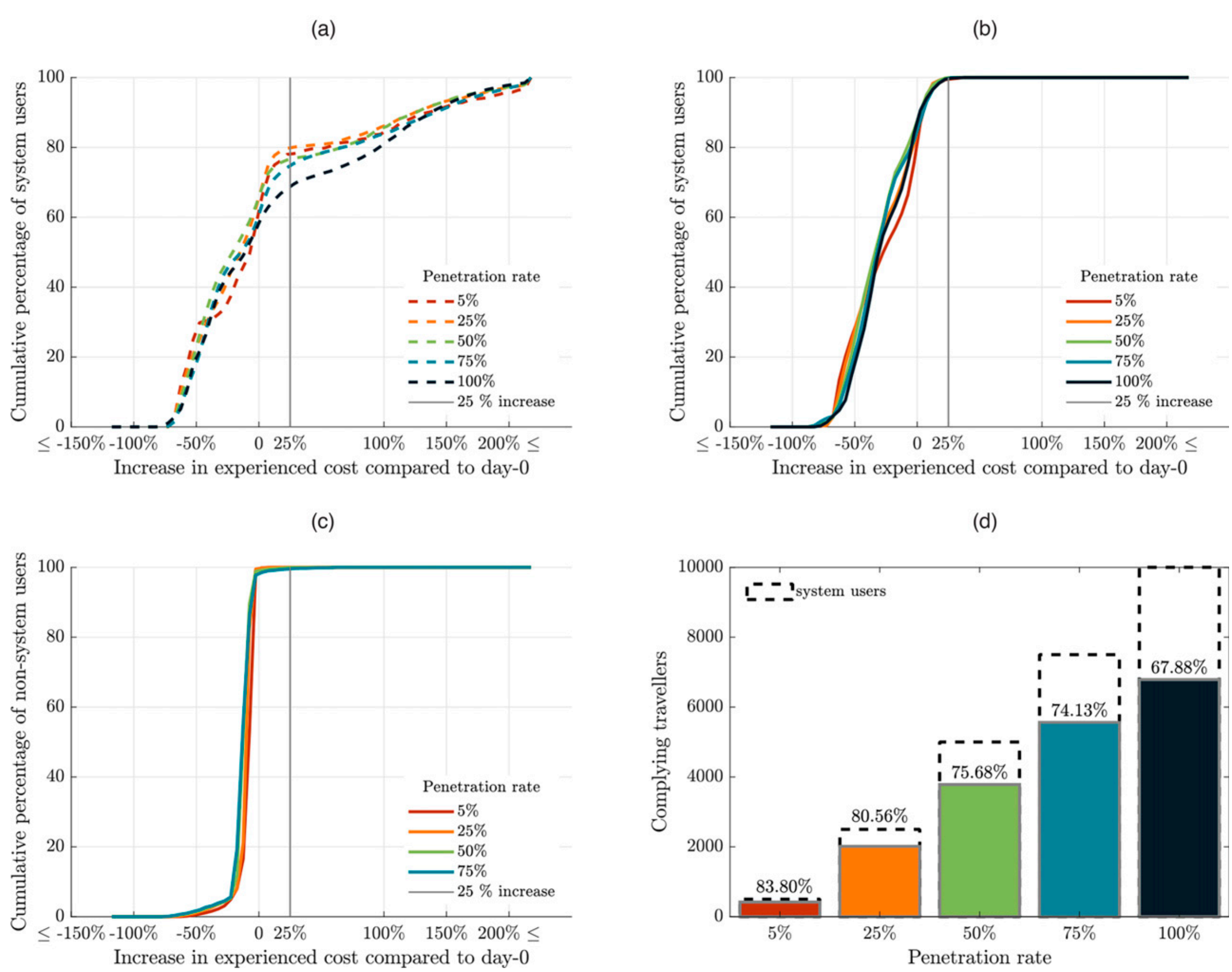
Whereas panels (a) and (b) of Figure 8 illustrate the sensitivity of secondary control for the compliance level, panels (c) and (d) of Figure 8 show the impact of the secondary control mechanism with different penetration levels, assuming $r\% = +25\%$. Here we assume that system users will not comply with the allocated departure time from the ATIS if their perceived travel cost (for the allocated departure time) imposes more than a 25% increase compared with their travel cost on day 0. The system user will adhere to the requested departure time if the allocated departure time is refused.

We see three different types of travelers in the partial compliance scenario: (1) complying system users, (2) noncomplying system users, and (3) nonsystem users. The TTSs for each group of users are shown in Figure 8(a). Although we expect a significant difference in performance levels, it was observed that the performance reduces only by 2%–3.5% for all penetration rates under partial compliance conditions (see Figure 8(a) and Figure 5(c)). The diminishing marginal benefit observed with an increase in penetration rate also exists in the partial compliance scenario. Further, we observe a slight reduction in percentage improvement when user penetration increases from 75% to 100%. The mismatch between the two traffic models and the change in compliance rate might have caused this discrepancy. A model mismatch occurs in our framework, as the accumulation-based model is used in the optimization framework, and the trip-based model is used in the test bed for performance evaluation. A noncomplying system user in the partial compliance scenario is analogous to a nonsystem user in the full compliance scenario. Nonetheless, noncomplying system users cause a higher mismatch between the model (i.e., accumulation-based MFD) and the plant (i.e., trip-based MFD), because the model assumes that all systems users would comply, whereas the compliance decision is taken at the individual level in the plant. As will be shown later (in Figure 9(d)), in this scenario, the compliance rate significantly decreases with an increasing user penetration rate. Due to the increase in the number of noncomplying users (from 75% to 100%), there is a higher mismatch in the proposed framework, which might cause the discrepancy in the network performance. Nevertheless, the drop in performance is only minimal. Compliance rules, in fact, serve as a secondary control mechanism and expose the robustness of the proposed scheme.

Figure 8(d) exhibits the day-to-day evolution of TTS values for different penetration rates in both full compliance (dashed lines) and partial compliance (solid lines) scenarios. We see an increase in the difference between TTS observed under full compliance and partial compliance conditions, which is a result of the increase in the number of noncomplying travelers with the penetration rate. Although the difference is rather negligible for low penetration rate scenarios, the 100% scenario exhibits a significant gap, as the number of noncomplying users is higher. The day-to-day variation of TTS sees consistency after day 10 for all penetration rates, which reveals the convergence of network traffic conditions to equilibrium.

The traffic conditions observed with full and partial compliance conditions behave close enough and demonstrate similar performance (with respect to TTS) levels; however, the user benefits present a significant difference. Figure 9 provides a detailed analysis of individual user experience in the partial and full

Figure 9. (Color online) Increase in Experienced Cost for (a) System Users Under Full Compliance Conditions; (b) System Users Under Partial Compliance Conditions; (c) Nonsystem Users; (d) Compliance Rate of System Users with Different Penetration Rates



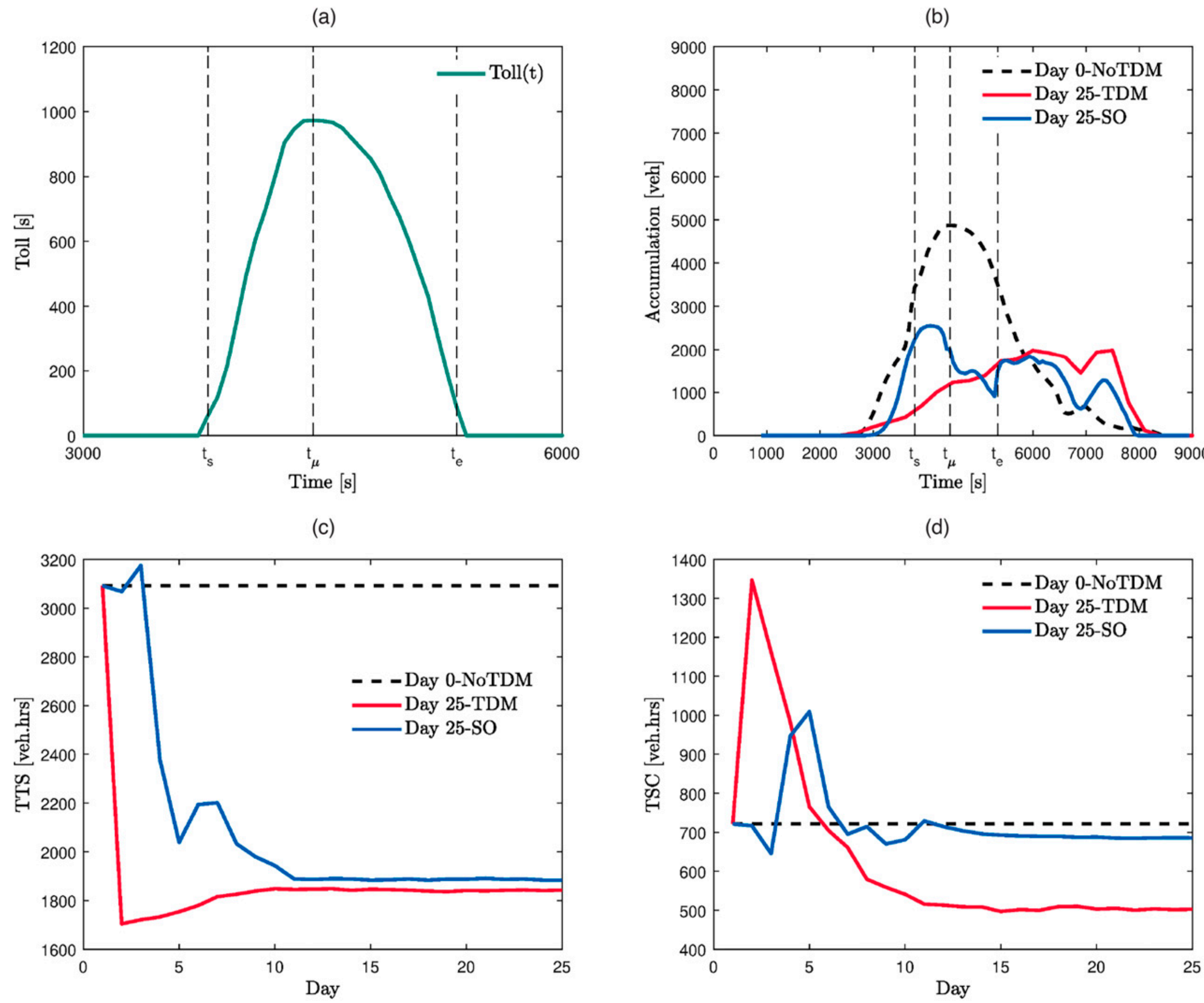
compliance scenarios. Figure 9(a) illustrates the cumulative distribution of the increase in the experienced cost on day 25 compared with the experienced cost in the no-TDM scenario (day 0) for only system users under full compliance conditions. For all penetration rates, we can see that 70%–80% of the system users do not experience more than a 25% increase. However, approximately 20%–30% of the system users undergo a higher increase. Hence, it will be a challenging task to ensure user retention (i.e., maintain daily active users), whereas complying travelers experience undesirable costs, unless they are compensated using other (monetary) incentives to comply. Here, we relax the full compliance assumptions and explore the system performance under a more flexible compliance setting.

Figure 9(b) shows the cumulative distribution of the increase in experienced cost on day 25 compared with the experienced cost in the no-TDM scenario (day 0) for system users under partial compliance conditions. Now, all system users experience an increase of less than 25%. Approximately 80%–90% of the system users experience a reduction in their cost, whereas 10%–20% confront a slight increase. Considering also the results in Figure 5(c), this shows that user benefits significantly improve in the partial compliance scenario, while the system performance is not compromised. Figure 9(c) indicates how nonsystem users are affected due to the operation of the TDM strategy in the partial compliance scenario. Nonsystem users enjoy the reduction in travel times that is brought by the TDM strategy. Almost all nonsystem users experience lower costs than on day 0 in all penetration rate scenarios. Figure 9(d) shows the bar chart of system users (black dashed lines) and the complying travelers at each penetration rate (color-filled). We observe that compliance rate decreases with an increase in penetration rate. A similar behavior is observed in the cumulative percentage of system users who experience an increase in costs of less than 25% under full compliance conditions (see Figure 9(a)). For example, about 68% of travelers experience a cost increase of less than 25% under the full compliance conditions, and we see a similar compliance rate under the partial compliance conditions. The overall benefit gained by both system users and nonsystem users at all penetration rates shows that the proposed TDM strategy can significantly improve traffic conditions and that it is robust with respect to varying penetration rates and compliance issues.

4.4. Performance of TDM Compared with Pricing Strategy

The performance of the TDM strategy is compared with user equilibrium conditions (no-TDM) in Figure 4, and it is important to understand the behavior of TDM compared with optimum system conditions. A vast literature has developed on congestion pricing following Wardrop (1952), who provided a theoretical approach in which a user equilibrium being converted to an optimum system equilibrium by introducing a toll structure (externality). Marchand (1968) further investigated and proposed the second-best tolling method, which many subsequent works followed (see De Palma and Lindsey 2011 for a summary). In this study, we resort to cordon-based pricing strategies, where we aim to introduce a time-dependent toll for travelers based on departure time. The theoretical background for cordon-based pricing for single-region MFD networks is presented elegantly in the works of Geroliminis and Levinson (2009), Arnott (2013), Fosgerau (2015), Liu and Geroliminis (2016), Lamotte and Geroliminis (2018), Amirgholy and Gao (2017), and Zheng and Geroliminis (2020). After all, applying a toll in multiregion networks is a less-explored path due to the complexities observed in multiregion traffic dynamics. Briefly, cordon-based pricing strategies focus on an optimal pricing strategy that charges users a dynamic toll equal to the difference between the generalized cost of the trips in the system optimum and user equilibrium (see Amirgholy and Gao 2017 for more details). The excess delay and schedule deviation cost of the system are essential components to be estimated in formulating the pricing strategy. However, considering the challenges of observation with respect to desired arrival times and schedule deviation, in this comparison, we consider only the excess delay to determine the time-dependent toll and aim to provide a comparative assessment of schedule change (retiming) versus congestion toll. We present the time-dependent toll (excess travel delay) as $Toll(t) = \tau_r(t) - \tau_r^{\mathrm{f}}$, where $\tau_r(t)$ is the travel time of region r at departure time $t(= t_{i,d}^{\mathrm{dep}})$, given by $\tau(t) = L_r/v_r(n_r(t))$, and τ_r^{f} is the travel time at the critical accumulation condition given by $\tau_r^{\mathrm{f}} = L_r/v_r(n_r^{\mathrm{cr}})$, which will also be the observed travel time when the region operates under system optimum conditions. We assume that we can apply the time-dependent toll to all travelers, since they are connected via the ATIS. Figure 10 presents the summary and comparison of applying the time-dependent toll in a two-region traffic network. Figure 10(a) indicates the time-dependent toll for travelers who depart within the $(t_s - t_e)$ time interval, where the accumulation in the inner region goes beyond critical accumulation in the No-TDM (user equilibrium) scenario. Time step t_μ reports the highest toll applied based on the peak accumulation observed. Geroliminis and Levinson (2009) and Amirgholy and Gao (2017) idealize the delay (toll) as a piecewise linear function for a single region system, but our scenario exhibits a parabolic pattern in the delay curve ($Toll(t)$ in Figure 10(a)). Figure 10(b) shows the resulting accumulation curves with three different

Figure 10. (Color online) (a) Time-Dependent Toll; (b) Comparison of Accumulation Curves; (c) Comparison of Total Time Spent; (d) Comparison of Total Schedule Cost



strategies: No-TDM (user equilibrium), TDM, and system optimum (SO). Convergence to uncongested traffic conditions in both SO and TDM reveal the applicability in both strategies, but the traffic patterns indicate that two approaches converge to substantially different traffic scenarios. Figure 10(c) indicates that TDM slightly outperforms SO in terms of the total time spent (TTS). However, Figure 10(d) reveals that the schedule cost of TDM is much lower than that of SO. We observe lower schedule costs in TDM than SO or No-TDM. Considering travelers' flexibility/willingness to comply via the secondary control mechanism (traveler compliance conditions) in TDM has brought this advantage. However, it should be noted that, cordon-based pricing strategies are lacking in mechanisms to incorporate willingness to pay and generate additional inertia, discouraging travelers' compliance. Further, special infrastructure or an ATIS similar to the TDM framework would require the implementation of a pricing strategy. Hence, we see equal competitiveness in the TDM strategy and practicable application when compared with cordon-based pricing strategies.

5. Conclusion

This paper has presented a travel demand management method for efficient temporal coordination of demand through the retiming of departure times within a limited time window in a large-scale urban network. The proposed TDM method requires an ATIS platform with cross-communication capacity to encourage/enforce departure time changes to users. The strategy assigns departure time shifts within a limited time window (e.g., 10 minutes early or late) and formulates an optimization problem that minimizes the TTS in the network with limited schedule change constraints. The proposed framework makes use of a model-plant approach to model a large-scale network with two regions; two MFD-based traffic models are jointly implemented to represent the model and the plant. The accumulation-based MFD model, which operates with aggregate traffic states, is used for demand optimization (model). The developed framework accounts for both observed and unobserved demand components, representing ATIS users and others, respectively. Actual traffic conditions in the network are evaluated by the trip-based MFD model (plant). The plant operates with individual traveler characteristics and is conjugated with a detour ratio model to incorporate the impact of route choice into the aggregated network level. The results demonstrate a substantial improvement in the network performance as a result of minor schedule changes. Further, the proposed TDM strategy generates significant benefits, even at low penetration rates, which adds to its applicability in real settings where a full penetration rate is not possible. Development of incentive schemes to maintain daily active users, validating the detour ratio model with more granular data types such as trajectory data, developing macroscopic demand estimation methods to estimate the unknown demand, and incorporating long-term demand variations (induced demand) in the modeling framework are alluring future research directions.

References

Aalipour A, Kebriaei H, Ramezani M (2019) Analytical optimal solution of perimeter traffic flow control based on MFD dynamics: A Pontryagin's maximum principle approach. *IEEE Trans. Intelligent Transportation Systems* 20(9):3224–3234.

Agatz N, Erera A, Savelsbergh M, Wang X (2012) Optimization for dynamic ride-sharing: A review. *Eur. J. Oper. Res.* 223(2):295–303.

Alfa AS (1986) A review of models for the temporal distribution of peak traffic demand. *Transportation Res. Part B* 20(6):491–499.

Amirgholy M, Gao HO (2017) Modeling the dynamics of congestion in large urban networks using the macroscopic fundamental diagram: User equilibrium, system optimum, and pricing strategies. *Transportation Res. Part B Methodological* 104:215–237.

Andersson J (2013) A general-purpose software framework for dynamic optimization. Unpublished doctoral dissertation, Faculty of Engineering Science, KU Leuven, Heverlee, Belgium, https://lirias.kuleuven.be/handle/123456789/418048.

Arnott R (2013) A bathtub model of downtown traffic congestion. *J. Urban Econom.* 76:110–121.

Batista SFA, Leclercq L (2019) Regional dynamic traffic assignment framework for macroscopic fundamental diagram multi-regions models. *Transportation Sci.* 53(6):1563–1590.

Batista SFA, Leclercq L, Geroliminis N (2019) Estimation of regional trip length distributions for the calibration of the aggregated network traffic models. *Transportation Res. Part B Methodological* 122:192–217.

Bebber DP, Hynes J, Darrah PR, Boddy L, Fricker MD (2007) Biological solutions to transport network design. *Proc. Royal Soc. B Biological Sci.* 274(1623):2307–2315.

Boscoe FP, Henry KA, Zdeb MS (2012) A nationwide comparison of driving distance vs. straight-line distance to hospitals. *Professional Geographer* 64(2):188–196.

Buisson C, Ladier C (2009) Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Res. Record* 2124(1):127–136.

Cantarella GE, Cascetta E (1995) Dynamic processes and equilibrium in transportation networks: Toward a unifying theory. *Transportation Sci.* 29(4):305–329.

Cao J, Menendez M, Waraich R (2019) Impacts of the urban parking system on cruising traffic and policy development: The case of Zurich downtown area, Switzerland. *Transportation* 46:883–908.

Cardillo A, Scellato S, Latora V, Porta S (2006) Structural properties of planar graphs of urban street patterns. *Physical Rev. E* 73:66107.

Cassidy MJ, Jang K, Daganzo CF (2011) Macroscopic fundamental diagrams for freeway networks: Theory and observation. *Transportation Res. Record* 2260(1):8–15.

Chen R, Mahmassani HS (2004) Travel time perception and learning mechanisms in traffic networks. *Transportation Res. Record* 1894(1):209–221.

Chen PST, Srinivasan KK, Mahmassani HS (1999) Effect of information quality on compliance behavior of commuters under real-time traffic information. *Transportation Res. Record* 1676(1):53–60.

Chiu YC (2014) Active traffic and demand management system. U.S. Patent 8744734B2, https://patents.google.com/patent/US8744734B2.

Cole JP, King CAM (1968) Quantitative Geography: Techniques and Theories in Geography (Wiley, New York).

Daganzo CF (1985) The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Sci.* 19(1):29–37.

Daganzo CF (2007) Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Res. Part B Methodological* 41(1):49–62.

Daganzo CF, Lehe LJ (2015) Distance-dependent congestion pricing for downtown zones. *Transportation Res. Part B Methodological* 75:89–99.

De Palma A, Lindsey R (2011) Traffic congestion pricing methodologies and technologies. *Transportation Res. Part C Emerging Tech.* 19(6):1377–1399.

De Palma A, Ben-akiva M, Lefèvre C, Litinas N (1983) Stochastic equilibrium model of peak period traffic congestion. *Transportation Sci.* 17(4):430–453.

Fosgerau M (2015) Congestion in the bathtub. *Econom. Transportation* 4(4):241–255.

Fosgerau M, Small KA (2013) Hypercongestion in downtown metropolis. *J. Urban Econom.* 76:122–134.

Fox JE, Boehm-Davis DA (1998) Effects of age and congestion information accuracy of advanced traveler information systems on user trust and compliance. *Transportation Res. Record* 1621(1):43–49.

Geroliminis N, Daganzo CF (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Res. Part B Methodological* 42(9):759–770.

Geroliminis N, Levinson DM (2009) Cordon pricing consistent with the physics of overcrowding. Lam WHK, Wong SC, Lo HK, eds. *Transportation and Traffic Theory 2009: Golden Jubilee* (Springer, Boston), 219–240.

Geroliminis N, Sun J (2011) Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Res. Part B Methodological* 45(3):605–617.

Geroliminis N, Zheng N, Ampountolas K (2014) A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. *Transportation Res. Part C Emerging Tech.* 42:168–181.

Godfrey JW (1969) The mechanism of a road network. *Traffic Engrg. Control* 11:323–327.

González Ramírez H, Leclercq L, Chiabaut N, Becarie C, Krug J (2019) Unravelling travellers' route choice behaviour at full-scale urban network by focusing on representative OD pairs in computer experiments. *PloS one* 14(11):e0225069.

Gu Z, Shafiei S, Liu Z, Saberi M (2018) Optimal distance- and time-dependent area-based pricing with the network fundamental diagram. *Transportation Res. Part C Emerging Tech.* 95:1–28.

Guo RY, Yang H, Huang HJ, Li X (2018) Day-to-day departure time choice under bounded rationality in the bottleneck model. *Transportation Res. Part B Methodological* 117:832–849.

Hendrickson C, Kocur G (1981) Schedule delay and departure time decisions in a deterministic model. *Transportation Sci.* 15(1):62–77.

Herman R, Prigogine I (1979) A two-fluid approach to town traffic. *Science* 204(4389):148–151.

Horowitz JL (1984) The stability of stochastic equilibrium in a two-link transportation network. *Transportation Res. Part B Methodological* 18(1):13–28.

Huang L, Li J (2011) Applicability of staggered work hours for urban traffic: Case of Guangzhou. *Proc. 3rd Internat. Conf. Transportation Engrg.* (American Society of Civil Engineers, Reston, VA), 404–409.

Ingole D, Mariotte G, Leclercq L (2020) Perimeter gating control and citywide dynamic user equilibrium: a macroscopic modeling framework. *Transportation Res. Part C Emerging Tech.* 111:22–49.

Jahn O, Möhring RH, Schulz AS, Stier-Moses NE (2005) System-optimal routing of traffic flows with user constraints in networks with congestion. *Oper. Res.* 53(4):600–616.

Lamotte R, Geroliminis N (2018) The morning commute in urban areas with heterogeneous trip lengths. *Transportation Res. Part B Methodological* 117:794–810.

Leclercq L, Paipuri M (2019) Macroscopic traffic dynamics under fast-varying demand. *Transportation Sci.* 53(6):1526–1545.

Leclercq L, Parzani C, Knoop VL, Amourette J, Hoogendoorn SP (2015) Macroscopic traffic dynamics with heterogeneous route patterns. *Transportation Res. Part C Emerging Tech.* 59:292–307.

Lindsney R, Verhoef E (2001) Traffic congestion and congestion pricing. Button KJ, Hensher DA, eds. *Handbook of Transport Systems and Traffic Control* (Emerald, Bingley, UK), 77–105.

Little JDC (1961) A proof for the queuing formula: $L = \lambda W$. *Oper. Res.* 9(3):383–387.

Liu W, Geroliminis N (2016) Modeling the morning commute for urban networks with cruising-for-parking: An MFD approach. *Transportation Res. Part B Methodological* 93:470–494.

Loder A, Ambühl L, Menendez M, Axhausen KW (2017) Empirics of multi-modal traffic networks—using the 3D macroscopic fundamental diagram. *Transportation Res. Part C Emerging Tech.* 82:88–101.

Mahmassani H, Williams JC, Herman R (1984) Investigation of network-level traffic flow relationships: Some simulation results. *Transportation Res. Record* (971):121–130.

Marchand M (1968) A note on optimal tolls in an imperfect environment. *Econometrica* 36(3–4):575–581.

Mariotte G, Leclercq L (2019) Flow exchanges in multi-reservoir systems with spillbacks. *Transportation Res. Part B Methodological* 122:327–349.

Mariotte G, Leclercq L, Laval JA (2017) Macroscopic urban dynamics: Analytical and numerical comparisons of existing models. *Transportation Res. Part B Methodological* 101:245–267.

Mohajerpoor R, Saberi M, Vu HL, Garoni TM, Ramezani M (2019) $H\infty$ robust perimeter flow control in urban networks with partial information feedback. *Transportation Res. Part B Methodological* 137:47–73.

Nakayama S, Kitamura R, Fujii S (1999) Drivers' learning and network behavior: Dynamic analysis of the driver-network system as a complex system. *Transportation Res. Record* 1676(1):30–36.

Newell GF (1987) The morning commute for nonidentical travelers. *Transportation Sci.* 21(2):74–88.

NYC OpenData (2018) New York City taxi trip data 2009–2015, https://opendata.cityofnewyork.us/.

O'Flaherty C (1997) *Transport Administration and Planning* (Wiley, New York).

Orski CK (1990) Can management of transportation demand help solve our growing traffic congestion and air pollution problems? *Transportation Quart.* 44:483–498.

Paipuri M, Xu Y, González MC, Leclercq L (2020) Estimating MFDs, trip lengths and path flow distributions in a multi-region setting using mobile phone data. *Transportation Res. Part C Emerging Tech.* 118:102709.

Papageorgiou M (2004) Overview of road traffic control strategies. *IFAC Proc.* 37(19):29–40.

Peeta S, Ziliaskopoulos AK (2001) Foundations of dynamic traffic assignment: The past, the present and the future. *Networks Spatial Econom.* 1(3):233–265.

Ramezani M, Nourinejad M (2018) Dynamic modeling and control of taxi services in large-scale urban networks: A macroscopic approach. *Transportation Res. Part C Emerging Tech.* 94:203–219.

Ran B, Hall RW, Boyce DE (1996) A link-based variational inequality model for dynamic departure time/route choice. *Transportation Res. Part B Methodological* 30(1):31–46.

Saffari E, Yildirimoglu M, Hickman M (2020) A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks. *Transportation Res. Part C Emerging Tech.* 119:102743.

Sheffi Y (1984) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods* (Prentice-Hall, Englewood Cliffs, NJ).

Shoup DC (1997) The high cost of free parking. *J. Planning Education Res.* 17(1):3–20.

Sirmatel II, Geroliminis N (2018) Economic model predictive control of large-scale urban road networks via perimeter control and regional route guidance. *IEEE Trans. Intelligent Transportation Systems* 19(4):1112–1121.

Small KA, Chu X (2003) Hypercongestion. J. Transport Econom. Policy 37(3):319–352.

Smeed RJ (1966) The road capacity of city centers. *Highway Res. Record* 169:22–29.

Srinivasan KK, Mahmassani HS (2000) Modeling Inertia and Compliance Mechanisms in Route Choice Behavior Under Real-Time Information. *Transportation Res. Record* 1725(1):45–53.

Thomson JM (1967) An evaluation of two proposals for traffic restraint in central London. *J. Royal Statist. Soc. Ser. A* 130(3):327–377.

Vickrey WS (1969) Congestion theory and transport investment. *Amer. Econom. Rev.* 59(2):251–260.

Wächter A, Biegler LT (2006) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Programming* 106(1):25–57.

Wardman M (2004) Public transport values of time. *Transport Policy* 11:363–377.

Wardrop JG (1952) Some theoretical aspects of road traffic research. *Proc. Institution Civil Engineers* 1(5):767–768.

Wardrop JG (1968) Journey speed and flow in central urban areas. *Traffic Engrg. Control* 9:528–532.

Yang H (2005) *Mathematical and Economic Theory of Road Pricing* (Emerald, Bingley, UK).

Yang H, Ke J, Ye J (2018) A universal distribution law of network detour ratios. *Transportation Res. Part C Emerging Tech.* 96:22–37.

Yang K, Menendez M, Zheng N (2019) Heterogeneity aware urban traffic control in a connected vehicle environment: A joint framework for congestion pricing and perimeter control. *Transportation Res. Part C Emerging Tech.* 105:439–455.

Yang K, Zheng N, Menendez M (2018) Multi-scale perimeter control approach in a connected-vehicle environment. *Transportation Res. Part C Emerging Tech.* 94:32–49.

Yildirimoglu M, Geroliminis N (2014) Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams. *Transportation Res. Part B Methodological* 70:186–200.

Yildirimoglu M, Kahraman O (2018) Searching for empirical evidence on traffic equilibrium. *PloS One* 13(5):e0196997.

Yildirimoglu M, Ramezani M (2020) Demand management with limited cooperation among travellers: A doubly dynamic approach. *Transportation Res. Part B Methodological* 132:267–284.

Yildirimoglu M, Ramezani M, Geroliminis N (2015) Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. *Transportation Res. Part C Emerging Tech.* 59:404–420.

Yildirimoglu M, Sirmatel II, Geroliminis N (2018) Hierarchical control of heterogeneous large-scale urban road networks via path assignment and regional route guidance. *Transportation Res. Part B Methodological* 118:106–123.

Yin YY, Yang H (2003) Simultaneous determination of the equilibrium market penetration and compliance rate of advanced traveler information systems. *Transportation Res. Part A Policy Practice* 37(2):165–181.

Yu H (2008) The geography of transport systems. *Professional Geographer* 60:435–437.

Zahavi Y (1972) Traffic performance evaluation of road networks by the alpha-relationship: Part 1. *Traffic Engrg. Control* 14:228–231.

Zhang X, Adamatzky A, Chan FTS, Deng Y, Yang H, Yang XS, Tsompanas MAI, Sirakoulis GC, Mahadevan S (2015) A biologically inspired network design model. *Sci. Rep.* 5:10794.

Zheng N, Geroliminis N (2016) Modeling and optimization of multimodal urban networks with limited parking and dynamic pricing. *Transportation Res. Part B: Methodological* 83:36–58.

Zheng N, Geroliminis N (2020) Area-based equitable pricing strategies for multimodal urban networks with heterogeneous users. *Transportation Res. Part A Policy Practice* 136:357–374.