# Dynamic Modeling of Large-Scale Urban Transportation Systems

Guilhem Mariotte

N° d'ordre NNT: 2018LYSET010



THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de
l'École Nationale des Travaux Publics de l'État

École Doctorale N° 162
MEGA (Mécanique, Energétique, Génie Civil et Acoustique)

Spécialité / discipline de doctorat : Génie Civil

# Modélisation dynamique des grands réseaux de transports

Soutenue publiquement le 14 novembre 2018 par :

**Guilhem Mariotte**

Devant le jury composé de :

| | | |
|---|---|---|
| Nikolas GEROLIMINIS | Prof. (EPFL) | Rapporteur |
| Monica MENENDEZ | Prof. (New York University Abu Dhabi) | Rapporteure |
| Christine SOLNON | Prof. (INSA Lyon) | Présidente du jury |
| Arnaud DE LA FORTELLE | Prof. (Mines Paris Tech) | Examinateur |
| Serge HOOGENDOORN | Prof. (TU Delft) | Examinateur |
| Ludovic LECLERCQ | Prof. HDR (Univ. Lyon, ENTPE, IFSTTAR) | Directeur de thèse |

Thèse préparée au LICIT (Laboratoire Ingénierie Circulation Transports)

Thesis National Number: 2018LYSET010




A THESIS OF THE UNIVERSITY OF LYON

Prepared at
École Nationale des Travaux Publics de l'État

Doctoral school No. 162
MEGA (Mechanics, Energy, Civil Engineering and Acoustics)

To obtain the graduation of
PhD in Civil Engineering

---

# Dynamic Modeling of Large-Scale Urban Transportation Systems

---

Defended on November 14, 2018 by:

**Guilhem Mariotte**

In front of the following examination committee:

| | | |
|---|---|---|
| Nikolas GEROLIMINIS | Prof. (EPFL) | Reviewer |
| Monica MENENDEZ | Prof. (New York University Abu Dhabi) | Reviewer |
| Christine SOLNON | Prof. (INSA Lyon) | Committee chair |
| Arnaud DE LA FORTELLE | Prof. (Mines Paris Tech) | Examiner |
| Serge HOOGENDOORN | Prof. (TU Delft) | Examiner |
| Ludovic LECLERCQ | Prof. HDR (Univ. Lyon, ENTPE, IFSTTAR) | Supervisor |

LICIT lab (Transport and Traffic Engineering Lab)

*"Scientific knowledge advances haltingly and is stimulated by contention and doubt"*

— Claude Lévi-Strauss

*French anthropologist and ethnologist*

# Abstract

Congestion in urban areas has become a major issue in terms of economic, social or environmental impact. For short or mid term, using dynamic road traffic simulation can help analyzing and providing guidelines to optimization policies of existing infrastructures. Today, because of the complexity of transport systems, classical modeling tools are limited to small geographical areas (of a district size). Computational time, together with simulation calibration, are notably very constraining at large scales. However, a new generation of models designed for metropolitan areas has arisen over the past decades. These models are based on a phenomenological relationship between travel production and the number of vehicles in a given spatial area of a road network, known as the Macroscopic Fundamental Diagram (MFD). This relationship, supported by empirical evidences from several cities around the world, has allowed the study of different traffic control schemes at a whole city scale, but was rarely used for traffic state forecasting.

The aim of this PhD is to propose an efficient modeling tool, based upon the concept of MFD, to simulate and analyze traffic states in large metropolitan areas. The theoretical framework of this tool must be consistent and applicable for traffic state forecasting, development of new control policies, traffic emission estimation, etc. There are two major contributions in this PhD. The first one is analyzing the mathematical and physical properties of existing models, and formalizing the dynamics of several trip lengths inside the same urban zone. In particular, this formalization distinguishes between internal trips and trips crossing the zone. Flow merging and diverging issues are also addressed when congestion propagates from one zone to another. The second contribution is proposing a new trip-based model based on individual traveled distance. This approach allows to treat users independently (previously represented with continuous flows), and thus to define their characteristics more precisely to couple their trips with assignment models on different paths. Finally, examples of application from various collaborations are given in the last part of this thesis. It includes a simulation study of the Grand Lyon urban area (France), as well as new modules to simulate search-for-parking or perimeter control. This PhD is part of a European ERC project entitled MAGnUM: Multiscale and Multimodal Traffic Modeling Approach for Sustainable Management of Urban Mobility.

# RÉSUMÉ

La congestion en milieu urbain est un enjeu majeur que ce soit d'un point vue économique, social ou environnemental. À court et moyen terme, l'utilisation de la simulation dynamique du trafic routier peut permettre d'analyser et de guider des politiques d'optimisation des infrastructures existantes. Aujourd'hui, du fait de la complexité des systèmes de transport, les outils de modélisation classiques sont limités à des échelles géographiques peu étendues (de l'ordre du quartier). À grande échelle, le temps de calcul devient rapidement un facteur limitant tout comme le calibrage et la scénarisation. Néanmoins les dernières décennies ont vu l'apparition d'une nouvelle génération de modèles bien adaptés aux métropoles urbaines. Ceux-ci sont basés sur une relation phénoménologique entre la production de déplacements et le nombre de véhicules dans une zone spatiale d'un réseau routier, appelée Diagramme Fondamental de Zone (*Macroscopic Fundamental Diagram*, MFD). Cette relation, validée empiriquement sur de nombreuses villes, a permis d'étudier différentes méthodes de contrôle du trafic pour une ville entière, mais a été peu utilisée à des fins de prévision de la congestion.

L'objectif de cette thèse est de proposer un premier outil opérationnel de simulation et d'analyse des grands réseaux de métropoles, en utilisant et développant les modèles de trafic basés sur la relation MFD. Cet outil doit posséder un cadre théorique cohérent qui puisse convenir à des applications telles que la prévision d'états de trafic, le développement de nouvelles politiques de contrôle, l'estimation de pollutions liées au trafic, etc. Les contributions de la thèse portent sur deux aspects. Le premier est l'analyse des propriétés mathématiques et physiques des modèles existants, en incluant une formalisation complète de la gestion de plusieurs longueurs de parcours au sein d'une même zone urbaine. En particulier, cette formalisation traite de la distinction des trajets internes à la zone et des problèmes de flux convergents et divergents pour les trajets traversant la zone lorsque la congestion se propage d'une zone à l'autre. Le deuxième aspect est la proposition d'un nouveau modèle basé sur la distance individuelle parcourue à l'intérieur d'une zone urbaine (trip-based). Cette approche permet d'individualiser les usagers (auparavant représentés sous forme de flux continus) et donc de définir plus finement leurs caractéristiques, en vue de coupler leurs déplacements à des modèles d'affectations sur différentes routes. Enfin, des exemples d'application illustrant diverses collaborations sont donnés en dernière partie de la thèse. La simulation du trafic sur l'aire urbaine du Grand Lyon (France) y est présentée, ainsi que de nouveaux modules de modélisation de la recherche de parking ou de contrôle périphérique. Cette thèse est partie intégrante d'un projet européen ERC intitulé MAGnUM : Approche multi-échelle et multimodale de la modélisation du trafic pour une gestion durable de la mobilité urbaine.

# ACKNOWLEDGEMENTS

First of all, I would like to thank Prof. Ludovic Leclercq, deputy director of the LICIT lab (Univ. Lyon, ENTPE, IFSTTAR), for supervising this PhD thesis and supporting me during these three years. While being a highly recognized expert in his domain and having great human management qualities, Ludovic was always there at any time for any support I could require: help to overcome scientific locks, arrange administrative issues, organize meetings and workshops to make the lab members share ideas and work together, and so on with many other activities. These three years of research have been a very enriching experience, both for inspiring scientific discussions and entertaining collaborations at social events. My memories about these years not only include MFD simulation but also playing the ukulele and practicing for the next competition of the Voice. Merci infiniment Ludo, I was very fortunate to work with you!

I would like to also express my deep and warm acknowledgements to Prof. Monica Menendez from NYU, Abu Dhabi, and Prof. Nikolas Geroliminis from EPFL, Lausanne, who both thoroughly reviewed my thesis and raised detailed comments that improved this revised version of the manuscript. I am also very thankful to all the other members of the examination committee: Prof. Serge Hoogendoorn from TU Delft, Prof. Arnaud De la Fortelle from Mines Paris Tech and Prof. Christine Solnon from INSA Lyon who chaired the committee. Together with Prof. Menendez and Prof. Geroliminis, they all addressed critical questions during the PhD defense that expanded the discussion well beyond the scope of my work, and thus made me think in a larger perspective. The quality of their scientific comments made me enjoy a very enriching experience. I find myself honored of having had such high level and internationally recognized experts in my committee.

I am very glad to Prof. Nour-Eddin El Faouzi, head of the LICIT lab, for accepting me as a member of the great LICIT family, and providing me a nice work environment during my stay in the lab. I have special thanks to Sonia Cenille and Anne-Christine Demanny, the two secretaries of the lab for their respective help in many and diverse administrative tasks.

Moreover, these years spent at the LICIT lab wouldn't have been so exciting without the help and support of all my colleagues. To my fellow former and current PhD students, Étienne, Péa, Maxime, Raphaël, Nicole, Clélia, Humberto, Aurélien, Johanna, Mostafa and Negin, to the post-doc Mahi, Ruiwei, Deepak and Omar, to the research engineers Arthur, Jean, Anastasia and Charlotte, to the master students Alexis, Alméria, Victor and Géraldine, many thanks for having spent very good times together! The experiences that we had during conferences, social events and more were resourcing and relaxing as well, no doubt they

played an important part for the years I spent in the lab ☺. I also wish good luck to the latest PhD adventurers who just joined the team: Anna, Manon, Cyril, Carlos, Céline, Élise and Loïc. I am very grateful to all the other lab members as well: Cécile, Nico, Christine, Delphine, Nathalie at ENTPE, and Aurélien, Colette, Angelo, Fouad, Andres, Bernard, Jean-Luc at Ifsttar Bron. It was a real pleasure to work (and more!) with you! Finally, I have special thanks to my office mate Sérgio, with who I shared my dreams and hope about MFD simulation along the numerous adventures we encountered at the famous D009 office. Muito obrigado Sérgio, o fim do mundo já passou!

Among all these experiences, this PhD was also a great opportunity to meet and collaborate with invited professors and students from all around the world. To Prof. Jorge Laval from Georgia Tech, Atlanta, muchas gracias for our scientific and musical exchanges; to Igor Dakic, Lukas Ambühl and Michael Kung from ETH, Zurich, vielen Danke und liebe Grüsse aus Lyon; to Chuanlin Zhao, Beihang Univ., Beijing, 很高兴认识你，祝你工作和生活愉快.

These times at ENTPE were full of good opportunities for meeting other people from the school, among other activities I particularly enjoyed the TAM and Café des Sciences organized at ENTPE. I have thus special cheers to Nadine, Maryse, Lolo and the team of ALDDENTPE.

Last but not least, my thoughts go for my family and friends who gave me the strength to overcome the difficulties I met and go on. Greetings to the Gouzou fellows, and many thanks to my parents, grand-parents, my brothers Gégé and Gégé, my lovely sons Pierrot and Marius, and of course my wife Xiaxia,谢谢你我的亲爱的，辛苦你帮了我那些三年，我很爱你!

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



# Table of Contents

# List of Figures

# List of Tables

# List of Notations

**Term definitions**

Macroscopic route: (also called regional path, and often simply referred to as route when there is no confusion with other naming) sequence of reservoirs, representing a given flow path from an origin reservoir to a destination reservoir in multi-reservoir systems, page 47

Oversaturation: congested regime of a reservoir, it corresponds to the traffic states where $n > n_c$ on the production-MFD, page 31

Reservoir: (also called neighborhood, region, zone, bin) urban area that should be quite compact and where a well-defined MFD exists, i.e. where traffic states are supposed homogeneous, page 27

Saturation: regime of a reservoir at capacity, it corresponds to the traffic states where $n = n_c$ on the production-MFD (or $n_{c1} < n < n_{c2}$ if there is a flat region in the MFD), page 31

Trip completion rate: rate at which users reach their destinations, page 29

Undersaturation: free-flow regime of a reservoir, it corresponds to the traffic states where $n < n_c$ on the production-MFD, page 31

**Acronyms**

BRUE Bounded Rational User Equilibrium, page 184

CTM Cell Transmission Model, page 37

DTA Dynamic Traffic Assignment, page 25

FD Fundamental Diagram, page 24

LWR Lighthill-Whitham-Richards model, page 24

MAGnUM Multiscale and Multimodal Traffic Modeling Approach for Sustainable Management of Urban Mobility, ERC project for which this thesis is a part of, page 5

MFD Macroscopic Fundamental Diagram (also called NFD, Network Fundamental Diagram), page 5

MPC Model Predictive Control, page 25

NEF Network Exit Function, page 30

OD Origin-Destination, page 32

RTUE Regret Theory User Equilibrium, page 184

SUE Stochastic User Equilibrium, page 47

UE User Equilibrium, page 184

**Main variables**

$G(n)$ outflow-MFD or reservoir trip completion rate, equal to $P(n)/L$ [veh/s], see equation (1.8), page 30

$G_i(n_i, n)$ reservoir trip completion rate of route $i$ [veh/s], see equation (4.14), page 131

$I(n)$ reservoir entry supply function, also called receiving or boundary inflow capacity [veh/s], page 39

$I_i(n_i, n)$ reservoir entry supply of route $i$ [veh/s], see equation (4.23), page 134

$K$ reservoir average or mean density [veh/m], see equation (1.4), page 28

$L$ average trip length in a reservoir [m], see equation (1.9), page 30

$L_{\text{net}}$ total network length in a reservoir [m], page 28

$L_i$ reservoir trip length of route $i$ [m], see equation (4.19), page 132

$\lambda(t)$ reservoir inflow demand, or inflow in case of no entry limitation [veh/s], see equation (1.11), page 38

$\lambda_i(t)$ reservoir inflow demand for route $i$ [veh/s], see equation (4.20), page 132

$\mu(t)$ reservoir outflow supply, or outflow in case of no exit limitation [veh/s], see equation (2.11), page 72

$\mu_i(t)$ reservoir outflow supply for route $i$ [veh/s], see equation (4.29), page 136

$n$ reservoir accumulation, number of circulating vehicles [veh], see equation (1.6), page 29

$N_{\text{in}}(t)$ reservoir entering cumulative count curve, equal to $\int_0^t q_{\text{in}}(t)dt$ [veh], page 64

$N_{\text{out}}(t)$ reservoir exiting cumulative count curve, equal to $\int_0^t q_{\text{out}}(t)dt$ [veh], page 64

$n_c$ reservoir critical accumulation, at which the maximum production $P_c$ is reached [veh], page 31

$n_i$ reservoir partial accumulation in route $i$ [veh], see equation (4.13), page 131

$n_j$ reservoir jam accumulation, maximum number of circulating vehicles [veh], page 31

$O(n)$ reservoir exit demand function, considered equal to the trip completion rate $G(n)$ in the literature, but a new definition is introduced in this manuscript [veh/s], see equation (2.1), page 62

$O_i(n_i, n)$ reservoir outflow demand of route $i$ [veh/s], see equation (4.15), page 132

$P$ reservoir production, total traveled distance by unit of time [veh.m/s], see equation (1.6), page 29

$P(n)$ production-MFD [veh.m/s], see equation (1.8), page 30

$P_c$ reservoir critical production, maximum throughput of the reservoir [veh.m/s], page 31

$P_d(n)$ reservoir exit production demand, corresponds to the exit demand function in production units [veh.m/s], see equation (4.7), page 126

$P_s(n)$ reservoir entry production supply, corresponds to the entry supply function in production units [veh.m/s], see equation (4.7), page 126

$Q$ reservoir average or mean flow [veh/s], see equation (1.4), page 28

$Q(n)$ flow-MFD, equal to $P(n)/L_{\text{net}}$ [veh/s], see equation (1.8), page 30

$q_{\text{in}}(t)$ reservoir effective inflow [veh/s], see equation (1.12), page 39

$q_{\text{out}}(t)$ reservoir effective outflow [veh/s], see equation (1.11), page 38

$q_{\text{in},i}(t)$ reservoir partial effective inflow into route $i$ [veh/s], see equation (4.13), page 131

$q_{\text{out},i}(t)$ reservoir partial effective outflow out of route $i$ [veh/s], see equation (4.13), page 131

$t$ time [s], see equation (1.10), page 38

$T(t)$ reservoir experienced travel time, time spent by vehicles exiting the reservoir at time $t$ [s], see equation (2.6), page 64

$T^*(t)$ reservoir exact predictive travel time, time that will be spent by vehicles entering the reservoir at time $t$ [s], see equation (4.11), page 128

$u$ reservoir free-flow speed, maximum mean speed experienced by circulating vehicles [m/s], page 31

$V$ reservoir mean speed, experienced by all circulating vehicles [m/s], see equation (1.7), page 29

$V(n)$ speed-MFD, equal to $P(n)/n$ [m/s], see equation (1.8), page 30

# 1. General introduction

## 1.1 Background

### 1.1.1 Hypercongestion in metropolitan areas

Over the past several decades, traffic conditions have dramatically gone worse in large metropolitan areas, so that some researchers sometimes use the notion of "hypercongestion" (Fosgerau & Small, 2013) to describe the situation. The impacts are negative on economics regarding the time wasted in jams, on social equity regarding the access to urban mobility, and on environment regarding the pollution generated in congested areas. According to a study from the CEBR research institute, road users from the UK, Germany, France and the USA spend, on average, 36 hours in gridlock every year in metropolitan areas. It is estimated that in 2013, direct costs (fuel consumption and time wasted) and indirect costs (loss of productivity in business) due to congestion accounted for 22.5 billions dollars in France, and for 124 billions dollars in the USA. These figures could increase of 31% (France) and 50% (USA) by 2030. During the same year, the fuel wasted in congestion generated extra amount of $CO_2$ equivalent emissions up to 1.9 megatons in France, and up to 8.6 megatons in the USA. CEBR also expects these amounts to increase of 13% (France) and 21% (USA) by 2030 if no major change occurs in the transportation systems studied (CEBR, 2014).

The widespread use of individual cars together with the increase of the number of inhabitants in large cities is the main cause for hypercongestion. Reducing congestion for the long term implies clear political choices in infrastructure development, urban space allocation, massive mode share in favor of public transport, cycling and walking, and deployment of new mobility services. In short-mid term, optimizing the use of existing transportation modes and infrastructures is probably an efficient management policy to improve traffic conditions in urban environment.

### 1.1.2 Traffic flow simulation

In this context, traffic simulation has become a key element for traffic managers to understand, analyze and forecast mobility in a given network. For now, the physics of traffic at link scale is quite well studied and understood in the research community. Depending on their description of vehicle interactions, the existing modeling approaches are classified as microscopic or macroscopic traffic flow models. Microscopic models represent traffic flow

as a succession of discrete vehicles. The resulting flow is then the consequence of vehicle-to-vehicle interactions, mostly described by car-following models (e.g. Gazis *et al.*, 1961, Gipps, 1981, Treiber *et al.*, 2000) and lane-changing models (e.g. Gipps, 1986a, Laval & Leclercq, 2008). On the other hand, macroscopic models represent traffic as a continuous flow. They are based on a simple conservation equation coming from hydromechanics, which describes the evolution of traffic flow and density. One of the best-known frameworks is the Lighthill-Whitham-Richards (LWR) model (Lighthill & Whitham, 1955, Richards, 1956). In these models, vehicle-to-vehicle interactions are implicitly represented by a phenomenological relationship, called Fundamental Diagram (FD), that relates link flow as a function of link density. This brief overview on major traffic flow models is not intended to be exhaustive, the reader can refer to van Wageningen-Kessels *et al.* (2015) for a more detailed bibliography review on this topic. Despite still being investigated and extended by researchers, link-scale models are already implemented in commercial softwares and are thus widely used for traffic operations.

A brief overview of traffic flow models can be seen in Figure 1.1, with corresponding examples of commercial and open source softwares. The models are classified regarding their traffic representation (microscopic to macroscopic), and their operating scale (vehicle level, link level or network level).

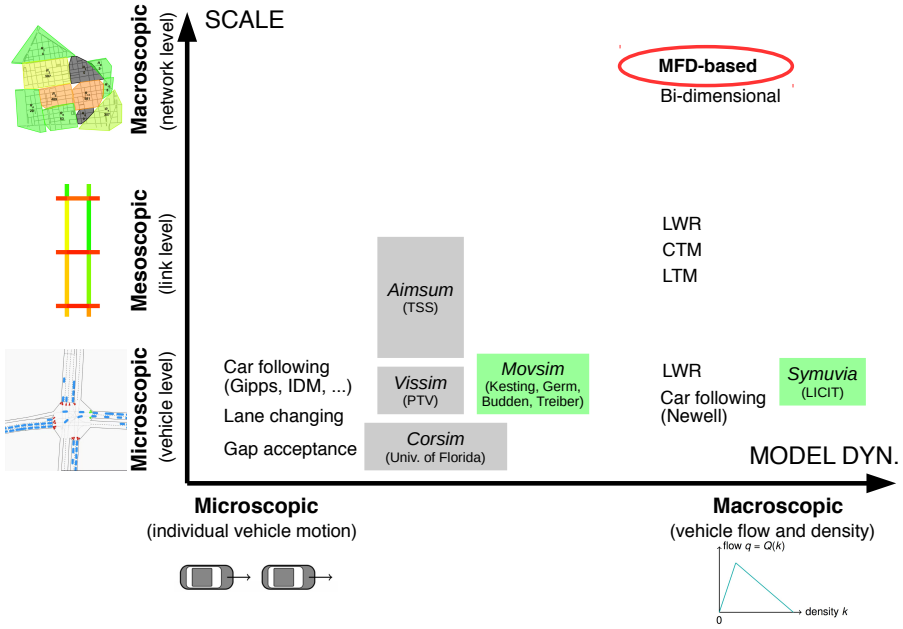



FIG. 1.1 – *Classification of traffic flow models regarding their traffic representation and operating scale, with examples of commercial (gray) and open source (green) softwares*

These models have allowed the development of efficient local control schemes to alleviate congestion. Well-known examples of these systems are the Split Cycle Offset Optimization Technique (SCOOT, Robertson & Bretherton, 1991) and the Sydney Coordinated Adaptive Traffic System (SCATS, Lowrie, 1982) that are used in many cities worldwide. However, these control algorithms only apply locally (at the scale of an arterial and a few neighboring intersections) and do not account for vehicle queues that would be induced outside the controlled area. More generally, it has become clear that large-scale transportation networks are extremely complex systems: they imply interactions between path flows

and mode share resulting from users behavior. This can only be studied from a global perspective at the whole city scale. Nevertheless, running link flow simulations on very large networks often requires high computational time, which prevents us from using real-time simulations for traffic states forecasting or Model Predictive Control (MPC) schemes. Moreover, a key component in the implementation of link-scale models is the demand scenarization, which requires extensive knowledge about travelers' motivations to choose across large amounts of paths. Accessing or modeling users' individual choices is a very tedious and laborious task to undertake at very large scales. While simulating traffic flows in large cities can still be achieved thanks to the thorough calibration of Dynamic Traffic Assignment (DTA) models—see examples for Chicago, USA (Jayakrishnan *et al.*, 1994, Mahmassani *et al.*, 2013b), Melbourne, Australia (Shafiei *et al.*, 2018) or Beijing, China (Ben-Akiva *et al.*, 2002, 2012)—there is a clear need for a new generation of models especially designed for large-scale simulation.

### 1.1.3 New modeling approaches at large scale

Modeling traffic states at an aggregate level is not a new question, various approaches have been proposed since the 1960s. Smeed (1967) investigated the flow capacity of whole urban areas such as city centers. Then Godfrey (1969) conjectured that there must be an optimum in the density of vehicles in a given urban zone, similarly to the existence of a maximum flow in the Fundamental Diagram (FD) of a single link. Later, Herman & Prigogine (1979) went further in assuming a relationship between mean speed and mean density in towns. All these conjectures have been verified with microscopic simulation studies first (Mahmassani *et al.*, 1984, 1987), and quite recently with real data from Yokohama, Japan (Geroliminis & Daganzo, 2008). Since then, this relationship between network average density and flow has been called Network or Macroscopic Fundamental Diagram (NFD or MFD), by analogy with the link Fundamental Diagram concept. A typical shape of MFD is presented in Figure 1.2. After Yokohama, empirical MFDs were estimated for various cities: Toulouse, France (Buisson & Ladier, 2009), Nairobi, Kenya (Gonzales *et al.*, 2009), Changsha, China (Shoufeng *et al.*, 2013), Shenzhen, China (Ji *et al.*, 2014), Brisbane, Australia (Tsubota *et al.*, 2014) and Zurich, Switzerland (Ambuhl *et al.*, 2017, Loder *et al.*, 2017). Other simulation studies also estimated the MFDs of Amsterdam, the Netherlands (Ji *et al.*, 2010), Chicago and Salt Lake City, USA (Saberi *et al.*, 2014). The empirical evidence of the MFD, again recently confirmed by data from numerous cities around the world (Loder *et al.*, 2018), has opened up a whole new field of research in traffic flow modeling, referred to as MFD-based models.

These models describe traffic states in a given urban region with a simple flow conservation equation, first proposed by Daganzo (2007). They were extended to two regions (Geroliminis & Daganzo, 2007, Haddad & Geroliminis, 2012) and then multiple regions (Geroliminis, 2009, Aboudolas & Geroliminis, 2013, Geroliminis, 2015) that exchange flows with each other. Figure 1.3 shows a representation of the principle of MFD-based modeling with multiple regions. The simple formulation of traffic dynamics within a region, together with the limited number of parameters to calibrate, have allowed the development of efficient control schemes. The main idea is to protect a given area by limiting inflow at its periphery and thus keeping a maximum throughput in it (in practice, the inflow limitation can be done via tuning traffic lights or road pricing). This method, known as perimeter control or gat-

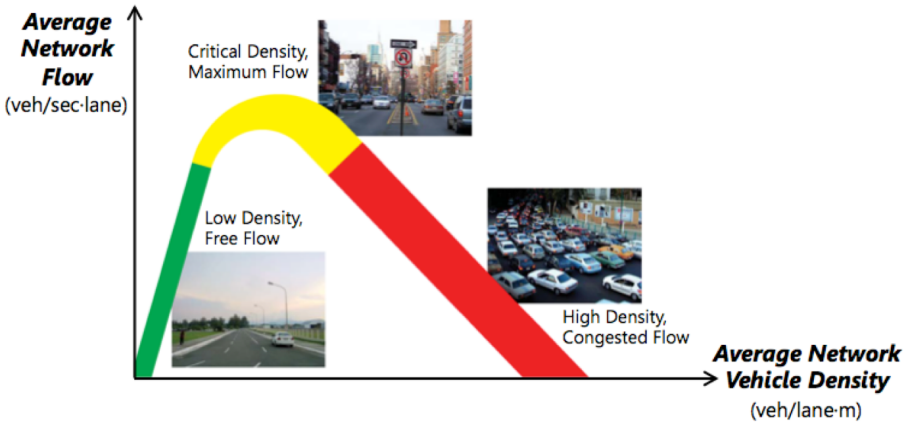



FIG. 1.2 – *Typical shape of a MFD. Credits: E. Gonzales, University of Massachusetts*

ing, has been successfully implemented in MFD-based models (Keyvan-Ekbatani *et al.*, 2012, Geroliminis & Boyaci, 2012), and has shown promising results to alleviate congestion in city centers. This was achieved in simulation studies only, but the city of Zürich in Switzerland already designed an efficient gating system, even before the existence of the MFD was proved (Ambuhl *et al.*, 2018). More recent works showed that route guidance at large scale by using MFD is also a viable option to improve traffic conditions in cities (Yildirimoglu *et al.*, 2015, Sirmatel & Geroliminis, 2017a). The increasing number of studies since the early work of Daganzo (2007) clearly shows that MFD-based models have grown more and more popular in the traffic flow community.

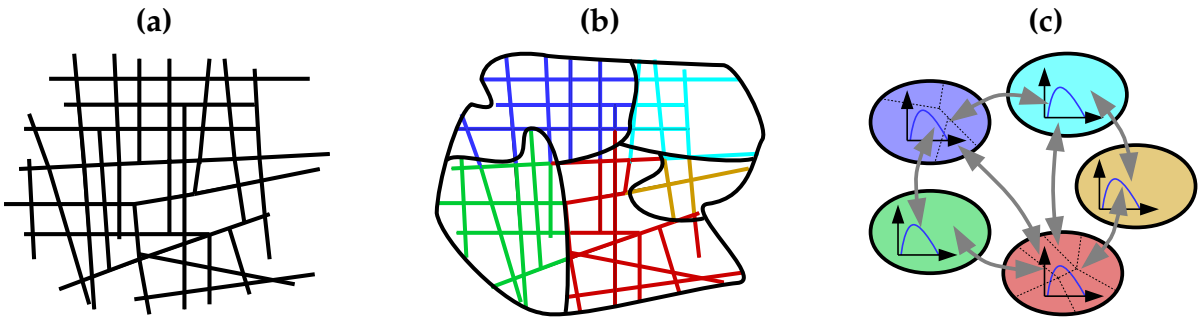



FIG. 1.3 – *Principle of MFD-based multi-regional simulation. (a) A given network, (b) partitioning in different urban regions, and (c) flow exchange representation between the regions*

To a lesser extend, other approaches have been developed to simulate traffic states at large scales. While MFD-based models rely on a partitioning of the network in several homogeneous regions that exchange flow, these approaches consider a two-dimensional continuous flow propagation over the city urban surface. They are consequently often called bi-dimensional traffic flow models. They more or less consist in an extension of the LWR model in two space dimensions, but differ from each other in the formulation of this extension. In the approach of Saumtally *et al.* (2011), further developed by Sossoe *et al.* (2015), Sossoe & Lebacque (2018), an urban area is discretized into two-dimensional cells which exchange flow with neighboring cells. In each cell, a system of unidimensional conservation equations corresponding to major flow directions are then solved to estimate the density of the cell. On the other hand, Du *et al.* (2017) solves the two-dimensional formulation of

the continuum conservation equation by using local travel cost functions including road characteristics and assuming isotropic flows. Mollier *et al.* (2018) recently proposed a two-dimensional directional field to characterize flow propagation and solve the conservation equation. But their approach is still limited to static propagation patterns, which are also hard to link with the classical definition of demand input flows. A detailed review about bi-dimensional traffic models can be found in Aghamohammadi & Laval (2018).

In this PhD thesis, it has been chosen to focus on MFD-based models to investigate their properties deeply. The main reasons for this choice are their empirical foundations, and their advanced level of development and implementation with appealing control strategies at a large scale. Nevertheless, there is still a lot of research to be done in the field of MFD-based simulation. MFD models have been mainly used for control applications, which does not necessarily require an accurate modeling on a long-time horizon as the control feedback keeps the system in the optimal region. One research direction that received little attention for now is the application of MFD models for urban traffic simulation over a long-time horizon (e.g. a day). This PhD thesis is thus devoted to contribute at the understanding, investigation of current approaches and proposition of new ones, together with their validation.

## 1.2 Review on traffic simulation based on the Macroscopic Fundamental Diagram (MFD)

This section presents the state-of-the-art of MFD-based modeling when this PhD thesis started in September 2015. Such a review will bring a list of questions that will define the research objectives for this PhD.

### 1.2.1 Aggregated variable definition and MFD concept

The concept of the MFD (sometimes called NFD, Network Fundamental Diagram) is generally attributed to Godfrey (1969) who first conjectured that there must be an optimum density of vehicles to reach a maximum throughput in a given urban area. This concept was investigated later by Herman & Prigogine (1979), Mahmassani *et al.* (1984, 1987), and the concept of MFD was then formalized by Daganzo (2007) until its existence was proved on a real city network (Geroliminis & Daganzo, 2008). The urban area on which the MFD relationship is defined was called "reservoir" or "neighborhood" by Daganzo (2007). Other names like "region", "zone" or "bin" are still used in the literature. In this manuscript, we mostly use the word "reservoir" which refers to the modeling of traffic dynamics at the network level.

#### Average flow and density

The MFD basically consists in a relationship between aggregated variables in a given urban network (a reservoir), such as average flow and density, mean speed, trip completion rate, perimeter outflow. These variables are defined as follows. Let us assume that the network considered has $N_{\text{link}}$ links where each link $i$ has a length $l_i$ (in [m]), and is defined by its flow $q_i(x_i, t)$ (in [veh/s]) and a density $k_i(x_i, t)$ (in [veh/m]) that may evolve with space $x_i$ along link length and time $t$. Here each link is assumed to be one-lane, hence multi-lane roads are

composed by several links. The network aggregated flow $Q_{\Delta T}$ and density $K_{\Delta T}$ are often defined as (Geroliminis & Daganzo, 2008):

$$Q_{\Delta T} = \frac{1}{L_{\text{net}}} \sum_{\text{link } i=1}^{N_{\text{link}}} l_i \langle q_i \rangle_{l_i, \Delta T} \tag{1.1}$$

$$K_{\Delta T} = \frac{1}{L_{\text{net}}} \sum_{\text{link } i=1}^{N_{\text{link}}} l_i \langle k_i \rangle_{l_i, \Delta T} \tag{1.2}$$

where $L_{\text{net}} = \sum_{i=1}^{N_{\text{link}}} l_i$ is the total length of the network, and $\langle q_i \rangle_{l_i, \Delta T}$ and $\langle k_i \rangle_{l_i, \Delta T}$ are the mean values of flow and density on link $i$ over the link length and over a given observation period $\Delta T$ (in [s]), also referred to as the aggregation period. The aggregated flow $Q_{\Delta T}$ is often called circulating, space-mean or average flow, because it aims at representing the vehicles actually traveling in the network. It can be measured through loop detectors in the links that count passing vehicles. The average density $K_{\Delta T}$ can also be measured indirectly through loop data, via the estimation of occupancy. In practice, it is not necessary to have Eulerian observations for all the links in the network. A limited set of observations is actually sufficient, providing the fact that the loop detectors are homogeneously distributed across the network (see e.g. Ortigosa *et al.*, 2014). However, this is rarely the case in real cities, as loops are often located near traffic signals for signal control purposes. As shown by Leclercq *et al.* (2014), the loop position on links influences the estimation of average flow and density.

Another definition for these aggregated variables was presented in Edie (1963), it synthesizes time- and space-mean calculations in avoiding undesirable border effects of the space-time analysis window for the estimation of $Q_{\Delta T}$ and $K_{\Delta T}$:

$$Q_{\Delta T} = \frac{1}{L_{\text{net}} \Delta T} \sum_{\text{veh } i=1}^{N_{\text{veh}}} td_i = \frac{TTD}{L_{\text{net}} \Delta T} \tag{1.3}$$

$$K_{\Delta T} = \frac{1}{L_{\text{net}} \Delta T} \sum_{\text{veh } i=1}^{N_{\text{veh}}} tt_i = \frac{TTT}{L_{\text{net}} \Delta T} \tag{1.4}$$

In these formulas, the aggregation is made on the number $N_{\text{veh}}$ of vehicles circulating in the network within the period $\Delta T$. The variable $td_i$ is the distance (in [m]) traveled by vehicle $i$, and $tt_i$ is the time (in [s]) spent by vehicle $i$ during this period. The sum of $td_i$ is often written as $TTD$, the Total Travel Distance, and the sum of $tt_i$ as $TTT$, the Total Travel Time. By construction, this estimation method is suitable when the dataset consists in Lagrangian observations, i.e. individual vehicle trajectories. Like Eulerian observations, only a limited set of probe vehicles is sufficient to get reliable estimation of the average flow and density (Nagle & Gayah, 2013). Leclercq *et al.* (2014) showed that a more accurate estimation can be obtained in using both loop and probe data. Fusion methods between these two sources of data are also discussed in Ambuhl & Menendez (2016).

These two definitions are actually equivalent if the link mean flow $\langle q_i \rangle_{l_i, \Delta T}$ and density $\langle k_i \rangle_{l_i, \Delta T}$ are estimated with Edie's formula, i.e. $l_i \Delta T \langle q_i \rangle_{l_i, \Delta T} = \sum_{\text{veh } j} td_j^i$ and $l_i \Delta T \langle k_i \rangle_{l_i, \Delta T} = \sum_{\text{veh } j=1} tt_j^i$, where $td_j^i$ and $tt_j^i$ are respectively the distance traveled and time spent by vehicle $j$ on link $i$ during $\Delta T$.

### Production and accumulation

Interestingly, the average flow and density can be easily transformed into additive variables, known as the production $P_{\Delta T}$ (in [veh.m/s]) and the accumulation $n_{\Delta T}$ (in [veh]):

$$P_{\Delta T} = L_{\text{net}}.Q_{\Delta T} = \frac{TTD}{\Delta T} \tag{1.5}$$

$$n_{\Delta T} = L_{\text{net}}.K_{\Delta T} = \frac{TTT}{\Delta T} \tag{1.6}$$

The advantage of these new variables is that the aggregation of different regions is straightforward: the new region characteristics will be the sum of its subregion ones. The accumulation $n_{\Delta T}$ is the mean number of circulating vehicles during $\Delta T$. The travel production $P_{\Delta T}$ is a measure of the quantity of trips that can be performed by units of time. It can be seen as a number of vehicles multiplied by a speed; for instance the production of 100 vehicles driving at 20 km/h is equivalent to the one of 50 vehicles driving at 40 km/h. It can also be seen as a trip distance multiplied by a flux; for example the production of 400 veh/hr along 1 km of road network is equivalent to the one of 200 veh/hr along 2 km. This corresponds to a total traveled distance of 400 km during one hour.

### Space-mean speed

The mean or average speed $V_{\Delta T}$ is the space-mean speed of all the vehicles traveling during the period $\Delta T$. Although it could be accessed through double loop detectors in links, in practice the mean speed is evaluated by estimating the vehicle travel distance and travel time. By definition, the mean speed verifies the following relationships:

$$V_{\Delta T} = \frac{Q_{\Delta T}}{K_{\Delta T}} = \frac{P_{\Delta T}}{n_{\Delta T}} = \frac{TTD}{TTT} \tag{1.7}$$

### Exit rate

The network perimeter outflow $q_{\text{out},\Delta T}$ is the time-mean total flow of users exiting the network through its nodes at periphery. This variable is important as it gives an idea of the network performance for travelers who are heading to neighboring regions. In practice, this quantity is however hardly measurable with loop data, as it would require a detector at each exit node of the network. Using only a subset of exit nodes is not possible, as a significant part of the total outflow could be missed.

The network trip completion rate $G_{\Delta T}$ is the mean rate at which users reach their destinations. It includes the perimeter outflow and the arrival rate to destinations inside the network. This quantity is even more difficult to assess, as the locations of inner destinations may be various and numerous. A convincing measure of the trip completion rate can only be achieved by tracking a significant amount of vehicles.

All these aggregated variables have been defined for a given aggregated period $\Delta T$, but note that by choosing an infinitesimal period, these formulations are still theoretically valid to define the instantaneous flow $Q(t)$, density $K(t)$, production $P(t)$, accumulation $n(t)$, mean speed $V(t)$, outflow $q_{\text{out}}(t)$, and trip completion rate $G(t)$ at time $t$.

### Several MFD definitions

Different definitions of MFD exists, depending of the variables involved. The conceptual formalism of the MFD which is still used today was first established by Daganzo (2007). He proposed that the accumulation $n$ (or the average density $K$) could be seen as a state variable, and that both the average flow $Q$, the production $P$ and the trip completion rate $G$ would be expressed as a function of accumulation or density. To avoid confusion, each relationship is given a different name following the notation in Lamotte & Geroliminis (2018): $Q = \mathbf{Q}(K)$ is the flow-MFD, $P = \mathbf{P}(n)$ is the production-MFD, and $G = \mathbf{G}(n)$ is the outflow-MFD. The speed-MFD $V = \mathbf{V}(n) = \mathbf{P}(n)/n$ can also be defined thanks to the previous relationships. In this section, we use bold letters to distinguish the MFD functions from the field measurements, but only normal letters will be used in the next sections, since the existence of a well-defined MFD will be often assumed. Note that the flow-MFD and production-MFD are linked through the total network length: $\mathbf{P}(n) = L_{\text{net}}\mathbf{Q}(n/L_{\text{net}})$. Hence, these three MFD definitions (mean flow, production and speed) are equivalent and represent an estimation of traffic states in the reservoir considered. Their relationships are summarized as follows:

$$\mathbf{P}(n) = n\mathbf{V}(n) = L_{\text{net}}\mathbf{Q}\left(\frac{n}{L_{\text{net}}}\right) \tag{1.8}$$

The relationship between these MFD definitions and the outflow-MFD is quite different, because a network with the same mean flow (same aggregated traffic state) can exhibit different values of trip completion rate. Total outflow (at perimeter and inside the reservoir) actually depends on users' routing in the network, but a simple steady-state relationship with mean flow can still be defined. Based on theoretical considerations on a ring road with endogenous traffic, Daganzo (2007) showed that the outflow-MFD is proportional to the production-MFD if we assume that all users travel the same distance $L$:

$$\mathbf{G}(n) = \frac{\mathbf{P}(n)}{L} \tag{1.9}$$

One should remember that this relationship constitutes a steady-state approximation, if one considers the existence of a production-MFD in a given reservoir (Lamotte *et al.*, 2018). But note that a lot of studies assume the existence of the outflow-MFD without referring to a production- or flow-MFD in their modeling framework (see e.g. Aboudolas & Geroliminis, 2013, Haddad & Shraiber, 2014, Hajiahmadi *et al.*, 2015). In this case, the outflow-MFD is seen as an intrinsic property of the network, it is also called Network Exit Function (NEF) by some authors (e.g. Ortigosa *et al.*, 2015, Wang *et al.*, 2017).

The distance $L$, often called the average trip length, is an essential component of MFD-based modeling. Equation 1.9 can also be seen as the application of the queuing formula of Little (1961). Assuming that a reservoir basically consists in a queuing system, the total queue outflow is equal to the average number of users in the queue divided by the mean time they spend in the queue, according to this formula. If users travel a distance of $L$ at speed $V$ on average, the mean travel time can be estimated as $L/V$. It follows that the queuing formula corresponds to $G = n.V/L = P/L$.

### Characteristic values

Several characteristic values are particularly important in the MFD shape. Let us take the

production-MFD and speed-MFD as examples, knowing that the other definitions can be applied as well by linear transformations. Figure 1.4 presents a simple typical shape of a production-MFD and its corresponding speed-MFD with the following characteristic values: the jam accumulation $n_j$ (in [veh]), the free-flow speed $u$ (in [m/s]), the critical accumulation $n_c$ (in [veh]) and the maximum production $P_c$ (in [veh.m/s]).

The jam accumulation $n_j$ corresponds to the maximum number of vehicles that can possibly travel in the reservoir, this theoretical state is known as "global gridlock" where the mean speed becomes null: all the reservoir is jammed so that no one can move anymore. Such a situation is rarely observed in reality, but can arise locally in dense urban areas. The theoretical value of $n_j$ is $L_{\text{net}}k_j$, where $k_j$ is the average link maximum density. But $n_j$ is generally expected to be lower than this maximum value, as a gridlock situation does not necessarily imply that all the links are full of vehicles in a network. More details and insights about this can be found in Mahmassani *et al.* (2013b).

The free-flow speed $u$ is the maximum speed that can be experienced by travelers in the reservoir on average. Its upper bound is the average free-flow speed in the network links. The speed $u$ is generally lower than this bound, because even in free-flow conditions during a whole trip, on average the cruising speed is reduced by traffic lights at intersections. In the production-MFD, $u$ corresponds to the slope at origin $dP(n=0)/dn$.

The critical or maximum production $P_c$ corresponds to the maximum throughput of the reservoir. Perimeter control strategies aim at keeping traffic states as close as possible to this point during demand peak periods. This ensures that the network delivers the best conditions to travelers when the demand is high. By analogy with traffic on road or highway, $P_c$ is also called the production capacity of the reservoir. Apart from getting it from the whole MFD shape, it is generally very difficult to estimate this parameter. Laval & Castrillon (2015) proposed an analytical method based on some network intrinsic features (mean block length and traffic signal red-to-green ratio) to estimate $P_c$. Their method seems promising as they validated it with the field data from Yokohama, Japan.

The critical accumulation $n_c$ distinguishes the free-flow or undersaturation regime ($n < n_c$) from the congestion or oversaturation regime ($n > n_c$). By definition, the maximum production is reached at this point: $P_c = \mathbf{P}(n_c)$. In some cases the MFD shape of a given reservoir exhibits a flat region around $P_c$ (case of a trapezoidal shape for instance). Then there are actually two critical accumulation values that define the range of the flat region. The critical accumulation (or the maximum of the critical accumulation range) serves as the set point in simple Proportional-Integral (PI) controllers to keep traffic states close to $P_c$. To our best knowledge, there is no direct way to get an order of magnitude of this parameter, it is generally measured after the shape of the MFD has been estimated (estimation methods are detailed further).

### Existence of the MFD

For the existence of the MFD in a reservoir, Daganzo (2007) made the following assumptions: traffic states within the reservoir must be homogeneous, i.e. the network must be loaded uniformly so that congestion is evenly distributed among its links. This means that at a given time, the links must have a similar mean density, or that all the vehicles travel at almost the same speed. Such a situation was empirically exhibited by Geroliminis & Daganzo (2008). Thanks to loop data from a 10 km$^2$ district of Yokohama (Japan), the authors showed a low-scattered and well-defined flow-MFD, as presented in Figure 1.5. The proportionality

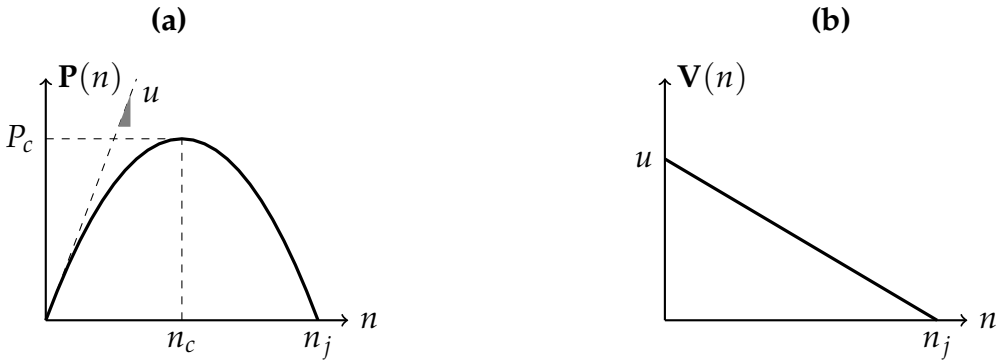



FIG. 1.4 – *Typical shape of MFD (Greenshield's parabolic model) with its characteristic values. (a) production-MFD and (b) speed-MFD*

relationship between trip completion rate and average flow was also verified empirically in the same study, see Figure 1.6.

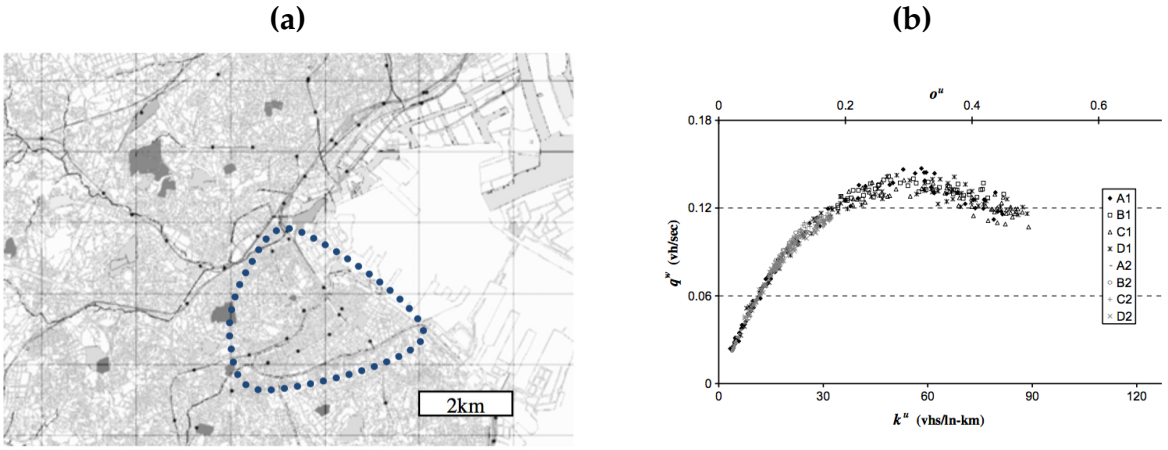



FIG. 1.5 – *Experimental MFD from Yokohama, Japan (Geroliminis & Daganzo, 2008, Fig. 2). (a) Map of Yokohama, (b) aggregated flow vs occupancy (measure of aggregated density)*

According to the early studies (e.g. Daganzo, 2007, Daganzo & Geroliminis, 2008, Geroliminis & Daganzo, 2007, 2008), the MFD is supposedly an intrinsic feature of a network. These authors assume that it only depends on the network supply: road layout and geometry, traffic signal timings, road capacity, etc. On the other hand, the MFD should be independent from the travel demand, i.e. the routing of travelers and the Origin-Destination (OD) matrix. However, we will see in the next sections that this assumption may be far from what is observed in reality, real networks have a lot of heterogeneities which could impact the shape, the scatter and even the existence of the MFD of a given area. OD matrix and user route choices may also have significant effects on its definition.

### Estimation of the MFD

As detailed in Leclercq *et al.* (2014), there are two kinds of estimation methods for the MFD: analytical and experimental methods. The first one is based on the Variational Theory and is widely presented in Daganzo & Geroliminis (2008), Leclercq & Geroliminis (2013). In short the MFD is approximated with different "cuts" in the flow-density plane, which are upper

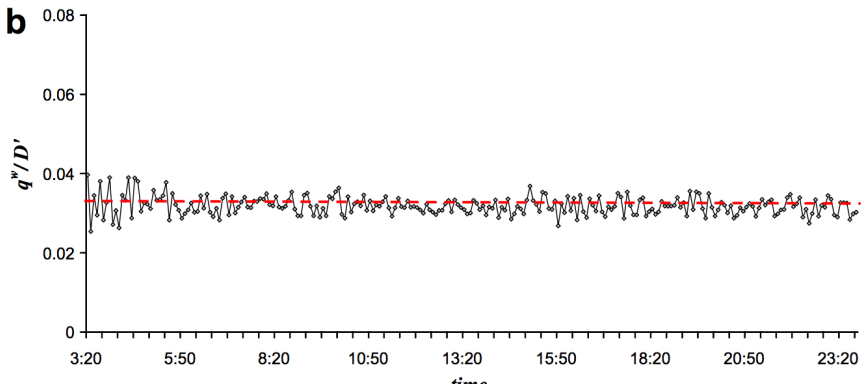



FIG. 1.6 – *Time series of the ratio of the circulating flow over the exit flow, data from Yokohama (Geroliminis & Daganzo, 2008, Fig. 8)*

bounds of the MFD. These cuts can be determined thanks to the Variational Theory and the characteristics of all single links within the network (length, number of lanes, free-flow speed, congested wave speed, saturation flow, traffic signal times). Under certain regularity sufficient conditions (slow-varying and distributed demand, redundant network i.e. many route choices, homogeneous network, links with FD independent from turning movements), the upper bound constituted by the cuts is a good approximation of the network MFD. The analytical method is only applicable for the MFD of arterials.

The second kind of estimation methods concerns experimental studies based on Eulerian and/or Lagrangian observations. Both are aimed to estimate flow and density for all links and then aggregate data to obtain a scatter plot approximating the real MFD. Eulerian observations may come from loop detectors whereas Lagrangian data consists of individual vehicle trajectories from probe vehicles for instance. Shoufeng *et al.* (2013) used both kinds of data, while Nagle & Gayah (2013) focused on the Lagrangian approach only, showing that limited mobile probes are sufficient to get accurate estimation results.

There is also a third method recently proposed by Laval & Castrillon (2015), who developed a new way to estimate MFD thanks to stochastic approximations. Their method is based on the Variational Theory in which stochastic processes are included. The main finding is that on the whole, the MFD of an urban corridor should depend on two dimensionless parameters: the mean block length to green ratio and the mean red to green ratio.

### Consistency and application domain for the MFD

Further studies on the MFD have allowed to exhibit some features influencing its shape or scatter. These have shed some light about restrictions of its definition domain and thus allowed to precise the sufficient conditions to observe a well-defined MFD. It has indeed been observed that in simulated or real situations the main hypothesis for the MFD to be defined—uniform distribution of congestion—is not verified; most of the time congestion appeared to be a local phenomenon (Mazloumian *et al.*, 2010, Geroliminis & Sun, 2011b). Buisson & Ladier (2009) were the first to investigate the impact of heterogeneity on MFD, that is, heterogeneity in network structure (arterial vs local streets) and in data measurement (position of loop detectors). They concluded that this heterogeneity greatly influences the MFD shape and scatter. Later on, the impact of road layout on the MFD definition is also highlighted by Knoop *et al.* (2014) with simulations on different networks. Mazloumian

*et al.* (2010) brought out a clearer understanding about spatial distribution of vehicles and how it affects the scatter and existence of the MFD. Their simulation results showed that a well-defined and low-scattered MFD can be related to a given level of spatial heterogeneity, measured as the standard deviation of link densities. This finding was then confirmed by Geroliminis & Sun (2011b) who proposed a general condition to observe a well-defined MFD: the spatial distribution of vehicles must be similar for times when average network occupancy is about the same. Heterogeneity in networks and its consequence on network capacity was also examined by Geroliminis & Boyaci (2012) who focused on the effect of turning movements.

At the same time, the existence and definition of MFD for freeways is notably discussed by Geroliminis & Sun (2011b,a), Cassidy *et al.* (2011). While Geroliminis & Sun (2011a) showed that freeway MFDs have high scatter and hysteresis thanks to real data, Cassidy *et al.* (2011) argued that for a restrained domain area and with an appropriate filtering method, well-defined MFD could arise for periods when a single regime (free-flow or congestion) is observed in all links.

The time of the day has also an influence on the MFD shape: hysteresis phenomena can be observed during morning peaks as illustrated in Figure 1.7, showing a lower network capacity for congestion recovery than for congestion onset. This was mentioned and analyzed by many authors (Buisson & Ladier, 2009, Gayah & Daganzo, 2011, Daganzo *et al.*, 2011, Geroliminis & Sun, 2011a, Muhlich *et al.*, 2015, Saeedmanesh & Geroliminis, 2015). Basically, the hysteresis phenomenon is explained by the fact that when congestion recovers, reducing queuing vehicles entails more heterogeneities in spatial distribution and thus a reduction of the network production, so that a lower MFD appears compared to the first onset of congestion (Gayah & Daganzo, 2011). While analyzing hysteresis patterns for different network design, Muhlich *et al.* (2015) found interesting conclusions: the presence of arterial streets has a significant influence on spatial distribution of congestion, and arterial ring roads seem to distribute congestion more evenly with better traffic performance than arterial grids.

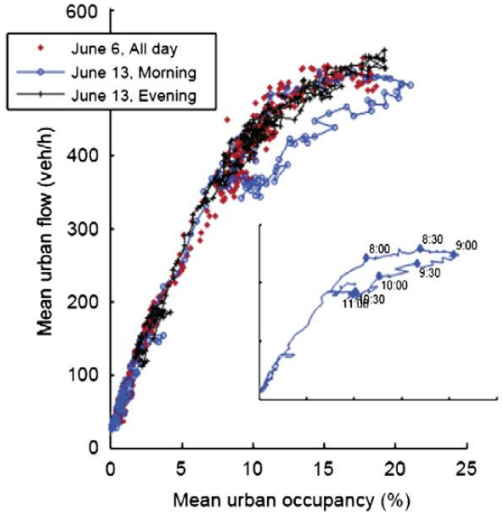



FIG. 1.7 – *Hysteresis phenomenon in the MFD, data from Toulouse, France (Buisson & Ladier, 2009, Fig. 5)*

Other factors influencing congestion distribution and MFD, like traveler information, have also been studied by Zhao *et al.* (2014). Thanks to numerical simulations on simple

arterial grids, it was shown that real-time information about congested areas for a part of the drivers has a dramatic impact on the network MFD. The method used was to introduce stochastic process in the route choice for travelers. Leclercq & Geroliminis (2013) also examined the consequences of driver route choices on the MFD shape, using an improved estimation method based on the Variational Theory.

### Developments and extensions, the 3D-MFD

Geroliminis *et al.* (2014) created the concept of Three-Dimensional MFD to characterize traffic flow in bimodal networks. The idea is to describe interactions between cars and buses as the road space, and see how the density of vehicles in each mode can affect the mean flow in the road. With simulation experiments based on a real network (3 km$^2$ area of Downtown San Fransisco), the authors defined traffic states by means of a point in a 3D-space (flow of vehicles vs concentration of cars and concentration of buses) instead of a 2D-space for the standard MFD concept presented earlier. Moreover, a distinction was made between vehicles and passengers, resulting in two models: the 3D-vMFD for vehicles, and the 3D-pMFD for passengers (see also Chiabaut, 2015) to take into account the fact that a bus carries more travelers than a car, see Figure 1.8. The relationship between vehicle and passenger flow is described by $p = h_c.q_c + h_b.q_b$, where $p$ is the total passenger flow, $q_c$ and $q_b$ are respectively the car and bus flows, and $h_c$ and $h_b$ the respective average numbers of on-board passengers. This new concept allowed to find out the proportion of cars and buses maximizing the total vehicle flow. The existence of the 3D-MFD was recently shown in the city center of Zurich (Switzerland) with data from loop detectors and automatic vehicle location devices (Loder *et al.*, 2017).

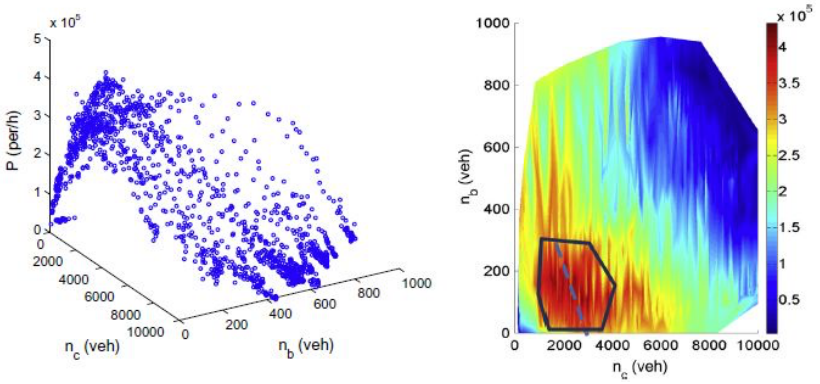
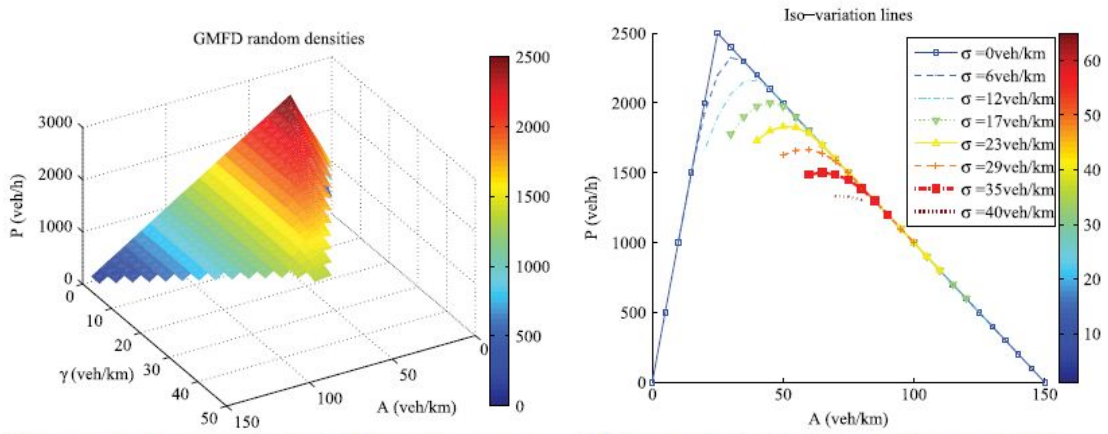



FIG. 1.8 – *3D-pMFD, passenger flow vs car accumulation $n_c$ and bus accumulation $n_b$, results from simulation (Geroliminis et al., 2014, Fig. 2)*

### Developments and extensions, the Generalized MFD (GMFD)

Knoop *et al.* (2015) proposed another kind of extension to take the spatial distribution of congestion into account. This extension is based on previous studies that addressed the influence of spatial heterogeneity on the MFD, as presented above when discussing the consistency and application domain for the MFD. Buisson & Ladier (2009) first introduced the question of spatial heterogeneity and how it affects the MFD definition. Then Mazloumian *et al.* (2010) showed with simulations on a regular grid network that there exists a relationship between the network average flow and the standard deviation of link densities. This

idea is also studied later by Mahmassani *et al.* (2013b). The effect of spatial variability in vehicle densities on the MFD is explored in more details by Geroliminis & Sun (2011b,a) with the data from downtown Yokohama, Japan, and the Twin Cities freeway network, USA. The model developed by Knoop *et al.* (2015) is called the Generalized MFD (GMFD), and relates average flow to average density and spatial inhomogeneity of density, as shown in Figure 1.9. As first introduced by Mazloumian *et al.* (2010), the spatial inhomogeneity is defined as the standard deviation of the link densities in the reservoir. The authors showed with simulation results that the GMFD is able to (i) describe the hysteresis patterns previously observed in the literature, and (ii) predict the loss of performance due to spatial heterogeneity.




(a) The generalized macroscopic fundamental diagram. To improve visibility of the surface, the graph is rotated such that the axis of density increases from right to left.

(b) Macroscopic fundamental diagrams with constant levels of spatial inhomogeneity of density.

FIG. 1.9 – *GMFD, results from simulation (Knoop et al., 2015, Fig. 1)*

### Application to city perimeter control

One of the direct applications of the previous models is the control of perimeter flows to restrain congestion development and try to keep accumulation near the maximum of production for a given zone. Haddad & Geroliminis (2012) first investigated on the stability analysis of traffic control for two-region systems, and dealt with issues about optimality vs stability in using system phase portraits. The two-reservoir or two-bin model was then deeply studied by Geroliminis *et al.* (2013). Aboudolas & Geroliminis (2013), Ampountolas *et al.* (2014), Hajiahmadi *et al.* (2015) further developed the perimeter control approach with multi-reservoir systems. In their models, a city is partitioned into $N$ reservoirs in which a well-defined MFD is assumed. Based on flow conservation equations in which a control parameter is added to regulate inflows, they obtained a set of coupled equations describing traffic dynamics. Because the resolution of such systems may be very complex, the authors used linearizing methods around the equilibrium state. Ampountolas *et al.* (2014) also used the 3D-MFD model for a mixed bimodal traffic, and the control policy they developed has allowed to reduce the overall congestion and improve traffic performance. Hajiahmadi *et al.* (2015) used two types of controller, one on the boundary to control transfer flows and one inside each reservoir which affects signal timing to control each MFD shape.

The application of multi-reservoir models implies spatial partitioning of cities into homogeneous regions, which remains a challenge and has been particularly studied by Ji & Geroli-

minis (2012), Zhou *et al.* (2012), Saeedmanesh & Geroliminis (2016). Another key point of the aggregated modeling of cities is the description of flow transfer between reservoirs. Whereas some authors like Aboudolas & Geroliminis (2013) mainly use flow conservation equations including inflow parameters (which represent control variables), Knoop & Hoogendoorn (2014) have based their model on the Cell Transmission Model (CTM) of Daganzo (1994) to build a Network Transmission Model. By determining transit flow with a supply/demand model at each border of reservoirs, this method seems reliable for the modeling of traffic dynamics, but has not been validated yet by experimental studies or microscopic simulations. Haddad *et al.* (2013) have explored an interesting case of an urban network with a freeway connecting two parts of it. While the urban region is modeled with the multi-reservoir approach, the authors have taken advantage of the CTM to describe the flow in the freeway between the two parts of the urban network. A new approach about transfer flow was also presented by Leclercq *et al.* (2015). In order to observe the effect of route patterns on the MFD, they modeled explicit macroscopic routes into a reservoir which is a promising approach to consider flow transmission at the interface of different models. This method should also provide a better estimation of the average trip length within a reservoir.

#### Application to route guidance

An alternative or complementary option to perimeter control is to divert users from congested areas in advising them longer but free-flow routes. Knoop *et al.* (2012) first presented the problem of routing strategies based on MFD as these ones may have an impact on the initial MFD in return. Hajiahmadi *et al.* (2013b), Yildirimoglu *et al.* (2015) have taken advantage of the multi-reservoir approach to propose optimal dynamic route guidance.

As first noticed by Knoop *et al.* (2012), boundary control and routing advice are interdependent. Yildirimoglu & Geroliminis (2014) provided an interesting analysis about dynamic equilibrium for multi-reservoir systems, in which a methodology was presented to assess the influence of control on route choice behavior.

### 1.2.2 Large-scale simulation using a multi-reservoir framework

Multi-reservoir modeling of large-scale urban areas consists in partitioning the network in $N$ reservoirs. In this way traffic conditions are assumed to be homogeneous in each region, so that a well-defined and low-scattered MFD can describe traffic dynamics within each region. In this section we present and discuss the existing formulations of MFD-based modeling, from a single reservoir to a multi-reservoir system. As mentioned in section 1.1.3, the purpose of our work is to investigate and challenge these approaches, together with proposing new ones to overcome some of the limitations that are identified.

#### The single reservoir model

The dynamics of a single reservoir was described in Daganzo (2007), Geroliminis & Daganzo (2007). It corresponds to a given part of an urban network where the traffic states are characterized by a well-defined production-MFD $P(n)$ (in [veh.m/s]), where $n$ (in [veh]) is the accumulation (number of circulating vehicles in the reservoir). The production-MFD is notably defined by its characteristic values: the jam accumulation $n_j$, the critical accumulation $n_c$ and the maximum production $P_c$. All users are assumed to travel the same average trip length $L$. The reservoir entry (also conceptually called "upstream boundary")

is the aggregation of all individual entry nodes of the network; similarly the reservoir exit (or "downstream boundary") aggregates all the exit nodes. Through the entry is defined the total effective inflow $q_{\text{in}}(t)$, and through the exit the total effective outflow $q_{\text{out}}(t)$. See Figure 1.10 for a schematic representation of the single reservoir model with its production-MFD.

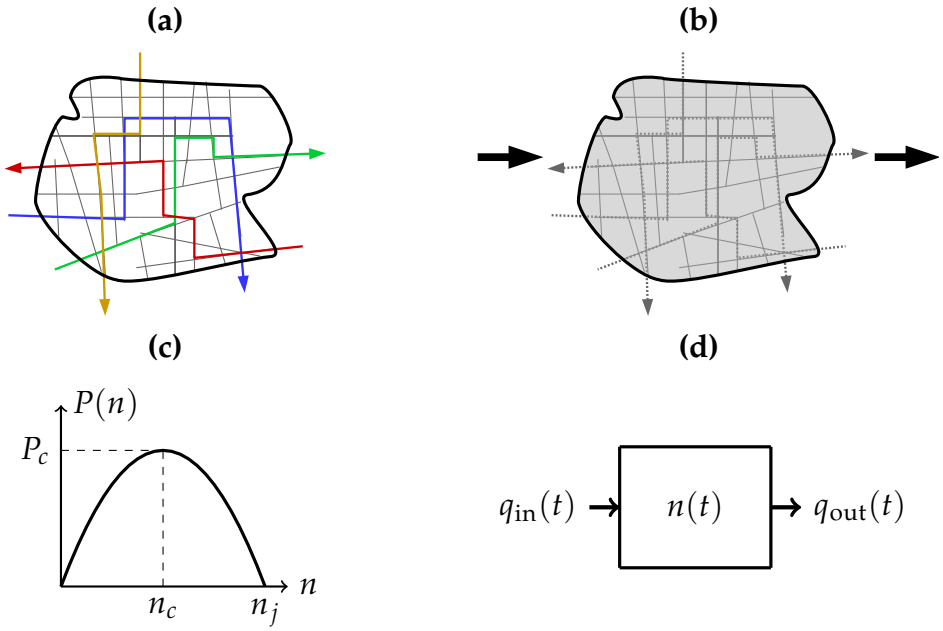



FIG. 1.10 – *The single reservoir model. (a) Network with link-scale trips, (b) aggregation into one reservoir, (c) production-MFD and (c) queuing model*

The accumulation $n(t)$ is the state variable of the system. Its evolution is governed by the following flow conservation equation:

$$\frac{dn}{dt} = q_{\text{in}}(t) - q_{\text{out}}(t) \tag{1.10}$$

In the most simple approach, the total inflow equals the inflow demand $\lambda(t)$ (departure rate of vehicles that want to enter the network or that are created inside), and the total outflow corresponds to the trip completion rate $G(n)$:

$$\left| \begin{array}{l} q_{\text{in}}(t) = \lambda(t) \\ q_{\text{out}}(t) = G(n(t)) = \dfrac{P(n(t))}{L} \end{array} \right. \tag{1.11}$$

Note that the space dimension is not explicitly considered in this formulation. The distance traveled is only characterized by the average trip length included in the outflow-MFD. In equation 1.11, we clearly notice that a sudden modification of the demand will have an instantaneous effect on the accumulation and thus on the outflow via $G(n)$. That is why, such modeling framework is suitable for quasi-static conditions only, i.e. when the demand evolves slowly (Daganzo, 2007, Geroliminis & Daganzo, 2007).

However, such a system is not stable in congestion as mentioned by Daganzo (2007), because of the shape of the outflow-MFD. In oversaturated conditions, $G(n)$ decreases with

$n$, so that an rise in accumulation entails a positive feedback effect on $n$ via the outflow, see equation 1.10. Thus the accumulation keeps increasing towards global gridlock. This is not the case of undersaturated situations, where $G(n)$ increases with $n$ and therefore induces a negative feedback mechanism on the system to keep it stable. As a consequence, there are two possibilities to reproduce congested states in the single reservoir model. The first one is to use a demand peak scenario that temporarily exceeds the MFD capacity. In this case, the system outflow is unable to sustain the departure rate of travelers, so that accumulation can reach oversaturated states. However, the reservoir can converge to gridlock if the demand peak lasts for a long time. The second option consists in designing an entrance function $I(n)$ that automatically restricts inflow for large values of accumulation. This function was first suggested by Geroliminis & Daganzo (2007), based on simulation results from the network of Downtown San Fransisco, USA. The advantage of including $I(n)$ in the single reservoir modeling is that the system is now stable for any demand scenario, because the effective inflow would be rewritten as:

$$q_{\text{in}}(t) = \min\left[\lambda(t); I(n)\right] \tag{1.12}$$

With this formulation, a queue at the reservoir entry must be included in the model as well, to take the waiting vehicles into account when the demand is not satisfied. If a point queue model is used, the exact calculation of inflow is then:

$$q_{\text{in}}(t) = \begin{cases} \min\left[\lambda(t); I(n)\right] & \text{if } n_q = 0 \\ \max I(n) & \text{if } n_q > 0 \end{cases} \tag{1.13}$$

$$\frac{dn_q}{dt} = \lambda(t) - q_{\text{in}}(t) \tag{1.14}$$

where $n_q(t)$ is the number of vehicles at $t$ queuing before entering the reservoir.

Such a function is ignored in perimeter control studies, as input flow controllers play a similar role (i.e. preventing the reservoir from reaching oversaturated states). But some authors like Knoop & Hoogendoorn (2014), Lentzakis *et al.* (2016) systematically used this function, and even proposed a theoretical shape like the one presented in Figure 1.11(b). Note that no clear consensus exists about $I(n)$ in the literature. For instance, the simulation study from Geroliminis & Daganzo (2007) suggests that the "critical" point in $I(n)$ is actually different from $(n_c, P_c/L)$.

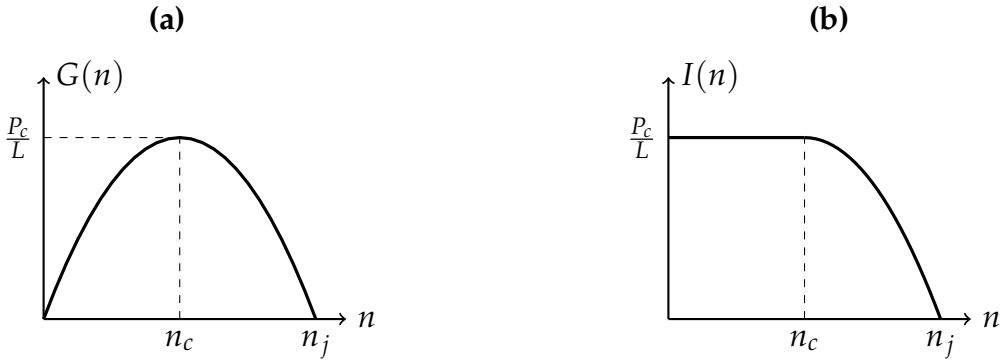



FIG. 1.11 – *Typical shape of MFD. (a) outflow-MFD and (b) entrance function*

In this formulation, another important assumption is that the system outflow exclusively depends on the reservoir inner conditions. Its possible reduction in oversaturation only

comes from internal congestion, but in some cases, exogenous limitations at exit could be included for users who want to leave the network. This is of critical importance when connecting several reservoirs together, because the outflow from one reservoir is the inflow for a neighboring one. And this inflow could be limited if entrance functions are used. Different solutions to handle flow exchanges are proposed in two-reservoir and multi-reservoir models.

### The two-reservoir or two-bin model

Different versions of the two-reservoir model (also called "two-bin" model) were presented in Geroliminis & Daganzo (2007), Gayah & Daganzo (2011), Haddad & Geroliminis (2012), Tsekeris & Geroliminis (2013), Geroliminis *et al.* (2013). Most of them consists in the basic modeling of a city split into two concentric areas, the center and the periphery. Figure 1.12 shows two representations of this model, where region $R_1$ is the periphery and region $R_2$ is the city center. We first present the approach of Geroliminis & Daganzo (2007) which is the most general one found in the literature. The flows mentioned on this figure are of two kinds for each reservoir:

- Endogenous inflow/outflow
  - $q_{\text{in}11}(t)$, $q_{\text{out}11}(t)$: allowed inflow into $R_1$ (resp. outflow from $R_1$) for trips ending in $R_1$
  - $q_{\text{in}12}(t)$: allowed inflow into $R_1$ for trips ending in $R_2$
  - $q_{\text{in}22}(t)$, $q_{\text{out}22}(t)$: allowed inflow into $R_2$ (resp. outflow from $R_2$) for trips ending in $R_2$
  - $q_{\text{in}21}(t)$: allowed inflow into $R_2$ for trips ending in $R_1$
- Exogenous inflow/outflow
  - $q_{12}(t)$: allowed transfer flow from $R_1$ to $R_2$
  - $q_{21}(t)$: allowed transfer flow from $R_2$ to $R_1$

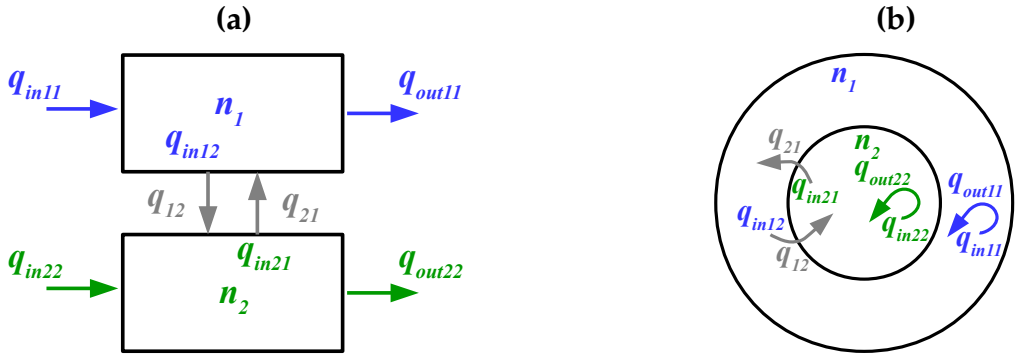



FIG. 1.12 – *Two possible representations of the two-reservoir model, (a) adjacent blocks and (b) concentric zones*

No flow from outside the city is considered by the above-mentioned authors (that is why inflow and outflow are called "endogenous"). However this could be easily taken into account by including the inflow from outside to $R_1$ in $q_{\text{in}11}(t)$ and the outflow from $R_1$ to

outside in $q_{\text{out}11}(t)$. In this section the endogenous flow may correspond to either internal or exchange flow with the city outside.

The state variables are the accumulations in each region $n_1(t)$ and $n_2(t)$, decomposed by trip destination:

$$n_1(t) = n_{11}(t) + n_{12}(t) \tag{1.15}$$
$$n_2(t) = n_{22}(t) + n_{21}(t) \tag{1.16}$$

where:

- $n_{11}(t)$ is the number of vehicles at $t$ in $R_1$ whose destination is in $R_1$
- $n_{12}(t)$ is the number of vehicles at $t$ in $R_1$ whose destination is in $R_2$
- $n_{22}(t)$ is the number of vehicles at $t$ in $R_2$ whose destination is in $R_2$
- $n_{21}(t)$ is the number of vehicles at $t$ in $R_2$ whose destination is in $R_1$

The whole dynamics of the system is then described by the following equations, a better representation of the flow exchange is given in figure 1.13:

$$\begin{cases} \dfrac{dn_{11}}{dt} = q_{\text{in}11}(t) + q_{21}(t) - q_{\text{out}11}(t) \\ \dfrac{dn_{12}}{dt} = q_{\text{in}12}(t) - q_{12}(t) \\ \dfrac{dn_{22}}{dt} = q_{\text{in}22}(t) + q_{12}(t) - q_{\text{out}22}(t) \\ \dfrac{dn_{21}}{dt} = q_{\text{in}21}(t) - q_{21}(t) \end{cases} \tag{1.17}$$

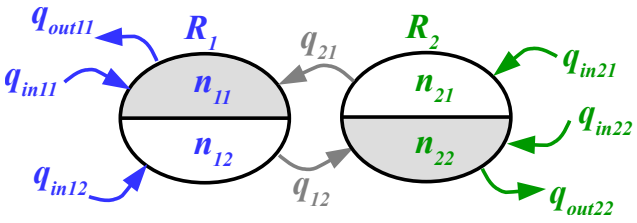



FIG. 1.13 – *Flow transfers in the two-reservoir model*

With the assumption of constant trip lengths, each region $R_i$ has a well-defined outflow-MFD represented by the trip completion flow $G_i(n_i)$ (where $i$ = 1,2) which is the sum of transfer flow plus endogenous flow (see equations 1.18 and 1.19).

$$G_1(n_1) = \frac{n_{12}}{n_1} G_1(n_1) + \frac{n_{11}}{n_1} G_1(n_1) \tag{1.18}$$
$$G_2(n_2) = \underbrace{\frac{n_{21}}{n_2} G_2(n_2)}_{\text{transfer flow}} + \underbrace{\frac{n_{22}}{n_2} G_2(n_2)}_{\text{endogenous flow}} \tag{1.19}$$

To determine the flow that can actually enter or exit the two reservoirs, one must take into account constraints given by the entrance and exit functions of each reservoir on the one

hand, and the input demand or supply on the other hand. In free-flow conditions it can be assumed that all the demand is satisfied, it is however more difficult to arbitrate between the different inflows and outflows when a reservoir becomes congested. One solution among others could be to use the merge model of Daganzo (1995). In the literature, authors manipulating this two-reservoir model are more interested by control strategies and prefer simplifying flow exchange.

Geroliminis & Daganzo (2007) assume that there are no restrictions on the demand $q_{\text{in},i1}$ and $q_{\text{in},i2}$ ($i$ = 1,2) and that the outflow is given by the exit function with respect to the corresponding ratio of accumulation as presented in equations 1.18 and 1.19. Only the transfer flows $q_{12}$ and $q_{21}$ are a mix between the entrance function $I_i(n_i)$ of the destination region and the exit function $G_i(n_i)$ of the departure region ($i$ = 1,2). Therefore we have:

$$\left|\begin{array}{l} q_{\text{in}11}(t) = \lambda_{11}(t) \\ q_{\text{in}12}(t) = \lambda_{12}(t) \\ q_{\text{in}22}(t) = \lambda_{22}(t) \\ q_{\text{in}21}(t) = \lambda_{21}(t) \\ q_{\text{out}11}(t) = \dfrac{n_{11}}{n_1} G_1(n_1) \\ q_{\text{out}22}(t) = \dfrac{n_{22}}{n_2} G_2(n_2) \\ q_{12}(t) = \min\left(u_{12}(t).I_2(n_2(t)); \dfrac{n_{12}}{n_1} G_1(n_1)\right) \\ q_{21}(t) = \min\left(u_{21}(t).I_1(n_1(t)); \dfrac{n_{21}}{n_2} G_2(n_2)\right) \end{array}\right. \tag{1.20}$$

The expressions of $q_{12}$ and $q_{21}$ include the control parameters $u_{12}$ and $u_{21}$ which are set dynamically by the author to regulate entrance flow from one reservoir to another (these parameters are between 0 and 1, their sum always being 1), these are introduced for perimeter control purpose only. To solve the kind of system presented in equation 1.17 applied to $N$ reservoirs, authors like Aboudolas & Geroliminis (2013) implement linearization methods around the equilibrium state. They consider that traffic evolves near the maximum of production in the city center, as their studies are devoted to perimeter control near this optimal traffic state. This justifies the simplification choice adopted by Geroliminis & Daganzo (2007) and other authors working on boundary control as for flow merging problems.

Other versions or extensions of this approach can be found in the literature. Tsekeris & Geroliminis (2013) used exactly the same model assumptions on a simplified case: a convergent flow from the periphery to the center of the city, that is, from $R_2$ to $R_1$. There is one point where their modeling approach differs from Geroliminis & Daganzo (2007). As there is only one possible destination in this model, accumulation decomposition is based on trip origins instead of trip destinations. More precisely, they write the accumulation in $R_1$ as $n_1 = n_{11} + n_{12}$, the number of vehicles originated from $R_1$ plus the number of vehicles originated from $R_2$.

Geroliminis *et al.* (2013) also founded their modeling on the same dynamic equations, but without considering any entrance functions. In their study the transfer flows $q_{12}$ and $q_{21}$ are simply the corresponding fractions of exit flow coming from the origin reservoir, multiplied

by a control parameter (equation 1.21):

$$\begin{aligned} q_{12}(t) &= u_{12}(t)\frac{n_{12}}{n_1}G_1(n_1) \\ q_{21}(t) &= u_{21}(t)\frac{n_{21}}{n_2}G_2(n_2) \end{aligned} \tag{1.21}$$

Haddad & Geroliminis (2012) focused on a very simplified model, with flow from $R_1$ to $R_2$ and internal flow in $R_2$ only. Accumulation is not decomposed on trip destination or origin. In this case traffic dynamics are described by:

$$\begin{cases} \dfrac{dn_1}{dt} = q_{\text{in}1}(t) - q_{12}(t) - q_{\text{out}1}(t) \\ \dfrac{dn_2}{dt} = q_{\text{in}2}(t) + q_{12}(t) - q_{\text{out}2}(t) \end{cases} \tag{1.22}$$

where:

$$\left| \begin{aligned} q_{\text{in}1}(t) &= \lambda_1(t) \\ q_{\text{in}2}(t) &= \lambda_2(t) \\ q_{\text{out}1}(t) &= 0 \\ q_{\text{out}2}(t) &= G_2(n_2) \\ q_{12}(t) &= u_{12}(t)G_1(n_1) \end{aligned} \right. \tag{1.23}$$

The flow $q_{\text{out}1}$ is null, meaning that all the outflow from the periphery $R_1$ is intending to enter the city center $R_2$. This approach helps the authors studying the dynamics of such a two-reservoir system. In particular they were able to draw its phase portrait based on the initial accumulation in each region.

#### Multi-reservoir models

Multi-reservoir models are generalizations of the different versions of the two-reservoir model. The present state-of-the-art figures out three main approaches with a decreasing level of aggregation, that is, a more and more refined view of traffic dynamics. These approaches are detailed below:

- The first approach is described in Aboudolas & Geroliminis (2013), there is no decomposition of accumulation in each reservoir (highest level of aggregation) and a constant average trip length is assumed
- The second approach is presented in Hajiahmadi *et al.* (2013a) which is a generalization of Geroliminis *et al.* (2013), accumulation is decomposed by trip destination and constant trip length is also assumed
- The third approach is defined in Yildirimoglu & Geroliminis (2014), accumulation is decomposed by trip OD and traveled route (in this model a route is seen as a succession of reservoirs), and the trip length is also decomposed in the same way and can vary over time

Although not included in multi-reservoir models, a last approach is worth mentioning because of its potential application. It is presented in Leclercq *et al.* (2015), where one reservoir

is studied with accumulation decomposed by macroscopic OD pairs. This means several macroscopic routes may be defined for a reservoir (a route is associated to a given OD pair). The originality of this approach is that it reintroduces the space dimension in the system dynamics. This system consists in flow-density conservation equations (one per OD pair), similar to the LWR model applied on several links, but where all the equations are coupled through the speed-MFD. The extension to a multi-reservoir model could be the distinction of OD pairs not only by reservoir as proposed in Yildirimoglu & Geroliminis (2014) but by reservoir border (a more refine view of macroscopic OD pairs).

#### Multi-reservoir approach 1: Reservoirs with constant trip length and without accumulation decomposition

This modeling approach is first detailed in Aboudolas & Geroliminis (2013) and then extended in Ampountolas *et al.* (2014). For a city partitioned in $N$ reservoirs, this model describes the temporal evolution of the accumulation $n_i(t)$ in each region $R_i$ ($1 \leq i \leq N$). The main hypothesis is time-invariant regional trip lengths $L_i$.

Aboudolas & Geroliminis (2013) defines the MFD of a reservoir $R_i$ as its outflow-MFD $G_i(n_i)$. The outflow $G_i(n_i)$ includes the trips ending in region $R_i$ and the trips exiting the reservoir $i$. The author also considers a linear relationship between outflow and circulating flow $Q_i(n_i)$ due to the above-mentioned assumption:

$$G_i(n_i) = \frac{L_{\text{net},i}}{L_i} Q_i(n_i) = \frac{P_i(n_i)}{L_i} \tag{1.24}$$

where $L_{\text{net},i}$ is the total network length and $P_i(n_i)$ is the production-MFD in reservoir $R_i$.

The dynamics of the whole system is represented as follows:

$$\forall i \in [1, N], \quad \frac{dn_i(t)}{dt} = q_{\text{in},i}(t) - q_{\text{out},i}(t) + d_i(t) \tag{1.25}$$

where:

- $q_{\text{in},i}(t)$ is the global effective inflow in $R_i$
- $q_{\text{out},i}(t)$ is the global effective outflow of $R_i$
- $d_i(t)$ is the uncontrolled demand, which represents internal flows (i.e. trips starting and ending in the region, thus invisible for reservoir boundary control)

There is no entrance function in this model, the reservoir inflow is defined as the sum of transfer flows from other reservoirs to $R_i$. The outflow is not limited either, it is precisely equal to the trip completion rate:

$$\left| \begin{array}{l} q_{\text{in},i}(t) = \sum_{j \in S_i} q_{ji} = \sum_{j \in S_i} \beta_{ji}(t - \tau_{ji}).G_j(n_j(t)) \\ q_{\text{out},i}(t) = G_i(n_i(t)) \end{array} \right. \tag{1.26}$$

where:

- $S_i$ is the set of origin reservoirs (including reservoir $R_i$ if reachable from the outside perimeter of the network)

- $q_{ji}$ is the transfer flow from $R_j$ to $R_i$ (for $j = i$, $q_{ii}$ represents the inflow from outside entering $R_i$)
- $\beta_{ji}$ is the fraction of flow rate in region $R_j$ allowed to enter the region $R_i$
- $\tau_{ji}$ is the travel time needed for vehicles to approach region $R_i$ from the origin reservoir $R_j$

Figure 1.14 displays the model overview for flow exchanges. The internal flow $d_i(t)$ is not particularly characterized, as part of a perimeter control problem it is more or less considered as system disturbances by the authors.

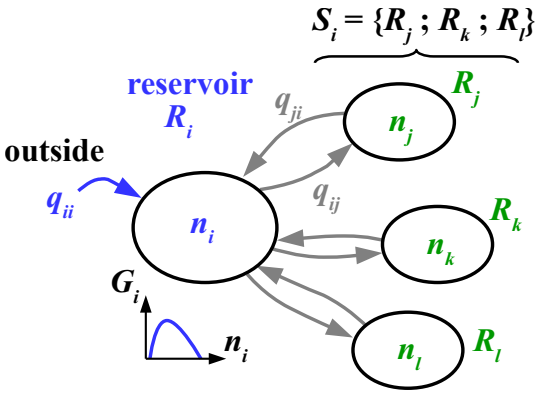



FIG. 1.14 – *Reservoir representation for modeling approach 1*

To get the global solution, the authors used the Taylor first order linearization around the equilibrium state (accumulation in each reservoir near the capacity, i.e. maximum of outflow). Every internal travel time is moreover assumed equal to zero, $\tau_{ji} = 0$. As no entrance function is used in practice regulating the region inflow is ensured by limitations on the fractions of inflows $\beta_{ji}$ and on accumulations:

$$\begin{aligned} &\forall j,i \;\; 0 < \beta_{ji\,\min} \leq \beta_{ji}(t) \leq \beta_{ji\,\max} \\ &\forall j \;\; \sum_{i=1}^{N} (\beta_{ji}(t) + \epsilon_i(t)) \leq 1 \\ &\forall i \;\; 0 \leq n_i(t) \leq n_{i\,\max} \end{aligned} \tag{1.27}$$

where $\epsilon_i(t)$ is the portion of uncontrolled (internal) flow entering reservoir $R_i$.

Other versions or extensions of this modeling approach exist. Ampountolas *et al.* (2014) applied the same modeling for a bi-modal control problem in including the concept of three-dimensional MFD developed by Geroliminis *et al.* (2014) to the reservoir dynamics. Accumulation in $R_i$ is decomposed by mode:

$$n_i(t) = n_{i\,\text{car}}(t) + n_{i\,\text{bus}}(t) = \delta_i(t).n_i(t) + \big(1 - \delta_i(t)\big).n_i(t) \tag{1.28}$$

where $\delta_i(t)$ is the proportion of cars in $R_i$. The MFD of $R_i$ is then replaced by the corresponding 3D-MFD $Q(n_i, \delta_i)$ (see example on Figure 1.8), thus the outflow $G_i(n_i)$ in equation 1.24 is replaced by:

$$G_i(n_i, \delta_i) = \frac{L_{\text{net},i}}{L_i} Q_i(n_i, \delta_i) \tag{1.29}$$

In Ampountolas *et al.* (2014), $L_{\text{net},i}$ is not the total network length of $R_i$ but the average link length in $R_i$.

**Multi-reservoir approach 2: Reservoirs with constant trip length and accumulation decomposition by trip destination to adjacent reservoirs**

The second modeling approach is first detailed in Hajiahmadi *et al.* (2013a) and then in Haddad (2015), this method is a generalization of the two-reservoir model established in Geroliminis *et al.* (2013) extended to a city partitioned into $N$ reservoirs. The main hypotheses are the same as in the previous approach, that is, trip lengths for all travelers inside a region are similar (independent from the OD of the trip), and traffic dynamics within a region $R_i$ is well-described by the outflow-MFD $G_i(n_i)$ $(1 \leq i \leq N)$.

The accumulation $n_i$ in $R_i$ is decomposed by trip destination to its adjacent reservoirs:

$$n_i(t) = n_{ii}(t) + \sum_{j \in S_i} n_{ij}(t) \tag{1.30}$$

where $S_i$ is the set of adjacent regions to $R_i$, i.e. which are directly reachable from $R_i$, and $n_{ij}(t)$ is the number of vehicles in $R_i$ with direct destination to $R_j$. One should notice that this modeling approach only considers trips between two adjacent reservoirs. This means there is no tracking on vehicles which could travel between distant regions as no "transferring" accumulation is taken into account in each region (e.g. vehicles in $R_i$ traveling from $R_k$ to $R_j$). Consequently, the definition of trip destination in this model must not be confused with the one of the real trip OD. Figure 1.15 displays the model overview for flow exchanges and accumulation decomposition.

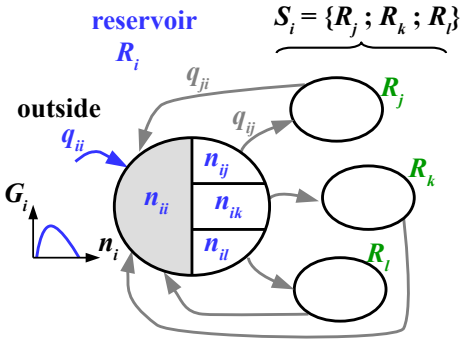



FIG. 1.15 – *Reservoir representation for modeling approach 2*

The trip completion flow $G_i(n_i)$ is the sum of exiting flows plus internal flow:

$$G_i(n_i) = \sum_{j \in S_i} \frac{n_{ij}}{n_i} G_i(n_i) + \frac{n_{ii}}{n_i} G_i(n_i) \tag{1.31}$$

The systems dynamics is then described by the following equations:

$$\begin{cases} \forall i \in [1, N], \quad \dfrac{dn_{ii}(t)}{dt} = q_{\text{in},ii}(t) + \displaystyle\sum_{j \in S_i} q_{ji}(t) - q_{\text{out},ii}(t) \\ \forall i \in [1, N], \ \ \forall j \in S_i, \quad \dfrac{dn_{ij}(t)}{dt} = q_{\text{in},ij}(t) - q_{ij}(t) \end{cases} \tag{1.32}$$

where:

– $S_i$ is the set of regions directly reachable from $R_i$

- $q_{\text{in},ii}(t)$ is the effective inflow (allowed flow to enter/circulate in the region) for internal trips within $R_i$
- $q_{\text{in},ij}(t)$ is the effective inflow for transfer flow from $R_i$ to $R_j$
- $q_{\text{out},ii}(t)$ is the effective outflow for trips ending in $R_i$
- $q_{ij}(t)$ and $q_{ji}(t)$ are respectively the effective transfer flows from $R_i$ to $R_j$ and from $R_j$ to $R_i$

The same approach as in Geroliminis *et al.* (2013) for the two-reservoir model is used: there is no entrance function in each region, the effective inflows are precisely the demand for internal flows and transfer flows, the effective outflow is the reservoir trip completion rate, and the transfer flows are the corresponding fractions of exit flows from departure regions. That is:

$$\left|\begin{aligned} q_{\text{in},ii}(t) &= \lambda_{ii}(t) \\ q_{\text{in},ij}(t) &= \lambda_{ij}(t) \\ q_{\text{out},ii}(t) &= \frac{n_{ii}}{n_i} G_i(n_i) \\ q_{ij}(t) &= \frac{n_{ij}}{n_i} G_i(n_i) \\ q_{ji}(t) &= \frac{n_{ji}}{n_j} G_j(n_j) \end{aligned}\right. \tag{1.33}$$

In Hajiahmadi *et al.* (2013a), Haddad (2015), regulating reservoir inflow is in practice based on boundary control parameters $u_{ij}(t)$ and $u_{ji}(t)$ that are added to the system dynamics to restrict transfer flows. Thus in the previous equations, $q_{ij}(t)$ and $q_{ji}(t)$ must be respectively replaced by $u_{ij}(t).q_{ij}(t)$ and $u_{ji}(t).q_{ji}(t)$. There are moreover other constraints:

$$\begin{aligned} &\forall i,j \;\; 0 < u_{ij\,\min} \leq u_{ij}(t) \leq u_{ij\,\max} \\ &\forall j,i \;\; 0 < u_{ji\,\min} \leq u_{ji}(t) \leq u_{ji\,\max} \\ &\forall j,i \;\; u_{ij}(t) + u_{ji}(t) = 1 \\ &\forall i \;\; 0 \leq n_i(t) \leq n_{i\,\max} \end{aligned} \tag{1.34}$$

**Multi-reservoir approach 3: Reservoirs with time-dependent trip length and accumulation decomposition by trip OD and chosen route**

This third approach is the latest evolution of the MFD-based simulators for multi-reservoir systems. This one has been proposed by Yildirimoglu & Geroliminis (2014), it includes a DTA procedure in order to better comply with dynamic Stochastic User Equilibrium (SUE) conditions. In this section however, we only deal with the modeling of reservoir dynamics. The main change regarding the previous approaches is that the average trip length becomes time-dependent, and differences in distance traveled by vehicles from different OD are dynamically taken into account.

Assuming a city partitioned in $N$ reservoirs, $n_r(t)$ is the total accumulation in region $R_r$. For all origin reservoirs $R_o$ and destination reservoirs $R_d$, $n_r(t)$ is decomposed into $n^r_{od}(t)$ which is the number of vehicles in $R_r$ traveling from $R_o$ to $R_d$. For all regional paths $p$ (or macroscopic routes, defined as sequences of reservoirs crossed) between origin and destination reservoir, $n^r_{od}(t)$ is itself decomposed into $n^{pr}_{od}(t)$ which is the number of vehicles in $R_r$

traveling from $R_o$ to $R_d$ using path $p$. We can write:

$$n_r(t) = \sum_{o=1}^{N} \sum_{d=1}^{N} \sum_{p=1}^{N_{od}} n_{od}^{pr}(t) \tag{1.35}$$

where $N_{od}$ is the number of paths between origin $R_o$ and destination $R_d$. Each region $R_r$ is supposed to exhibit a well-defined production-MFD $P_r(n_r)$, or equivalently a speed-MFD $V_r(n_r) = P_r(n_r)/n_r$. In case of a constant average trip length $L_r$, the trip completion rate is simply $P_r(n_r)/L_r$. However in this approach, the average trip length is made variable and depends on the path used for a given OD pair. Therefore the trip completion rate $G_{od}^{pr}(t)$ for vehicles in $R_r$ traveling from origin $R_o$ to destination $R_d$ using path $p$ is defined by:

$$G_{od}^{pr}(t) = \frac{n_{od}^{pr}(t)}{n_r(t)} \frac{P_r(n_r(t))}{L_{od}^{pr}(t)} \tag{1.36}$$

where $L_{od}^{pr}(t)$ is the average trip length crossed in region $R_r$ by travelers with origin $R_o$ and destination $R_d$ using path $p$. Figure 1.16 displays a model overview for flow exchange and accumulation decomposition.

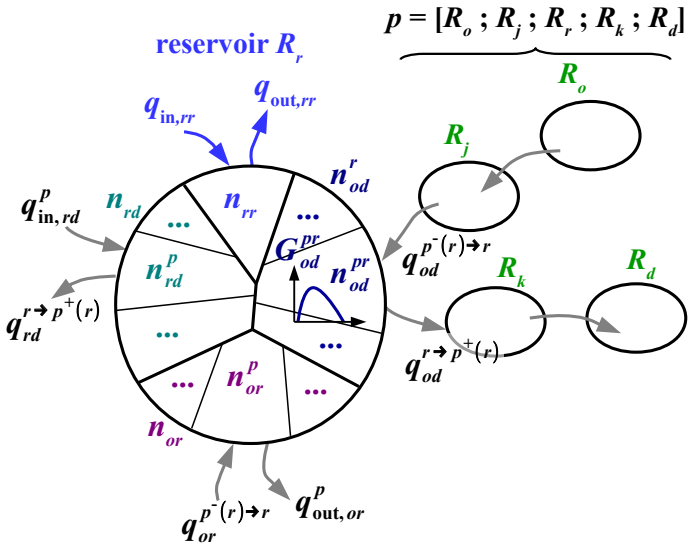



FIG. 1.16 – *Reservoir representation for modeling approach 3*

The reservoir dynamics is described as follows:

$$\begin{cases} \forall r \in [1,N], & \dfrac{dn_{rr}(t)}{dt} = q_{\text{in},rr}(t) - q_{\text{out},rr}(t) & \text{case } r = o = d \\ \forall r \in [1,N], \ \forall d, \ \forall p, & \dfrac{dn_{rd}^{p}(t)}{dt} = q_{\text{in},rd}^{p}(t) - q_{rd}^{r\to p^+(r)}(t) & \text{case } r = o \neq d \\ \forall r \in [1,N], \ \forall o, \ \forall p, & \dfrac{dn_{or}^{p}(t)}{dt} = q_{or}^{p^-(r)\to r}(t) - q_{\text{out},or}^{p}(t) & \text{case } r = d \neq o \\ \forall r \in [1,N], \ \forall (o,d), \ \forall p, & \dfrac{dn_{od}^{pr}(t)}{dt} = q_{od}^{p^-(r)\to r}(t) - q_{od}^{r\to p^+(r)}(t) & \text{case } r \neq o,d \end{cases} \tag{1.37}$$

where:

- $q_{\text{in},rr}(t)$ is the effective inflow for internal trips within $R_r$
- $q^p_{\text{in},rd}(t)$ is the effective inflow for trips generated in $R_r$ with destination to $R_d$ using path $p$
- $q_{\text{out},rr}(t)$ is the effective outflow for internal trips within $R_r$
- $q^p_{\text{out},or}(t)$ is the effective outflow for trips originated in $R_o$ ending in $R_r$ using path $p$
- $q_{od}^{p^-(r)\to r}(t)$ and $q_{od}^{r\to p^+(r)}(t)$ are respectively the effective transfer flows from previous region $R_{p^-(r)}$ to $R_r$ and from $R_r$ to next region $R_{p^+(r)}$ in path $p$ for the considered OD pair ($od$, $rd$ or $or$)

To calculate these different values of flow, the following assumptions are adopted: entrance functions depending on the reservoir sending the inflow are used to regulate transfer flows between reservoirs, sending flow from a reservoir is given by the corresponding trip completion rate for the considered OD pair and path (see equation 1.36), inflow for an OD demand on a given path is not limited, and outflow for trips ending in a reservoir is also not bounded. This gives:

$$\left|\begin{array}{rl} q_{\text{in},rr}(t) &= \lambda_{rr}(t) \\ q^p_{\text{in},rd}(t) &= \lambda^p_{rd}(t) \\ q_{\text{out},rr}(t) &= G_{rr}(t) \\ q^p_{\text{out},or}(t) &= G^p_{or}(t) \\ q_{od}^{p^-(r)\to r}(t) &= \min\left[I_{p^-(r),r}(n_r); G_{od}^{p,p^-(r)}(t)\right] \\ q_{od}^{r\to p^+(r)}(t) &= \min\left[a_{od}^{r\to p^+(r)}(t).I_{r,p^+(r)}(n_{p^+(r)}); G_{od}^{pr}(t)\right] \end{array}\right. \tag{1.38}$$

where:

- $\lambda_{rr}(t)$ is the demand for internal flow in $R_r$
- $\lambda^p_{rd}(t)$ is the demand for trips generated in $R_r$ with destination $R_d$ using path $p$
- $I_{p^-(r),r}(n_r)$ and $I_{r,p^+(r)}(n_{p^+(r)})$ are respectively the entrance functions in $R_r$ for sending flow from $R_{p^-(r)}$ and in $R_{p^+(r)}$ for sending flow from $R_r$
- $a_{od}^{r\to p^+(r)}(t)$ is a coefficient used to take into account the fact that inflow into $R_{p^+(r)}$ may have several origins if this reservoir is included in many routes crossing $R_r$, its formulation given by equation 1.39 is inspired by the queuing formula of Little (1961)

$$a_{od}^{r\to p^+(r)}(t) = \frac{n^{pr}_{od}(t)}{L^{pr}_{od}(t)} \frac{1}{\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{N_{od}} \delta^{t^+(r)}_{p^+(r)} \frac{n^{tr}_{ij}}{L^{tr}_{ij}}} \tag{1.39}$$

where $\delta^{t^+(r)}_{p^+(r)}$ is the Kronecker's symbol, this term is null unless the next region in path $t$ and the next one in path $p$, both paths crossing $R_r$, are identical.

As regards the entrance functions mentioned in equation 1.38, no further details are provided by the authors on their shape or particularity.

## 1.3 Research objectives and major contributions

### 1.3.1 Research questions

Based on the previous literature review, it appears that MFD-based models have been widely used in several studies presenting promising applications. While interesting extensions or simplifications were proposed for these models (from the single reservoir to the multi-reservoir case), on the other hand some crucial elements of traffic flow modeling have been poorly investigated. In general, it seems that each author developed or adapted his own model for his needs, so that a global and consistent framework designed for any simulation or control study is missing. Moreover, as most of the works using MFD focus on control strategies, some strong assumptions are often made to keep the model mathematical formulation quite simple. But these hypotheses have been rarely discussed for other applications like traffic state estimations at large scale. All these concerns are summed up in the following list of questions related to short-term and long-term traffic simulation:

- The mathematical formulation of traffic dynamics in the single reservoir model does not explicitly account for the space dimension. As a result, variation on the inflow affects the outflow instantaneously: the propagation of information in a reservoir may appear inconsistent with traffic flow theory in some cases. How can we introduce delays in the reservoir dynamics to mitigate this effect?

- The single reservoir model is assumed to be valid for slow-varying conditions only. However, demand surge is expected in many situations (e.g. during peak hours). Moreover, the regulation of inflow due to perimeter control is far from operating at a slow rate, as most of the cited works show. To what extend can we pretend that the single reservoir model is accurate? Can we relax the slow-varying hypothesis?

- The MFD has been supposed to be an intrinsic feature of the reservoir, but a lot of studies actually show that heterogeneities in the demand, routing, congestion distribution, etc, have significant impacts on its shape. How can we account for these heterogeneities in MFD-based simulation? What are their effects on simulation results?

- The assumption of a constant average trip length shared by all the travelers has already been questioned, although still adopted in a lot of studies. Irregular shapes of reservoirs are very likely to generate various trip lengths experienced by travelers in real networks. What is the impact of this hypothesis on aggregated traffic simulation? For the multi-reservoir frameworks that account for multiple trip lengths inside a reservoir, what are the implications of these different lengths on traffic dynamics?

- Several studies have shown that usually, congestion does not propagate uniformly in a network. Is there a way to fairly reproduce congestion spilling back in multi-reservoir systems? What are the theoretical implications of the homogeneity assumption on traffic states while simulating oversaturated conditions in multi-reservoir systems?

- The design of the entrance function is still a matter of debate. What is the inflow capacity faced by users who want to enter a reservoir? What is the shape of the entrance function that can fairly reproduce congestion propagation in a reservoir?

- In case several flows (e.g. coming from different neighboring reservoirs) want to enter the same reservoir, how is the available entry capacity shared among these incoming flows? How can we design a proper inflow merging scheme that can apply in any situation?

- The concept of trip completion rate, widely used in multi-reservoir formulations, mixes travelers ending inside the reservoir with others willing to exit the network. However, these two groups of users experience different situations in congestion. The second group can be blocked and thus have to queue at the reservoir border to enter a neighboring reservoir, while the first group can only experience a reduction of mean speed. How can we adjust the trip completion rate modeling to describe the interactions between internal trips and trips that cross the reservoir?

- Finally, a lot of theoretical MFD-based frameworks were developed, but only a few simulation comparisons were carried on with microscopic models or real data. How can we validate MFD-based models? What kind of errors can we expect while predicting aggregated traffic states?

We see that these concerns can be classified in two categories. The first one encompasses the questions concerning the inner dynamics of the single reservoir model itself: instantaneous variation of outflow, slow-varying hypothesis, effect on heterogeneities on simulation results, impact of the constant average trip length assumption. The second category comprises the extension to multi-reservoir system and the proper modeling of flow exchanges: accurate reproduction of congestion propagation between reservoirs, accounting for local bottlenecks, design of the reservoir entry capacity, inflow merging scheme, outflow restriction modeling. The common methodological aspect between these two categories is the validation of the different modeling approaches through simulation comparisons.

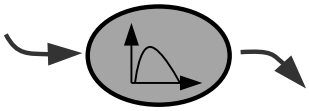
### 1.3.2 Thesis outline

The objective of this PhD thesis is to address the aforementioned research questions. Hence, the thesis outline is built upon the classification discussed previously. The manuscript is divided into three main parts: the first one deals with the internal dynamics of the single reservoir modeling, the second one investigates flow exchanges in multi-reservoir systems, and finally the third part presents the complete MFD-based simulation platform together with applications on real test cases and other extensions. The thesis outline is illustrated in Figure 1.17.

In the first part, Chapter 2 investigates analytical and numerical methods to solve the single reservoir model. In particular, a promising alternative modeling approach is investigated. This new approach, referred to as "trip-based" in contrast to the traditional one called "accumulation-based", received little attention in the literature. The physical properties of both models are compared, and it is shown that the trip-based framework provides a more reliable representation of traffic states by construction in undersaturated conditions. On the other hand, the limitations of the widely used accumulation-based model are clearly identified, and the inaccuracies due to the slow-varying hypothesis are quantified.

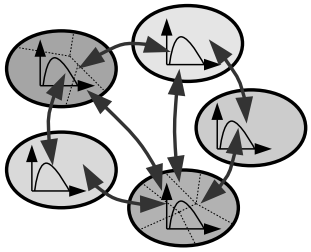
Chapter 3 presents a simulation comparison between the accumulation-based single reservoir model and a mesoscopic link-scale simulator on a simple Manhattan network. The

**PART I: Investigations on internal dynamics of a single reservoir**



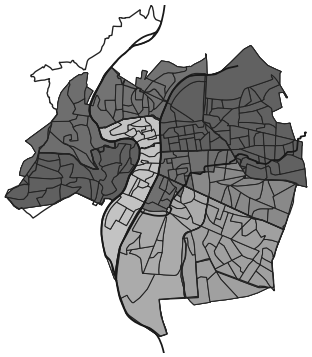
**Chapter 2**: Analytical and numerical comparisons of existing models for a single reservoir
**Chapter 3**: Impact of heterogeneous perimeter flow distributions on a single reservoir simulation

**PART II: Investigations on flow exchanges in multi-reservoir systems**



**Chapter 4**: Investigations on flow exchanges in multi-reservoir and multi-route systems
**Chapter 5**: Comparison between flow exchange modeling approaches and validation against microscopic simulation

**PART III: Developments and applications**



**Chapter 6**: Model implementation and first application to a real network

FIG. 1.17 – *Thesis outline*

latter link-scale model is intended to reproduce local bottlenecks and heterogeneities in the network due to local differences in the boundary conditions (entry demand and exit supply exogenous flow distributions). It is shown that these variations in perimeter flow settings, invisible to the single reservoir model because of aggregation, can impact the MFD shape and the average trip length traveled by the vehicles. The errors from the MFD-based model

in traffic state predictions are quantified, and correction methods such as the recalibration of the MFD and the trip length are proposed to reduce the errors.

In the second part, Chapter 4 investigates flow exchanges in multi-reservoir and multi-route systems. Based on theoretical considerations from well-known principles in traffic flow theory, general and consistent entrance functions are proposed, together with a unique inflow merge scheme that complies with flow and production constraints at the reservoir entry border. Moreover, a proper treatment of outflow is proposed to account for supply flow limitation at the reservoir exit border. It is notably shown that such a treatment implies drastic constraints on the outflow diverge scheme in case of several macroscopic routes crossing the reservoir. Our new multi-reservoir framework is compared to existing ones from the literature, and great discrepancies are observed on a simple test case. Another contribution of this study is the extension of the trip-based model to multi-reservoir systems and congested situations.

Chapter 5 is intended to validate the theoretical findings from the previous chapter, and also bring new insights about the calibration of the entrance function based on link-scale information from the microscopic network (e.g. link capacities). The methodology of the validation study is to compare traffic states simulated by a microscopic platform with the prediction from an aggregated multi-route modeling. The case of two main macroscopic trips is investigated: inflow and outflow share in congestion, flow and accumulation steady state after the network loading, evolution of flow and accumulation per macroscopic trip when congestion propagates. It is shown that the main findings of the previous chapter are corroborated by the simulation comparison. A new framework is also proposed to account for local bottlenecks at the reservoir border. It appears that this extension of the multi-route model is efficient to handle various demand and supply settings at the reservoir perimeter, such as the ones studied in Chapter 3.

Chapter 6 presents the architecture of the MFD simulation platform, together with the merge algorithm used. Several applications and extensions that result from various collaborations with PhD, post-docs and researchers are also included. This comprises the implementation of a DTA model, the extension of a search-for-parking module in the trip-based approach, and a study on the effects of perimeter control and routing on a simple test case. A simulation study on the Lyon (France) metropolitan area also illustrates the application and the potential of this MFD simulation platform. All these collaborations are part of the ERC project **MAGnUM** held by Prof. L. Leclercq (Multiscale and Multimodal Traffic Modeling Approach for Sustainable Management of Urban Mobility). This PhD thesis is one of the contributions of the project, as it connects traffic modeling with control applications at a large scale.

Most of the chapters in this manuscript are updated versions of papers published or submitted in peer-reviewed journals or conference proceedings. They all consists of original works first-authored by the PhD candidate.

### 1.3.3 Publication list

#### Peer-reviewed journal papers

- **Mariotte, G.** & Leclercq, L. (2019). Flow exchanges in multi-reservoir systems with spillbacks. *Transportation Research Part B: Methodological*, accepted

- **Mariotte, G.**, Leclercq, L. & Laval, J. A. (2017). Macroscopic urban models: Analytical and numerical investigations of existing models. *Transportation Research Part B: Methodological*, 101:245–267, doi:10.1016/j.trb.2017.04.002

- Leclercq, L., Sénécat, A. & **Mariotte, G.** (2017). Dynamic macroscopic simulation of on-street parking search: a trip-based approach. *Transportation Research Part B: Methodological*, 101:268–282, doi:10.1016/j.trb.2017.04.004

### Peer-reviewed conference proceedings

- **Mariotte, G.**, Paipuri, M. & Leclercq, L. (2019). Flow exchanges in multi-trip MFD-based systems: A validation study against microscopic simulation, In *Transportation Research Board 98th Annual Meeting*, Washington DC, USA

- **Mariotte, G.** & Leclercq, L. (2018). MFD-based simulation: Spillbacks in multi-reservoir networks, In *Transportation Research Board 97th Annual Meeting*, Washington DC, USA

- **Mariotte, G.**, Leclercq, L. & Laval, J. A. (2017). Dual expression of macroscopic urban models: Analytical and numerical investigations, In *Transportation Research Board 96th Annual Meeting*, Washington DC, USA

- Sénécat, A., Leclercq, L. & **Mariotte, G.** (2017). Macroscopic simulation of on-street parking search: a trip-based approach, In *Transportation Research Board 96th Annual Meeting*, Washington DC, USA

### International conference presentations

- **Mariotte, G.**, Paipuri, M. & Leclercq, L. (2018). Modeling local flow restriction at boundaries in multi-reservoir systems: An hybrid approach, *hEART 7th Symposium*, Athens, Greece

- **Mariotte, G.** & Leclercq, L. (2017). The MFD trip-based approach applied to multi-reservoir systems, *hEART 6th Symposium*, Haifa, Israel

- **Mariotte, G.** & Leclercq, L. (2016). Impact of the local demand-supply perimeter distribution and the internal reservoir trip length on MFD-based traffic simulation, *hEART 5th Symposium*, Delft, the Netherlands

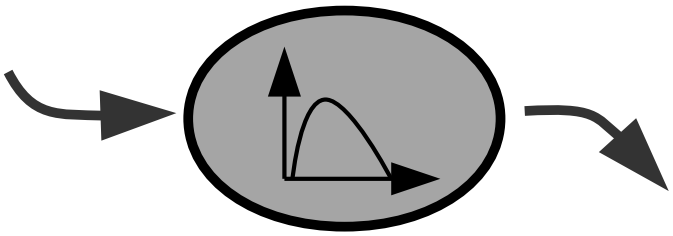
### Working papers

- **Mariotte, G.** & Leclercq, L. (2018). Heterogeneous perimeter flow distributions and MFD-based traffic simulation, submitted to *Transportmetrica B*, under review

- Ingole, D., **Mariotte, G.** & Leclercq, L. (2018). Macroscopic traffic dynamics with dual-perimeter control and external user equilibrium discipline, submitted to *IEEE Transactions on Control Systems Technology*

# Part I

# Investigations on internal dynamics of a single reservoir



# Introduction

This first part focuses on the description and the understanding of the internal traffic dynamics of a single reservoir. Because this model constitutes the building block of any multi-reservoir systems, the analysis of its dynamics will serve as a basis for the development of more complex modeling approaches in the second part.

In the present part, the investigations on internal dynamics are divided into two different studies. Chapter 2 introduces two existing modeling approaches based on the MFD, namely the accumulation-based and trip-based models. Theoretical and numerical comparisons between them allow to address the question of the description of the space dimension in MFD-based modeling. The validity domain of the widely used accumulation-based model is also discussed thanks to these comparisons. Then, chapter 3 proposes to challenge the estimation of the key elements in the accumulation-based modeling of a single reservoir, namely the MFD itself and the average trip length. Thanks to a comparison with link-scale simulation, this chapter aims at finding to which extend the MFD and the average trip length can be considered as intrinsic properties of the network.

# Contents

# 2. Analytical and numerical comparisons of existing models for a single reservoir

Large-scale network modeling using the Macroscopic Fundamental Diagram (MFD) is widely based on the single-reservoir model, where the variation of the accumulation of circulating vehicles in the reservoir equals inflow minus outflow. However, inconsistent lags for information propagation between boundaries may be observed with this single accumulation-based model. For example, outflow is reacting too fast when inflow varies rapidly, whereas this information should be carried by vehicles that are never driving faster than the free-flow speed. To overcome this limitation, a trip-based model has been recently proposed, but whose solution cannot be obtained analytically.

In this chapter we compare both models under piecewise linear MFD and a piecewise constant demand. These assumptions allow to establish the exact solution of the accumulation-based model, and continuous approximations of the trip-based model at any order using Taylor series. Moreover, a flexible event-based simulation framework is implemented to solve the latter model, making it a promising tool to account for heterogeneity in distance traveled. Thanks to these resolution schemes we are able to measure the inaccuracy of the accumulation-based approach when the demand varies rapidly, and propose a validity domain for this model. Other applications with different trip lengths and supply limitations are also discussed.

This chapter is an updated version of the paper:

> Mariotte, G., Leclercq, L. & Laval, J. A. (2017). Macroscopic urban models: Analytical and numerical investigations of existing models. *Transportation Research Part B: Methodological*, 101:245–267, doi:10.1016/j.trb.2017.04.002

## 2.1 Notations for this chapter

TAB. 2.1 – *Specific notations in this chapter*

| Notation | Definition [units] |
|---|---|
| $t_{\text{free}}$ | Reservoir free-flow travel time = $L/u$ where $u$ is the free-flow speed [s] |
| $\mu_c$ | Outflow-MFD capacity, maximum outflow = $P_c/L$ where $P_c$ is the production-MFD capacity [veh/s] |
| $P_i(n)$ | $i$th branch of the piecewise linear production-MFD $P(n)$ (linear function) [veh.m/s] |
| $w_i$ | Slope of the $i$th branch $P_i(n)$ [m/s] |
| $\eta_i$ | $n$-intercept of the $i$th branch $P_i(n)$ [veh] |
| $n_i$ | Accumulation value at the intercept of the $i$th and the $(i+1)$th branches of $P(n)$ [veh] |
| $t_i$ | Time at which there is a change in MFD branch [s] |
| $t_d$ | Time at which there is a change in demand value [s] |
| $n_{\text{ini}}$ | Initial accumulation at $t = 0$ [veh] |
| $T_{\text{ini}}$ | Initial travel time at $t = 0$ [s] |
| $\lambda_{\text{ini}}$ | Initial demand at $t = 0$ [veh/s] |
| $n_0$ | $n(t_0)$ = Accumulation at time $t_0$ [veh] |
| $T_0$ | $T(t_0)$ = Travel time at time $t_0$ [s] |
| $\lambda_0$ | Constant demand value, or $\lambda(t_0)$ = inflow at time $t_0$ [veh/s] |
| $\mu_0$ | Constant supply value, or $\mu(t_0)$ = outflow at time $t_0$ [veh/s] |
| $\delta t$ | Time step for the Euler-type finite difference resolution method [s] |
| $\Delta\lambda$ | Inflow gap in the evolution of the demand $\lambda(t)$ [veh/s] |
| $\Delta t$ | Duration of the inflow gap in the evolution of the demand $\lambda(t)$ [s] |
| $Q_\lambda$ | Demand variation rate = ratio of $\Delta\lambda$ over $\Delta t$ [veh/s$^2$] |
| $h_0$ | Height of the travel time drop in the evolution of $T(t)$ [s] |
| $d_0$ | Duration of the travel time drop in the evolution of $T(t)$ [s] |
| $N$ | Vehicle index [veh] |
| $\delta N$ | Vehicle fraction or increment (may be below 1) [veh] |
| $t_0 = t_N^{\text{in}}$ | Entering time of vehicle $N$ [s] |
| $t_1 = t_N^{\text{out}}$ | Exiting time of vehicle $N$ [s] |
| $t_{N+\delta N}^{\text{in}}$ | Entering time of vehicle $N + \delta N$ [s] |
| $t_{N+\delta N}^{\text{out}}$ | Exiting time of vehicle $N + \delta N$ [s] |
| $\delta t$ | Elapsed time between the exits of vehicles $N$ and $N + \delta N$ = $t_{N+\delta N}^{\text{out}} - t_N^{\text{out}}$ [s] |
| $n_0$, $T_0$, $\lambda_0$, $\mu_0$, $V_0$ | Respectively accumulation [veh], travel time [s], inflow [veh/s], outflow [veh/s] and mean speed [m/s] at time $t_0$ [veh] |
| $n_1$, $T_1$, $\lambda_1$, $\mu_1$, $V_1$ | Respectively accumulation [veh], travel time [s], inflow [veh/s], outflow [veh/s] and mean speed [m/s] at time $t_1$ [veh] |

## 2.2 Motivations

Large-scale network modeling based on the Macroscopic Fundamental Diagram (MFD) has advanced significantly in recent years. It appears as a viable option for congestion management applications such as perimeter control (see e.g. Haddad & Geroliminis, 2012, Aboudolas & Geroliminis, 2013, Ramezani *et al.*, 2015, Haddad, 2017) and for modeling large cities based on the multi-reservoir framework presented in Hajiahmadi *et al.* (2013b), Knoop & Hoogendoorn (2014), Yildirimoglu *et al.* (2015), Kouvelas *et al.* (2017). The theoretical foundations of all these approaches have been established in Daganzo (2007), Geroliminis & Daganzo (2007), where the dynamics of a single reservoir is represented by a conservation equation where outflow is determined by the MFD function. This approach will be referred to as the "accumulation-based MFD model". Though it has proved to be an attractive description of large urban areas, this approach still relies on strong hypotheses, which may have non negligible impacts. For example, Xue *et al.* (2016) show that simplifying an urban system by two reservoirs may lead to a different optimal perimeter-control strategy than when the system is described using classical traffic flow theory. Also, it has been shown in Leclercq *et al.* (2015) that the accumulation-based model suffers from significant numerical viscosity even when the time step is small. Outflow (respectively inflow) may then overreact to sudden demand surge (respectively supply drop) leading to inconsistent propagation of information between opposite perimeter boundaries. The simplest illustration is an empty reservoir and a demand that starts increasing. The outflow instantaneously also increases creating an immediate reaction that may be interpreted as information that propagates from one boundary to the other at an infinite speed.

Another crucial point concerns the representation of vehicle trip length within the reservoir. It is often assumed constant for all vehicles for the sake of simplicity (see e.g. Haddad & Geroliminis, 2012, Aboudolas & Geroliminis, 2013, Hajiahmadi *et al.*, 2013b), but this is not consistent with what is observed when the local dynamics are taken into account (Leclercq *et al.*, 2015). Trip lengths not only depend on the OD (Origin-Destination) matrix but also on the traffic conditions within the reservoir. Yildirimoglu & Geroliminis (2014) first highlight the error made in simulation when using the standard formulation of Daganzo (2007) in comparison with an improved description taking into account the variability of the trip length with respect to OD pairs. Such a description requires that the original single conservation equation that describes the behavior of the reservoir must be split into different vehicle classes, for example per OD pairs, with constant travel distances. The influence of the trip length for the accumulation-based MFD model will be thoroughly investigated in chapter 3. Leclercq *et al.* (2015) propose an alternative modeling framework to Yildirimoglu & Geroliminis (2014) to consider different trip lengths by defining macroscopic routes and jointly solving the related system of conservation equations. Based on an idea of Arnott (2013) and Fosgerau (2015), Daganzo & Lehe (2015) and then Lamotte & Geroliminis (2016) elaborate a simple and elegant reformulation of the single-reservoir dynamics to address the question of variable trip lengths. This reformulation will be further referred to as "trip-based MFD model" since the main idea is to guaranty that all vehicles cover their travel distance within the reservoir by adjusting their instantaneous speed to the current reservoir mean speed defined by the MFD.

In this chapter, we compare the accumulation-based and the trip-based MFD models by investigating their analytical and numerical solutions for a piecewise linear MFD and

piecewise constant inflows. The analytical solutions of the accumulation-based MFD model can be determined piece by piece while the trip-based MFD model require a continuous approximation based on the entering vehicle discretization. A major insight of our analytical developments is that the accumulation-based has no memory while the trip-based approach accounts for a reaction time. The comparison of the solutions on simple test cases will permit to further analyze the different model properties. In all cases, the trip-based approach appears superior as it better represents wave propagation and travel time evolution within the reservoir. On the other hand, an attempt to define a validity domain for the accumulation-based model application has been proposed. Numerical schemes are also derived for both approaches. Notably, we propose here an event-based solution method for the trip-based model that provides the exact solution when the inflow cumulative curve is kept as a step-function related to the entry of each individual vehicles. This solution method can be easily extended to account for different trip lengths for each driver, which is a significant improvement compared to the accumulation-based model. This means that the trip-based MFD model would be even easier to implement with multiple reservoirs.

The layout of this chapter is as follows: we first present the analytical and numerical resolution of the accumulation-based reservoir model. Then we introduce the numerical resolution of the trip-based model and compare its properties with the original model. The event-based numerical scheme for the trip-based approach is also presented there. Finally, we compare the results with the previous method and the accumulation-based model on three case studies: a typical peak-hour demand profile, a supply restriction at the reservoir exit, and a peak demand case with various trip lengths inside the reservoir.

## 2.3 Accumulation-based approach

We present in this section the seminal *accumulation-based* approach for describing the dynamics of a single-reservoir system.

### 2.3.1 The single-reservoir dynamics

The single-reservoir system dynamics has been extensively presented in Daganzo (2007), Geroliminis & Daganzo (2007). The basic principle is that an urban network can be described in an aggregate manner with a reservoir characterized by its accumulation $n(t)$, the number of circulating vehicles at time $t$ (in [veh]). The traffic state within the reservoir is given by a well-defined relationship between the travel production $P$ (in [veh.m/s]) and the accumulation $n(t)$. Note that the mean speed $V$ of travelers (in [m/s]) is given by $V(n(t)) = P(n(t))/n(t)$ at every time $t$ (Geroliminis & Daganzo, 2007). As in Lamotte & Geroliminis (2016), $P(n)$ will be referred to as the "production-MFD", and $V(n)$ as the "speed-MFD". According to Geroliminis & Daganzo (2007), the trip length $L$ is the same for all travelers and satisfies the following equation:

$$O(n) = \frac{P(n)}{L} \tag{2.1}$$

where $O(n)$ is the reservoir outflow (in [veh/s]). It may represent either trips ending in the reservoir or trips exiting the area. This equation can be viewed as the application of Little's

formula at the reservoir scale (Little, 1961). In this chapter, we first add some further but non restrictive assumptions in order to facilitate analytical calculations. The initial state of the reservoir is free-flow so that the outflow is not restricted by the supply at the network exits. The total demand $\lambda(t)$ for internal or passing-by trips is always below the network capacity and therefore not limited by the reservoir entrance or supply function. With these hypotheses the evolution of the accumulation $n(t)$ is given by:

$$\frac{dn}{dt} = \lambda(t) - O(n(t)) \tag{2.2}$$

Moreover we choose a piecewise linear shape for $P(n)$, in which each branch $P_i$ of $P(n)$ is defined by equation 2.3:

$$P_i(n) = w_i(n - \eta_i) \tag{2.3}$$

where $w_i$ is the slope of the branch $P_i$ (in [m/s]) and $\eta_i$ the $n$-intercept (in [veh]), see Figure 2.1(c). This choice is motivated by the analytical method in Leclercq *et al.* (2014). This assumption is not restrictive as any MFD can be approximated by such a family of linear envelope curves. Our analytical solution also considers that the demand $\lambda(t)$ is piecewise constant and thus the cumulative demand count curve is piecewise linear.

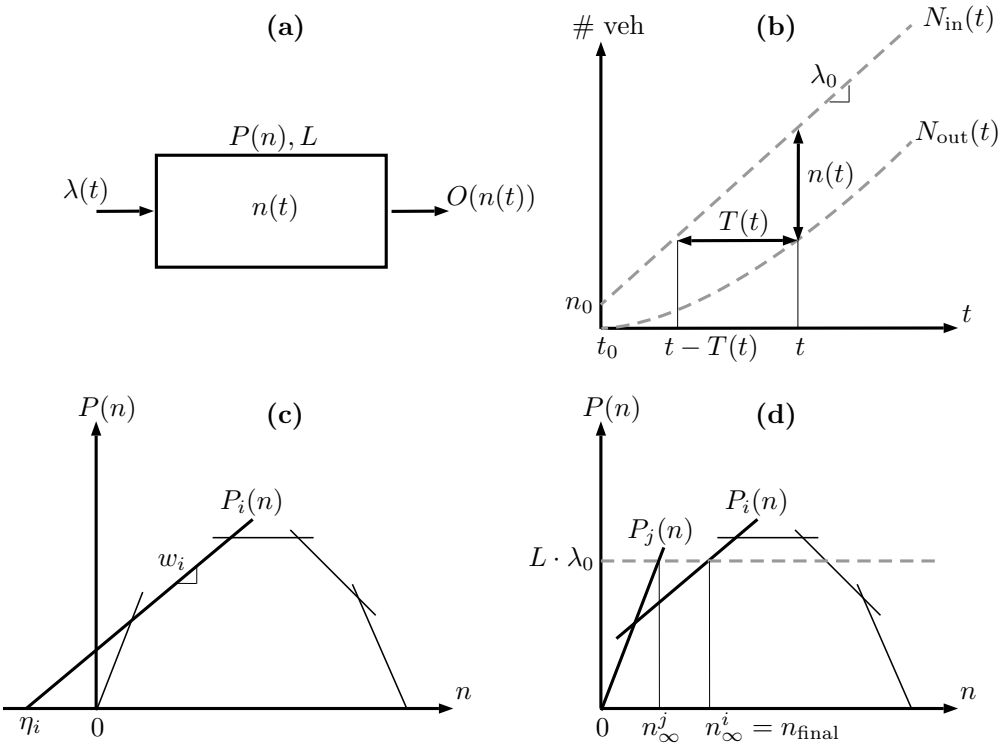



FIG. 2.1 – *(a) Schematic representation of the single-reservoir dynamics, (b) entering and exiting curves with the definition of $n(t)$ and $T(t)$, (c) piecewise linear production-MFD $P(n)$, and (d) example of two successive applications of our analytical formula with a given demand $\lambda_0$ (transition from the MFD branch $P_j$ to $P_i$)*

## 2.3.2 Analytical solution

Here we focus on the accumulation $n(t)$ and the *experienced* travel time $T(t)$, see Figure 2.1(b). While accumulation evolves on the same branch $P_i(n)$ of the production-MFD and the de-

mand remains constant, $\lambda(t) = \lambda_0$, the conservation equation 2.2 becomes:

$$\frac{dn}{dt} = \lambda_0 - \frac{w_i}{L} \cdot (n(t) - \eta_i) \tag{2.4}$$

which corresponds to a first-order linear differential equation with constant coefficients. With the initial condition $n(t_0) = n_0$ at a given time $t_0$, the solution of $n(t)$ is:

$$n(t) = (n_0 - n_\infty)\mathrm{e}^{-(t-t_0)/\tau} + n_\infty \tag{2.5}$$

where $\tau = L/w_i$ is the characteristic time of the system, and $n_\infty = \tau\lambda_0 + \eta_i$ is the asymptotic value of $n(t)$ when $t \to +\infty$.

In this study $T(t)$ is the travel time experienced by the vehicle that exits the reservoir at time $t$. It corresponds to the time spent by the vehicle in the reservoir. The travel time evolution results from the model dynamics and can be determined from the integration of the outflow function, see Figure 2.1(b).

The accumulation can be written as $n(t) = N_{\text{in}}(t) - N_{\text{out}}(t)$, where $N_{\text{in}}(t)$ and $N_{\text{out}}(t)$ are the cumulative count curves for entering and exiting vehicles respectively. Knowing that under FIFO $N_{\text{out}}(t)$ equals to $N_{\text{in}}(t - T(t))$ by definition, this gives directly $n(t) = \int_{t-T(t)}^{t} \lambda(s)ds = \lambda_0 T(t)$; see Figure 2.1(b). Therefore the travel time is directly obtained by dividing equation 2.5 by $\lambda_0$:

$$T(t) = (T_0 - T_\infty)\mathrm{e}^{-(t-t_0)/\tau} + T_\infty \tag{2.6}$$

where $T_0 = n_0/\lambda_0$ is the initial travel time at $t = t_0$ and $T_\infty = n_\infty/\lambda_0$ is the asymptotic value of $T(t)$ when $t \to +\infty$. Note that this equation is only valid if the demand remains constant during $T(t)$ at least, otherwise the general relation $n(t) = \int_{t-T(t)}^{t} \lambda(s)ds$ must be used. The application of such a formula is straightforward as $\lambda(t)$ is piecewise constant.

### 2.3.3 Step-by-step analytical resolution in the general case

Because our initial condition in accumulation might take any value, for a given problem implying more than one branch of the piecewise production-MFD and different levels of demand evolving with time, a complete solution can be obtained by successive applications of the previous formulas in equations 2.5 and 2.6. Note that our method can adapt to any demand variations since these are represented by step functions, and that this scheme does not require steady state to be reached before the next change in demand. Its principle is the following:

1. Initialize $t_0 = 0$, $n_0 = n_{\text{ini}}$ (the initial accumulation for the whole problem)
2. Save the current level of demand $\lambda_0$ and directly calculate the entering vehicle curve $N_{\text{in}}(t)$ until the next change in demand
3. Determine the current branch of the production-MFD $(w_i, \eta_i)$
4. Update $n(t)$ and $T(t)$ using equations 2.5 and 2.6 until the minimum $t_{\min}$ between the time $t_d$ of the next change in demand and the time $t_i$ at which there is a change in the

production-MFD branch. The latter can be determined by: $t_i = \tau \cdot \ln\left(\frac{n_0 - n_\infty}{n_i - n_\infty}\right) + t_0$, where $n_i$ is the accumulation at the branch change defined in the production-MFD. Note that just after a change in demand the expression of $T(t)$ must be calculated using the general relationship $n(t) = \int_{t-T(t)}^{t} \lambda(s)ds$ as mentioned before.

5. Update the initial condition: $t_0 = t_{\min}$, $n_0 = n(t_{\min})$. If $t_{\min} = t_i$, a change in MFD branch happens, go to 3. Else if $t_{\min} = t_d$, a change in demand happens, go to 2.

It is worth mentioning that during the warm-up period, i.e. when $N_{\text{out}}(t) < n_{\text{ini}}$ at the beginning of the simulation, we need additional hypotheses to define properly the travel time—in particular we need to know the previous demand level $\lambda_{\text{ini}}$ at $t = 0$, see point 4 of the process. To this end we suppose that the system is in a steady state before $t = 0$ characterized by the initial accumulation $n_{\text{ini}}$. For $t < 0$ we have then $n_{\text{ini}} = \lambda_{\text{ini}} T_{\text{ini}}$ where $T_{\text{ini}}$ is the initial travel time equal to $L/V(n_{\text{ini}})$. This yields the demand level that should be defined for $t < 0$: $\lambda_{\text{ini}} = n_{\text{ini}} V(n_{\text{ini}})/L$.

With our step-by-step process any part of the global solution can be obtained at any precision. This resolution scheme provides by construction the exact solution of the accumulation-based model under our hypotheses.

### 2.3.4 Application to a simple demand step case

Figures 2.2(a) and (b) show a simple case of a one-staircase demand profile and a six-branch piecewise MFD, describing a 5-km network. The trip length is set to $L$ = 2.5 km, the free-flow speed to $u$ = 15 m/s, and the maximum production to $P_c$ = 3000 veh.m/s. The initial number of vehicles $n_{\text{ini}}$ has been chosen to correspond to the initial value of demand, so that the system is in steady state at $t$ = 0. These numerical values as well as the chosen MFD shape will be the same for all the numerical applications in this chapter. Although the instantaneous increase in demand at time $t_{d1}$ = 700 s is rather unrealistic, this simple profile reveals a first drawback of the accumulation-based model: the analytical solution for the travel time exhibits a drop of $T(t)$ right when the demand increases. This phenomenon has also been noticed by Lamotte & Geroliminis (2016). The accumulation-based formulation in equation 2.2 does not include any features that bound or more precisely delay outflow variations when demand changes are observed. Hence an increase of the inflow leads to an instantaneous modification (increase) of the outflow, see Figure 2.2(e). It means that information (changes of the flow between the reservoir entry and exit perimeter) is traveling between boundaries at a much higher speed than the maximal vehicle speed, and this information speed is even infinite at the time when the inflow changes. Travel time has to adapt to this immediate variation of the outflow to guaranty vehicle conservation. That means some vehicles have to drive faster to sustain the outflow rate, which explains the decrease of the travel time function at time $t_{d1}$.

Figure 2.3 presents the evolution of accumulation and travel time for different initial conditions $n_{\text{ini}}$ and different levels of demand $\lambda_0$. Here the system is in steady state defined by $n_{\text{ini}}$ for $t < 0$, and directly enters in transient state at $t = 0$ because $\lambda_0 \neq \lambda_{\text{ini}}$. The same non-intuitive drops in travel time are encountered. They are all the result of too fast propagation of the information between the reservoir boundaries. Figure 2.2(f) illustrates this problem by exhibiting wave propagation within the reservoir. Waves are here defined similarly as

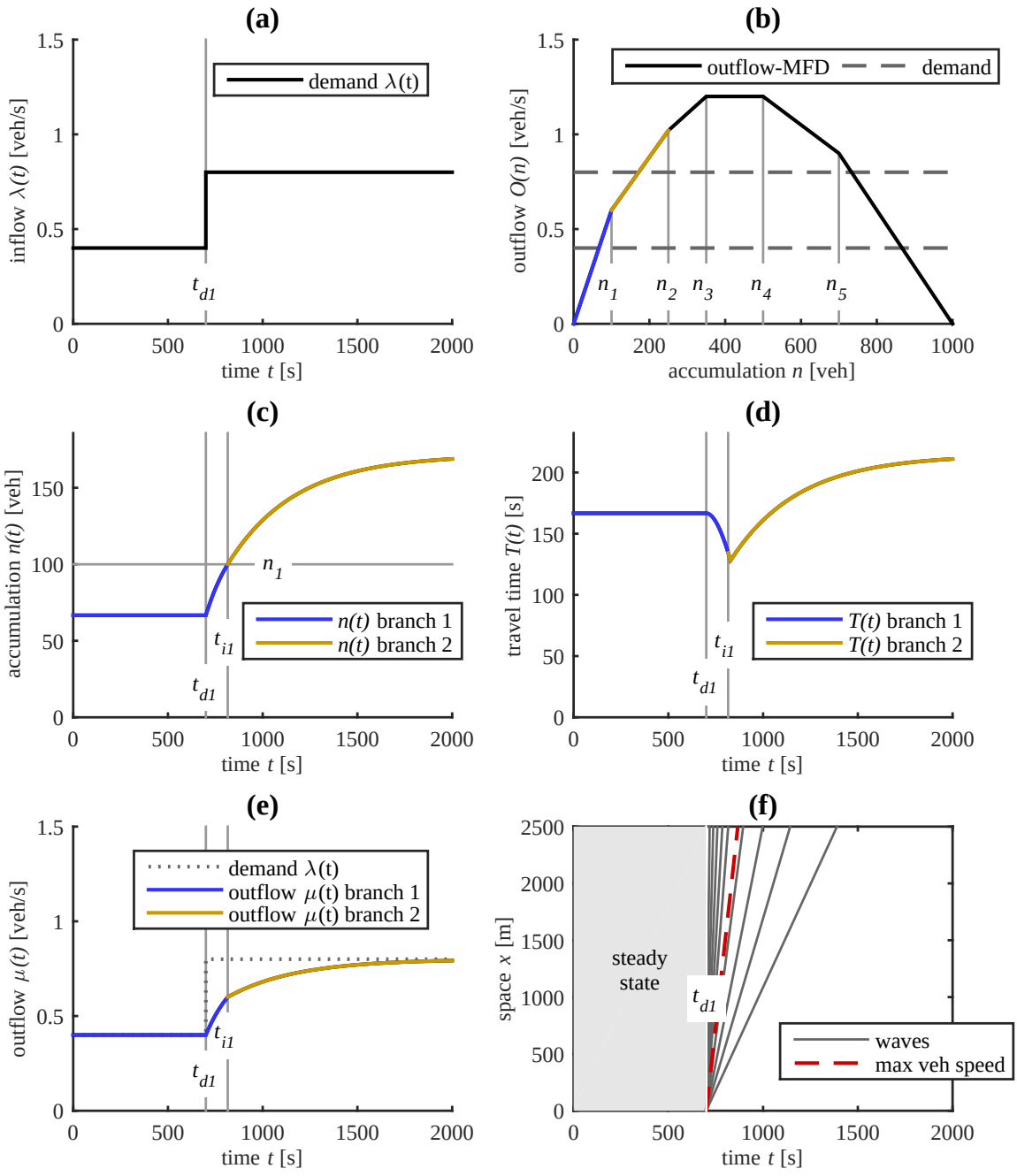



FIG. 2.2 – *(a) One-staircase demand profile $\lambda(t)$, (b) outflow-MFD with the particular accumulation values $n_{i,\ 1\leq i\leq 5}$ indicating the branch intersections and with the levels of demand, (c) evolution of accumulation $n(t)$ and (d) evolution of travel time $T(t)$ where the time $t_{d1}$ corresponds to a change in demand, and $t_{i1}$ to a change in MFD branch. (e) Evolution of outflow $\mu(t)$ and (f) representation of wave propagation in the reservoir, a wave being defined as a curve connecting an inflow value at upstream to the same outflow value at downstream. The space axis represents the distance to travel along a trip. Here the maximum vehicle speed is the free-flow speed*

kinematic waves in link traffic flow theory, i.e. as locus in the $(x, t)$ plane where identical

flow (or speed) values may be observed. Such kinematic waves represent the propagation of information in traffic stream as a change in flow value at one particular location will then propagate following such curves. Classically information is propagating forward in free-flow (positive kinematic wave speed) and backward in congestion. For reservoir models, there is no explicit representation of the $x$-axis but we can define wave by analogy to the link traffic flow theory by connecting identical flow values observed as inflow and outflow, i.e. between the two boundaries of the reservoir that define the starting and ending locations of trips. This permits to investigate how information is propagating inside the reservoir. Figure 2.2(f) clearly shows that some waves are faster than the vehicle speed (and even tends towards infinity when time tends towards $t = t_{d1}$), because the outflow starts being modified before the first vehicle experiencing the demand surge joins the exit. This does not mean that some vehicles experience zero travel time because we are speaking here about the propagation of information. More precisely, only an infinitesimal small quantity of vehicles will indeed carry such a super-fast (and even infinite) information propagation. However, these inconsistent waves and output modifications induce the drop in travel times. This is why it is currently admitted in link traffic flow theory that waves should not travel faster than vehicle speeds. We believe that such a requirement is also important for reservoir models and we will show that the trip-based model in section 2.4 overcomes this problem.

### 2.3.5 Congestion case and full numerical resolution

In congestion when the outflow is restricted to a given supply value $\mu_0$, equation 2.2 transforms to $dn/dt = I(n(t)) - \mu_0$, where $I(n(t))$ is the reservoir inflow. In this case $I(n(t))$ may indeed depend on $n(t)$ because for high values of accumulation the demand is limited by the reservoir entrance function. The expression of $I(n(t))$ is assumed to be equal to the reservoir capacity in free-flow and to correspond to the congested branch of the MFD in congestion (Geroliminis & Daganzo, 2007). This means this differential equation will lead to a quite similar solution as presented in equation 2.5 in switching the role of inflow and outflow. This shows that our approach is relevant to examine congested situations too, though not presented here. With this shape of $I(n)$, intermediate situations where both inflow and outflow depend on $n(t)$ lead automatically to equilibrium, as $I(n) = O(n)$ on the congested part. Figure 2.3(e) illustrates the reservoir dynamics in the ($\lambda_0$, $n$, $t$) space. We clearly see that the free-flow part of the outflow-MFD is attractive, whereas the congested part is repulsive: once an initial condition ($\lambda_0$, $n_{\text{ini}}$) is below the congested branches the solution will automatically converge to the corresponding point on the free-flow branches. Our choice here for $I(n)$ makes the solutions with ($\lambda_0$, $n_{\text{ini}}$) above the congested branches stable and not converging to gridlock, because inflow is instantaneously restricted by $I(n)$ equal to outflow $O(n)$ in congestion.

In order to ease the implementation of the accumulation-based model, a full numerical resolution method can be used. To illustrate, we used a first-order Euler-type finite difference method, where the evolution of accumulation is obtained by the following iterative process:

$$n(t + \delta t) = n(t) + \delta t \cdot \Big( \min[\lambda(t), I(n(t))] - \min[\mu_0, O(n(t))] \Big) \tag{2.7}$$

where $\delta t$ is the time step, and $\mu_0$ (in [veh/s]) is a given supply value at the reservoir exit, which may eventually depend on time. Note that if the inflow is restricted, i.e. $\min[\lambda(t), I(n(t))] =$

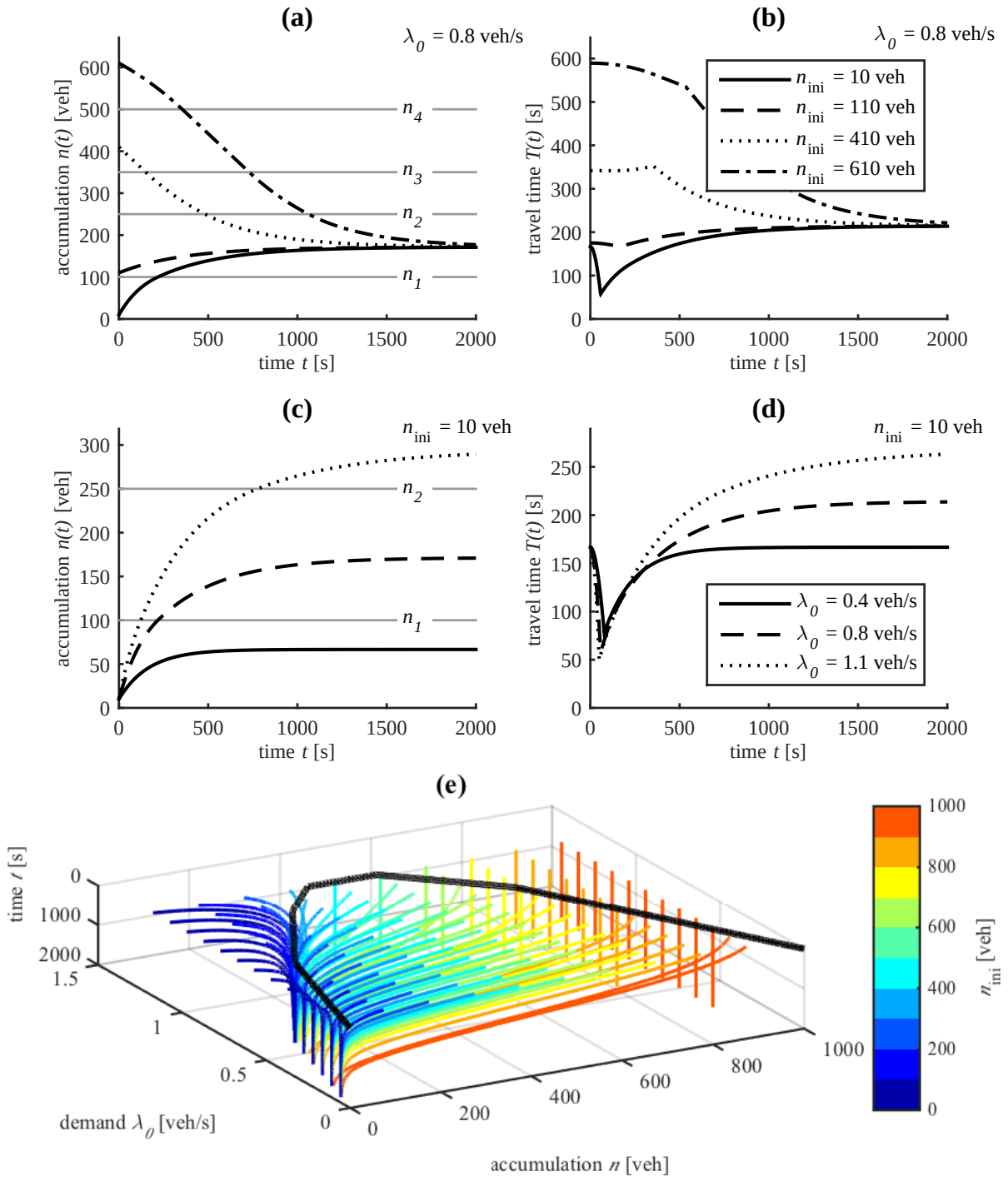



FIG. 2.3 – *(a) Different initial conditions $n_{\text{ini}}$ with demand $\lambda_0 = 0.8$ veh/s, evolution of accumulation $n(t)$ and (b) travel time $T(t)$. (c) Different demand levels $\lambda_0$ with initial condition $n_{\text{ini}} = 10$ veh, evolution of accumulation $n(t)$ and (d) travel time $T(t)$. (e) Phase diagram in the ($\lambda_0$, $n$, $t$) space, one line corresponds to the solution with initial conditions ($\lambda_0$, $n_{\text{ini}}$), the black line being the outflow-MFD. Initial conditions above the congested part of the MFD lead to steady state solutions with our choice for the reservoir entrance function $I(n)$*

$I(n(t))$, the vehicles waiting at the reservoir entry must be taken into account in case of demand evolution. In terms of computational time, the semi-analytical method is about ten

times faster than the full numerical one if only a few branches are needed to define the MFD and if the demand scenario has only a few different levels. But the fully numerical method becomes the fastest one in case of any given shape for the MFD and continuous variations of demand. In that case the semi-analytical process will need numerous calculation loops to approximate both the shape of the MFD and the demand evolution. The influence of the number of MFD branches and demand steps for the semi-analytical method, and the time step choice for the full numerical resolution is analyzed in A.2. Results show that 8 MFD branches and a 40-step demand profile are sufficient to get accurate approximations of the exact solution. For the full numerical method, a time step equal to 10 % of the reservoir free-flow travel time is small enough to get acceptable numerical errors. The exact solution used to make these comparisons has been obtained in *Mathematica*®, as further discussed in Laval *et al.* (2018).

### 2.3.6 Validity domain of the accumulation-based model

Based on the first limitations highlighted in this section for the accumulation-based model, a validity domain can be established regarding the travel time drop phenomenon. Our hypothesis is that this undesirable artifact is due to the demand curve varying too fast. Daganzo (2007) already conjectured that the accumulation-based model is suitable for slow varying boundary conditions only. What we propose here is to characterize the demand change rate threshold below which the model can be considered accurate. For this purpose we perform numerous simulations with a logistic "S-shaped" demand defined by two parameters: the demand gap $\Delta\lambda$ and the gap duration $\Delta t$. The demand variation rate is described by $Q_\lambda = \Delta\lambda/\Delta t$. We observe then the evolution of the travel time drop height $h_0$ and its duration $d_0$ as presented in Figures 2.4(a) and (b). If we compare the blue curve ($\Delta t$ = 0.2 s) with the yellow one ($\Delta t$ = 800 s), we see that the faster the demand varies, the higher the travel time drop $h_0$ is, but the shorter its duration $d_0$ is. Therefore we could expect $h_0$ and $d_0$ evolving in opposite directions as $Q_\lambda$ increases. That is exactly what Figures 2.4(c) and (d) show, where $h_0$ and $d_0$ are plotted in the $(\Delta t, \Delta\lambda)$ plane. White lines represent the iso-$Q_\lambda$ curves. Both effects of the drop height and drop duration can be visualized in Figure 2.4(e) where the product $h_0 \cdot d_0$ is plotted.

These graphs point out that the relationship between the travel time drop phenomenon and the demand variation rate is not obvious, and only considering the value of $Q_\lambda$ to define an acceptable validity domain of the model may be too restrictive. Considering the global effect, i.e. the product $h_0 \cdot d_0$, it seems that the demand gap $\Delta\lambda$ alone is relevant for the characterization of the validity domain. For example we could claim that a demand increase below 0.1 veh/s ensures the accumulation-based model accuracy in travel time description, whatever the time it takes. However, using only the indicator $h_0 \cdot d_0$ makes these two situations equivalent: $h_0$ = 60 s during $d_0$ = 1000 s, and $h_0$ = 6 s during $d_0$ = 10,000 s, which may not be appropriate for cases when very small differences in travel time are considered negligible. So if we suppose the drop height is more prejudicial than its duration, only the indicator $h_0$ should be used. In that case, the validity domain would be $\Delta\lambda < 0.1$ veh/s *or* $Q_\lambda < 10^{-4}$ veh/s$^2$. Note that these values are completely dependent on our numerical application. An attempt to provide a validity domain applicable for any single reservoir problem is presented in A.2.

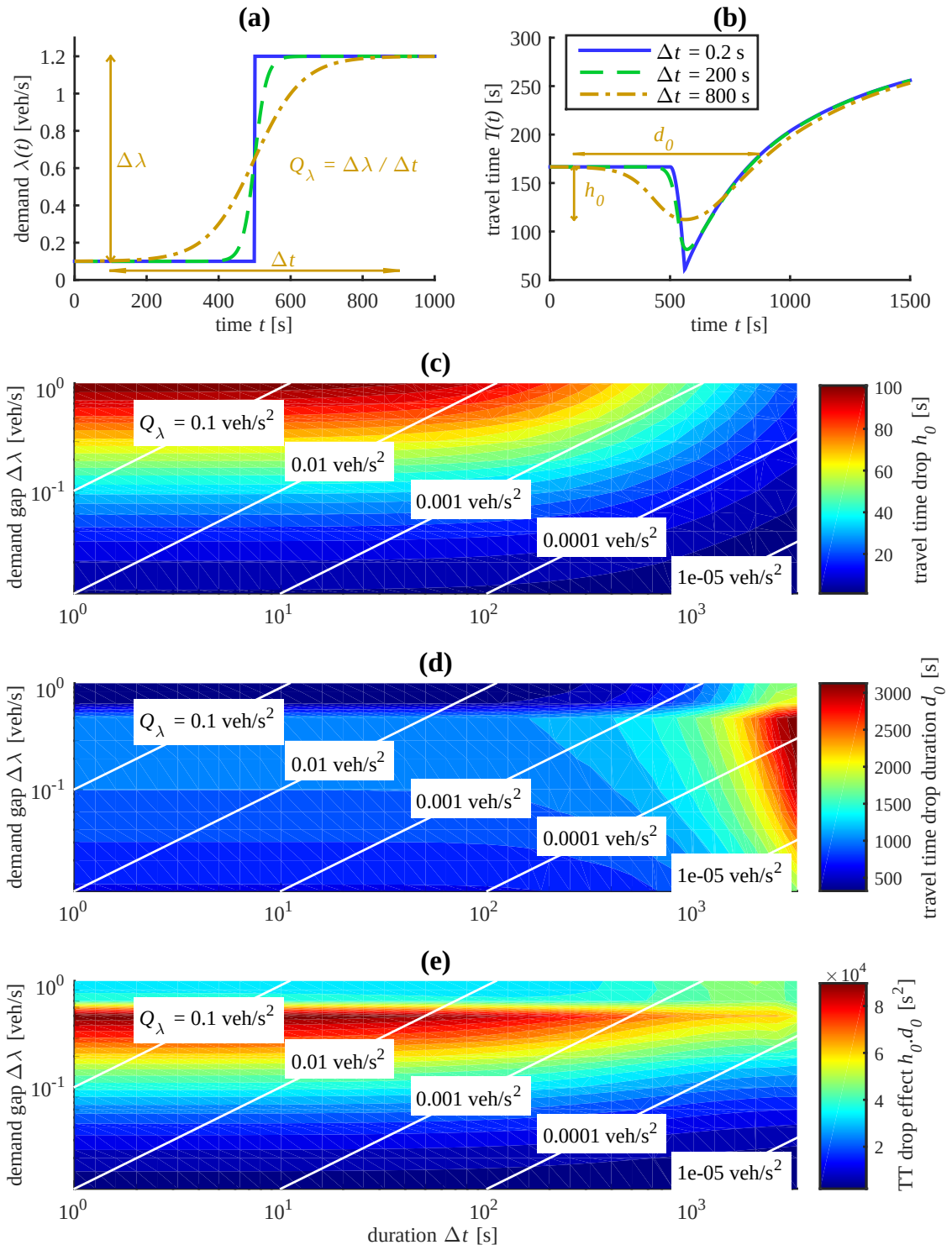



FIG. 2.4 – *Influence of the demand variation rate on the accumulation-based model accuracy for travel time. This accuracy is defined regarding the size of the travel time drop artifact. (a) Demand profiles with different variation rates, (b) resulting travel time profiles where two inaccuracy factors are measured: the drop height $h_0$ and the drop duration $d_0$, (c) drop height value in the $(\Delta t, \Delta\lambda)$ plane, (d) drop duration value, and (e) the product height-duration value in the same plane*

## 2.4 Trip-based approach

### 2.4.1 Theoretical basis

The general principle of the trip-based approach is to derive the inflow and outflow curves noting that the travel distance $L$ by a driver entering at time $t - T(t)$ should satisfy:

$$L = \int_{t-T(t)}^{t} V(n(s))ds \tag{2.8}$$

This definition is mentioned in the 11$^{\text{th}}$ footnote of Arnott (2013) first, and used in the studies of Fosgerau (2015) and Lamotte & Geroliminis (2016) then. Daganzo & Lehe (2015) also made use of such an approach, but without explicitly mentioning equation 2.8. Note that equation 2.8 also resort on the notion of MFD to define the common speed of vehicles within the reservoir at time $t$. Note also that this formulation focuses on the distance $L$ traveled by the vehicle entering at $t - T(t)$ and exiting at $T(t)$, which is defined as the integration of the reservoir speed during the vehicle traveling period. The evolution of the accumulation $n(t)$ is then only defined in an implicit way, i.e. it results from vehicles entering and exiting the reservoir. With this formulation of the reservoir system, it appears clearly that the definition of the individual trip length is no longer a problem as it becomes part of the input parameters. In this section, we will first consider that $L$ is the same for all travelers. This assumption allows us to develop a proper resolution scheme, and ensures the validity of the comparisons we will make with the accumulation-based model. This hypothesis will be relaxed in section 2.4.4 when presenting our event-based numerical resolution method.

To see the fundamental difference between the trip- and the accumulation-based models, we rewrite the definition of $L$ for the latter. If we use the formulation of the travel time based on outflow, namely $n(t-T(t)) = N_{\text{out}}(t) - N_{\text{out}}(t-T(t)) = \int_{t-T(t)}^{t} O(n(s))ds$, and combine it with equation 2.1 we find: $n(t-T(t)) = \int_{t-T(t)}^{t} \frac{P(n(s))}{L} ds$. As a consequence, we get the following expression for $L$ in the accumulation-based model framework:

$$L = \frac{1}{n(t-T(t))} \int_{t-T(t)}^{t} P(n(s))ds \qquad \text{(accumulation-based)} \tag{2.9}$$

This has to be compared with equation 2.8 for the trip-based model, in which $V(n(s))$ is replaced by $P(n(s))/n(s)$:

$$L = \int_{t-T(t)}^{t} \frac{P(n(s))}{n(s)} ds \qquad \text{(trip-based)} \tag{2.10}$$

It becomes apparent that the accumulation-based model considers the accumulation $n(s)$ ($s \in [t-T(t), t]$) constant and equal to $n(t-T(t))$ during the travel duration $[t-T(t), t]$ of the vehicle exiting at $t$ for the trip length calculation. On the other hand, the trip-based model takes the accumulation evolution into account during the same period, which is more realistic. In particular, this comparison shows that the two approaches are equivalent in steady state, or when the accumulation varies slowly which is consistent with slow demand variations. In other cases, e.g. as in peak-hour periods, we expect the accumulation-based model to be inaccurate during transition phases.

While being the first to introduce the trip-based model formulation, Arnott (2013) mentioned that its analytical resolution is almost impossible considering the current advances in mathematics. So, he adopted the accumulation-based framework in the following of his chapter. Fosgerau (2015) proposed some analytical developments of equation 2.8 but defined the departure time rate as a function of the trip length (two users sharing the same trip distance have necessarily the same departure time), which seems too restrictive for us. Daganzo & Lehe (2015) and then Lamotte & Geroliminis (2016) presented an interesting resolution method for the trip-based model that consists in assigning to each user a trip length and let him/her travels at the global mean speed $V(n(t))$ until he/she completes his/her trip and exits the reservoir. As the departure rate is considered as an input of the simulation, this method provides the arrival rate, i.e. the outflow of the system. Here, we propose an elegant way to get numerical solutions from the trip length equation 2.8 while controlling the accuracy of the approximation by using Taylor developments. This confirms that the methods originally proposed by Daganzo & Lehe (2015) and Lamotte & Geroliminis (2016) provide accurate results when discretizing the entry flow in small vehicle fractions. Our numerical approximation is presented in the following sections 2.4.2 and 2.4.3. Finally, it appears that Daganzo & Lehe (2015) are using a time step to perform the simulation. We will show in section 2.4.4 that like in Lamotte & Geroliminis (2016) we can get rid of time steps to achieve a full event-based simulation process when discretizing the entry flow in units of vehicles.

### 2.4.2 Numerical approximation

In a first time, we solve the problem under free-flow conditions. The congested case will be studied latter. In a free-flow situation, i.e. when inflow and outflow are not restricted, the behavior of the system is governed by the demand and solving the problem consists in determining the outflow $\mu(t)$. Here the basic idea is to use equation 2.8 to express the latter. From the definition of the experienced travel time $T(t)$ we know that $N_{\text{out}}(t) = N_{\text{in}}(t - T(t))$ for the cumulative count curves $N_{\text{in}}(t)$ and $N_{\text{out}}(t)$, see Figure 2.1(b). The derivative of $N_{\text{out}}(t)$ with respect to time leads to: $\mu(t) = (1 - dT/dt)\lambda(t - T(t))$. On the other hand, as $L$ is considered constant here, the time derivative of equation 2.8 is: $0 = V(n(t)) - (1 - dT/dt)V(n(t - T(t)))$. Combining these two relationships gives an explicit expression of $\mu(t)$:

$$\mu(t) = \lambda(t - T(t)) \cdot \frac{V(n(t))}{V\big(n(t - T(t))\big)} \tag{2.11}$$

This result has already been mentioned by Arnott (2013) as an equivalent problem to equation 2.8. Introducing this expression into the state equation 2.2 still remains an intractable mathematical problem because it corresponds to a differential equation with endogenous delay. However this expression can be transformed into an efficient numerical approximation by discretizing vehicles into fractions $\delta N$.

Let us assume that all the reservoir states are known until a given time $t_1$: the demand $\lambda(t)$, the outflow $\mu(t)$, the accumulation $n(t)$ and the travel time $T(t)$. We then want to solve the reservoir dynamics for $t > t_1$. The vehicle exiting at $t_1$ is denoted by the continuous index $N$, its entering and exiting times are written $t_N^{\text{in}} = t_0$ and $t_N^{\text{out}} = t_1$ respectively. The following vehicle is denoted by the index $N + \delta N$ and its entering and exiting times are $t_{N+\delta N}^{\text{in}}$ and $t_{N+\delta N}^{\text{out}}$. Physically speaking $\delta N$ should be equal to 1 and represents one vehicle.

However, we deal here with a continuous model, so we can fix $\delta N$ to very small values to investigate numerical accuracy. Since we are in free-flow conditions the entering vehicle time is determined by the demand, thus we have also $t^{\text{in}}_{N+\delta N} = t^{\text{in}}_N + \delta N/\lambda(t^{\text{in}}_N)$. Solving the problem for this vehicle is to determine its exit time $t^{\text{out}}_{N+\delta N} = t_1 + \delta t$, that is to say, to determine the time increment $\delta t$ after the last exit, see Figure 2.5(a). With these notations equation 2.11 can be approximated by:

$$\frac{\delta N}{\delta t} = \lambda(t^{\text{in}}_N) \cdot \frac{V(n(t^{\text{out}}_N))}{V(n(t^{\text{in}}_N))} \tag{2.12}$$

which gives directly an estimation for $\delta t$. Actually this corresponds to a first-order approximation on the exiting vehicle curve as the slope $\mu(t)$ is estimated by $\delta N/\delta t$. However we have found an interesting way to perform the calculation of $\delta t$ at higher orders. This method is detailed in A.1, results show that in fact the estimation of $\delta t$ at first order is largely sufficient if $\delta N$ is small enough (around 0.1 veh).

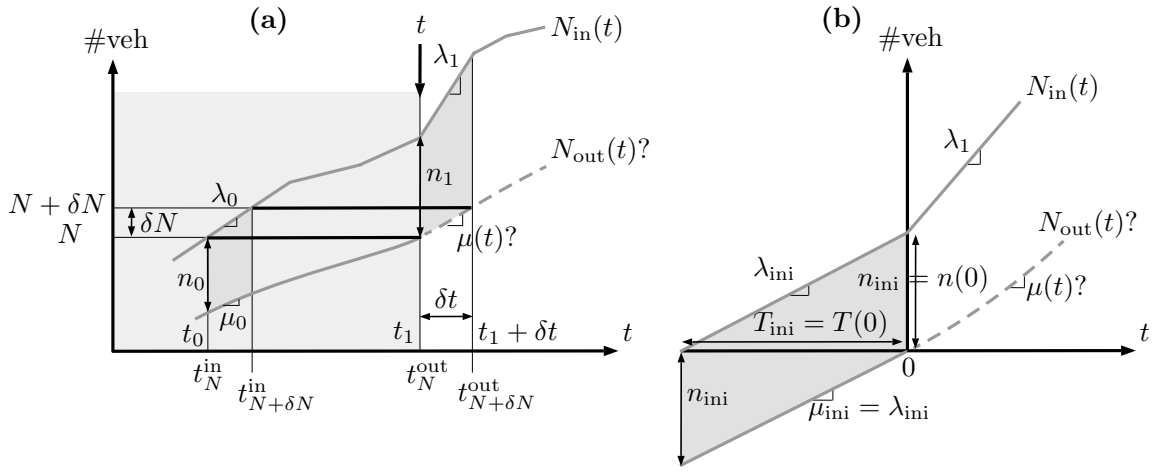



FIG. 2.5 – *(a) Representation of our resolution scheme on the cumulative count curves, the system evolution is supposed known until $t_1 = t^{\text{out}}_N$, (b) explanation for the first step of the resolution process, the system is supposed to be in a steady state before $t = 0$ determined by the choice of initial accumulation $n_{\text{ini}}$*

### 2.4.3 Step-by-step numerical resolution in the general case

We present now the complete algorithm to solve any problem that complies with our study hypotheses stated in section 2.3. It is based on the method previously described and summarized in Figure 2.5(a). Given a choice of a vehicle increment $\delta N$, we estimate the exiting vehicle curve by calculating the time step $\delta t$ after which the next vehicle $N + \delta N$ will exit. Since all the travelers have the same trip length $L$, this method is consistent with the *First-In, First-Out* (FIFO) rule which ensures the validity of this process. Note that unlike the accumulation-based model which has no memory, solving the system dynamics at any time $t_1$ for the trip-based model requires the knowledge of the past of the system. This may be highlighted as another fundamental difference between the two modeling approaches. In practice, this will be a problem for the simulation initialization, when the first vehicles start to exit the reservoir without knowing their entering times. This issue is tackled by simply

considering that the system is in a steady state completely defined by the choice of the initial accumulation $n_{\text{ini}}$ for $t < 0$. This is illustrated in Figure 2.5(b), the demand in the steady state before $t = 0$ is given by $\lambda_{\text{ini}} = n_{\text{ini}} V(n_{\text{ini}})/L$. In fact, the same consideration was also observed with the accumulation-based model to define properly the travel time evolution during the warm-up period. Hence for the simulation initialization the two modeling approaches are consistent with each other.

The numerical resolution process of the trip-based model is then summarized below:

1. Choose a vehicle increment $\delta N$
2. Initialize $t_1 = 0$, $n_1 = n_{\text{ini}}$, $V_1 = V(n_1)$, $T_1 = L/V_1$ and for the past of the system $t_0 = -T_1$, $n_0 = n_1$, $V_0 = V_1$, $\lambda_0 = n_0 V_0/L$, $\mu_0 = \lambda_0$
3. Save the current level of demand $\lambda_1$
4. if $t_0 > 0$, update at $t = t_0$: $n_0 = n(t_0)$, $V_0 = V(n_0)$, $\lambda_0 = \lambda(t_0)$, $\mu_0 = \mu(t_0)$
5. Determine the current branch of the production-MFD $(w_i, \eta_i)$
6. Update the mean speed $V_1 = V(n_1)$
7. Calculate the next time step $\delta t$ at the chosen order, using equation A.6 for order 1 or equation A.7 for order 2
8. Update at $t = t_1$: $t_1 = t_1 + \delta t$, $N_{\text{in}} = N_{\text{in}} + \lambda_1 \delta t$, $N_{\text{out}} = N_{\text{out}} + \delta N$, $\mu = \delta N/\delta t$, $n_1 = N_{\text{in}} - N_{\text{out}}$, $t_0 = t_0 + \delta N/\lambda_0$, $T_1 = t_1 - t_0$
9. If $t_1$ has not reached the next change in demand yet, go to 4. Else go to 3.

The application of this process is illustrated on the simple one-staircase demand profile that we used in section 2.3.4. The calculation of $\delta t$ is computed at order 1 with $\delta N = 0.1$ veh. Figures 2.6(a) and 2.7(b) show the accumulation and travel time evolution in solid line, the dashed lines represent the previous solutions obtained with the accumulation-based model. This example immediately points out the benefit of the trip-based approach, as the results for the travel time are now consistent. In this case, the absence of drop in travel time may be explained by the fact that the trip-based model keeps the memory of the past of the system, and thus provides more reliable results when the boundary conditions are changing. Moreover we notice that this model is more reactive than the accumulation-based one, the latter suffers indeed from intrinsic viscosity phenomenon due to the form of equation 2.4. Figure 2.6(d) shows the wave pattern associated to the trip-based model. This pattern has been constructed similarly as the one for the accumulation-based model in Figure 2.2(f). It appears that now no wave is traveling faster than the vehicles, and thus that travel times are properly reproduced by the trip-based approach. Finally, it should be mentioned that the trip-based model may exhibit local inconsistencies with respect to the outflow curve, see Figure 2.6(c). When the demand surge is observed at time $t_{d1}$ the outflow is constant as expected, but then slightly decreases before the surge in outflow is observed. The reason is that the trip-based model, while providing a much better description of vehicle trips, still resorts to an homogeneous speed for all the vehicles within the reservoir. Therefore, even the vehicles that enter the reservoir before the demand surge are slow down by the increase of accumulation resulting from this event.

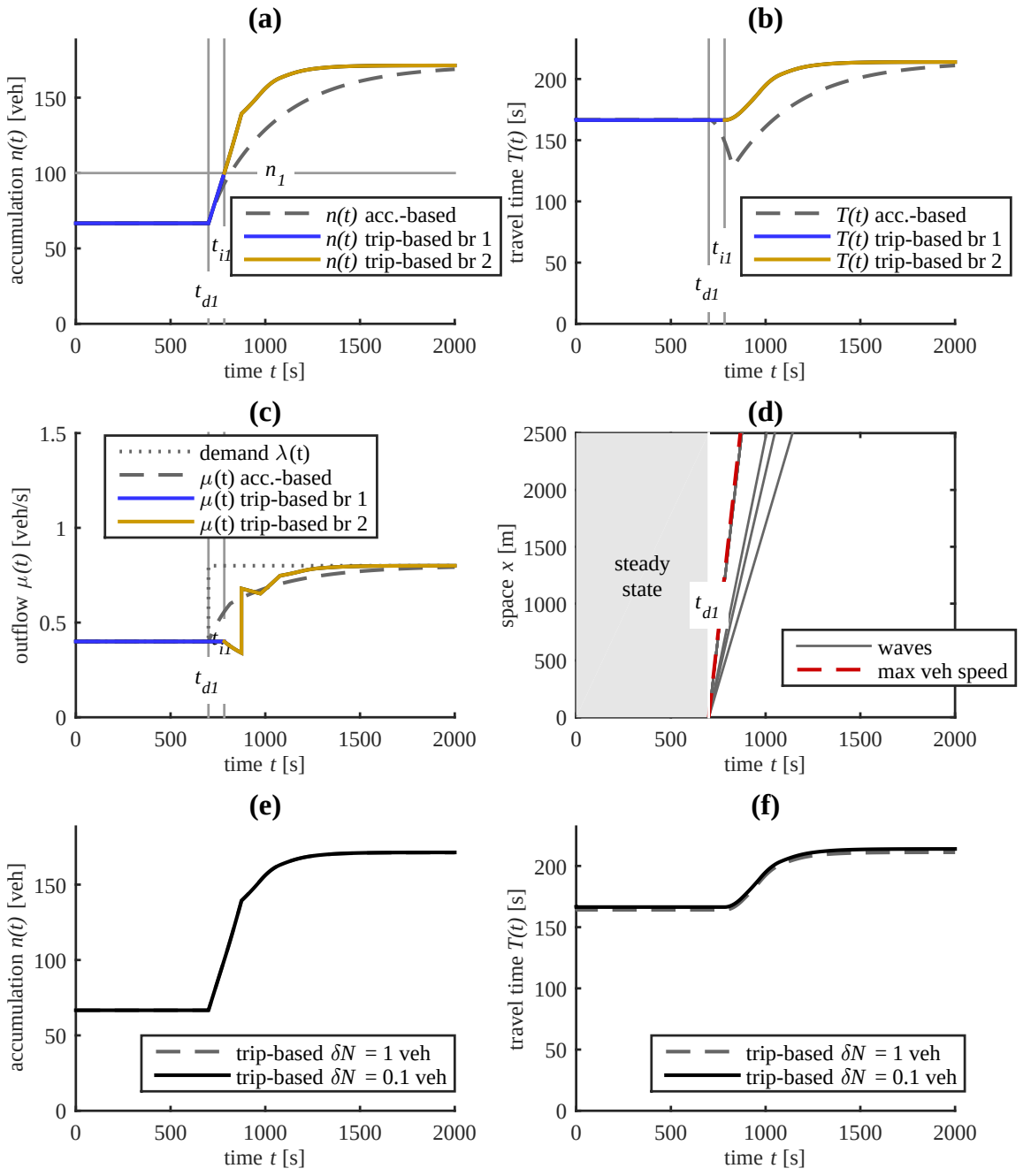



FIG. 2.6 – *Comparison between accumulation-based and trip-based models, the latter is performed with $\delta N$ = 0.1 veh. (a) Evolution of accumulation $n(t)$ and (b) evolution of travel time $T(t)$ where the time $t_{d1}$ corresponds to a change in demand, and $t_{i1}$ to a change in MFD branch, also indicated by the change in color ("br" stands for "branch"). (c) Evolution of outflow $\mu(t)$ and (d) representation of wave propagation in the reservoir for the trip-based model, a wave being defined as a curve connecting an inflow value at upstream to the same outflow value at downstream. The maximum vehicle speed is the free-flow speed. (e) Comparison for the trip-based model with two values of $\delta N$, evolution of accumulation $n(t)$ and (f) evolution of travel time $T(t)$*

Figure 2.7 presents, similarly as in Figure 2.3, the evolution of accumulation and travel

time for different initial conditions $n_{\text{ini}}$ and different levels of demand $\lambda_0$. Again because $\lambda_0 \neq \lambda_{\text{ini}}$, the system enters the transient phase at $t = 0$. The same conclusions are observed, namely the trip-based model gives more realistic travel times in transient phases, and is also more reactive than the accumulation-based one. In particular we can notice the very bad description of $T(t)$ in the accumulation-based framework compared with the trip-based one when the gap of demand is high. For example in Figures 2.7(b) or (d) for $n_{\text{ini}} = 10$ veh (which corresponds to $\lambda_{\text{ini}} = 0.06$ veh/s), and for $\lambda_0 = 0.8$ veh/s, the travel time described by the accumulation-based model exhibits a high drop during the whole transition period of the travel time provided by the trip-based model. On the other hand when the value of $\lambda_0$ is close to the initial state $\lambda_{\text{ini}}$, the effect of the travel time drop is minimized. An example is the case $n_{\text{ini}} = 110$ veh (which corresponds to $\lambda_{\text{ini}} = 0.66$ veh/s) with $\lambda_0 = 0.8$ veh/s in Figure 2.7(b). In this situation the solutions for $T(t)$ are quite similar in both approaches. This corroborates our analysis on the validity domain of the accumulation-based model in section 2.3.6, for which we notice that the height of the demand gap $\Delta\lambda$ seem to have the most influence on the travel time drop, in comparison with the gap duration $\Delta t$.

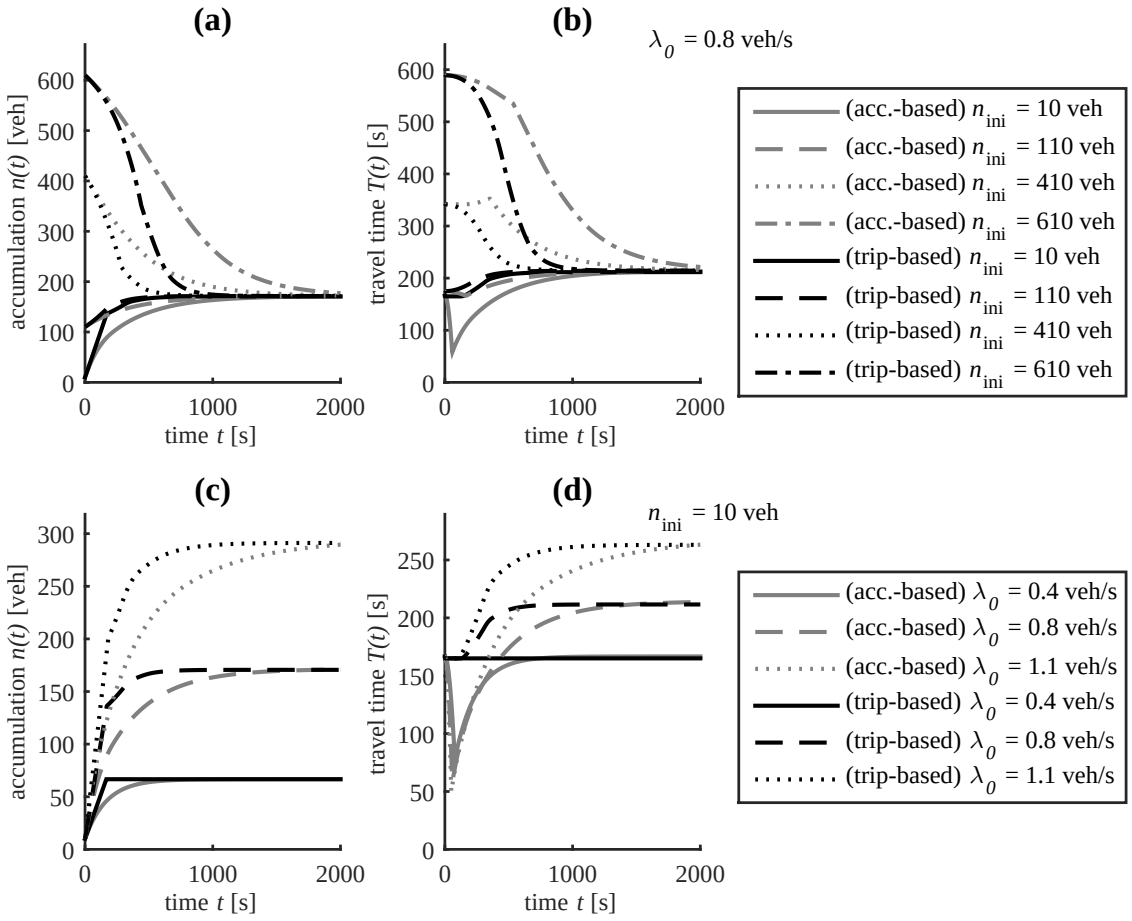



FIG. 2.7 – *Comparison between accumulation-based and trip-based models. (a) Different initial conditions $n_{\text{ini}}$ with demand $\lambda_0 = 0.8$ veh/s, evolution of accumulation $n(t)$ and (b) travel time $T(t)$. (c) Different demand levels $\lambda_0$ with initial condition $n_{\text{ini}} = 10$ veh, evolution of accumulation $n(t)$ and (d) travel time $T(t)$*

### 2.4.4 The event-based resolution method

This solver uses the same discretization of entering vehicles into vehicle units like in Daganzo & Lehe (2015) and Lamotte & Geroliminis (2016): we create the vehicles entering the system in accordance with a given demand level, and assign them a chosen trip length, which may be eventually different for each vehicle. The users travel at the current global mean speed $V(n(t))$ given by the speed-MFD of the reservoir for the current accumulation $n(t)$, and exit the reservoir once they complete their assigned trip length. This principle is illustrated in Figure 2.8. Like in Lamotte & Geroliminis (2016), our event-based simulator does not update the system on a regular time basis, but each time a vehicle enters or exits the network. Our method takes advantage of the fact that the mean speed is the same for all travelers, so that the update of $V(n)$ is only needed when the accumulation $n$ changes, i.e. at a vehicle entry or exit. This speeds up computation as the resulting method is fully event-based without any variable update on a fixed basis (no need for time step). Our resolution scheme adapts by itself to the evolution of the system and is always numerically accurate by construction. When the trip length is the same for all users this approach corresponds precisely to the discretization of the trip-based model, and is similar to the previous resolution method computed at order 1 with $\delta N = 1$. Comparisons between $\delta N = 0.1$ and $\delta N = 1$ for the previous method show similar results in Figures 2.6(c) and (d).

The main structure of the event-based method is detailed below:

1. Entry of the initial vehicles in the reservoir with the same trip lengths. Like the previous resolution methods, a steady state characterized by this initial accumulation is considered before $t = 0$. The traveled distance of these vehicles is then updated at $t = 0$ in accordance with this hypothesis.

2. Determine the next event which is the minimum between the next vehicle entry (according to the current demand level), and the next vehicle exit. This gives the next time step $\delta t$.

3. Update the traveled distance for all the circulating vehicles—i.e. which have not finished their trip yet—during $\delta t$ using the mean speed $V(n)$. Actually this update can be done one for all for the next exiting vehicles by storing the speed evolution, because all the users travel at the same speed within the reservoir. This can speed up calculations.

4. If the next event is a vehicle entry, create a new vehicle in the reservoir and assign it a given trip length. Else if it is a vehicle exit, remove the corresponding vehicle from the reservoir. Update the accumulation $n$.

5. Update the speed of all the circulating vehicles, evaluated at $V(n)$.

6. Go to 2.

The procedure stops when the simulation time reaches a given simulation duration.

In closing, it is important to point out that under congested conditions the two resolution schemes proposed for the trip-based model apply with minimal changes. We propose a three-step procedure, where (i) the problem is solved assuming no supply restriction, (ii) the outflow is reduced according to the supply profile, and (iii) Newell's minimum principle (Newell, 1993) is applied to take the congestion wave impact on inflow into account.

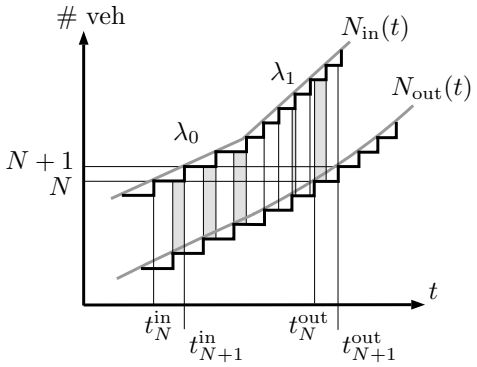



FIG. 2.8 – *Representation of the event-based resolution method in case of a unique trip length for all travelers (FIFO process)*

## 2.5 Comparison between the accumulation-based and the trip-based models on three case studies

### 2.5.1 Typical peak-hour demand profile

We apply now the event-based method on a peak-hour demand case. We compare the results with the resolution scheme of the trip-based model established in the previous section and the accumulation-based model. Two simulations are performed: the first one is considering a slow-varying demand with a simulation duration of 10,000 s, whereas the second one is characterized by much faster demand variations as the whole simulation lasts only 2000 s. The two demand profiles are shown in Figures 2.9(a) and 2.9(e), they are identical in shape and reach the same levels. The continuous aspect of these profiles is actually approximated by successive small stairs to comply with the piecewise constant hypothesis.

Figures 2.9(b)-(d) present the results for the first simulation, and Figures 2.9(f)-(h) the results for the second simulation. It appears clearly that the accumulation-based model is more suitable for slow-varying demand situations, as already mentioned in Daganzo (2007). In case of fast variations in demand, the slow reaction time of this model, already highlighted in section 2.4, makes it underestimate the accumulation peak in Figure 2.9(f). The drop phenomenon in travel time is also noticed in both simulations, but the estimation of $T(t)$ is clearly worse in the second case. For the latter the drop size could have been forecasted thanks to our study in section 2.3.6. Given an increase of 1 veh/s over 500 s for the demand brings us to a drop of 90 s during 500 s. This is indeed what is observed in Figure 2.9(g), although the drop height is a bit overestimated. Another phenomenon worth mentioning is the counter-clockwise hysteresis loop occurring for the trip-based model when displaying the traffic states in the (accumulation, outflow) plane, see Figures 2.9(d) and 2.9(h). The same phenomenon has been also observed by Lamotte & Geroliminis (2016). On the other hand, experimental (e.g. Buisson & Ladier, 2009, Geroliminis & Sun, 2011a), theoretical (e.g. Gayah & Daganzo, 2011) and simulation (e.g. Mahmassani *et al.*, 2013b) studies suggested that clockwise hysteresis loops are more likely to occur in the production-MFD. These authors gave useful insights on heterogeneity in traffic density being one of the causes of hysteresis loops. For instance, Gayah & Daganzo (2011) highlighted hysteresis in a simplistic two-bin system due to density heterogeneity

between the two bins, their modeling approach being grounded on the accumulation-based model. Some authors like Daganzo (2011) explained these phenomena by a growing effect of instabilities in the network responsible for internal capacity drops during the recovery of congestion. In our framework this effect is not modeled, nor potential internal heterogeneity in density. The counter-clockwise loop we observe in the trip-based approach has a completely different explanation, and to the authors' best knowledge, this has never been exhibited in previous works on hysteresis in MFD, on the notable exception of Lamotte & Geroliminis (2016). This phenomenon is a direct consequence of the delay between the evolution of accumulation and the one of outflow which is reproduced by the trip-based model. This generates traffic states which are below the outflow-MFD during the onset of congestion (accumulation varies faster than outflow due to the time needed for inflow information to travel from upstream to downstream), and then traffic states that are above the outflow-MFD during the offset of congestion (the reduction of accumulation is faster than the one of outflow for the same reason). It should be also noticed that our observations apply to the outflow-MFD and not to the production-MFD. This is another major difference compared with the above-mentioned studies, which either make observation on the production-MFD or consider that outflow- and production-MFD are simply proportional (use of the accumulation-based model). More data and especially reliable outflow data are required to have a complete picture of which effects are the most significant. All effects can also be combined for real observations.

### 2.5.2 Temporary supply restriction at exits

This case study aims to analyze the effect of a temporary supply restriction at the reservoir exit. This supply restriction is followed by a demand decrease to retrieve an equilibrium in free-flow conditions. Two scenarios described in Figures 2.10(a) and (e) are compared. The first one concerns a short supply reduction, and we see that the accumulation-based and the trip-based models are almost equivalent in these conditions, see Figures 2.10(a)-(d). The second one is about a longer supply reduction, and results show very different conclusions depending on the hypotheses made for the accumulation-based model in congestion.

In Figures 2.10(f)-(h), the green curves ("model 1") correspond to our basic assumption of an entrance function $I(n)$ equal to the outflow-MFD in congestion, which is consistent with a CTM-based approach for inflow management. In this case the accumulation-based model reaches an equilibrium point on the congested part of the outflow-MFD and remains here even when the boundary conditions become more favorable after 5000 s. With this approach the system is very sensitive to a supply reduction: if it is short enough the reservoir will return to a free-flow state after the demand has decreased like in Figures 2.10(a)-(d), but if it is longer than a critical value, i.e. the time required to reach the congested part of the outflow-MFD, then the system state is permanently degraded. This situation may be even worse if no inflow limitation $I(n)$ is taken into account. The results in that case are represented by the red curves ("model 2"), after hitting the congested branch the accumulation-based model immediately converges to global gridlock. The last option corresponds to the yellow curves ("model 3"), it assumes an outflow-MFD which is not restricted in congestion, i.e. $O(n)$ is constant equal to the outflow capacity for high values of accumulation. This hypothesis allows a quasi-perfect matching between the accumulation-based and the trip-based models. In fact, the way we implemented congestion management in the trip-based approach in us-

ing Newell's method makes our reservoir behaving like a freeway in congestion. Then it is no surprise that its equivalent for the accumulation-based model is a CTM-like modeling approach.

### 2.5.3 Case of different trip lengths

This last case study is taking advantage of the event-based resolution method for the trip-based model to account for different groups of travelers having their own trip length. This is compared with an approach proposed by Geroliminis (2015) to include different trip length classes in the accumulation-based framework. This author considers $N$ classes of travelers, each class having a current accumulation $n_i(t)$ and a trip length $L_i$, $1 \leq i \leq N$. The total accumulation is $n = \sum_{i=1}^{N} n_i$. A distinction is made for the outflow of each class, which means the system dynamics is described by the following equations in free-flow conditions:

$$\forall i \in \{1, \ldots, N\}, \qquad \frac{dn_i}{dt} = \lambda_i(t) - O_i(n_i(t)) \tag{2.13}$$

where $\lambda_i(t)$ and $O_i(n_i(t))$ are respectively the demand and the outflow for class $i$. We suppose the sum of all demands is below the global outflow capacity and no supply at exit is limiting the sum of all outflows. Otherwise a merge function has to be employed to manage inflows and outflows.

To calculate the average trip length $L$, note that it satisfies $L = nV(n)/O(n)$ and that $O(n) = \sum_{i=1}^{N} O_i(n_i)$, where each outflow function $O_i(n_i)$ is determined by applying Little's formula (Little, 1961): $O_i(n_i) = n_i V(n)/L_i$. Combining these equations, gives:

$$L(t) = \frac{n(t)}{\sum_{i=1}^{N} \frac{n_i(t)}{L_i}} \tag{2.14}$$

which is time-dependent since $n_i = n_i(t)$. An alternative definition can be obtained by noting that $\sum_{i=1}^{N} O_i L_i = nV(n)$ and therefore:

$$L(t) = \frac{\sum_{i=1}^{N} O_i(n_i(t)) L_i}{O(n(t))} \tag{2.15}$$

which means that the average trip length is also the arithmetic mean of all $L_i$ weighted by their corresponding outflow $O_i$.

Figures 2.11(a)-(d) present a situation with two classes: $L_1$ = 1000 m and $L_2$ = 2000 m with the same initial demand. A peak for the first class occurs at $t$ = 500 s. For the trip-based model the evolution of the average trip length is calculated by successive means of individual trip lengths over a given time period. It is the discretized version of $L(t)$, seen as the mean of all $L_i$ weighted by outflows. We expect a temporary drop in the evolution of $L(t)$, due to the demand peak for the smallest trip distance $L_1$. This should result in a higher outflow for this class, and thus a greater impact in the mean calculation for $L(t)$ compared to those traveling a longer distance $L_2$. On the whole the two modeling approaches reproduce this effect quite well, although the accumulation-based model emphasizes it more than the trip-based one, see Figure 2.11(b). As regards the accumulation and travel time evolution, there is a clear overestimation of the demand peak for the accumulation-based model.

This study has been compared to a demand peak case with a unique trip length presented in Figures 2.11(e)-(h). The demand profile $\lambda(t)$ is exactly the global demand of the previous case: $\lambda(t) = \lambda_1(t) + \lambda_2(t)$. Two situations are examined, one considers a constant trip length $L_0$ = 1500 m equal to the mean of $L_1$ and $L_2$ in steady state, whereas the other concentrates on a time-varying trip length $L(t)$ which corresponds to the instantaneous mean of $L_1$ and $L_2$ at each time, i.e. $L(t) = (\lambda_1(t)L_1 + \lambda_2(t)L_2)/\lambda(t)$.

We observe in Figure 2.11(f) that $L(t)$ reproduces almost the same evolution as the previous one in Figure 2.11(b). However, the behavior of accumulation and travel time is completely different in this case. As seen in Figures 2.11(g) and (h), using either a constant or dynamic trip length $L(t)$ leads to an underestimation of the demand peak for the accumulation-based model as it was already noticed in section 2.5.1. However, if we assume that the most accurate modeling for this two-class problem is the trip-based approach with $L_1$ and $L_2$, then the accumulation-based model paradoxically gives better results in accumulation with $L_0$ or $L(t)$ than with $L_1$ and $L_2$. In this case, we have to compare the evolution of total $n(t)$ in Figure 2.11(c) with $n(t)$ in Figure 2.11(g). We observe that the overestimation of the accumulation peak given by the accumulation-based model is lower in Figure 2.11(g) than in Figure 2.11(c).

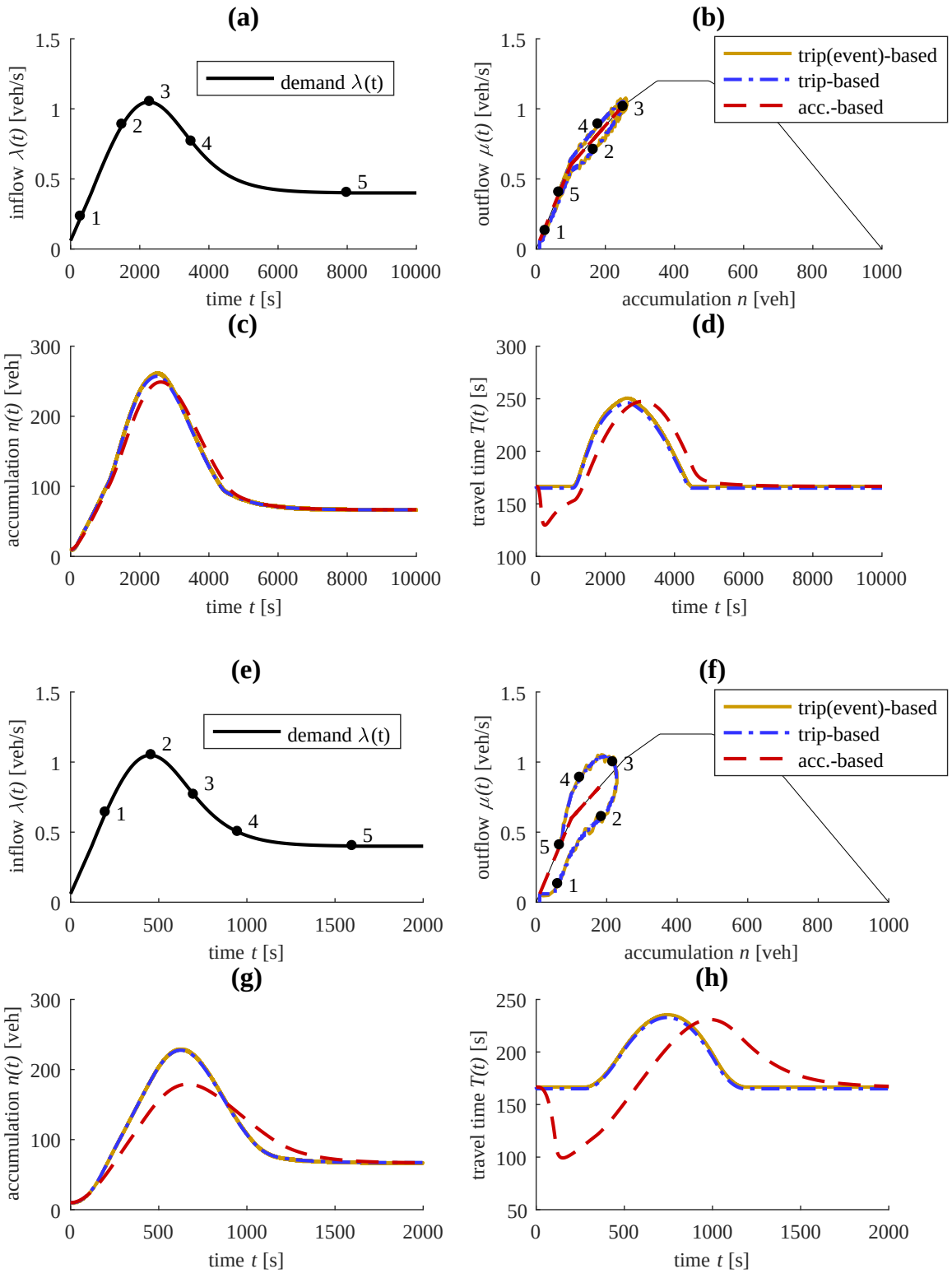



FIG. 2.9 – *Comparison between models on two simulations. Simulation 1 with a slow-varying demand over 10,000 s: (a) peak demand profile, (b) outflow vs accumulation at each time, (c) evolution of accumulation and (d) evolution of travel time. Simulation 2 with a fast-varying demand over 2000 s: (e) peak demand profile, (f) outflow vs accumulation at each time, (g) evolution of accumulation and (h) evolution of travel time.*

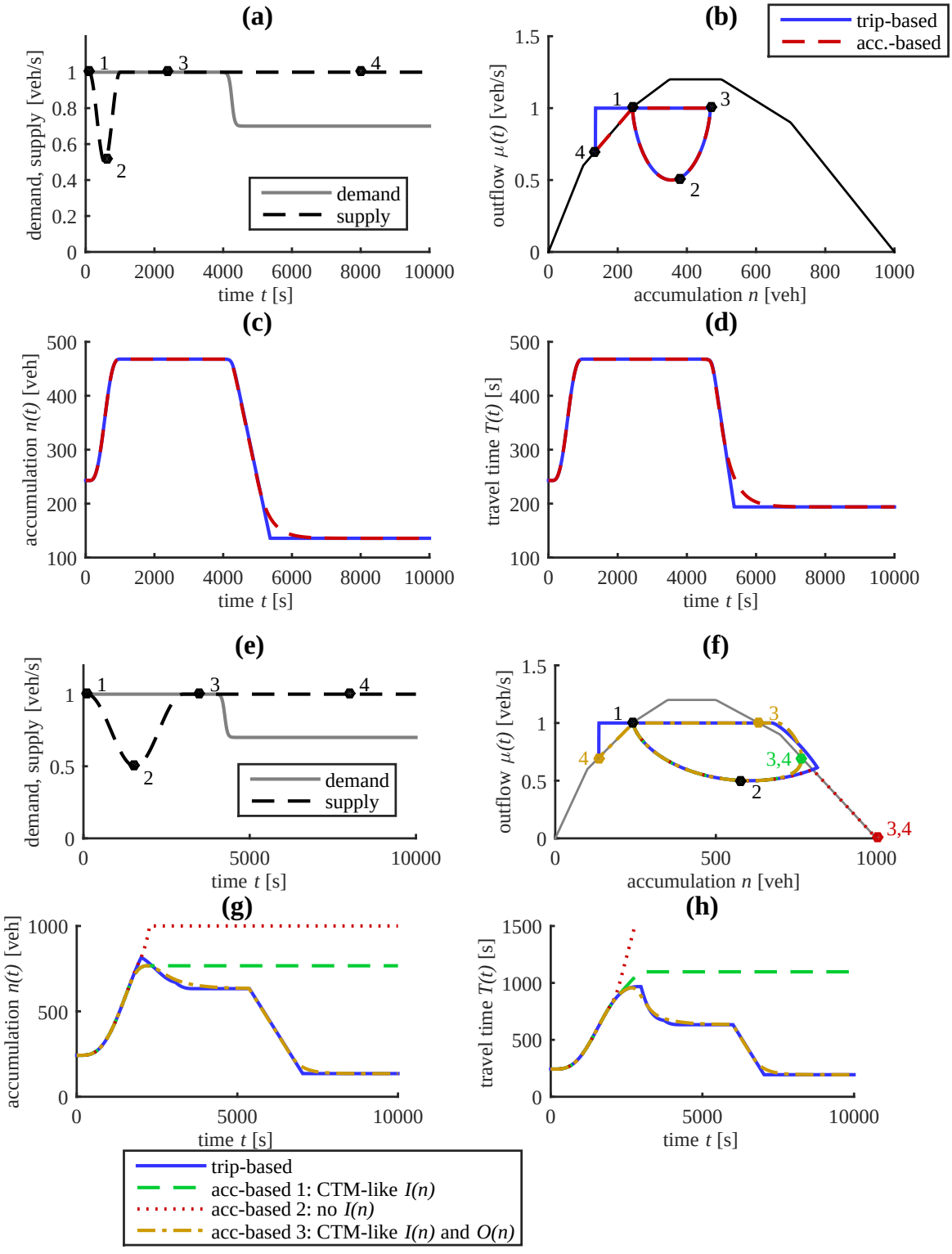



FIG. 2.10 – *Comparison between models on two simulations. Simulation 1 with supply reduction over 1000 s: (a) demand and supply profiles, (b) outflow vs accumulation at each time, (c) evolution of accumulation and (d) evolution of travel time. Simulation 2 with supply reduction over 3000 s, three cases for the accumulation-based model: 1, CTM-like $I(n)$, 2, no $I(n)$ and 3, CTM-like $I(n)$ and $O(n)$. (e) Demand and supply profiles, (f) outflow vs accumulation at each time, (g) evolution of accumulation and (h) evolution of travel time.*

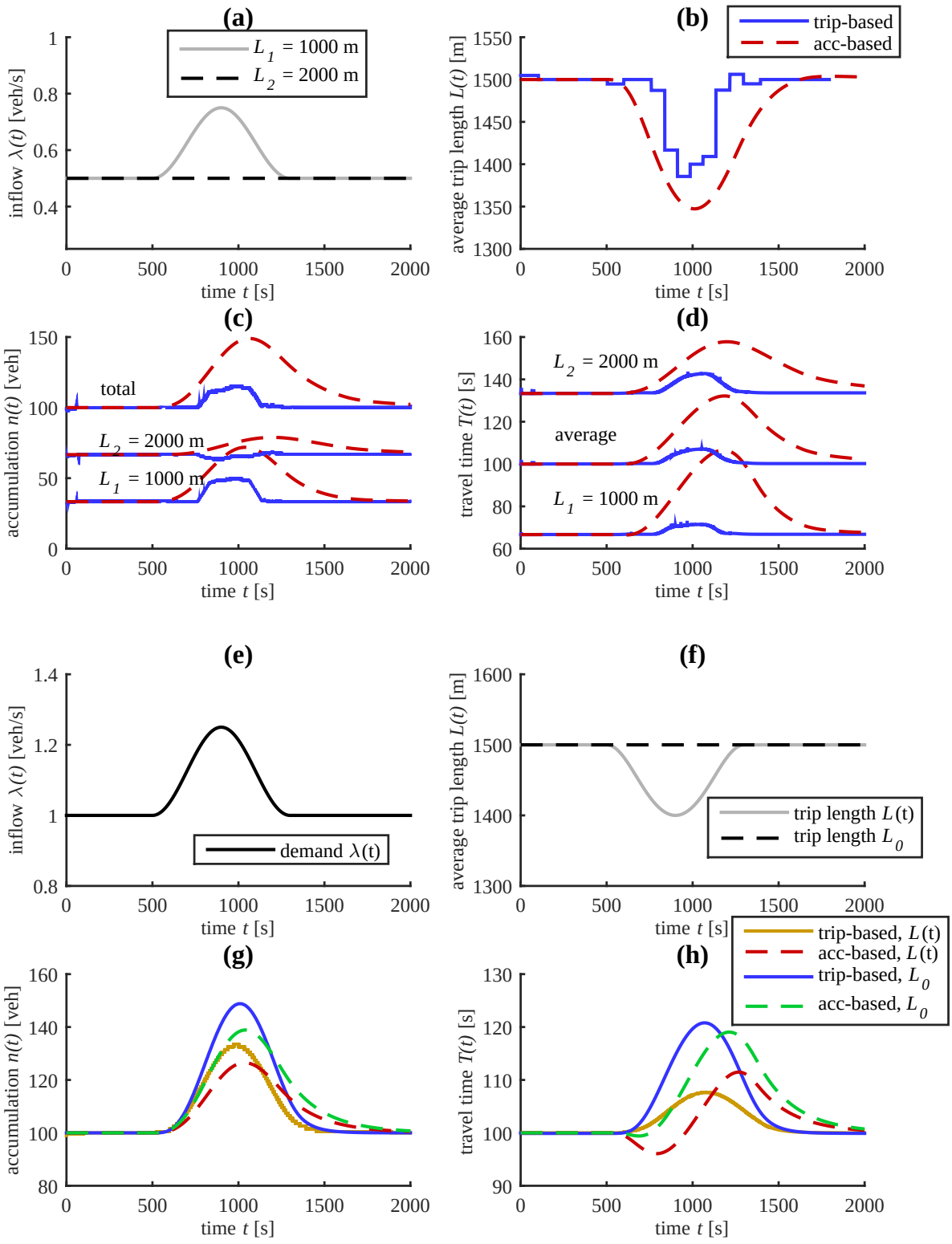



FIG. 2.11 – *Comparison between models on two simulations. Simulation 1 with traveler trip lengths separated by class: (a) demand profile for each class, (b) evolution of the average trip length, (c) evolution of accumulation and (d) travel time for each class. Simulation 2 with a unique trip length varying or not through time: (e) demand profile, (f) evolution of the trip length, (g) evolution of accumulation and (h) travel time.*

## 2.6 Discussion and conclusion

In this chapter we proposed different resolution schemes for the accumulation-based and the trip-based models. For the first one, exact solutions have been established with piecewise linear functions thanks to the semi-analytical method, although this process may be limited when changes in demand are numerous. The full numerical scheme proved to have acceptable numeric errors as well, and is more flexible than the semi-analytical method. Its basic formulation has also been extended to solve the approach of Geroliminis (2015) integrating multiple trip lengths. As for the trip-based framework, in comparison with Arnott (2013), Fosgerau (2015), Daganzo & Lehe (2015), Lamotte & Geroliminis (2016) we developed a numerical resolution method that can approximate the model solution in an efficient way, and can offer some theoretical support to our second resolution scheme. While having some similarities with the methods presented by the previous authors, the latter is event-based like in Lamotte & Geroliminis (2016), and does not introduce any numerical bias by the choice of a time step. With encouraging applications and results, it proved to be a very simple and convenient way to account for varying trip lengths among travelers, and thus may be considered as a promising tool for further applications like multi-reservoir modeling, studies including on-street parking search, or integration with route choice problems.

Apart from these consideration about resolution methods, this chapter made a large comparison between the trip-based model and the accumulation-based one, which is still widely used for the development of more complex applications like multi-reservoir simulation and dynamic control. Our literature review allows to draw a synthesis about the hypothesis history of both approaches. Their differences are then clearly stated: unlike the trip-based model the accumulation-based one (i) mostly ignores traveled distances, (ii) has no memory of the past of the system, and (iii) considers that accumulation remains constant when a user travels in the reservoir. These assumptions are responsible for the incoherence in the model results during transient phases: drop in travel time when the demand increases and slow response of the system in fast-varying conditions are notably observed. In steady state however, the two models give the same results. Regarding the travel time drop phenomenon, we have then proposed a validity domain of the accumulation-based model with respect to the demand variation rate. We have also provided error estimation depending on the size of the demand variations. It should be noted that such errors correspond to simulation where the demand is determined from the beginning. Control applications for which the accumulation-based model is widely used introduce a feedback on this demand with respect to the reservoir state. Such feedback loop may certainly help to correct the mistakes of the accumulation-based model and then reduce the inaccuracies. On the other hand, multi-reservoir simulation grounded on the accumulation-based framework may be more obviously affected by the artifacts we highlighted.

# 3 Impact of heterogeneous perimeter flow distributions on a single reservoir simulation

The concept of Macroscopic Fundamental Diagram (MFD) appears to be an attractive option for the large-scale simulation of urban networks. However, many authors have pinpointed that the implementation of an MFD in a given region can be jeopardized by network or traffic heterogeneities. So far, most studies have focused on internal heterogeneities and their impacts on the MFD shape. Little attention has been paid to their impacts on the accuracy of MFD-based simulation outputs. Heterogeneities not only affect the MFD shape but also the distance traveled in an area. Furthermore, heterogeneities can be observed not only inside the area but also at its border.

This chapter investigates how network and traffic heterogeneities influence the accuracy of the MFD-based model with a single reservoir. To this end, a simple network is simulated using a mesoscopic kinematic wave model and extensive local demand and supply settings at the boundaries. All such settings generate homogeneous or heterogeneous network loadings both within and at the border of the region studied. The simulation results are then compared to the MFD model outputs when steady conditions are reached. This shows significant discrepancies between both approaches, in flow and accumulation, when heterogeneous distributions of demand (or supply) are applied to the local entries (or the exits) of the mescocopic model. Discrepancies can be drastically reduced for heterogeneous demand settings, when the MFD shape and the average trip length are calibrated accordingly. However, no simple solution exists for heterogeneous supply settings because they may drive very different internal congestion patterns in the network due to local spillbacks. We propose a correction method in this chapter to adjust the MFD-based model outputs in such a case.

This chapter is an updated version of the paper:

Mariotte, G. & Leclercq, L. (2018). Heterogeneous perimeter flow distributions and MFD-based traffic simulation, submitted to *Transportmetrica B*, under review

## 3.1 Notations for this chapter

TAB. 3.1 – *Specific notations in this chapter*

| Notation | Definition [units] |
|---|---|
| $q_{\text{in}}^D$ | Nominal inflow demand at each entry network link [veh/s] |
| $q_{\text{out}}^S$ | Nominal outflow supply at each exit network link [veh/s] |
| $a_1$ | heterogeneity coefficient for the network entry, slope of the demand distribution along links. $a_1 = 0$: homogeneous distribution, $a_1 = 1$: highly heterogeneous distribution [–] |
| $a_2$ | heterogeneity coefficient for the network exit, slope of the supply distribution along links. $a_2 = 0$: homogeneous distribution, $a_2 = 1$: highly heterogeneous distribution [–] |
| $Q_{\text{in}}^D$ | Reservoir inflow demand = 6 $q_{\text{in}}^D$ [veh/s] |
| $Q_{\text{out}}^S$ | Reservoir outflow supply = 6 $q_{\text{out}}^S$ [veh/s] |
| $Q_{\text{in}}(t)$ | Reservoir effective inflow [veh/s] |
| $Q_{\text{out}}(t)$ | Reservoir effective outflow [veh/s] |
| $L_{\text{simul}}$ | Estimation of the average trip length for a given mesoscopic simulation [m] |
| $L_{\text{MFD}}$ | Estimation of the average trip length with a linear regression on steady state outflows from a set of simulations [m] |
| $T_c$ | Time at which a slope break appears in $N_{\text{out}}(t)$ [s] |
| $Q_{\text{out}}^{\text{eff}}$ | Total effective outflow obtained in the mesoscopic simulation after the slope break [veh/s] |

## 3.2 Motivations

The Macroscopic Fundamental Diagram (MFD) has received much attention in the literature over the last decade, in particular for control applications, (e.g. Haddad & Mirkin, 2017, Sirmatel & Geroliminis, 2017b, Ampountolas *et al.*, 2017). Partitioning a city into a multi-reservoir system characterized by MFDs also appears to be an appealing and computationally efficient method for simulating urban multimodal networks at a large scale (e.g. Knoop & Hoogendoorn, 2014, Yildirimoglu *et al.*, 2015). The concept of MFD was first introduced by Godfrey (1969), Herman & Prigogine (1979) and Mahmassani *et al.* (1984). However, Daganzo & Geroliminis (2008) were the first authors to present a well-defined MFD based on real data. Since then, numerous studies have investigated the conditions for good MFD definition. Network and/or traffic heterogeneities are strong sources of MFD scattering. Heterogeneities can be driven by: (i) non-consistent network definition (Buisson & Ladier, 2009), (ii) data sources and processing methods (Leclercq *et al.*, 2014), (iii) the spatial distribution of vehicles within the reservoir (Mazloumian *et al.*, 2010, Geroliminis & Sun, 2011b, Saberi & Mahmassani, 2012), (iv) time of the day (Gayah & Daganzo, 2011), (v) network layout and configuration (Knoop *et al.*, 2014, Ortigosa *et al.*, 2015, Muhlich *et al.*, 2015), (vi) traffic signal settings (de Jong *et al.*, 2013, Zhang *et al.*, 2013, Gayah *et al.*, 2014, Girault *et al.*, 2016), (vii) the definition of the Origin-Destination (OD) matrix (Doig *et al.*, 2013, Leclercq *et al.*, 2015) and (viii) traveler information (Zhao *et al.*, 2014).

Most studies dealing with heterogeneities in traffic networks mainly focused on the shape of the MFD. Geroliminis & Sun (2011a), Gayah & Daganzo (2011), Daganzo (2011), Saeedmanesh & Geroliminis (2015) presented hysteresis loops during congestion on- and offset. Leclercq & Geroliminis (2013) showed that different route flow distributions on parallel networks lead to different MFD shapes, in particular when the network is close to its maximal capacity. Networks with uneven and inconsistent distribution of congestion may exhibit traffic states that are well below the upper bound of an MFD, and that are too much scattered to define the MFD precisely (Ji *et al.*, 2010, Daganzo *et al.*, 2011, Geroliminis & Sun, 2011b, Mahmassani *et al.*, 2013b, Knoop *et al.*, 2015).

Surprisingly, little attention has been paid in the literature to assess the impacts of heterogeneous network loadings on the accuracy of MFD-based simulation outputs. As heterogeneities not only affect the MFD shape but also the distance traveled in an area (Leclercq *et al.*, 2015), the impacts on simulation results may be even more considerable. Furthermore, heterogeneities may be observed not only inside the area but also at its border. This may trigger major simulation issues as borders correspond to boundary conditions not only for a reservoir but also for its neighborhood. A simple way to assess the accuracy of MFD-based simulation results is to compare them with those of more refined traffic flow models using microscopic (e.g. Gipps, 1986b, Barcelo *et al.*, 2006) or mesoscopic approaches (e.g. Ben-Akiva *et al.*, 2002, Mahut *et al.*, 2003, Burghout *et al.*, 2005, Leclercq & Becarie, 2012). Geroliminis & Daganzo (2007) were the first to compare the single reservoir model with a microscopic simulator on a 2.5 square mile area of Downtown San Francisco. They showed that the results from the reservoir model fit the micro-simulation outputs quite well, mainly because the network studied exhibits a low-scattered MFD which seems independent from OD tables. But this low scattering is not universally expected, as suggested by the above-mentioned studies on network heterogeneities. Yildirimoglu & Geroliminis (2014) also compared the evolution of the perimeter flow estimated by a multi-reservoir MFD model with

microscopic simulation. The results are more mitigated as consistency is obtained by integrating a feedback mechanism that adjusts the accumulation in the MFD model from time to time, depending on the current state predicted by the microscopic model. In this chapter, the authors also clearly pinpointed that assuming a single and constant trip length for all vehicles within a reservoir may lead to strong bias with the microscopic simulation results. Such an assumption was first conjectured by Daganzo (2007). Since then, most papers that have used MFD simulation (mostly for control applications) have taken advantage of this simplification. However, recent results (e.g. Yildirimoglu & Geroliminis, 2014, Leclercq *et al.*, 2015, Lamotte & Geroliminis, 2018) are now questioning this hypothesis. They show that not only traffic conditions but also the OD matrix may significantly affect the mean travel distance in a reservoir.

In this study we aim to characterize how the MFD model is sensitive to network heterogeneities generated by uneven demand and supply settings at the local entries and exits that define the perimeter of a single reservoir. In the following, for the sake of simplicity we use the terms "exit supply" or sometimes "supply" to mean the capacity of the exit nodes (expressed in [veh/s]). Such capacities can simulate downstream link-level restrictions due to spillbacks from an adjacent zone for instance. MFD simulation results are compared with a mesoscopic simulator (Leclercq & Becarie, 2012) capable of reproducing local dynamic effects in the network and in particular spillbacks between intersections during congestion. This link-level simulator is used in our study as the ground truth to identify possible limitations in the MFD model. We choose the same 2×4 network as presented in Leclercq *et al.* (2015) because this network, although very simplistic, permits generating a wide range of different heterogeneous loadings. Its features (asymmetrical shape, one-way streets, etc) are deliberately exaggerated to emphasize heterogeneities and thus identify possible shortcomings while simulating traffic states with the MFD. This network is not meant to be general, dense and homogeneous, but rather to represent a simplified version of the irregular and asymmetrical reservoir shapes often exhibited by partitioning algorithms (see e.g. Ji & Geroliminis, 2012, Saeedmanesh & Geroliminis, 2016). We focus on the ability of the MFD model to predict accurate traffic states in the steady phase, i.e. after the initial network loading. Comparisons are based on the dynamic evolution of the vehicle accumulation within the reservoir, and inflows and outflows. Our results show that in the case of perturbations in the boundary conditions (heterogeneous settings), the reservoir model fails to reproduce the network-level states predicted by the mesoscopic simulator. More interestingly, it appears that the effects of perturbations on the demand loadings are quite different from those of modifications of the supply settings. When the demand distribution is heterogeneous at the network boundary, the reservoir model gives accurate results only if both the MFD and the mean trip length are properly re-calibrated to adjust the modifications at the reservoir perimeter. But in the case of a heterogeneous distribution of supply at the reservoir boundary, such re-calibrations in the MFD model prove to be insufficient to properly reproduce the network-level steady state of the mesoscopic model. Indeed, heterogeneous supply distributions may trigger local bottlenecks that result in specific congestion patterns when congestion spills back. This reduces the total outflow at the perimeter compared to the total available capacity obtained by simply summing the local exit capacities. A thorough analysis of this phenomenon makes it possible to propose a framework to adjust the MFD simulation outputs to what is predicted at the mesoscopic scale. Although dealing with a simplistic network configuration, this study provides insight on how to apply MFD-based

simulation properly when heterogeneous boundary conditions are observed at the network perimeter. In practice, for applications based on the MFD, like perimeter control, this work suggests that both the MFD and the average trip length should be re-calibrated on a regular time basis to account for changes in OD patterns. In further studies, it would be interesting to investigate the time scale of the re-calibrations, i.e. how often we should re-calibrate. Both time scales for MFD and trip length might be different, but this aspect is out of the scope of the present chapter.

This chapter is organized as follows. Section 3.3 introduces the case study and the two modeling approaches (mesoscopic and MFD-based). Section 3.4 presents the MFD and average trip length estimations made with the link-level simulations for different boundary settings. Section 3.5 presents the comparison of the simulation outputs in steady state between the two modeling scales. Guidelines will be proposed on the best way to set up the MFD-based model. Section 3.6 discusses the major findings of this study while presenting new results obtained on a bigger network.

## 3.3 Presentation of our case study

### 3.3.1 Mesoscopic simulation settings

#### 3.3.1.1 Network configuration and geometry

In this study, we focus on a very simple $2 \times 4$ Manhattan-type network where all the links are one-way and have one lane. Thus, the network has 6 entries and 6 exits, see Figure 3.1(a). The West-East inner links are more than twice as long as the North-South ones. All the intersections are controlled by traffic lights with the same green and red times equal to 30 s. All the signals have the same offset. This asymmetrical network configuration has been designed to favor heterogeneous loadings. The geometry ensures highly variable trip lengths as a function of the OD matrix, as the West-East routes are much longer than the North-South ones (Leclercq *et al.*, 2015). Furthermore, uneven supply distributions at the exits will trigger many different congestion patterns. We keep the network configuration simple as a large number of mesoscopic simulations must be run in order to investigate the influence of the demand and supply distribution over the entries and exits.

#### 3.3.1.2 Origin-Destination matrix

Since our aim is to investigate the impact of boundary loadings on MFD-based simulation, we consider only transferring trips for the sake of simplicity. By ignoring internal trips, we decided to limit the number of potential parameters in the demand settings to keep the analysis simple. Moreover, the treatment of internal trips might differ from that of transferring trips in MFD-based simulation. Consequently, considering only the latter allows omitting this research question which is out of the scope of this study. Thus, in the following, origins and destinations refer to the 6 entries and 6 exits defining the network perimeter, see also Figure 3.1(a). In the whole study, the distribution of destinations among users starting from a unique origin remains constant, i.e. the coefficients of the normalized OD matrix in Figure 3.1(b) are time-invariant. Such a distribution guarantees that homogeneous demand loadings (same demand at all entries) will correspond to a homogeneous flow distribution

at the exits. This does not imply a homogeneous use of the network however. Here, the OD flow matrix depends only on the demand values assigned to each entry.

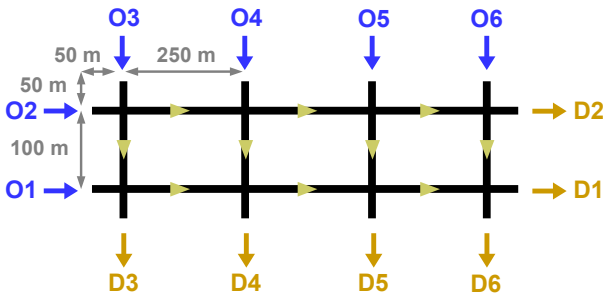
#### 3.3.1.3 Traffic simulation settings

The mesoscopic simulator in this study is fully described in Leclercq & Becarie (2012), Joueiai *et al.* (2015). This simulator is based on the space-Lagrangian resolution of the LWR model (Lighthill & Whitham, 1955, Richards, 1956) and provides the exact solutions of this model for links (Laval & Leclercq, 2013). At the link level, traffic flow is characterized by a triangular fundamental diagram (FD) with the following parameters: free-flow speed $u$ = 25 m/s, wave speed $w$ = 5 m/s, and jam density $k_j$ = 0.2 veh/m. Outputs of one simulation provide the passing times of all the vehicles at each node.

#### 3.3.1.4 Assignment and convergence

One simulation lasts 5000 s including a 1000 s warm-up period. For all local OD pairs, a maximum of the three shortest routes in distance may be used by vehicles. The flow distribution over the different routes is calculated by the mesoscopic simulator based on the first Wardrop principle (user equilibrium), i.e. all users travel on the paths with the minimum travel time. To distribute the users between the different routes, different simulations are run iteratively using the Method of Successive Averages (MSA). The travel time values are based on the free-flow speed for the first simulation, and on the previous simulation output for the next iterations. The traffic assignment process converges once the mean travel time of 95% of all the link lengths is almost constant (less than 5% variation over two successive iterations).

**(a)**




**(b)**

| | D1 | D2 | D3 | D4 | D5 | D6 |
|---|---|---|---|---|---|---|
| O1 | 1/6 | 0 | 1/3 | 1/4 | 1/4 | 0 |
| O2 | 0 | 1/6 | 1/3 | 1/6 | 1/6 | 1/6 |
| O3 | 0 | 0 | 1/3 | 1/4 | 1/6 | 1/4 |
| O4 | 1/4 | 1/4 | 0 | 1/3 | 1/6 | 0 |
| O5 | 1/4 | 1/4 | 0 | 0 | 1/4 | 1/4 |
| O6 | 1/3 | 1/3 | 0 | 0 | 0 | 1/3 |

FIG. 3.1 – *(a) Network configuration. (b) Normalized origin-destination matrix*

#### 3.3.1.5 Demand scenarios

A wide range of constant demand values at the entries and constant supply (capacity) values at the exits are applied to this network. The reference scenario, named SC0 hereon, corresponds to a uniform distribution between the entries for the total demand and uniform distribution between the exits for the total supply, see Figure 3.2(a). The demand at each entry is equal to $q_{\text{in}}^D$ (in [veh/s]) and the supply at each exit is $q_{\text{out}}^S$ (in [veh/s]). Heterogeneous network loadings will be obtained by either using uneven demand distribution over all the

entries (scenario SC1), see Figure 3.2(b), or uneven supply distribution at the exits (scenario SC2), see Figure 3.2(c). Demand and supply patterns are considered linear by the network directions. For a given value of $q_{in}^D$ (or $q_{out}^S$ in the case of scenario SC2), a single parameter $a_1$ (or $a_2$ in the case of scenario SC2) which ranges from 0 to 1 is thus sufficient to define the demand values at all the entries (or the supply values at all the exits for SC2), see Figure 3.2(b) and (c). Parameters $a_1$ and $a_2$ are called the heterogeneity coefficients. Note that $a_1$ and $a_2$ = 0 for SC0. Also note that the sum of the demand flow on all the entries (or the sum of the flow restrictions on all the exits) is always the same, i.e. independent from the heterogeneity coefficients. It is equal to $Q_{in}^D = 6q_{in}^D$ for the demand, and to $Q_{out}^S = 6q_{out}^S$ for the supply.

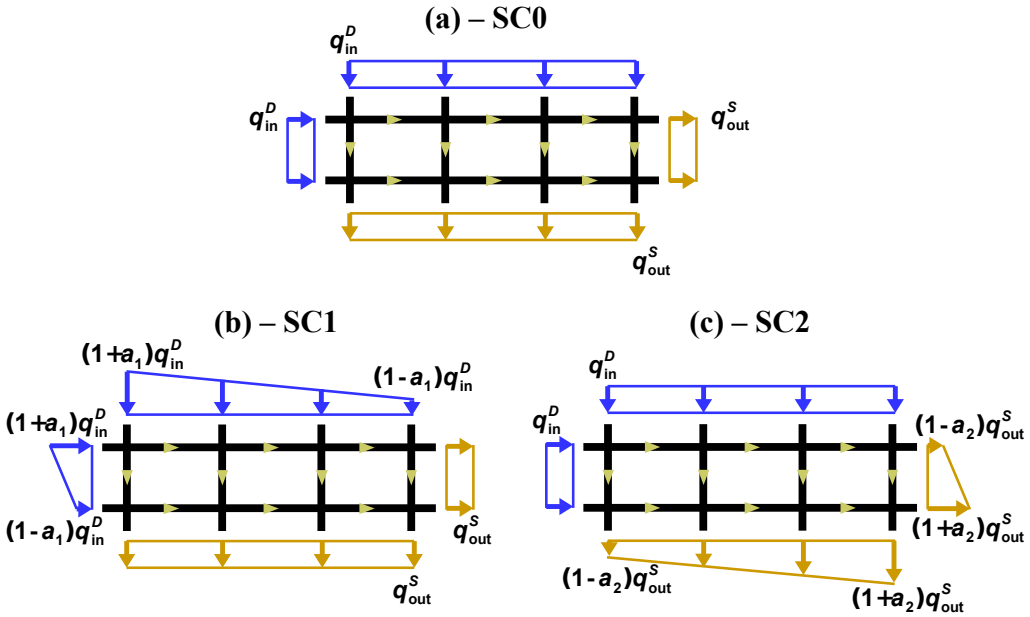



FIG. 3.2 – *(a) Scenario SC0: Homogeneous demand and supply distribution. (b) Scenario SC1: Heterogeneous demand distribution at the entries with homogeneous supply distribution at the exits. (c) Scenario SC2: Heterogeneous supply distribution at the exits with homogeneous demand distribution at the entries.*

### 3.3.2 Single reservoir implementation of the MFD-based model

In parallel to the mesoscopic simulations, the traffic states in the network are estimated using an MFD-based model (Daganzo, 2007, Geroliminis & Daganzo, 2007). Here, we consider that the whole network is described by a single reservoir with a total inflow $Q_{in}(t)$ and a total outflow $Q_{out}(t)$ (in [veh/s]). Traffic states within the reservoir are characterized by the vehicle accumulation $n(t)$ (in [veh]), which is the number of vehicles circulating at time $t$. Cumulative count curves $N_{in}(t)$ and $N_{out}(t)$ (in [veh]) are also calculated by respectively integrating inflows and outflows to monitor the reservoir perimeter. The network MFD is defined as the relationship between the travel production $P(n)$ (in [veh.m/s]) and the accumulation $n$. The existence of a well-defined MFD will be tested against the mesoscopic simulation results in the next section 3.4.1. According to Daganzo (2007), the evolution of $n(t)$ is given by the following conservation equation:

$$\frac{dn}{dt} = Q_{in}(t) - Q_{out}(t) \tag{3.1}$$

The same author makes the following crucial assumption: the outflow $Q_{\text{out}}(t)$ is proportional to the MFD travel production $P(n(t))$, where the proportionality factor is the inverse of the average trip length $L$ (in [m]) in the reservoir, as detailed in equation 3.2. This means that all vehicles have almost the same travel distance whatever their origins and destinations.

$$Q_{\text{out}}(t) = \frac{P(n(t))}{L} \tag{3.2}$$

When capacity restrictions are applied to the exits, the outflow formulation must be updated to account for the exogenous boundary condition (total supply) $Q^S_{\text{out}}$; see equation 3.3. Similarly, the total demand inflow $Q^D_{\text{in}}$ may be reduced to adjust the reservoir supply function $I(n(t))$, see equation 3.4. Note that $Q_{\text{in}}(t)$ is set at the maximal demand value during the whole period during which a stock of vehicles is waiting to enter the reservoir due to capacity supply regulation. Note also that we do not consider internal flow generation in the reservoir as this does not occur in the mescoscopic simulation.

$$Q_{\text{out}}(t) = \min\left[\frac{P(n(t))}{L}; Q^S_{\text{out}}\right] \tag{3.3}$$

$$Q_{\text{in}}(t) = \min\left[I(n(t)); Q^D_{\text{in}}\right] \tag{3.4}$$

The definition of the supply or entrance function $I(n(t))$ is introduced in Geroliminis & Daganzo (2007) and its shape is presented in Figure 3.3(b). The role of this function is to simulate spillbacks reaching the reservoir entry when studying transfer flows through a reservoir. However, considerations of its use and shape are still being discussed in the literature[1]. Thus, as there is no clear consensus on this question and to keep our model simple, we chose to adopt the definition of Knoop & Hoogendoorn (2014), Hajiahmadi *et al.* (2013b), Lentzakis *et al.* (2016), which is based on the supply function of a cell in the Cell Transmission Model (CTM) of Daganzo (1994), see equation 3.5. Since our goal is to study steady state flows in this chapter, this choice does not have any impact on the final results. We will actually see that the role of $I(n(t))$ is basically to ensure that the equilibrium flow is equal to the total supply $Q^S_{\text{out}}$ in congestion for the MFD-based simulation.

$$I(n) = \begin{cases} P_c/L = \max O(n) & \text{if } n < n_c = \operatorname{argmax} O(n) \\ O(n) & \text{else} \end{cases} \tag{3.5}$$

where $O(n) = P(n)/L$ is the outflow function, and $P_c$ the maximum network production estimated with the MFD $P(n)$. $O(n)$ may be also referred to as the outflow-MFD, while $P(n)$ is denoted the production-MFD to avoid confusion. Such a distinction was proposed in Lamotte & Geroliminis (2018).

Figure 3.4 synthesizes the MFD-based modeling approach. For a given and constant set of boundary conditions $\left(Q^D_{\text{in}}; Q^S_{\text{out}}\right)$, the network steady state can be directly calculated by solving $Q_{\text{in}}(t) = Q_{\text{out}}(t)$. Note that at this aggregated level, neither the total demand $Q^D_{\text{in}}$ nor the total supply $Q^S_{\text{out}}$ depends on the heterogeneity coefficients $a_1$ and $a_2$. This is because the MFD-based model considers homogeneous perimeter flows. Thus in this model,

[1]Most large-scale control applications do not even consider it because of the feedback loop introduced with the controller, thus protecting the reservoir from oversaturated states. But in our case, $I(n(t))$ is needed to simulate congested steady states

the network steady state depends only on $(Q^D_{\text{in}}; Q^S_{\text{out}})$, the production-MFD $P(n)$ and the average trip length $L$. However, a crucial question is whether $P(n)$ or $L$ may be affected by heterogeneous network loadings and therefore by $a_1$ or $a_2$. This will be investigated in the next section.

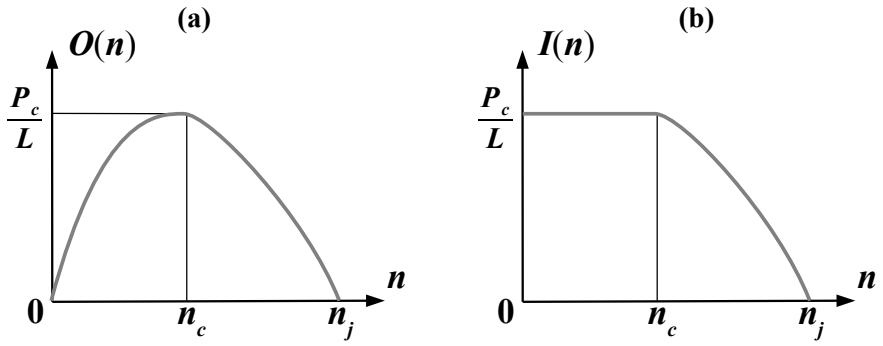



FIG. 3.3 – *(a) Typical shape of the outflow function $O(n)$ and (b) the supply or entrance function $I(n)$*

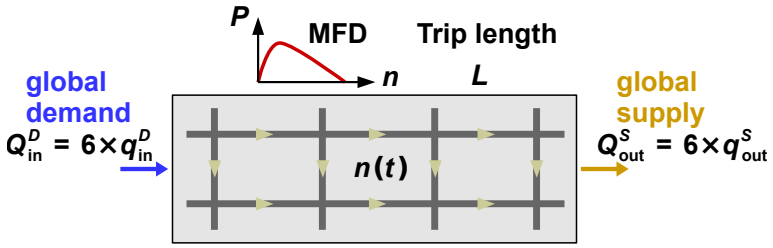



FIG. 3.4 – *MFD-based model definition*

## 3.4 Influence of the boundary settings on the MFD and the average trip length

### 3.4.1 MFD estimation based on the mesoscopic outputs

The production-MFD $P(n)$ is estimated by simulating different traffic loadings with the mesoscopic framework. One simulation corresponds to one particular loading, and the steady state reached after the warm-up period defines a single point in the (accumulation, production) plane. The different states for the free-flow part of the MFD are obtained from simulations with local demand values varying from $q^D_{\text{in}} = 0.1$ to 0.5 veh/s and a supply value fixed to $q^S_{\text{out}} = 1$ veh/s at each exit. On the other hand, the dots for the congested part of the MFD are obtained by varying the supply at the exits from $q^S_{\text{out}} = 0.1$ to 1 veh/s and fixing the demand at its highest level $q^D_{\text{in}} = 0.5$ veh/s at each entry. These two simulation settings allow us to almost completely estimate the network MFD.

For one simulation, the steady state accumulation $n$ is calculated as the mean difference between the entering and exiting cumulative count curves during the time period [1000; 5000 s]. The count curves are direct outputs from the simulation because the mesoscopic

simulator provides all the entering and exiting times for all vehicles. On the other hand, the steady state production $P$ is calculated with Edie's formula (Edie, 1963):

$$P = \frac{1}{\Delta T} \sum_{i=1}^{N_{\text{veh}}} td_i \tag{3.6}$$

where $\Delta T$ is the duration of the aggregation period [1000; 5000 s], $N_{\text{veh}}$ the total number of vehicles circulating during the same period, and $td_i$ the distance traveled by the $i$th vehicle during this period.

The MFD is estimated for homogeneous loadings (SC0) but also for different degrees of heterogeneous loadings (SC1 and SC2), i.e. when coefficient $a_1$ varies from 0 to 0.8, and $a_2$ from 0 to 0.6. The resulting MFDs and their fits by a semi-parabola and semi-linear function can be seen in Figures 3.5(a) and (b). A clear impact of the heterogeneity coefficients appears on the MFD shape. Figure 3.5(a) shows the results for SC1. It can be seen that the more heterogeneous the demand on entries is, the lower the MFD capacity is. This is because the flows concentrate on a few entries when the demand distribution becomes heterogeneous. In this case certain OD pairs are almost not used. This has no effect when the demand rates are low, but creates uneven density distribution between the links when the demand reaches its highest level. This uneven density distribution is therefore responsible for the decrease in the network's global capacity, as shown previously by Geroliminis & Sun (2011a), Knoop *et al.* (2015) in bigger networks. Zhang *et al.* (2013) also studied the influence of anisotropic demand settings at network boundaries on the MFD shape. Likewise, their simulations show a decrease in the network capacity when anisotropy exists in the demand distribution.

On the other hand, heterogeneity among the supply values at the exits does not affect the MFD capacity, see Figure 3.5(b). This is because the network capacity is reached for network loadings where the exit flows are not limited, whatever the heterogeneity coefficient $a_2$ is. However, a clear decrease of the average flow in the congested part is observed. This underlines that traffic states become even worse as congestion increases, when the heterogeneity coefficient $a_1$ is high. Reducing outflow at some exits may lead to severe congestion on several network links, which may also affect other exits when spillbacks occur.

Figures 3.5(c) and (d) show the increase of spatial heterogeneity among link traffic states when the spatial mean density increases for both scenarios SC1 and SC2. These two graphs are based on the same data points we used for the MFD estimations. For one simulation, link densities are calculated as the mean of link accumulation over the period [1000; 5000 s], divided by the link length. Then, the spatial mean density $k$ and the spatial link heterogeneity $\sigma_k$ respectively correspond to the arithmetic mean and the standard deviation of all these link densities. This definition of spatial link heterogeneity has already been used in Zhang *et al.* (2013), Zhao *et al.* (2014), Knoop *et al.* (2015). Figure 3.5(c) shows the results for SC1. The more heterogeneous the demand distribution is, the more heterogeneous the link traffic states. On the contrary, Figure 3.5(d) shows the results for SC2. It obviously appears that high heterogeneity coefficients for supply values lead to highly heterogeneous link densities in congestion. In free-flow however, the heterogeneity coefficient $a_2$ has no impact, see also Figure 3.5(b). In both figures, a linear trend can also be seen between the spatial mean density $k$ and the spatial heterogeneity $\sigma_k$. This trend was noticed by Knoop *et al.* (2012), and by Mahmassani *et al.* (2013a), Kim & Mahmassani (2015) about the variability in travel times.

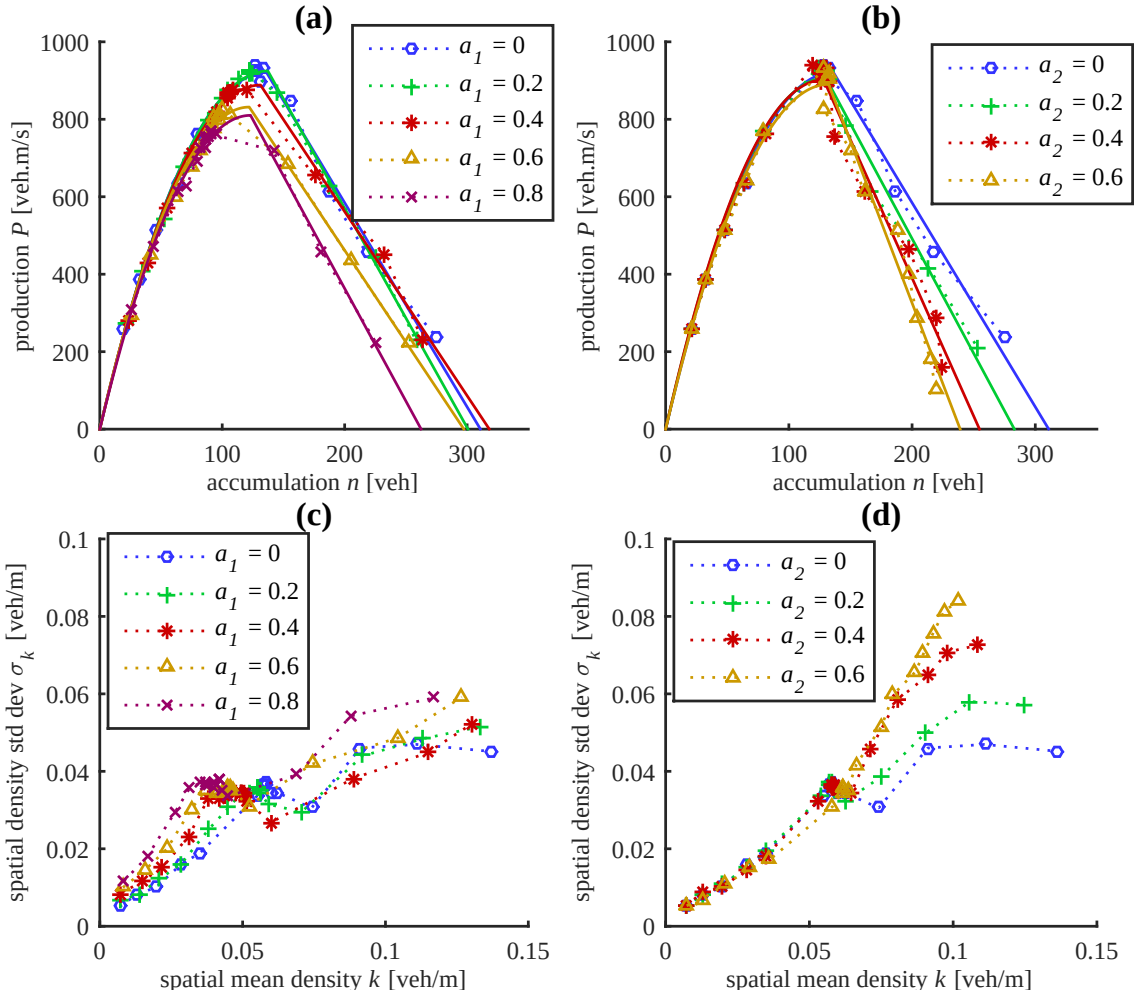



FIG. 3.5 – *(a) MFD estimation for SC1 (heterogeneous demand / homogeneous supply), $a_1$ = 0 to 0.8. (b) MFD estimation for SC2 (homogeneous demand / heterogeneous supply), $a_2$ = 0 to 0.6. (c) Spatial heterogeneity (standard deviation) of link densities vs spatial mean link density in SC1 with $a_1$ = 0 to 0.8, and (d) in SC2 with $a_2$ = 0 to 0.6*

### 3.4.2 Mean trip length estimation based on the mesoscopic outputs

The average trip length $L$ can be estimated in two ways. First, the mesoscopic model provides the trip distances directly for all exiting vehicles. The average trip length for a given simulation can then be calculated as the arithmetic mean of these distances during the observation period. This estimation is denoted $L_{\text{simul}}$ as it refers to one single simulation setting. The second method consists in determining the slope of the linear fit between the outflow and the production values; see equation 3.2. For each level of heterogeneity, i.e. a couple of $a_1$ and $a_2$ values, a linear regression is applied to all the simulation results to estimate the mean trip length, here referred to as $L_{\text{MFD}}$ because it corresponds to one MFD estimation.

$L_{\text{simul}}$ values are presented in Figures 3.6(a) and (b) for the two scenarios SC1 and SC2 respectively. For each simulation, $L_{\text{simul}}$ is given versus the corresponding steady state accumulation. Figures 3.6(c) and (d) show the same data points, but versus the heterogeneity coefficient $a_1$ (resp. $a_2$). The network state, i.e. free-flow or congested, related to the MFD estimation (see section 3.4.1) is mentioned in the figure. The second estimation $L_{\text{MFD}}$ is

also presented in these two graphs for each heterogeneity coefficient, including the 95% confidence interval of the slope estimation in the linear regression. It clearly appears that the average trip length greatly depends on the boundary conditions and may have significant variability (from 350 to 500 m). An interesting observation is that although the confidence intervals are small for the $L_{\text{MFD}}$ values, the second estimation method proved to be insufficient to capture the great variability that may exist between the different simulations. Other network-specific trends of lesser importance can also be noticed. Such observations are more specific to the network studied and should not be expected as universal results. First, one counter-intuitive phenomenon is that the average trip length decreases as congestion becomes heavier, because the inflow is more reduced drastically for the longest routes when congestion propagates. Second, the heterogeneity in the demand distribution at entries has a clear impact on the average trip length in free-flow: Figures 3.6(a) and (c) show that higher heterogeneity coefficient values increase the user flow on longer routes, see also Figure 3.2(b). On the other hand, in Figures 3.6(b) and (d) the average trip length is independent of the heterogeneity coefficient in free-flow ($a_2$ characterizes the heterogeneity on supply). This is because the demand is homogeneous in scenario SC2. The impact of these differences in MFD and/or trip length estimations on the reservoir simulation will be examined next.

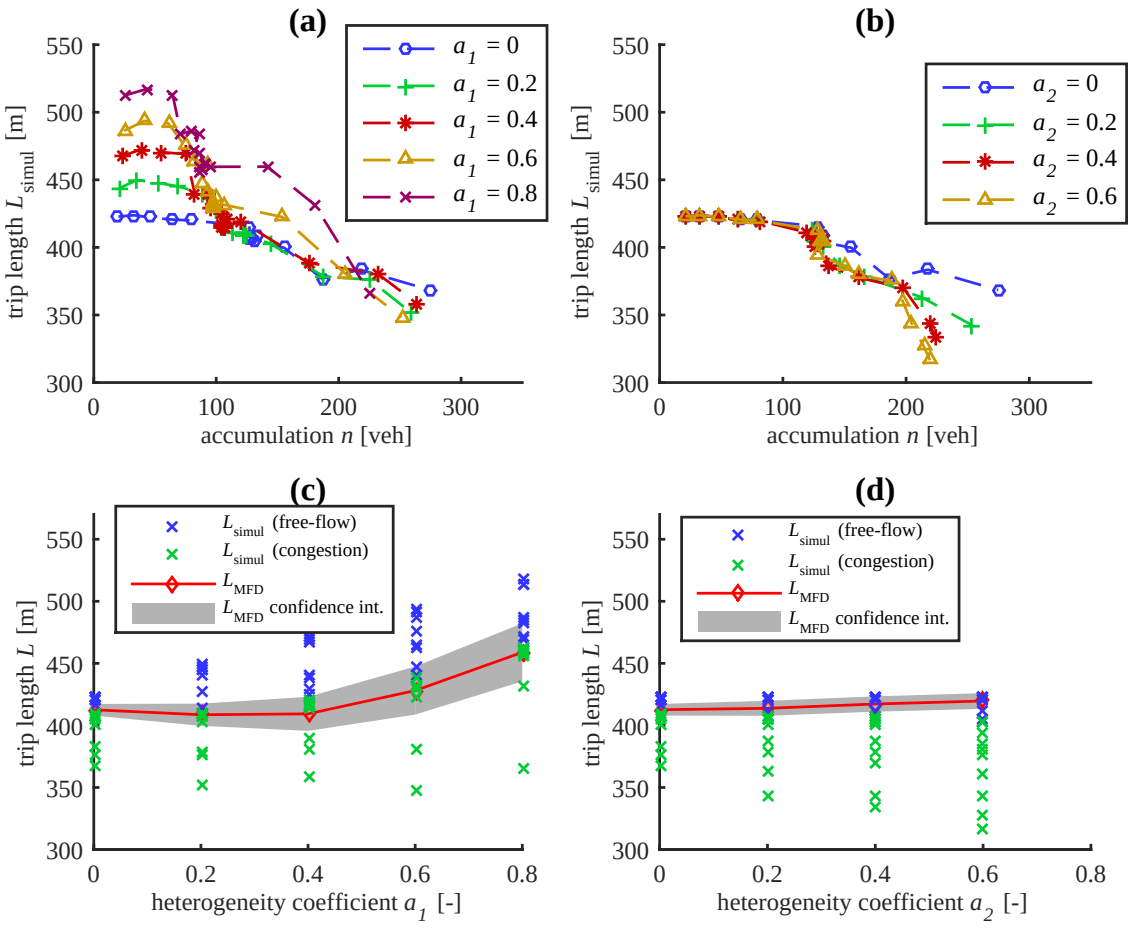



FIG. 3.6 – *(a) Average trip length $L_{simul}$ versus accumulation for SC1 (heterogeneous demand), $a_1$ = 0 to 0.8. (b) $L_{simul}$ vs accumulation for SC2 (heterogeneous supply), $a_2$ = 0 to 0.6. (c) $L_{simul}$ and $L_{MFD}$ vs $a_1$ in SC1. (d) $L_{simul}$ and $L_{MFD}$ vs $a_2$ in SC2*

## 3.5 Comparisons of MFD and mesoscopic simulation results for different network loading cases

### 3.5.1 Comparison between mesoscopic and MFD-based approaches for homogeneous loadings

In this section we present the simulation results on a simple network loading case: the simulation starts with an empty network and then reaches a steady state defined by the demand level in free-flow, or the supply level in congestion. Both levels are constant during the whole simulation period. We first focus on two typical network loadings: (i) the free-flow case, defined by $q_{\text{in}}^D = 0.2$ veh/s and $q_{\text{out}}^S = 1$ veh/s; (ii) the congested case, characterized by $q_{\text{in}}^D = 0.5$ veh/s and $q_{\text{out}}^S = 0.1$ veh/s. These are the reference values for one entry or one exit in SC0. The effective local inflows and outflows are then derived based on the $a_1$ and $a_2$ coefficients in SC1 and SC2, see section 3.3.1. At the aggregated (reservoir) level, we have $Q_{\text{in}}^D = 1.2$ veh/s and $Q_{\text{out}}^S = 6$ veh/s for free-flow, and $Q_{\text{in}}^D = 3$ veh/s and $Q_{\text{out}}^S = 0.6$ veh/s for congestion.

We first compare the simulation results between the mesoscopic and MFD-based approaches for the homogeneous case SC0, when $a_1$ and $a_2$ are set to 0. The evolution of accumulation and $N$-curves are presented in Figures 3.7(a) and (b) for the free-flow situation, and in Figures 3.7(c) and (d) for the congested situation. Overall, a good concordance is observed between the two modeling frameworks in steady state. Of course, the traffic description at the aggregated (reservoir) level cannot capture all the variability which is rendered at the lower (link) level by the mesoscopic simulator. Note that the scattering is considerable in congestion for the mesoscopic results.

### 3.5.2 Comparison between mesoscopic and MFD-based approaches for heterogeneous demand and homogeneous supply loadings – scenario SC1

We now focus on the impact of heterogeneity in SC1, by setting the $a_1$ value to 0.8 (highly heterogeneous distribution) and $a_2$ to 0 for both free-flow and congested situations. Several options are available for calibrating the reservoir model, i.e. the MFD shape and the mean trip length value. We can either consider that the reservoir model has been calibrated for the homogeneous case, i.e. when $a_1 = 0$ and that we are not aware of the change of the demand distribution; or we can re-calibrate the reservoir model using heterogeneous data observations. We consider updating the MFD shape and the trip-length separately, which makes four cases in total: the estimated MFD with $a_1 = 0$ (initial calibration), or with $a_1 = 0.8$ (re-calibration), and the estimated average trip length when $a_1 = 0$ (estimation by regression $L_{\text{MFD}}$), or $a_1 = 0.8$ (the estimation from the mesoscopic simulation results $L_{\text{simul}}$). These calibration cases are summarized in Table 3.2. The results for these four calibration options are presented in Figures 3.8(a) and (b) for the free-flow and congested cases respectively. It appears that the reservoir model fails to reproduce the mean steady state value given by the mesoscopic simulator if no re-calibration is performed on the MFD and the average trip length. On the other hand, the steady state is properly estimated when the two components of the model are fully re-calibrated. This demonstrates how sensitive the

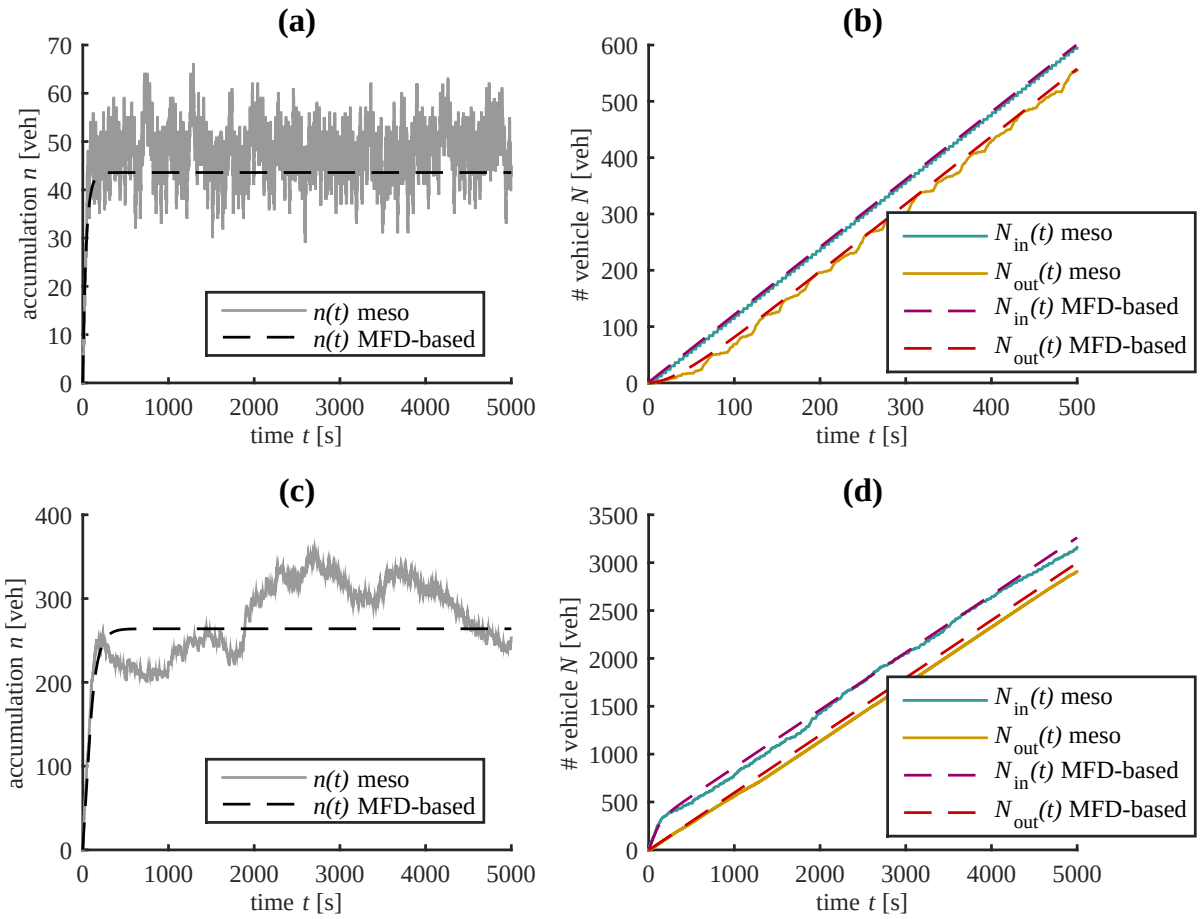



FIG. 3.7 – *Comparison between the mesoscopic and the MFD-based framework when $a_1 = 0$ and $a_2 = 0$. (a) Evolution of accumulation $n(t)$ for the free-flow case (b) $N_{\text{in}}(t)$ and $N_{\text{out}}(t)$ for the free-flow case (the time window was reduced to improve the graph's readability). (c) Evolution of accumulation $n(t)$ for the congested case (d) $N_{\text{in}}(t)$ and $N_{\text{out}}(t)$ for the congested case*

reservoir model is to local changes in demand distribution, which requires re-calibration. Figure 3.8(a) shows that the improvements gained by re-calibrating the trip length only are much better than re-calibrating the MFD only for this case. This suggests that the average trip length is indeed a key parameter that should be updated as soon as the mobility pattern changes in the reservoir. This can be seen in Figure 3.6(a), where $L_{\text{simul}}$ varies greatly with respect to $a_1$, whereas the free-flow part of the MFD is not really influenced by $a_1$; see also Figure 3.5(a). The opposite observation can be made for the congested case; see Figure 3.8(b). The main improvement comes from the MFD calibration. This can be explained when looking at Figure 3.5(a) and 3.6(a). The trip length is less sensitive than the MFD to $a_1$ for congested situations. A synthesis of the improvements due to the different possible calibration methods is presented in Table 3.3.

It should be noted that the relative error on the steady state accumulation predicted by the reservoir model can be directly calculated for a given MFD, $L$ and the boundary conditions $\left(Q_{\text{in}}^{D}; Q_{\text{out}}^{S}\right)$. In fact, the steady state related to equation 3.1 can be analytically determined without resorting to simulation. We use equation 3.2 to estimate the error propagation in the reservoir model. Assuming that there is an absolute error $\Delta P$ on the steady state

TAB. 3.2 – *The four calibration cases for the MFD model*

| Calibration case | **no calibration** | **MFD calibration** | $L$ **calibration** | **both calibrations** |
|---|---|---|---|---|
| MFD used | MFD $(a_1 = 0)$ | MFD $(a_1)$ | MFD $(a_1 = 0)$ | MFD $(a_1)$ |
| Trip length used | $L_{\text{MFD}}(a_1 = 0)$ | $L_{\text{MFD}}(a_1 = 0)$ | $L_{\text{simul}}(a_1)$ | $L_{\text{simul}}(a_1)$ |

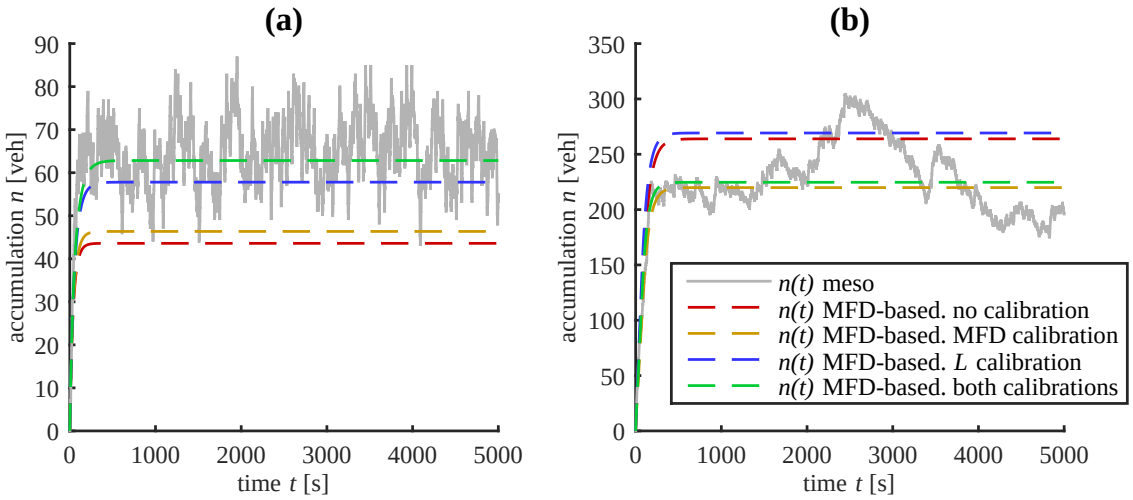



FIG. 3.8 – *Comparison between the mesoscopic and the MFD-based framework when $a_1 = 0.8$ and $a_2 = 0$ – SC1. Different calibrations are considered for the MFD-based model: MFD and/or mean trip length calibrated from data corresponding to $a_1 = 0$ or $a_1 = 0.8$. (a) Evolution of accumulation $n(t)$ in free-flow, and (b) evolution of accumulation $n(t)$ in congested conditions*

production $P$ and an absolute error $\Delta L$ on the average trip length $L$, the absolute error $\Delta O$ on the steady state outflow $O$ is:

$$\Delta O = \frac{1}{L}\Delta P - \frac{O}{L}\Delta L \tag{3.7}$$

This error is related to the absolute error $\Delta n$ on the steady state accumulation $n$ via the derivative $O'(n)$ of the outflow function, which satisfies $|O'(n)| = \Delta O/\Delta n$ in the first approximation. Then, for the free-flow situations, we can neglect $\Delta P$ and estimate $\Delta L$ at 100 m by referring to Figures 3.5(a) and 3.6(a). In free-flow steady state, the outflow value should be $O = Q_{\text{in}}^D = 1.2$ veh/s, and the value of $L$ corresponds to $L_{\text{simul}} = 515$ m given by the mesoscopic simulation, see Figure 3.6(a). The derivative $O'(n) = P'(n)/L = 0.012$ s$^{-1}$ is determined from the MFD plot for $n = n^{\text{meso}} = 66$ veh in Figure 3.5(a). As a result, the relative error $\Delta n/n^{\text{meso}}$ from the mesoscopic approach can be roughly estimated at 30%, which corresponds quite well to the error given in Table 3.3 when the MFD-based model is not re-calibrated. Similarly, for the congested situation, we can estimate $\Delta P = 230$ veh.m/s for $n = n^{\text{meso}} = 232$ veh, and neglect $\Delta L$. We obtain $L = L_{\text{simul}} = 370$ m in Figure 3.6(a) and $|O'(n)| = 0.016$ veh/s in Figure 3.5(a). Thus, the relative error $\Delta n/n^{\text{meso}}$ is estimated at 16%, which again provides a good prediction of the error given in Table 3.3 for congestion situations and the initial MFD-based model setting.

We now present the evolution of MFD-based errors for different heterogeneity coefficients varying from $a_1 = 0$ to 0.8 when $a_2 = 0$. The same free-flow and congested situations are studied, as described in section 3.5.1. The relative errors are calculated for the steady

TAB. 3.3 – *Comparison of accumulation values in steady state between the mesoscopic and the MFD-based approaches for scenario SC1 ($a_1$ = 0.8 and $a_2$ = 0). Lowest errors are indicated in bold face*

| value | mesoscopic | MFD-based | | | |
|---|---|---|---|---|---|
| | | no calibration | MFD calibration | $L$ calibration | both calibrations |
| $n$ (free-flow) [veh] | 66 | 44 | 46 | **58** | **63** |
| error from meso [%] | 0 | 33 | 29 | **12** | **4** |
| $n$ (congestion) [veh] | 232 | 264 | **220** | 269 | **225** |
| error from meso [%] | 0 | 14 | **5** | 16 | **3** |

state accumulation, and now also for the outflow. Note that the inflows and outflows have the same values in steady-state. The steady state outflow is estimated by calculating the mean slope of the exit count curve $N_{\text{out}}(t)$ over the analysis period [1000; 5000 s].

Figure 3.9 shows the relative errors in the MFD model in comparison with the mesoscopic one. It can be seen clearly that the higher the heterogeneity coefficient is, the higher the relative error. These results above all confirm that an accurate estimation of the mean trip length is essential in free-flow to obtain good simulation results from the MFD-based approach; see Figure 3.9(a). On the contrary, re-calibrating the MFD shape is crucial during congestion, see Figure 3.9(b). If the MFD shape and the mean trip length are not re-calibrated when the demand or supply distribution changes at the reservoir boundaries, simulation errors up to 30% can be observed. When re-calibration is performed for these two key elements, the errors fall below 5% for all levels of heterogeneity.

Figures 3.9(c) and (d) show very different patterns when considering outflow errors. They appear to be independent of the calibration methods. This is because the total flow in steady state is completely defined by the global demand $Q_{\text{in}}^D$ in free-flow, and the global supply $Q_{\text{out}}^S$ in congestion. Thus we would have expected a perfect estimation of the in- and outflows in steady state using the MFD-based approach whatever the calibration method applied. However, some small discrepancies (around or less than 5%) can still be observed in Figures 3.9(c) and (d). These are due to numerical approximations in the mesoscopic framework as the outflows are discretized in vehicle units.

### 3.5.3 Comparison between mesoscopic and MFD-based approaches for homogeneous demand and heterogeneous supply loadings – scenario SC2

We now focus on free-flow and congested network loadings with homogeneous demand distribution $a_1$ = 0 and heterogeneous supply distributions $a_2$ = 0.6, i.e. scenario SC2. The total demand and supply values are the same as in section 3.5.2. As in the previous section, mesoscopic and MFD-based simulation results are compared when the MFD and the mean trip length are re-calibrated or not with respect to the fully homogeneous case $a_1$ = 0 and $a_2$ = 0.

Figure 3.10 presents the evolution of accumulation for the mesoscopic and the MFD-based simulations with the four calibration cases. Figure 3.10(a) shows the results in the free-flow situation. It appears that the different calibration methods have no effect on the MFD-based simulation outputs, and that the steady state accumulation is rather well estimated

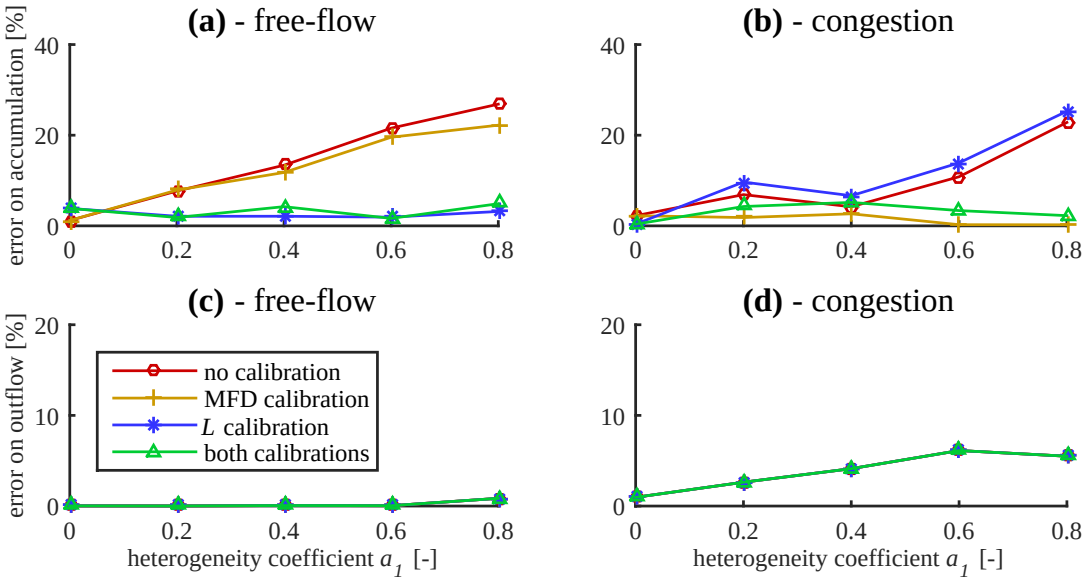



FIG. 3.9 – *Relative comparison of the MFD-based and the mesoscopic simulation results for SC1 ($a_2$ = 0). (a) Relative errors for steady state accumulation n for different heterogeneity levels, in free-flow and (b) in congestion. (c) Relative errors for steady state outflow $Q_{out}$ for different heterogeneity levels, in free-flow and (d) in congestion.*

in this case. This is not surprising because heterogeneous supply distributions only affect network loadings when exits act as bottlenecks, which is not the case in free-flow. This is confirmed by Figures 3.5(b) and 3.6(b), where neither the free-flow part of the MFD nor the estimated average trip length in free-flow depend on the heterogeneity coefficient $a_2$. The results are different in the congested situation; see Figure 3.10(b). The same conclusions are observed as in scenario SC1: re-calibrating the MFD allows reducing the error when estimating the steady state accumulation. The re-calibration of the average trip length has little impact in this case. This can be explained by the differences observed between cases $a_2$ = 0 and $a_2$ = 0.6 in Figure 3.5(b) for the MFD, and in Figure 3.6(b) for the mean trip length. A synthesis of the errors in steady state for SC2 is proposed in Table 3.4.

TAB. 3.4 – *Comparison of accumulation values in steady state between the mesoscopic and the MFD-based frameworks for scenario SC2 ($a_1$ = 0 and $a_2$ = 0.6). Lowest errors are indicated in bold face*

| value | mesoscopic | MFD-based | | | |
|---|---|---|---|---|---|
| | | no calibration | MFD calibration | *L* calibration | both calibrations |
| *n* (free-flow) [veh] | 49 | **43** | **44** | **45** | **46** |
| error from meso [%] | 0 | **11** | **10** | **8** | **7** |
| *n* (congestion) [veh] | 223 | 264 | **209** | 275 | **216** |
| error from meso [%] | 0 | 18 | **6** | 23 | **3** |

As in section 3.5.2, we now present the evolution of MFD-based errors for different heterogeneity coefficients varying from $a_2$ = 0 to 0.6 when $a_1$ = 0.

Figure 3.11 shows the relative errors in the MFD model in comparison with the mesoscopic one. The trends oppose each other as in scenario SC1 when considering free-flow and congested situations. In congestion, errors may increase up to 30% and almost vanish when

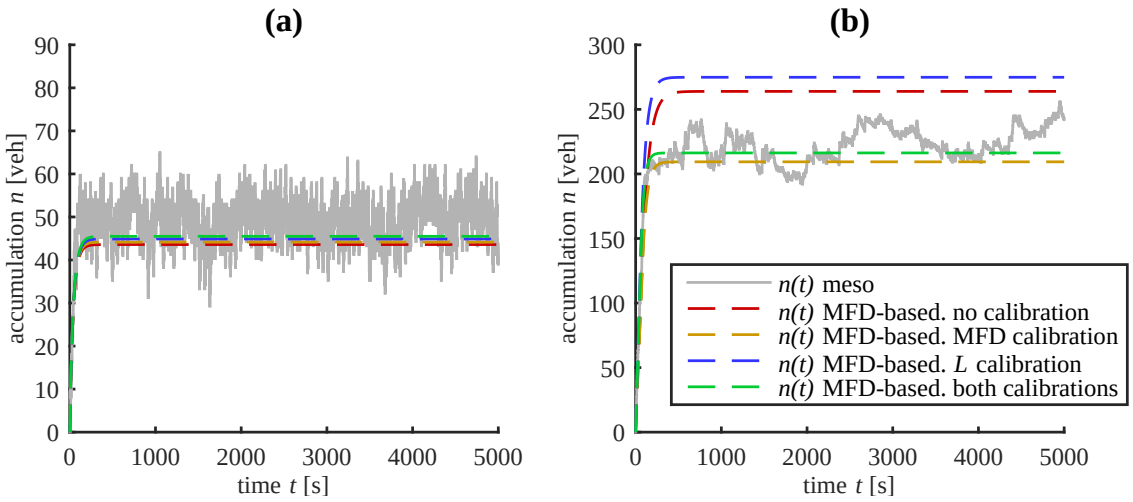



FIG. 3.10 – *Comparison between the mesoscopic and the MFD-based approaches for SC2 ($a_1 = 0$ and $a_2 = 0.6$). (a) Evolution of accumulation $n(t)$ in free-flow, and (b) evolution of accumulation $n(t)$ in congestion*

proper calibrations of both MFD and mean trip length are performed; see Figure 3.11(b). In free-flow, Figure 3.11(a) shows that there is almost no observable error because the heterogeneous supply distribution has no impact in this case.

As in SC1, almost no errors are observed for the outflow in free-flow conditions, see Figure 3.11(c). This is no longer the case for congestion as huge errors that increase with the heterogeneity coefficient (almost up to 150%) are observed in Figure 3.11(d). The phenomena behind this unexpected result will be described extensively in the next section. In brief, discrepancies between the mesoscopic and the MFD models can be explained by congestion patterns that propagate within the reservoir in the mesoscopic simulation. The total outflow $Q_{\text{out}}$ from the network falls below the global supply limitation $Q^S_{\text{out}}$ in the mesoscopic outputs, while $Q_{\text{out}}$ remains exactly equal to $Q^S_{\text{out}}$ in the MFD-based approach. The mesoscopic simulator is able to account for congestion spillbacks between internal intersections that induce flow reduction at several local exits. This phenomenon is not reproduced by the MFD-based approach as the perimeter flow is always considered as uniform. In fact, this phenomenon also influences the accumulation values in steady states, but this is hardly visible in Figure 3.11(b) as the differences in accumulation are relatively small for severe congestion. Moreover, the range of accumulation errors is of the same magnitude as the variability of the $n(t)$ outputs from the mesoscopic simulation. Unlike in SC1, this shows that, while being sufficient to accurately estimate the mean steady state accumulation, the proper calibration of the reservoir model (MFD shape and mean trip length) cannot ensure a good estimation of steady state flow in congestion.

### 3.5.4 Studying drops in total outflow due to internal congestion patterns

This section investigates in detail the congestion patterns that appear within the reservoir when using the mescoscopic simulator. We notably aim to propose a correction method capable of reducing the discrepancies observed in the outflow for the MFD simulation and that cannot be improved by simply re-calibrating the MFD or the trip length.

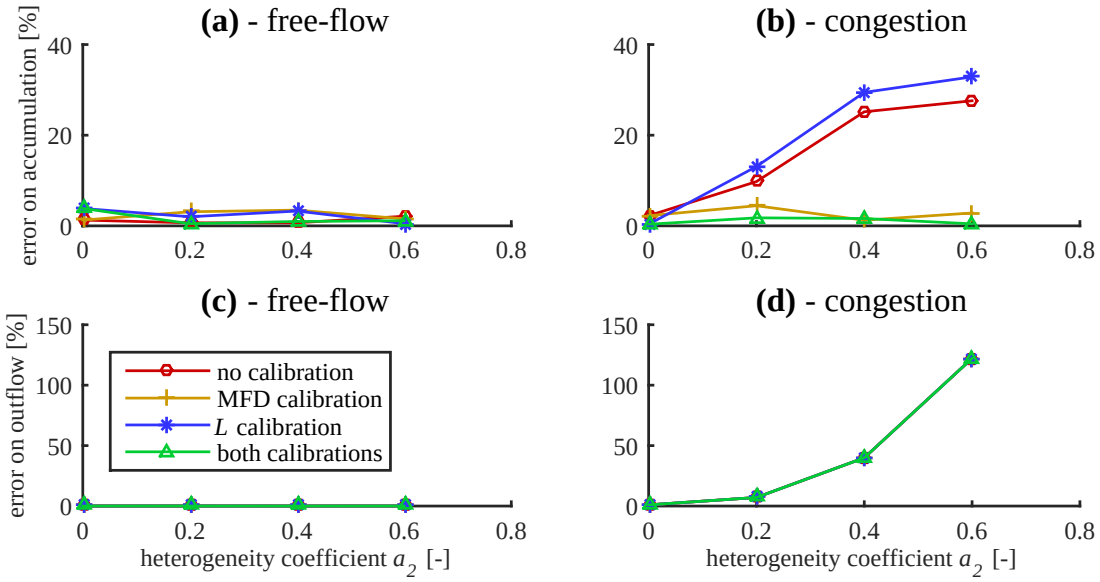



FIG. 3.11 – *Relative comparison of the MFD-based and the mesoscopic simulation results for SC2 ($a_1 = 0$). (a) Relative errors for steady state accumulation n for different heterogeneity levels, in free-flow and (b) in congestion. (c) Relative errors for steady state outflow $Q_{\text{out}}$ for different heterogeneity levels, in free-flow and (d) in congestion.*

Let us now focus on the evolution of the global outflow $Q_{\text{out}}(t)$ given by the mesoscopic simulator for different supply values at the exits. Both scenarios SC1 and SC2 are studied. Figure 3.12 shows the simulation results when the total demand $Q_{\text{in}}^D$ is equal to 3 veh/s and $Q_{\text{out}}^S$ = 0.6 to 6 veh/s. Figure 3.12(a) shows the results for SC1, and Figure 3.12(b) the results for SC2. In these figures, one point corresponds to the steady state global outflow $Q_{\text{out}}$ obtained in the mesoscopic simulation for a given global supply $Q_{\text{out}}^S$. Two reference lines have been added to the graphs: the MFD capacity (maximum outflow) obtained with homogeneous boundary conditions (SC0) and the first bisector that corresponds to total outflow equal to the total supply. Two very distinct situations are observed for scenarios SC1 and SC2. For SC1, the steady state outflow matches either the network capacity (gray area in the graph) or the available total supply $Q_{\text{out}}^S$. It should be recalled that the network capacity depends on the heterogeneity coefficient $a_1$ in SC1, see Figure 3.5(a). For SC2, the total outflow starts to decrease well before the supply becomes a constraint at the aggregated level. Furthermore, the difference between the total outflow and the available total supply increases with the level of heterogeneity. As already mentioned, this is the result of internal and local congestion patterns with spillbacks between intersections that reduce the local outflows even below the available local capacity.

A simulation example including internal congestion patterns that leads to a drop in outflow is presented in Figure 3.13. The $N$-curves corresponds to scenario SC2 with $a_2 = 0.6$. In the MFD-based modeling it can be seen that the steady state outflow corresponding to the $N$-curve slopes is $Q_{\text{out}}^S$ = 0.6 veh/s. However, the total outflow starts decreasing after $t$ = 1500 s in the mesoscopic simulator, and then stabilizes to a value of almost half of the total supply. This can be explained as follows. At the link level, due to heterogeneity in supply distribution, only a small fraction of vehicles can exit through D2; see Figure 3.2(c). This makes the route O2-D2 increasingly congested. Before the congestion propagates through

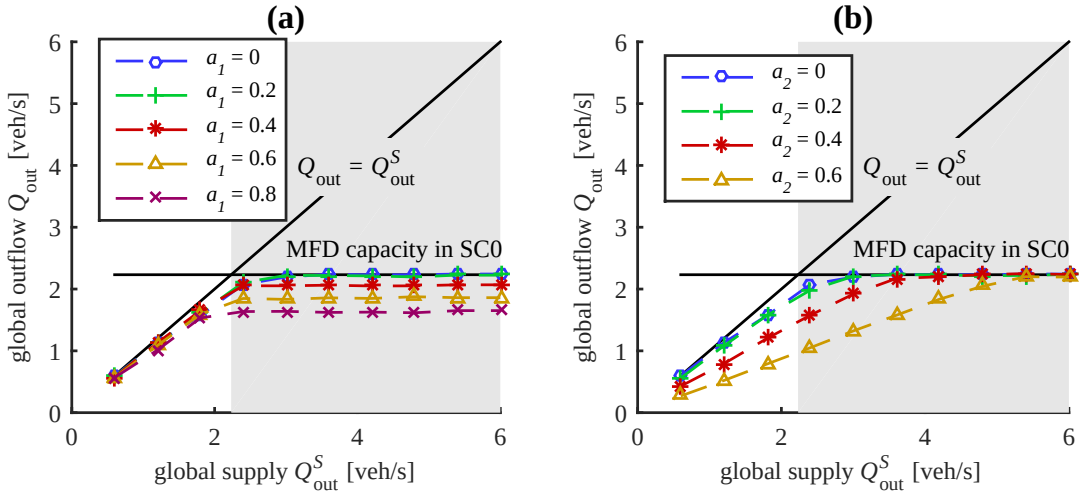



FIG. 3.12 – *Steady state total outflow versus total supply in the mesoscopic framework. (a) Results for scenario SC1 (heterogeneous demand distribution), and (b) scenario SC2 (heterogeneous supply distribution).*

the whole route, the outflow still corresponds to the maximum flow allowed for each exit. This is the situation at time $t_1$ for instance, for which the network state (density in each link) is described in figure 3.13(b). However, after a certain time (around 1500 s) route O2-D2 becomes fully congested. This makes all the users go to destinations other than D2, but travel on at least one link of the arterial O2-D2, thus exiting the network at a lower rate. As a result, the other exits D1 and D3 to D6 do not function at full capacity; see time $t_2$ in Figure 3.13(c). As a result, the total steady state outflow is below the sum of all the exit capacities.

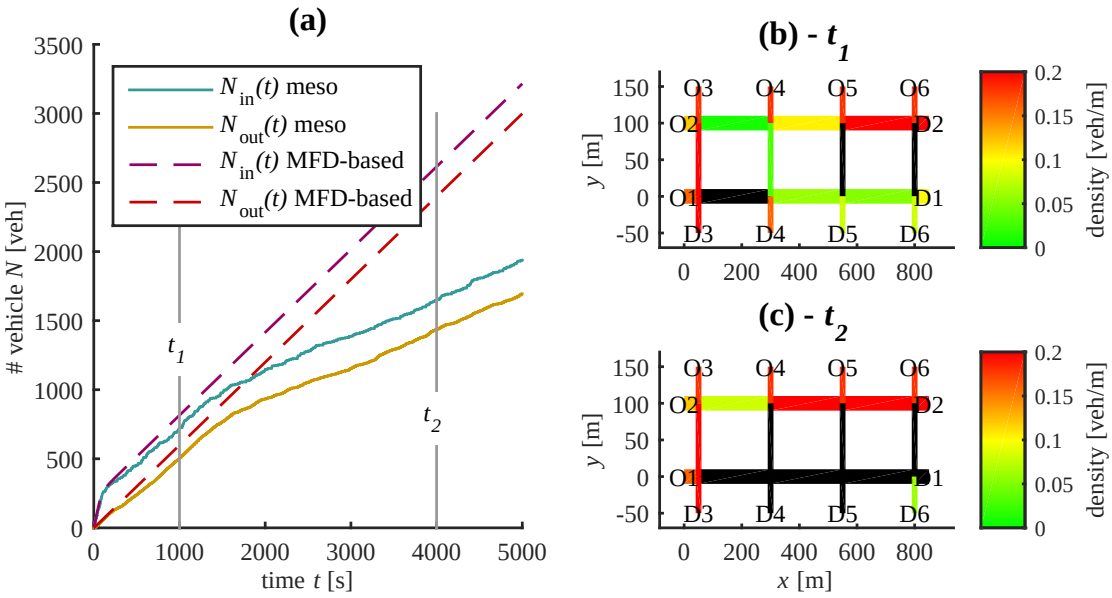



FIG. 3.13 – *An example of outflow drop due to a local internal congestion pattern in the mesoscopic framework – SC2 with $a_2 = 0.6$. (a) N-curves calculated by the mesoscopic and the MFD-based approaches, (b) traffic states at the link-level (density on each link calculated with the mesoscopic simulator, the black color meaning null density) at time $t_1$, and (c) at time $t_2$.*

We further analyze this internal congestion pattern to implement a possible correction in the MFD-based model. We observe that such a pattern is actually characterized by two

parameters: the time $T_c$ at which a slope break appears in $N_{\text{out}}(t)$, and the new value of the slope, i.e. the effective global outflow $Q_{\text{out}}^{\text{eff}}$, see figure 3.14(c). Note that this new value defines the steady-state situation. Several simulations were run in scenario SC2 with $a_2$ = 0 to 0.9 and $Q_{\text{out}}^S$ = 0.6 to 3 veh/s to study the evolution of these two parameters. The results are presented in Figure 3.14(a) for $T_c$ and Figure 3.14(b) for $Q_{\text{out}}^{\text{eff}}$. The results for $T_c$ are only given for $Q_{\text{out}}^S$ lower than 1.2 veh/s. This is because as the supply value increases, the drop in the total outflow occurs near $t = 0$ during the warm-up period, before the first equilibrium state has been reached. Thus for $Q_{\text{out}}^S > 1.2$ veh/s we can consider that $T_c$ is equal to 0.

When a drop in outflow is actually observed, we find that the two parameters $T_c$ and $Q_{\text{out}}^{\text{eff}}$ may be well approximated by a bilinear function of the heterogeneity coefficient $a_2$ and the global supply $Q_{\text{out}}^S$:

$$T_c\left(a_2, Q_{\text{out}}^S\right) = \begin{cases} \left(\alpha_t Q_{\text{out}}^S + \beta_t\right) a_2 + \gamma_t Q_{\text{out}}^S + \delta_t & \text{if } a_2 \geq 0.2 \text{ and } Q_{\text{out}}^S \leq 1.2 \text{ veh/s} \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

$$Q_{\text{out}}^{\text{eff}}\left(a_2, Q_{\text{out}}^S\right) = \left(\alpha_q Q_{\text{out}}^S + \beta_q\right) a_2 + \gamma_q Q_{\text{out}}^S + \delta_q \tag{3.9}$$

where $\alpha_t$ = 1.8×10$^3$ veh$^{-1}$.s$^2$, $\beta_t$ = -4.2×10$^3$ s, $\gamma_t$ = -2.1×10$^3$ veh$^{-1}$.s$^2$, $\beta_t$ = 4.9×10$^3$ s, $\alpha_q$ = -1.1, $\beta_q$ = 0.071 veh/s, $\gamma_q$ = 1.1, and $\delta_q$ = -0.051 veh/s. Note that the fit for $Q_{\text{out}}^{\text{eff}}$ is still consistent when $a_2$ = 0 (scenario SC0), as the bilinear function gives $Q_{\text{out}}^{\text{eff}}$ almost equal to $Q_{\text{out}}^S$.

As we now have a complete description of the effect of the congestion pattern on outflow, it is easy to implement a correction method in the MFD-based model to account for the outflow limitation $Q_{\text{out}}^{\text{eff}}\left(a_2, Q_{\text{out}}^S\right)$ after the time $T_c\left(a_2, Q_{\text{out}}^S\right)$. The result of this correction method is presented in Figure 3.14(d) for the same example we analyzed in Figure 3.13(a). It shows that the correction method greatly improves the MFD-based results compared to Figure 3.13(a).

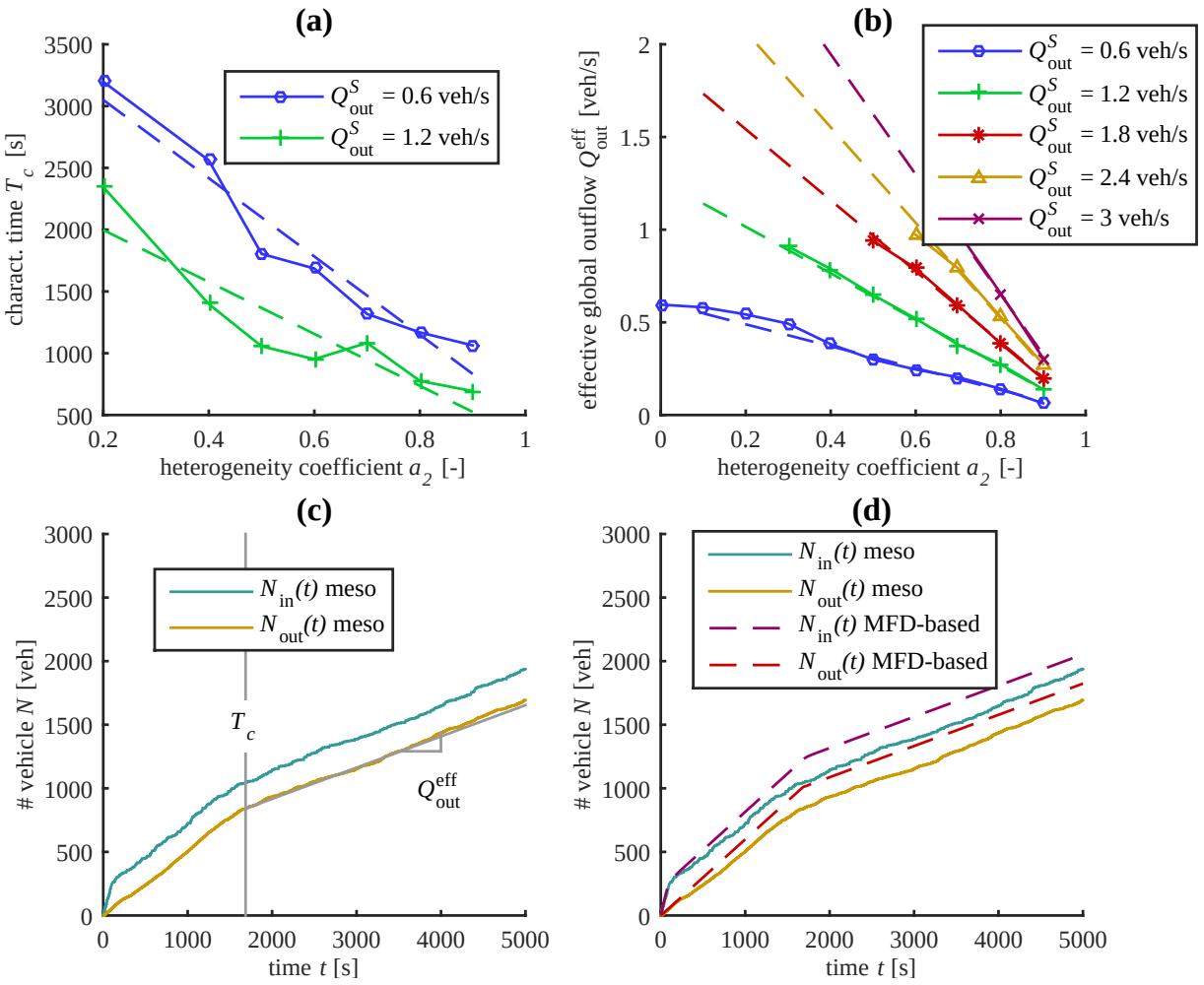



FIG. 3.14 – *(a) Evolution of time $T_c$ when an outflow drop is observed in the mesoscopic simulation, for different heterogeneity levels $a_2$ in scenario SC2 and different supply values $Q^S_{out}$. (b) Evolution of the effective steady state total outflow $Q^{eff}_{out}$ after the outflow drop in the mesoscopic simulation, for different $a_2$ and $Q^S_{out}$. (c) Definition of $T_c$ and $Q^{eff}_{out}$ on the mesoscopic N-curves. (d) Application of the correction method based on $T_c$ and $Q^{eff}_{out}$ to the MFD-based approach*

## 3.6 Discussion

In this study, we compared steady state simulation results for an MFD-based model and a mesoscopic model when considering a simple network. The MFD-based model describes traffic states at the aggregated (reservoir) level, while the mesoscopic tool provides a complete description of traffic dynamics at the link level inside the reservoir. The purpose of such a comparison was to test the robustness of the aggregated approach to different heterogeneous loadings at the network perimeter. Two heterogeneous scenarios were considered. SC1 assumes a heterogeneous demand distribution at the network entries, while the supply distribution is homogeneous. The reverse case, i.e. SC2, considers a heterogeneous supply distribution at the network exits, while the demand distribution is homogeneous. The results on the network studied showed that heterogeneity coefficients $a_1$ and $a_2$ have a strong impact on its MFD shape and its average trip length. While such impacts may not be universally expected for other network configurations, we showed that the proper calibration of these two crucial features of the MFD-based approach is essential to avoid errors (up to 30% in our case study) when predicting steady traffic states.

More interestingly and more generally, we highlighted a major difference between the two scenarios. In SC1, a proper re-calibration of the MFD and the mean trip length allows reducing the prediction errors completely (down to 5%). However, in SC2, the appearance of internal congestion patterns including spillbacks between intersections close to the reservoir perimeter makes the re-calibration almost inefficient. In this case, the MFD-based model fails to reproduce the network's internal dynamics. In particular, we noticed significant drops in the total outflow in the mesoscopic outputs that were not reproduced by the MFD-based approach. Analyzing congestion patterns with spillbacks using the mesoscopic simulator makes it possible to propose a correction method to account for the outflow drop in the MFD-based model.

With our simple network, it should be kept in mind that the congestion observed with highly heterogeneous supply values is emphasized by the constraints of our simple network, i.e. the users are forced to exit through a specific link. In a context of a multi-reservoir system, a dense and well-connected grid network would offer several re-routing options to avoid local congestion at the exit border of a reservoir. If such a border has many connections to the downstream reservoirs, the spillbacks and resulting outflow drops observed may be less severe in reality. Nevertheless, sparse connectivity between reservoirs is very likely to be observed in European type cities, in the case of bridges or major arterials that connect different urban areas. In such a configuration, the users are forced to exit through a specific connection point in their trip, unless major re-routing is imposed. Our study is specifically intended to question the robustness of MFD-based simulation when the capacities of such critical connection points are heterogeneously distributed along the reservoir border.

### 3.6.1 New results on a bigger network

As one of the most critical limitation of the present work is the relatively small size of the network, and thus the lack of redundancy and routing options, we provide additional results for a bigger grid network with the same kind of configuration, see Figure 3.15. This new topology consists of a 6×6 one-way Manhattan network with West-East links more than twice as long as North-South ones. The same scenarios SC1 and SC2 have been applied

to this network. Figure 3.16 shows clearly that the various boundary settings significantly influence the MFD shape. Nevertheless, a lesser influence on the trip length is noticed in this case of a more redundant network: the order of magnitude of its variations in free-flow is only around 7% for SC1.

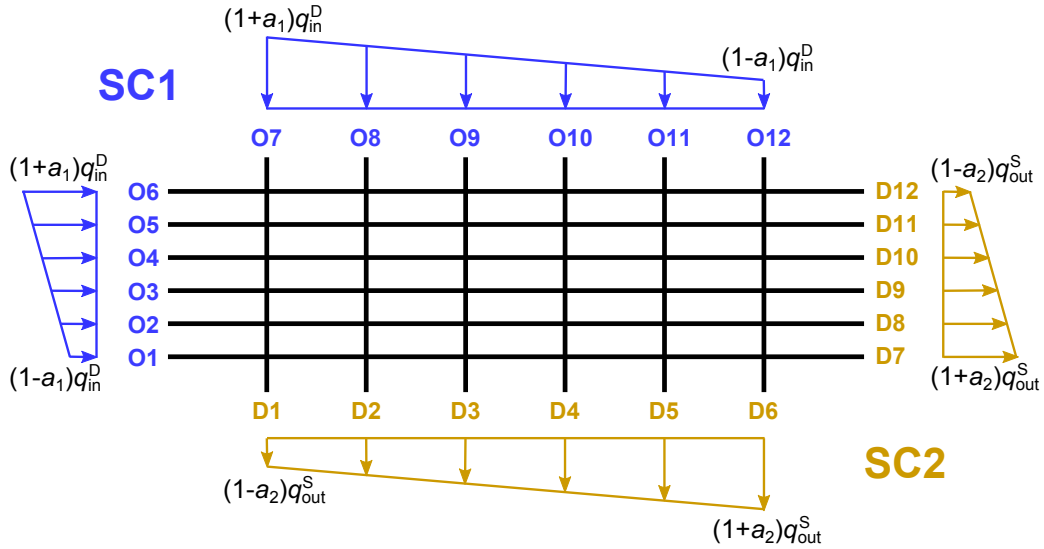



FIG. 3.15 – *New network configuration with scenarios SC1 and SC2 settings*

The relative errors in the MFD model compared with the mesoscopic simulations are presented in Figure 3.17 for SC1 and in Figure 3.18 for SC2. Regarding the steady state accumulation, both figures show that the errors can be reduced to acceptable values once the MFD and the trip length are well calibrated. Similar conclusions are found with this new network in free-flow, i.e. re-calibrating the trip length only improves the MFD simulation results much more than re-calibrating the MFD only. Regarding the steady state outflow, we observed huge errors for congested situations in both scenarios. In SC2, these errors can be explained by outflow drop phenomena, similar to what we observed in the 2×4 network, as Figure 3.19(b) suggests. However, the major difference with this 6×6 network is the significant error (more than 20%) in outflow made for the homogeneous case ($a_1$ = $a_2$ = 0). This is also due to an outflow drop, independent of the heterogeneity coefficient, as presented in Figure 3.19(a). This means that even homogeneous loadings can trigger internal congestion patterns in this bigger network. Such an outflow drop phenomenon was in fact impossible to present with our smaller network. While this new mechanism should be further investigated, a simple outflow correction method can still be implemented in the reservoir MFD model to account for it. Note that such a correction does not depend on the boundary settings in SC1, therefore the total effective outflow $Q_{\text{out eff}}$ can be estimated one for all as shown in figure 3.19(a), and then implemented in the MFD model to modify the outflow $Q_{\text{out}}$ during congestion. This is completely different from the heterogeneous supply case (SC2), where the correction method we propose actually depends on the boundary settings, see section 3.5.4.

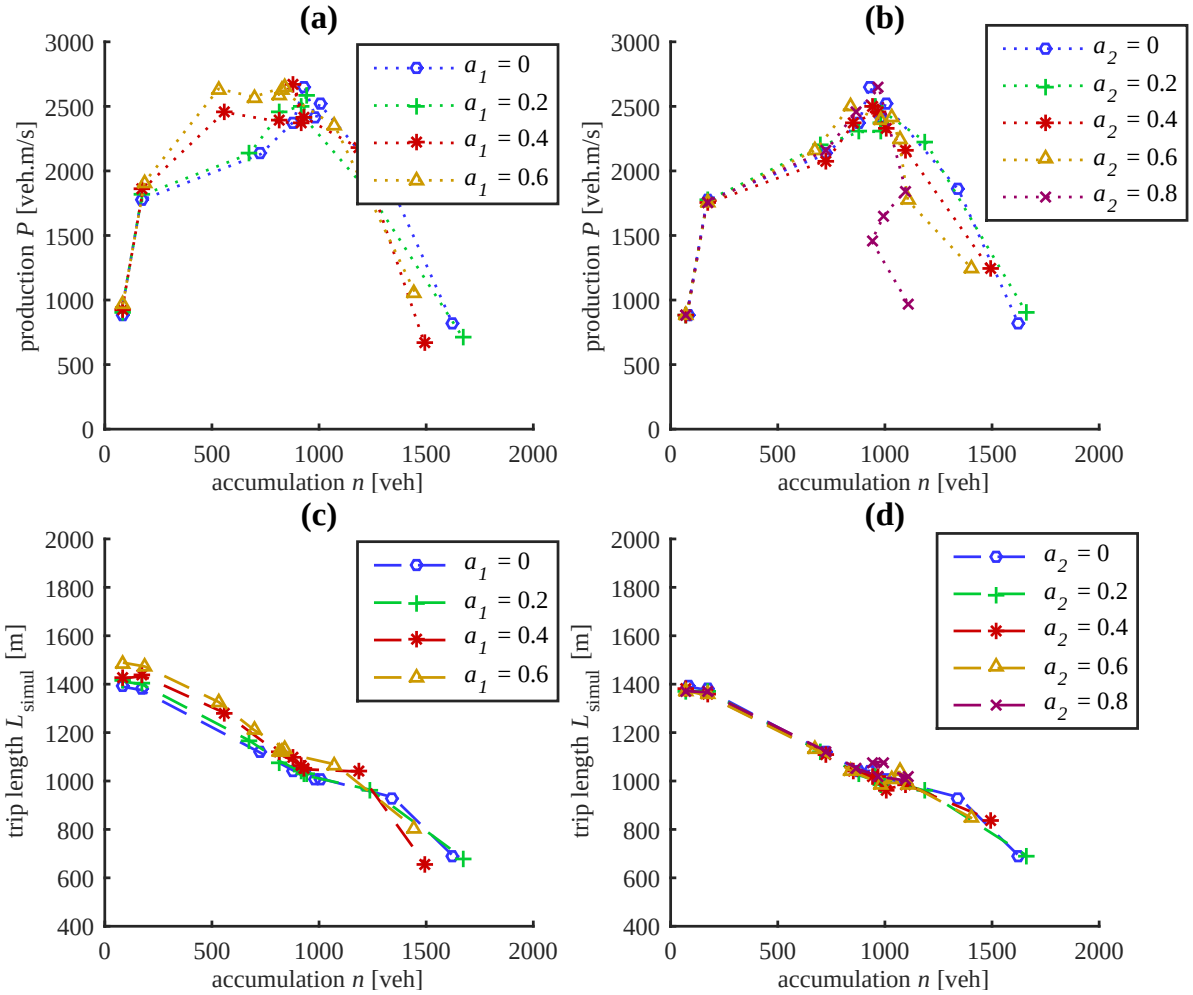



FIG. 3.16 – *(a) Network MFD for different coefficients $a_1$ in SC1 and (b) for different $a_2$ in SC2; (c) vehicle average trip length for different coefficients $a_1$ in SC1 and (d) for different $a_2$ in SC2*

### 3.6.2 General conclusion

While being specific to the kind of grid networks studied here, the results of this chapter permit highlighting possible shortcomings in MFD-based simulation due to heterogeneous link-level settings at reservoir boundaries. In particular, this chapter clearly shows that unlike the hypothesis made in the seminal paper on MFD, the OD matrix and the flow distribution between local entries and exits may matter when studying traffic dynamics at the network level. Large changes for these elements are likely to require re-calibrating both the MFD and the mean trip-length, especially for irregular and asymmetrical networks. Moreover, our conclusions show that in some cases, an accurate calibration may not be sufficient to reproduce the network dynamics well with the MFD-based model. More detailed knowledge of local capacity at exits, meaning a more detailed description of how congestion propagates from the links connected with the reservoir neighborhood, is required to catch the outflow drops that may occur. Note that this crucial issue has not been addressed when studying multi-reservoir systems (Yildirimoglu *et al.*, 2015) and thus will need further research. Interestingly, we showed that a simple correction method based on network observations in different configurations may be sufficient to update MFD-based simulation

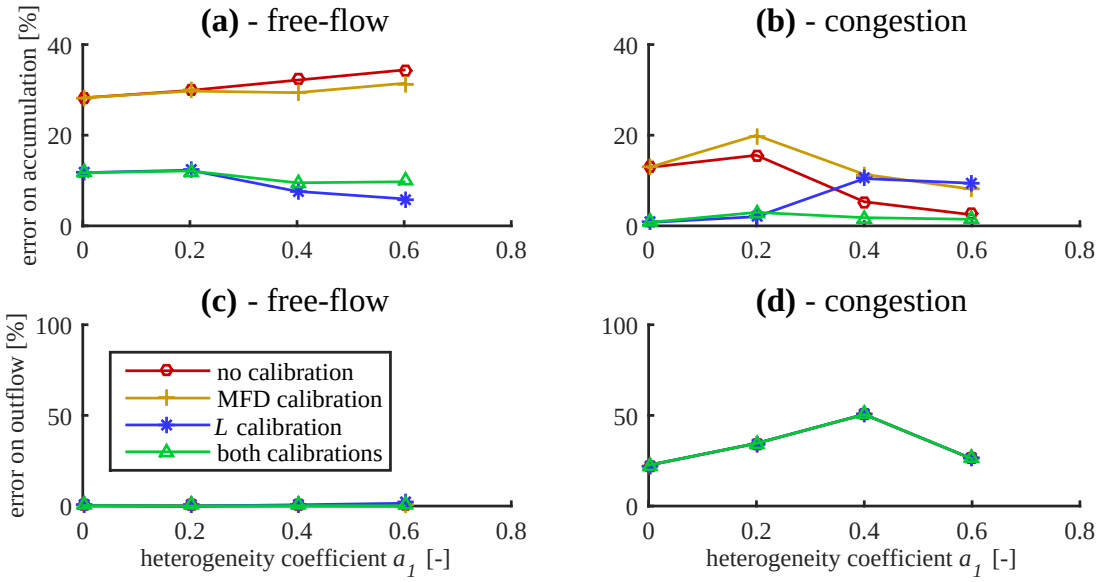



FIG. 3.17 – *Relative comparison of the MFD-based and the mesoscopic simulation results for SC1 ($a_2 = 0$). (a) Relative errors for steady state accumulation $n$ for different heterogeneity levels in free-flow and (b) in congestion. (c) Relative errors for steady state outflow $Q_{\text{out}}$ for different heterogeneity levels in free-flow and (d) in congestion.*

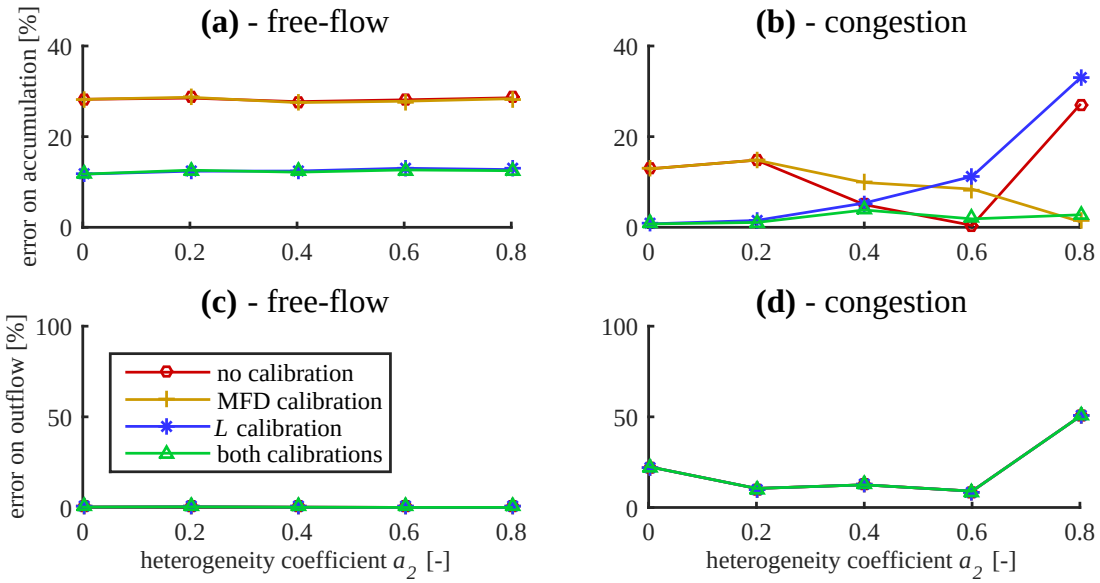



FIG. 3.18 – *Relative comparison of the MFD-based and the mesoscopic simulation results for SC2 ($a_1 = 0$). (a) Relative errors for steady state accumulation $n$ for different heterogeneity levels in free-flow and (b) in congestion. (c) Relative errors for steady state outflow $Q_{\text{out}}$ for different heterogeneity levels in free-flow and (d) in congestion.*

outputs.

Further research is needed to study the propagation of simulation errors in multi-reservoir systems to guaranty that such an approach can be a valid option for simulating large-scale urban areas and making traffic predictions. Efficient methods for re-calibrating the model on line should also be investigated.

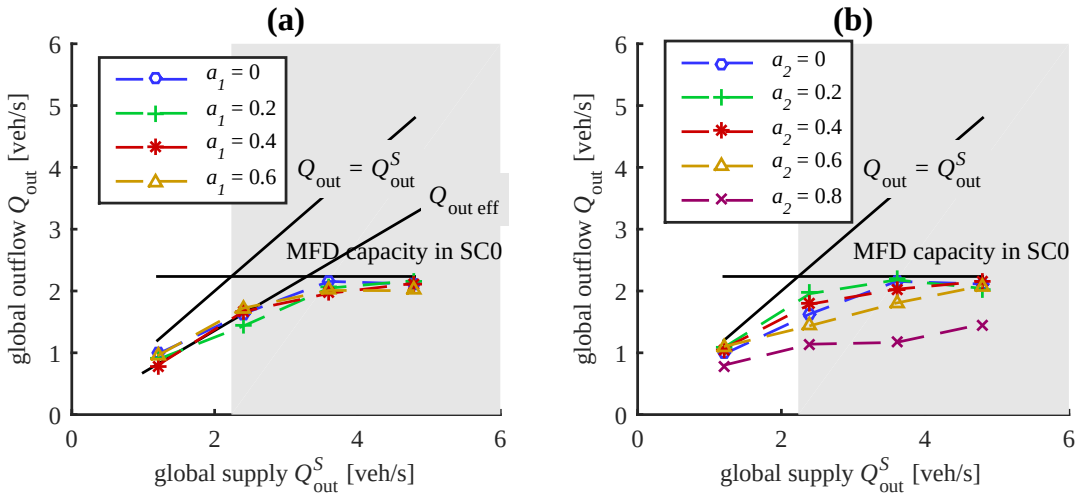



FIG. 3.19 – *(a) Total effective outflow vs total exit supply for different coefficients $a_1$ in SC1 and (b) for different $a_2$ in SC2*

# Conclusion of part I

In this first part, we investigated the internal traffic dynamics in the simplest modeling of a single reservoir (with one trip category). Regarding the theoretical modeling, thorough comparisons between the existing accumulation-based and trip-based approaches were made. While both models rely on the homogeneity hypothesis defined by the production- or speed-MFD, they fundamentally differ on the description of the reservoir outflow. In the accumulation-based model, the description of outflow is grounded on the outflow-MFD, whereas in the trip-based model, the outflow is implicitly defined by the mean speed and travel time evolution (the latter being notably defined through the average trip length). Because it accounts for the delay experienced when traveling from the reservoir entry to the exit, the trip-based model appears as a promising framework to run MFD simulation in fast-varying conditions (e.g. as during peak hours).

On the other hand, regarding the practical implementation of a simple MFD simulation, comparisons between link-scale mesoscopic simulations on a grid network and its accumulation-based modeling as a single reservoir were performed to assess the accuracy of the MFD simulation. It was shown that heterogeneous distribution of perimeter flows might generate uneven traffic conditions within the network that are critical for the MFD simulation. Traffic state estimated in the single reservoir model may notably exhibit significant bias as compared to the aggregation of link-level simulation outputs. In particular, it was confirmed that the MFD and the average trip length both depend on the perimeter demand and supply settings. Hence they are not pure intrinsic properties of the network only, but are also defined by the mobility patterns resulting from users own choices and adaptation to the network boundary conditions.

The contributions of this part can be summarized as follows:

- ✓ Proposition of numerical methods based on linear approximation to solve both the accumulation-based and trip-based models. In particular, an efficient event-based scheme is designed to solve the trip-based model with individual vehicle trajectories to represent flows in the reservoir.

- ✓ Identification of the validity domain for the accumulation-based description of traffic flow in a reservoir. This domain is characterized by the inaccurate representation of travel time in this model when traffic conditions evolve quickly.

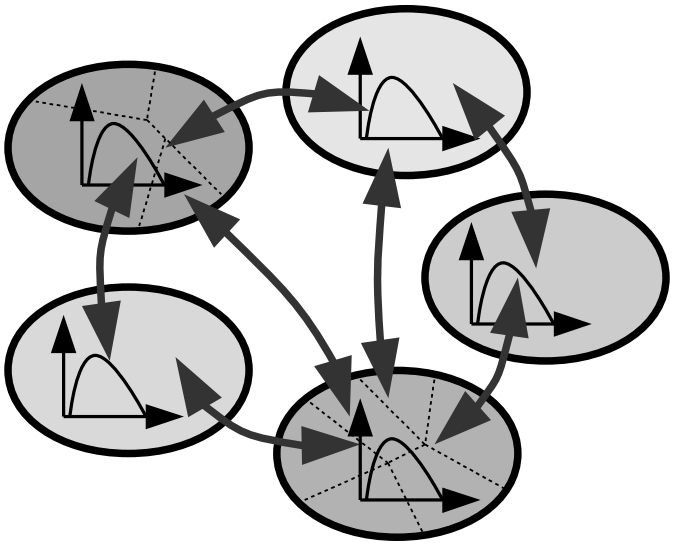
- ✓ Comparison between the accumulation-based modeling of traffic flow and the aggregation of link-level simulation outputs for a simple grid network. The study of different perimeter flow conditions shows the importance of the re-calibration of both the production-MFD and the average trip length to avoid significant bias in traffic states predicted by the MFD simulation.

The investigations conducted in this part have been focused on the single reservoir model with a unique average trip length. However, the literature on MFD models already provided new frameworks with multiple trip lengths inside the same reservoir. Thus, the further research directions should now investigate the traffic dynamics of a single reservoir with several trip lengths. Moreover, as the comparisons between accumulation-based and trip-based models were made in undersaturated conditions, the extension to oversaturation should also be included in our scope. All these research directions are then detailed below:

- Extend the investigation of the single reservoir dynamics to multiple trip categories, for both accumulation-based and trip-based models.
- Extend the study of MFD traffic dynamics in highly congested situations. This should include the extension of the trip-based framework to oversaturated conditions, and a proper understanding of inflow restriction and merging at the entry of a single reservoir.
- Further validate the proposed approaches by comparing MFD simulations with link-level simulation or real traffic data

# Part II

# Investigations on flow exchanges in multi-reservoir systems



# Introduction

This part aims at extending the investigations presented in the first part to the case of multiple trip categories in a single reservoir. Whereas still focusing on the single reservoir model, this part is heading towards the building of a consistent multi-reservoir simulation platform by representing the interactions between a reservoir and its neighbors with exogenous boundary conditions. In this perspective, close attention is paid to the modeling of flow exchanges at the perimeter (entry and exit) of the single reservoir, because in a multi-reservoir context all reservoirs interact with each other by sending or receiving flows.

This part is split into two different studies. The first one deals with the understanding of theoretical models including multiple trip categories that have been proposed in the literature. Several limitations are addressed, notably the fact that the proposed approaches are valid in undersaturation only, or not well applied in oversaturation. This first study thus develops a new framework to handle inflow merging and outflow diverging in a single reservoir. This framework is consistent for any traffic conditions, and applies for both accumulation-based and trip-based models. The second study is devoted to the validation of this framework and other approaches from the literature. To this end, the predictions of different MFD-based models are compared against link-level simulation results on a grid network.

# Contents

# 4. Investigations on flow exchanges in multi-reservoir and multi-route systems

Most of large-scale traffic flow models based on the Macroscopic Fundamental Diagram (MFD) are grounded on what has been called the "accumulation-based" approach by some authors. Recent studies have highlighted the limitations of such an approach in fast-varying conditions, and have focused on a better and more flexible representation of the reservoir inner dynamics, which has been referred to as "trip-based". However, when connecting several reservoirs together, the management of inflows and outflows to propagate congestion properly through the reservoirs is still a challenge in both modeling approaches. In particular, in saturated traffic conditions, handling inflow merge at a reservoir entry or outflow diverge at exit is of crucial importance as it governs the whole system evolution.

In this chapter, we propose a thorough analysis of the way congestion is usually handled in the accumulation-based framework. This serves as a basis to implement a proper congestion propagation model in the trip-based approach. Theoretical and simulation studies show that in case of several trip lengths in a reservoir, there exists only one form of inflow limitation at the reservoir entry that complies with the global constraints on flow and production. Moreover, the outflows from all trips exiting the same reservoir have to be all interdependent to satisfy the definition of the reservoir total exit production. This has strong implications when several reservoirs are connected together. We notably point out that the system can easily converge to global gridlock if an outflow management based on the most constrained exit of each reservoir is not adopted. Comparisons are given with the latest developments from the literature.

This chapter is a preliminary version of the paper:

> Mariotte, G. & Leclercq, L. (2019). Flow exchanges in multi-reservoir systems with spillbacks. *Transportation Research Part B: Methodological*, accepted

## 4.1 Motivations

Over the past decade, the Macroscopic Fundamental Diagram (MFD) has appeared to be a powerful tool to describe traffic states at the network level with few implementation and computational efforts. Many studies have notably used MFD-based traffic simulators for several promising applications, like traffic state estimation (where contributions can be found in Knoop & Hoogendoorn, 2014, Yildirimoglu & Geroliminis, 2014), perimeter control (Keyvan-Ekbatani *et al.*, 2012, Haddad, 2015, Haddad & Mirkin, 2017, Ampountolas *et al.*, 2017), route guidance at large scale (Hajiahmadi *et al.*, 2013b, Yildirimoglu *et al.*, 2015, Ding *et al.*, 2017), or analyzing cruising-for-parking issues (Leclercq *et al.*, 2017, Cao & Menendez, 2015, Zheng & Geroliminis, 2016). Their modeling approaches take advantage of the multi-reservoir representation of a city, where the dynamics of each urban reservoir (sometimes also called "neighborhood", "zone" or "region") is described by the single reservoir model of Daganzo (2007). This framework, also referred as the "accumulation-based" model, assumes that the reservoir outflow is proportional to the total circulating flow inside the zone if one consider a constant average trip length assigned to all travelers. Some authors have extended this approach to account for multiple trip lengths in a reservoir, either to develop new applications like modeling search-for-parking (Geroliminis, 2009, 2015) and macroscopic routing (with contributions in Yildirimoglu *et al.*, 2015, Ramezani *et al.*, 2015); or to highlight inaccuracies in MFD-based models due to the constant trip length hypothesis (Yildirimoglu & Geroliminis, 2014, Leclercq *et al.*, 2015). More recently, a "trip-based" formulation of the single reservoir model has gained a new interest in the community. Based on a idea of Arnott (2013), this approach has been exploited in Daganzo & Lehe (2015), then in Lamotte & Geroliminis (2018), Leclercq *et al.* (2017) and in chapter 2. The principle is that all users in a reservoir share the same space-mean speed (given by the MFD) at a given time, and exit once they have completed their individually assigned trip length. As shown in our thorough comparison with the accumulation-based model in chapter 2, the trip-based approach gives more accurate results during transient phases, especially in terms of travel time. Nevertheless, some authors like Haddad & Mirkin (2016), Haddad & Zheng (2017) suggest that the inaccuracies of the accumulation-based model can be taken into account directly by implementing delays either in the control inputs or in the state of the dynamic system. A proper investigation about the properties of these new modeling approaches is still missing, in particular to assess their differences with the trip-based model. While we choose to focus on the accumulation-based and trip-based models in this chapter, integrating these new approaches could be worth for further research.

However, from the modeler's perspective, despite all these recent advances in MFD-based simulation, congestion propagation in a multi-reservoir framework is not fully understood yet. Notably, the questions of if and how boundary flows should be limited when a reservoir is oversaturated, and how to distribute inflows and outflows have not been completely addressed. In details, as the wide majority of MFD-based simulators are developed for control applications, most authors argue, with reason, that the controllers will not allow the reservoir to reach highly congested states, so that the aforementioned concerns may be eclipsed in their traffic flow models (see e.g. Kouvelas *et al.*, 2017). However, other applications of MFD-based models should not ignore them. Actually some interesting works already propose incomplete but viable solutions to deal with congestion propagation. Hajiahmadi *et al.* (2013b), Lentzakis *et al.* (2016), whose simulator is based on the Network

Transmission Model (NTM) of Knoop & Hoogendoorn (2014), consider exogenous boundary capacities between reservoirs and a global entry supply function per reservoir, similar to the Cell Transmission Model (CTM) of Daganzo (1994). Their approach ensures a perfect protection of the reservoirs from global gridlock, nevertheless this one can hardly be extended to heterogeneous trip lengths within reservoirs. Yildirimoglu & Geroliminis (2014) certainly developed the more advanced tool in MFD-based simulation, as they account for different trip lengths and manage flow exchanges with a Dynamic Traffic Assignment (DTA) procedure on macroscopic routes (a route being a succession of reservoirs, sometimes called a regional path). However, they handle each boundary between two adjacent reservoirs separately with a pro-rata inflow merge, and do not provide any further information on how the global protection of each reservoir is ensured.

In this chapter, we investigate the constraints and the requirements to design a proper congestion propagation model, valid in any situations for an MFD-based multi-reservoir system. The main objective of this study is to build up a consistent simulation framework that properly addresses congestion spillbacks in a multi-reservoir system. In particular, we will ensure that flow exchanges at interfaces are consistent with classical kinematic wave theory principles (conservation, supply limitation, etc). For the accumulation-based model, we will refer to the framework of Yildirimoglu & Geroliminis (2014), Geroliminis (2015). For the trip-based model, we will pursue our effort in chapter 2 to provide a first attempt to handle spillbacks in this formulation. For both modeling approaches, it appears that elementary constraints on flow and production lead to a unique definition of the entry supply function for each accumulation or trip length category in one reservoir. Moreover, applying the same constraints for exit flows allow us to show that each partial outflow cannot be treated independently from the other trip outflows in a reservoir. It follows that if one wants to preserve the consistency of an MFD-based model with multiple trip lengths (i.e. that the users are traveling at the same mean speed at each time), one must let the inflow merge and outflow diverge allocation be endogenously defined by the reservoir state. Consequently, only a few degrees of freedom are left to the modeler to control flow exchanges in simulation. Comparisons between our framework and the initial approaches of Yildirimoglu & Geroliminis (2014) and Knoop & Hoogendoorn (2014) show significant variations in the simulation results.

This chapter is organized as follows: section 4.2 presents the case of a single reservoir with a unique trip length. This section allows to set the background of our study and its assumptions, and also discusses and proposes an efficient method to account for spillbacks in the trip-based approach. Section 4.3 deals with inflow and outflow allocation between multiple accumulation or trip length categories in one reservoir, especially when congestion propagates through the reservoir. Section 4.4 illustrates the differences in simulation between our approach and two existing models from the literature for a simple multi-reservoir system. Finally section 4.5 discusses the introduction of internal trips in our framework.

## 4.2 Flow transfer in a single reservoir with a unique trip length

### 4.2.1 Accumulation-based modeling

In this study, we focus on how a reservoir should interact with its neighbors in the context of a multi-reservoir representation of a city.

The concept of the single reservoir model has been first presented in Daganzo (2007), Geroliminis & Daganzo (2007). It corresponds to a given part of an urban network where the traffic states are characterized by a well-defined production-MFD $P(n)$ (in [veh.m/s]), or equivalently, a speed-MFD $V(n) = P(n)/n$ (in [m/s]), where $n$ (in [veh]) is the accumulation (number of circulating vehicles in the reservoir) . The reservoir entry (also conceptually called "upstream boundary") is the aggregation of all individual entry nodes of the network; similarly the reservoir exit (or "downstream boundary") aggregates all the exit nodes. Through the entry is defined the total effective inflow $q_{\text{in}}(t)$, and through the exit the total effective outflow $q_{\text{out}}(t)$. In a first approach, we do not consider internal trips (trips that start and end into the reservoir) and assume that traffic states result from "transfer" trips only (trips from the aggregated entry to the aggregated exit). The question of internal trips will come up in the discussion (section 4.5). In this framework, the reservoir dynamics are governed by the following vehicle conservation equation (Daganzo, 2007):

$$\frac{dn}{dt} = q_{\text{in}}(t) - q_{\text{out}}(t) \qquad \text{(reservoir dynamics)} \tag{4.1}$$

The accumulation and trip-based models differ from the definition of $q_{\text{in}}(t)$ and $q_{\text{out}}(t)$. We present here the accumulation-based model first. Most of the previous studies dealing with MFD-based aggregated dynamics actually do the distinction between internal and transfer trips, however their treatments are often mixed in the same modeling approach, so that a proper definition of inflow and outflow is sometimes missing. In control-oriented works, some authors like Aboudolas & Geroliminis (2013), Ampountolas *et al.* (2017), Kouvelas *et al.* (2017) split the inflow into the receiving flow from adjacent reservoirs (for which the controllers apply), and the "uncontrolled demand" which may be from inside or outside. Because we focus on a reservoir with no internal demand and interacting with its neighbors here, we propose a definition of flow exchange at boundaries inspired by the CTM (Daganzo, 1994) and the NTM (Knoop & Hoogendoorn, 2014). Thus, at the reservoir entry, the effective inflow results from the competition between a given demand $\lambda(t)$ from some sending reservoirs and an entry supply function $I(n)$ depending on the reservoir state and restraining the inflow when the reservoir becomes congested:

$$q_{\text{in}}(t) = \min[\lambda(t); I(n)] \qquad \text{(effective inflow)} \tag{4.2}$$

At the reservoir exit, the effective outflow balances a given supply $\mu(t)$ (inflow restriction into some destination reservoirs in case of congestion), and an outflow demand function $O(n)$ also depending on the reservoir state:

$$q_{\text{out}}(t) = \min[\mu(t); O(n)] \qquad \text{(effective outflow)} \tag{4.3}$$

The single reservoir model and its boundary conditions are represented in Figure 4.1(a). As there is no spatial extension in the reservoir, the simplest version of the accumulation-based

model assumes an average trip length $L$ for all travelers and applies the queuing formula of Little (1961) to define the "trip completion rate" $G(n) = n/L \cdot V(n) = P(n)/L$ (Daganzo, 2007, Geroliminis & Daganzo, 2007). This quasi-static approach has several limitations as detailed in chapter 2. To our best knowledge, all the studies from the literature consider that the system demand for outflow (or effective outflow if they do not apply supply limitations) always equals $G(n)$. In our opinion however, we believe that this is only true to model internal congestion for the outflow of internal trips, but that a distinction between $O(n)$ and $G(n)$ should be made for transfer trips. That is why we use a special definition of $O(n)$ to describe the demand for outflow. Note that $O(n)$ is not the effective outflow $q_{\text{out}}(t)$, which is the result of the competition between demand and supply as written in equation 4.3. Otherwise with the traditional approach, the reservoir can easily converge to heavily congested situations without any possibility to recover. This phenomenon was notably highlighted in chapter 2 with a supply reduction scenario at exit. This issue happens only if $O(n)$ describes the demand for outflow and not the effective outflow. In classical kinematic wave theory, the observed flow at the head of a queue is always equal to the available capacity. This is mathematically expressed by a demand value equal to the maximal capacity. We should observe the same behavior when a reservoir is discharging without external downstream constraints. If $O(n)$ always equals $G(n)$, the outflow demand is very low for a high accumulation $n$. The consequence is that the queue cannot empty, and thus the reservoir cannot retrieve a free-flow situation after a congestion period. This situation will be also illustrated in section 4.2.3. This can be avoided if we adopt the following definition of outflow demand:

$$O(n) = \begin{cases} \frac{n}{L} V(n) = G(n) & \text{if } n < n_c \\ \frac{P_c}{L} & \text{otherwise} \end{cases} \qquad \text{(outflow demand function)} \tag{4.4}$$

Without having access to the ground truth, we cannot claim that the latter expression is more or less realistic than the traditional approach $O(n) = G(n)$. Nevertheless, our proposition is specifically designed to handle flow exchanges during congestion periods in multi-reservoir systems, according to what is already well-known in link-scale traffic flow theory. Because temporarily supply limitations are very likely to occur at the reservoir perimeter when congestion propagates, assuming $O(n) = G(n)$ would often lead to gridlock, making such a model quite useless in practice to study oversaturated situations over a large-time horizon. The traditional approach is actually based on observations from either simulation studies (Geroliminis & Daganzo, 2007) or empirical evidences (Geroliminis & Daganzo, 2008), but what these authors observed is the effective outflow. Whereas our modeling framework is based on the demand for outflow, a quantity which is not measurable in simulation nor the real field. This concept proves to be efficient to avoid the extreme situation mentioned above and presented in section 4.2.3 when considering transfer trips only. Moreover, we show in the the discussion (section 4.5) that the traditional approach can be well observed and explained in our framework once internal trips are added.

At the reservoir entry, the general shape of the entry supply function $I(n)$ is first introduced in Geroliminis & Daganzo (2007), based on simulation results on the San Fransisco downtown network, but the authors do not mention any explicit formulation. Like in Hajiahmadi *et al.* (2013b), Knoop & Hoogendoorn (2014), Lentzakis *et al.* (2016), we define $I(n)$

in accordance with the basic principles of traffic flow theory:

$$I(n) = \begin{cases} \frac{P_c}{L} & \text{if } n < n_c \\ \frac{n}{L} V(n) = G(n) & \text{otherwise} \end{cases} \qquad \text{(entry supply function)} \tag{4.5}$$

It should be mentioned that the limitation $P_c/L$ in undersaturated regime certainly corresponds to the most restrictive conditions that can be applied when defining $I(n)$. $P_c/L$ corresponds to the reservoir capacity in a stationary state. When this capacity is driven by internal bottlenecks in the middle of the reservoir, it would not be surprising to observe higher inflow values during the reservoir loading. The maximal bound for inflow for low accumulations is certainly closer to the sum of all entry link capacities, especially when the reservoir loading is done with a sharp demand profile. The shape of $I(n)$ appears then as an important element when doing the model calibration. One proposition is to define $I(n)$ as a monotonic decreasing curve for $n < n_c$, starting from a point between $P_c/L$ and the total entry capacity, and ending at $P_c/L$ for $n = n_c$. The choice of the critical accumulation $n_c$ to distinguish under- and oversaturated conditions in $I(n)$ is also a matter of debate. Despite the above-mentioned studies adopted the same hypothesis, Geroliminis & Daganzo (2007) found with simulated data that the critical accumulation of $I(n)$ is higher than $n_c$. However, to keep the analysis simple, we choose to describe $I(n)$ as in equation 4.5 in this study. This choice does not change our overall methodology. The complete definition of the shape of $I(n)$ is out of the scope of the current study.

Thanks to the previous definitions of $O(n)$ and $I(n)$, we can also use the concept of "exit production demand" $P_d(n)$ and "entry production supply" $P_s(n)$ written as:

$$P_d(n) = \begin{cases} P(n) & \text{if } n < n_c \\ P_c & \text{otherwise} \end{cases} \qquad \text{(exit production demand)} \tag{4.6}$$

$$P_s(n) = \begin{cases} P_c & \text{if } n < n_c \\ P(n) & \text{otherwise} \end{cases} \qquad \text{(entry production supply)} \tag{4.7}$$

With this concept, the outflow demand and inflow supply functions are simply defined as:

$$O(n) = \frac{P_d(n)}{L} \qquad \text{and} \qquad I(n) = \frac{P_s(n)}{L} \tag{4.8}$$

The definitions of $P_d(n)$ and $P_s(n)$ do not consist in new concepts, they are simply the translation of outflow demand and inflow supply into production quantities. These notations are mostly useful for the following of this chapter. The two functions $O(n)$ and $I(n)$ are represented in Figures 4.1(c) and (d), for a typical shape of the production-MFD in Figure 4.1(b). Note that $O(n)$ and $I(n)$ are similar to the entry and exit functions of a cell in the CTM (Daganzo, 1994). When dealing with transfer trips only, such a formulation for the accumulation-based model is fully consistent to handle both free-flow and congested situations in the reservoir. Figures 4.1(c) and (d) show examples of equilibrium states reached in free-flow and congestion respectively with given boundary conditions.

### 4.2.2 Trip-based modeling

The theoretical background of the trip-based model has been first introduced by Arnott (2013). Let us consider a single reservoir with a unique trip length $L$. At each time $t$, all

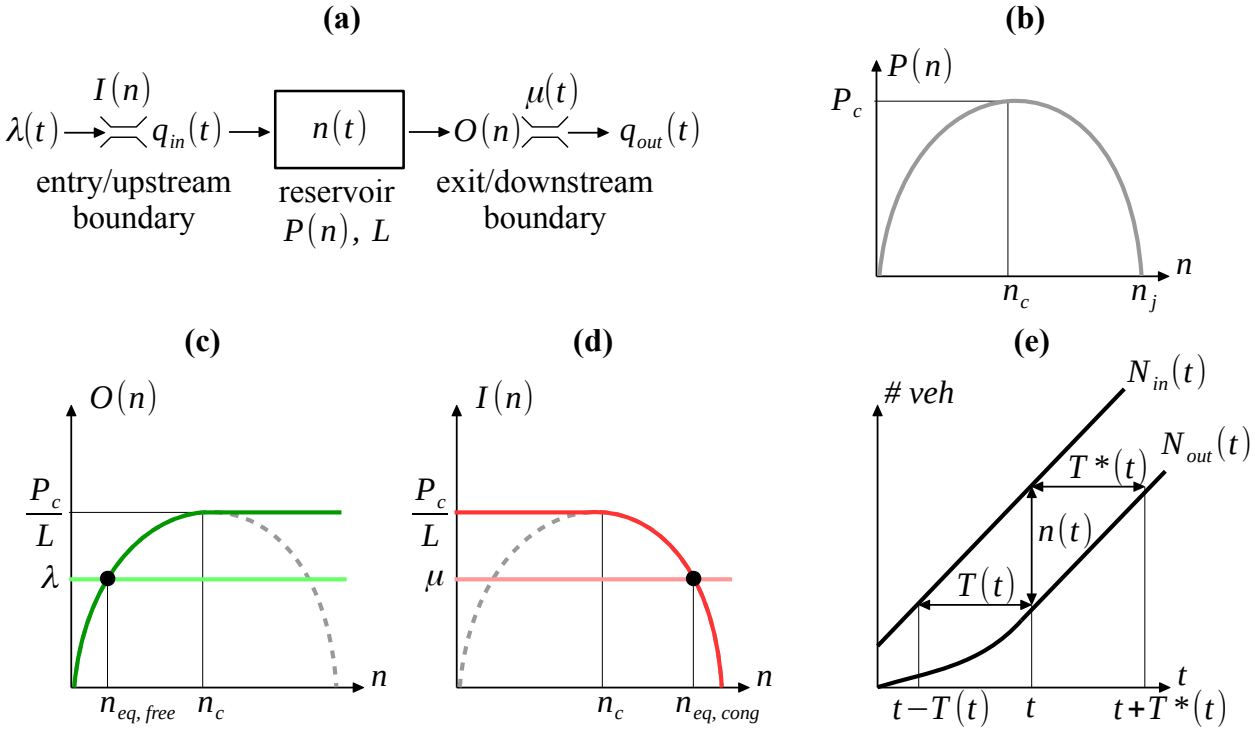



FIG. 4.1 – *(a) Representation of a single reservoir and its boundaries, (b) typical shape for the production-MFD, (c) outflow demand function $O(n)$: example of an equilibrium point for a given inflow demand $\lambda$ in free-flow conditions, (d) inflow supply function $I(n)$: example of an equilibrium point for a given exit restriction $\mu$ in congestion, and (e) cumulative count curves with the accumulation $n(t)$, the experienced travel time $T(t)$ and the exact predictive travel time $T^*(t)$*

the vehicles are traveling at the same speed $V(n(t))$. A user exiting the reservoir at $t$ has experienced a travel time of $T(t)$. This user thus entered the reservoir at $t - T(t)$, and his/her trip distance was $L$. The trip-based model considers that the accumulation and therefore the mean speed may change during the user's trip, which is mathematically expressed as:

$$L = \int_{t-T(t)}^{t} V(n(s))ds \tag{4.9}$$

By using basic relationships based on entering and exiting count curves, it can be shown that the derivative of equation 4.9 leads to (see e.g. Arnott, 2013):

$$q_{\text{out}}(t) = q_{\text{in}}(t - T(t)) \cdot \frac{V(n(t))}{V\big(n(t - T(t))\big)} \tag{4.10}$$

Using equation 4.10 to solve the conservation equation 4.1 leads to a differential equation with endogenous delay. Despite being mathematically intractable, this formulation of the outflow can allow the development of efficient numerical resolution schemes (continuous approximation on vehicle indexes or event-based resolution method), as shown in chapter 2. A simpler formulation integrating static delays has been proposed for the accumulation-based model by Haddad & Zheng (2017). However, a comparison study is currently missing to assess the differences between these both approaches, so that in this chapter we decided

to focus on the more refined trip-based formulation. These resolution methods of the latter work in free-flow only, where $q_{\text{in}}(t)$ is the input, equal to the inflow demand $\lambda(t)$, and where $q_{\text{out}}(t)$ is the consequence of the system evolution. In congestion however, the role of inflow and outflow are switched, as $q_{\text{out}}(t)$ becomes the given boundary condition, equal to the outflow supply $\mu(t)$, and $q_{\text{in}}(t)$ has now to adapt to the system evolution due to the restriction at exit. It can be shown that equation 4.10 can be reversed to express $q_{\text{in}}(t)$ as a function of $q_{\text{out}}(t)$:

$$q_{\text{in}}(t) = q_{\text{out}}(t + T^*(t)) \cdot \frac{V(n(t))}{V\big(n(t + T^*(t))\big)} \tag{4.11}$$

where $T^*(t)$ is the exact predictive travel time, i.e. the time during which the user entering at $t$ will travel, see also Figure 4.1(e). By construction we have: $T(t) = T^*(t - T(t))$. But equation 4.11 means that to calculate the effect of an exit flow limitation on the entry, one needs to know the future of the system, which is problematic. First, it is not possible to deduce the inflow when downstream supply restriction should apply, and second, if this was possible we have no clue on how to make the switch.

Thus in practice, this model needs to be coupled with another model for reproducing congestion propagation. A first attempt has been made in chapter 2. They assume a free-flow evolution of the system and then apply the outflow reduction and the minimum principle of Newell (1993) on the inflow. This method with off-line calculations is sufficient for the analysis of a single reservoir, but not suitable in a multi-reservoir context where traffic states in the reservoirs depend from each other all the time.

In our study, we propose a simple way to perform in-line computations of inflow limitations in the trip-based model. It consists in switching to the accumulation-based framework in congestion, by using the same entry supply function $I(n)$ which restrains inflow for high values of accumulation $n$. We show in the next section with simple simulation scenarios that such a method works well in practice.

### 4.2.3 Numerical implementation

In the following of this chapter, we will use the event-based scheme presented in chapter 2 to solve numerically the trip-based model. In congestion, the reservoir exit flow is limited to $\mu(t)$ at each time by retaining the vehicles inside the zone until the next exit time, even if they have already completed their trip length. At entry, the inflow limitation is ensured with the definition of a minimum or supply time for entering the reservoir:

$$t^{N^{\text{in}}}_{\text{entry supply}} = t^{N^{\text{in}}-1}_{\text{entry}} + \frac{1}{I(n)} \qquad \text{(entry supply time)} \tag{4.12}$$

where $t^{N^{\text{in}}}_{\text{entry supply}}$ is the supply time for the $N^{\text{in}}$th vehicle to enter the reservoir, $t^{N^{\text{in}}-1}_{\text{entry}}$ is the entering time of the previous vehicle, and $I(n)$ is the entry supply function of the accumulation-based model, see equation 4.5.

The application of this method is illustrated with two test cases. The first one is about a demand peak temporarily exceeding the exit supply, and the second one concerns a supply reduction at exit below the demand level at entry. These numerical examples consider a single reservoir with maximum accumulation $n_j$ = 1000 veh, average trip length $L$ = 2.5 km, free-flow speed $u$ = 15 m/s, and characterized by a production-MFD built with two

parabolic branches, with maximum production $P_c$ = 3000 veh.m/s and critical accumulation $n_c$ = 400 veh.

Figure 4.2(a1) shows the demand $\lambda(t)$ and supply $\mu(t)$ profiles for the demand peak case. The simulation scenario has been designed to let the congestion reach the entry before the demand decreases. The reservoir state evolution is presented in Figures 4.2(b1) and (c1) with the inflow/outflow and accumulation. The blue curves correspond to the accumulation-based model, the green ones to the trip-based model. Note that a queue at the reservoir entry is taken into account when vehicles are waiting to enter if the inflow is limited in both models, though not presented here. All graphs show similar results for both modeling approaches. This was actually expected, since the modeling of spillbacks is handled in the same manner in both models. This also proves that the switch to the accumulation-based model works well in the trip-based framework with few modifications in the event-based resolution scheme.

Figure 4.2(a2) shows the demand $\lambda(t)$ and supply $\mu(t)$ profiles for the supply reduction case. Similarly, the simulation scenario has been designed to let the congestion reach the entry before the exit supply increases again. In Figures 4.2(b2) and (c2), the red and yellow curves corresponds to the evolution of inflow/outflow and accumulation when $O(n)$ always equals $G(n)$ as it is traditionally assumed in the literature. In the accumulation-based model (in red), we observe that the system reaches an equilibrium point once inflow equals outflow shortly after 4000 s. Then, the reservoir state does not evolve anymore because after this point the outflow corresponds to the exit demand $O(n)$, and thus $q_{\text{out}}(t)$ is not impacted by an increase of $\mu(t)$, see equation 4.10. In the trip-based approach (in yellow), the users travel at a low mean speed after 4000 s to adapt the exit supply reduction. But when this limitation disappears, the vehicle exit rate is still the same because the mean speed remains low, and consequently the system cannot recover from congestion in this framework too. We can fix this problem if we keep the outflow demand $O(n)$ maximum during severe congestion periods. This can be modeled in the trip-based framework only if we force the travelers to complete their trip length at a pace that matches the exit time defined by the outflow when $n \geq n_c$. Theoretically, it implies that the related users will have a speed different from $V(n)$ during congested situations. This formulation happens to be equivalent to the outflow demand definition of equation 4.4. This is illustrated by the blue and green curves in Figures 4.2(b2) and (c2), which also show similar results for both models.

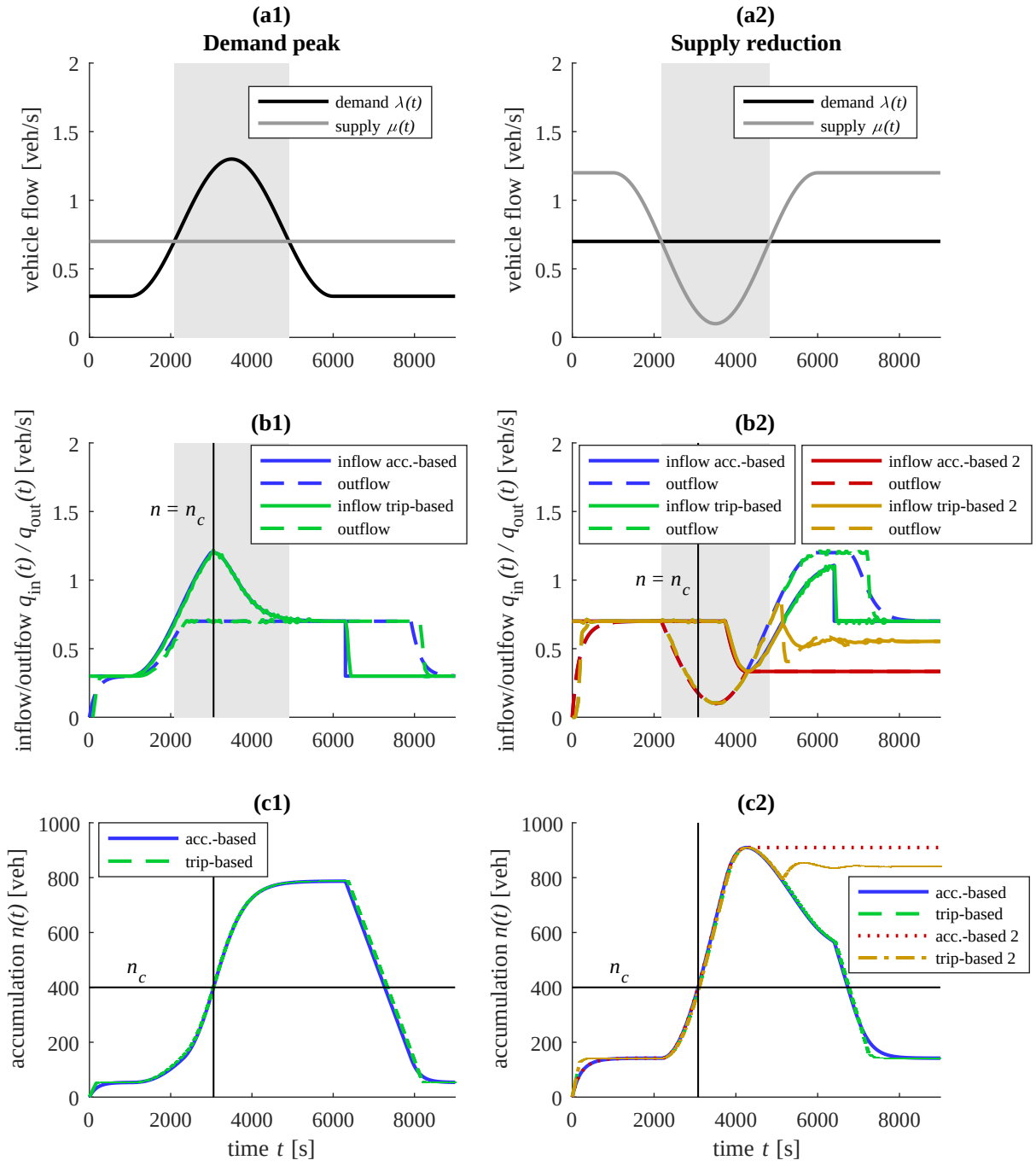



FIG. 4.2 – *(a1) Demand peak at the reservoir entry, demand $\lambda(t)$ and supply $\mu(t)$ profiles where the gray area indicates when demand exceeds supply, (b1) inflow $q_{\text{in}}(t)$ and outflow $q_{\text{out}}(t)$ and (c1) accumulation $n(t)$ for the accumulation and trip-based models. (a2) Supply reduction at the reservoir exit, demand and supply profiles where excess of demand compared to supply is indicated by the gray area, (b2) inflow and outflow and (c2) accumulation for the accumulation- and trip-based models, where model "2" corresponds to the situation when $O(n)$ always equals $G(n)$*

## 4.3 Flow transfer in a single reservoir with multiple trip lengths

### 4.3.1 Accumulation-based framework

The aim of our study is to propose a robust modeling framework for congestion propagation in a multi-reservoir environment. Like in the approach of Yildirimoglu & Geroliminis (2014), we consider that users are assigned to a set of given "macroscopic routes", i.e. successions of reservoirs, as illustrated in Figure 4.3(a), and that the system state can be described at the level of a macroscopic route, later simply referred as "route". As a reservoir can be crossed by different routes with different trip lengths, the thorough understanding of flow dynamics in one reservoir with heterogeneous trip lengths is essential to build a proper multi-reservoir simulation tool.

The extension of the single reservoir model with one trip length to several trip lengths has been first established in Geroliminis (2009, 2015). This theoretical framework has then been used in various studies with more complex multi-reservoir settings (e.g. Yildirimoglu & Geroliminis, 2014, Ramezani *et al.*, 2015). Note that the present study also applies for a reservoir with different accumulation categories (users are distinguished by their route or destination) but with a unique trip length (users are assumed to travel the same distance). This will lead to several simplifications in the following, nevertheless the conclusions will be the same.

Let us consider a single reservoir with $N$ trip length categories $L_i$, or $N$ routes with length $L_i$, as presented in Figure 4.3(b). All accumulations $n_i$ in each route $i$ should satisfy the following system (Geroliminis, 2015):

$$\forall i \in \{1, \dots, N\}, \qquad \frac{dn_i}{dt} = q_{\text{in},i}(t) - q_{\text{out},i}(t) \tag{4.13}$$

where $q_{\text{in},i}(t)$ and $q_{\text{out},i}(t)$ are respectively the effective inflow and outflow for route $i$. In this framework, the potential connection to neighboring reservoirs is represented by boundary conditions. As depicted in Figure 4.3(b), each inflow $i$ is induced by an exogenous demand $\lambda_i(t)$ coming from a given upstream reservoir, and each outflow $i$ may encounter an exogenous supply limitation $\mu_i(t)$ imposed by a given downstream reservoir.

Conceptually, the reservoir is split into "sub-reservoirs" governed by the accumulation $n_i$. These sub-reservoirs are coupled together by the mean speed $V(n)$ or the total production $P(n)$, where $n = \sum_{i=1}^{N} n_i$. It is assumed that in slow-varying conditions, the trip completion rate $G_i$ of each route $i$ satisfies the queuing formula of Little (1961):

$$G_i(n_1, \dots, n_N) = G_i(n_i, n) = \frac{n_i}{L_i} V(n) = \frac{n_i}{n} \frac{P(n)}{L_i} \qquad \text{(trip completion rate)} \tag{4.14}$$

Like in section 4.2, we distinguish the outflow demand $O_i$ from the trip completion rate $G_i$ for each class $i$. Likewise, note that $O_i$ is a demand for outflow, which may be different from the effective outflow $q_{\text{out},i}$. $O_i$ is supposed to be maximum in oversaturated situations ($n > n_c$):

$$O_i(n_i, n) = \frac{n_i}{n} \frac{P_d(n)}{L_i} = \begin{cases} \frac{n_i}{n} \frac{P(n)}{L_i} = G_i(n_i, n) & \text{if } n < n_c \\ \frac{n_i}{n} \frac{P_c}{L_i} & \text{otherwise} \end{cases} \qquad \text{(outflow demand)} \tag{4.15}$$

where $P_d(n)$ is the exit production demand previously defined in equation 4.6. We have moreover by definition:

$$n = \sum_{i=1}^{N} n_i \qquad \text{(total accumulation)} \tag{4.16}$$

$$G(n) = \sum_{i=1}^{N} G_i(n_i, n) \qquad \text{(total trip completion rate)} \tag{4.17}$$

$$O(n) = \sum_{i=1}^{N} O_i(n_i, n) \qquad \text{(total outflow demand)} \tag{4.18}$$

Note also that a dynamic average trip length $L(t)$ can be defined by applying Little's formula at the reservoir scale: $G(n) = n/L \cdot V(n)$. It comes (Geroliminis, 2009):

$$L(t) = \frac{n(t)}{\sum_{i=1}^{N} \frac{n_i(t)}{L_i}} = \frac{\sum_{i=1}^{N} G_i(n_i(t), n(t)) L_i}{G(n(t))} \qquad \text{(average trip length)} \tag{4.19}$$

The major difference with the unique trip length model is that here, each outflow demand $O_i$ depends not only on $n_i$, but also on the other accumulations $n_1, \ldots, n_N$ through the total accumulation $n$.

### 4.3.2 Calculating perimeter inflows

All the effective inflows $q_{\text{in},i}(t)$ for each route $i$ may be treated as independent variables. They are the result of the competition between the corresponding demand $\lambda_i(t)$ and an entry supply function $I_i(n_1, \ldots, n_N)$:

$$\forall i \in \{1, \ldots, N\}, \qquad q_{\text{in},i}(t) = \min[\lambda_i(t); I_i(n_1(t), \ldots, n_N(t))] \qquad \text{(effective inflow route } i\text{)} \tag{4.20}$$

There is no clear consensus on the definition of each entry supply function $I_i(n_1, \ldots, n_N)$ in the literature for multi-reservoir systems. Geroliminis (2009), Knoop & Hoogendoorn (2014), Yildirimoglu & Geroliminis (2014), Ramezani *et al.* (2015) consider a global supply function $I(n)$ at the reservoir entry that applies for all routes crossing the reservoir. Knoop & Hoogendoorn (2014) use the same definition as in equation 4.5 for this function, but they also define exogenous boundary capacities between adjacent reservoirs. Ramezani *et al.* (2015) simplify the shape of $I(n)$ for $n > n_c$ with a linear curve, however they give no further details on the maximum flow capacity they use for $n < n_c$. The general idea is to allocate portion of flows regarding a global inflow limitation for the whole reservoir to prevent it from gridlock[1]. Main issues are the definitions of: (i) the allocation scheme, (ii) the maximum allowed flow for each route (capacity at entry), and (iii) the distinction between under- and oversaturated states.

In this study, we assume that the functions $I_i(\cdot)$ must be designed to ensure that the total effective inflow and total entry production (i) never exceed the reservoir capacity and (ii) adapt to the reservoir state in oversaturated regime. As discussed earlier when presenting

[1]Our point is that gridlock may happen due to internal demand but not due to external loading at the perimeter

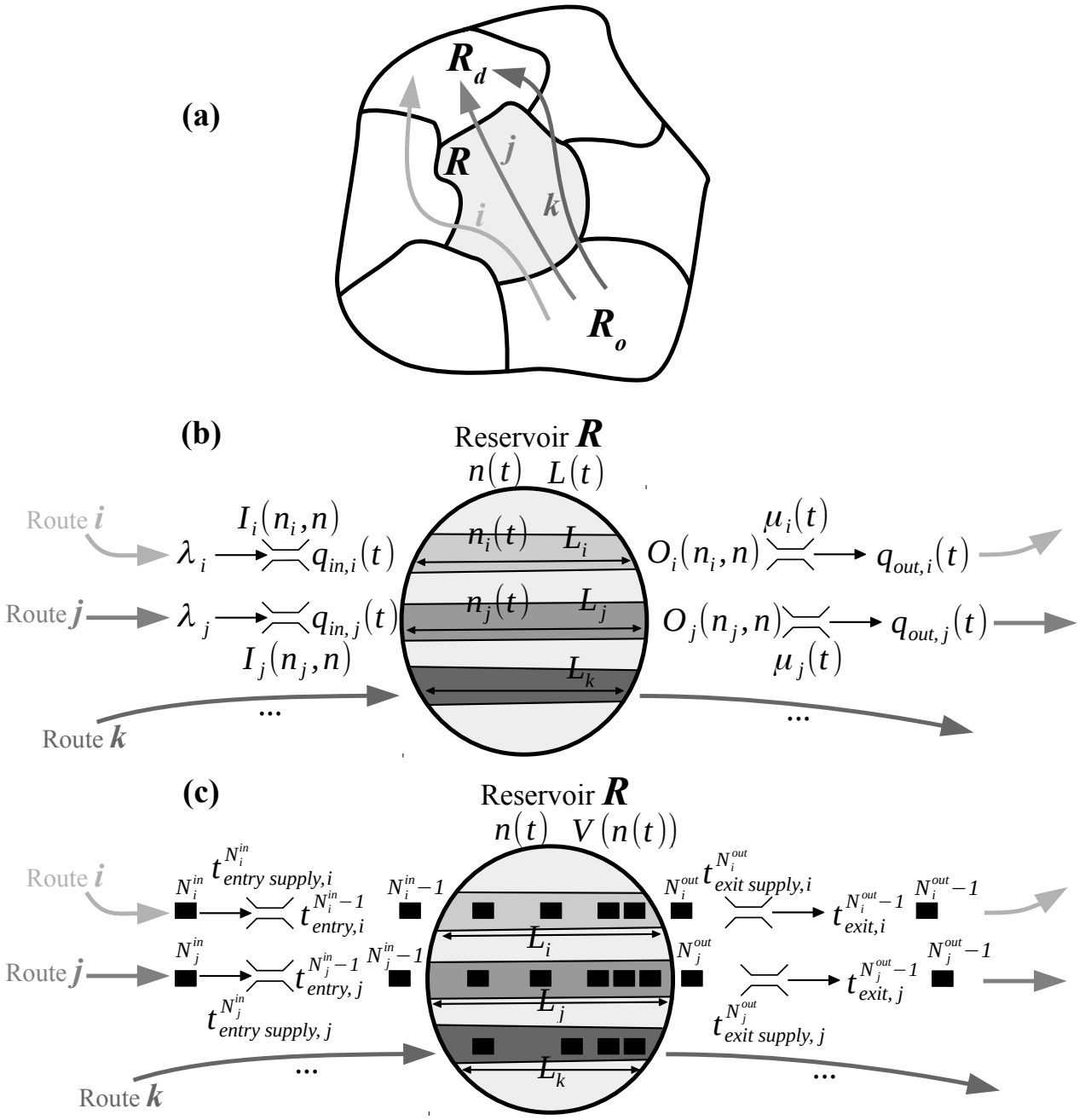



FIG. 4.3 – *(a) Examples of three routes $(i, j, k)$ for a macroscopic OD $(R_o, R_d)$ in a multi-reservoir system, (b) representation of the reservoir R crossed by the routes in the accumulation-based and (c) trip-based frameworks*

the entry flow function, see section 4.2.1, the separation between the under- and oversaturated regimes may happen at a different $n_c$ value than for the production-MFD (Geroliminis & Daganzo, 2007). Here, we simply assume that the $n_c$ values are the same but a particular attention to this question should be paid when undertaking the model calibration.

#### 4.3.2.1 Case when all the routes are congested

If all the routes have their inflow limited, $\forall i, q_{\text{in},i} = I_i(n_1, \ldots, n_N)$, at the reservoir scale the functions $I_i(\cdot)$ must comply with two global constraints, on flow and on production, described as follows. In free-flow, the total inflow and production cannot exceed the reservoir flow and production capacity respectively, and in congestion, the total entering flow and production must adapt to the reservoir current trip completion rate and production respectively (the $n_i$ variables are omitted for the sake of simplicity):

$$\sum_{i=1}^{N} I_i = \frac{P_s(n)}{L} \qquad \text{(flow constraint)} \tag{4.21}$$

$$\sum_{i=1}^{N} L_i I_i = P_s(n) \qquad \text{(production constraint)} \tag{4.22}$$

where $P_s(n)$ is the entry production supply defined in equation 4.7, and $L$ the average trip length defined in equation 4.19. We will demonstrate that under these two constraints, the unique possible definition of these functions $I_i(\cdot)$ is:

$$\forall i \in \{1, \ldots, N\}, I_i(n_i, n) = \frac{n_i}{n}\frac{P_s(n)}{L_i} = \begin{cases} \frac{n_i}{n}\frac{P_c}{L_i} & \text{if } n < n_c \\ \frac{n_i}{n}\frac{P(n)}{L_i} = \frac{n_i}{L_i}V(n) & \text{otherwise} \end{cases} \text{(entry supply function)} \tag{4.23}$$

*Proof.* Let assume without loss of generality that each function $I_i$ can be regarded as a portion of a global entry supply $I(n)$: $\forall i \in \{1, \ldots, N\}, I_i = \alpha_i I(n)$ where the $\alpha_i(\cdot)$ are functions of $(n_1, \ldots, n_N)$. As we have $\sum_{i=1}^{N} I_i = I(n)$, the functions $\alpha_i(\cdot)$ should verify $\sum_{i=1}^{N} \alpha_i = 1$. The application of the flow constraint (equation 4.21) forces $I(n) = P_s(n)/L$, whereas the application of the production constraint (equation 4.22) leads to $I(n)\sum_{i=1}^{N} L_i\alpha_i = P_s(n)$, and thus $\sum_{i=1}^{N} L_i\alpha_i = L$. Let us express the $\alpha_i(\cdot)$ functions as follows: $\forall i \in \{1, \ldots, N\}, \alpha_i = \beta_i L/L_i$, where $\beta_i(\cdot)$ are any functions of $(n_1, \ldots, n_N)$ and satisfy $\sum_{i=1}^{N} \beta_i = 1$. Knowing that $\sum_{i=1}^{N} \alpha_i = 1$, it results that $\sum_{i=1}^{N} \beta_i = 1/L = \sum_{i=1}^{N} n_i/(nL_i)$. Again, the $\beta_i(\cdot)$ functions can be expressed as follows: $\forall i \in \{1, \ldots, N\}, \beta_i = \gamma_i n_i/n$, where $\gamma_i(\cdot)$ are any functions of $(n_1, \ldots, n_N)$ and must verify $\sum_{i=1}^{N} \beta_i = \sum_{i=1}^{N} \gamma_i n_i/n = 1$. Because the last equality if true whatever the accumulations $(n_1, \ldots, n_N)$, it entails that $\forall i \in \{1, \ldots, N\}, \gamma_i = 1$, and therefore $\forall i \in \{1, \ldots, N\}, \alpha_i = n_i/n \cdot L/L_i$. Finally we have shown that $\forall i \in \{1, \ldots, N\}, I_i = n_i/n \cdot P_s(n)/L_i$. □

In the literature, the above-mentioned authors who use a global entry supply function with multiple trip lengths (Geroliminis, 2009, Yildirimoglu & Geroliminis, 2014, Ramezani *et al.*, 2015) only focus on the flow constraint. Despite some differences between their treatment of flow exchanges, they all apply a pro-rata merge at each reservoir entry to allocate the portions of flow in case of congestion. With our notations, this would be equivalent to define the entry supply functions as: $\forall i \in \{1, \ldots, N\}, I_i(n_i, n) = \alpha_i P_s(n)/L$ where the coefficients $\alpha_i$ would represent the inflow pro-rata merge, and thus depend on the corresponding demands $\lambda_i$ from each route $i$. They may take any values as soon as: $\sum_{i=1}^{N} \alpha_i = 1$. By definition, such an approach ensures a strict compliance with the flow constraint (equation 4.21). However, we demonstrated above that the production constraint is respected if and only

if $\forall i \in \{1,\dots,N\}, \alpha_i = n_i/n \cdot L/L_i$, which will be hardly the case in practice for any $\alpha_i$. Therefore, without any other constraints, a pro-rata merge is not likely to fulfill the production constraint (equation 4.22). This is one of the major differences between this current approach and the ones from the existing literature[2]. This will be illustrated in section 4.4.

#### 4.3.2.2 Case when only some of the routes are congested

When only a fraction of the routes are congested ($\lambda_i > I_i(n_i, n)$) while the others are not limited ($\lambda_j < I_j(n_j, n)$), the reservoir entry capacity may not be fully used if equation 4.20 is applied as it is. Following the idea of the fair merge of Daganzo (1995), the remaining part of the capacity can be then allocated to the congested routes. In this case, the entry capacity may be either the flow capacity described in equation 4.21 or the production capacity in equation 4.22. However, we show with a simple example of two routes in B.1 that it is impossible to ensure a strict compliance with both flow and production constraints at the same time in this case. Thanks to some simulation tests detailed in section 4.3.5, we observe that if we force the sum of all inflows to equal the flow capacity $P_s(n)/L$, then the sum of all entering productions may exceed the production capacity $P_s(n)$. However, in the reverse case, i.e. when the merge is applied to the entering productions instead of the inflows, the resulting total inflow almost never exceeds the flow capacity. This may actually happen during short periods, but then $P_s(n)/L$ adapts very fast to the total inflow, notably through the evolution of the average trip length $L(t)$. Thus in practice, we choose to apply the generalized merge of Leclercq & Becarie (2012) to the entering production:

$$\forall i, L_i q_{\text{in},i}(t) = \begin{cases} L_i \lambda_i & \text{if } \lambda_i \leq I_i(n_i, n) \\ \dfrac{n_i}{\sum_{q_{\text{in},j} > I_j(n_j,n)} n_j} \left( P_s(n) - \sum_{q_{\text{in},j} \leq I_j(n_j,n)} L_j q_{\text{in},j} \right) & \text{otherwise} \end{cases} \quad \text{(entry merge)} \tag{4.24}$$

By definition, this guaranties that the available production capacity is always fully reached when some routes are congested, i.e. the total entering production $\sum_{i=1}^{N} L_i q_{\text{in},i}$ always equals $P_s(n)$ during congestion. Nevertheless, the flow capacity may be under-used, i.e. the total inflow $q_{\text{in}} = \sum_{i=1}^{N} q_{\text{in},i}$ is often below $P_s(n)/L$ during the same period, although this flow capacity may be exceeded at some points as mentioned earlier. Due to the analytical complexity of the system, even with two routes, we cannot provide any further proof of our statements. These are only illustrated by simulation in the upcoming section 4.3.5.

### 4.3.3 Calculating perimeter outflows

Because they are the result of the reservoir inner dynamics, the situation may be different for the outflows $q_{\text{out},i}(t)$ of each route $i$. Without loss of generality we can consider that each outflow $q_{\text{out},i}$ is a fraction of the total outflow $q_{\text{out}}$ (time $t$ is omitted): $\forall i \in \{1,\dots,N\}, q_{\text{out},i} = \alpha_i q_{\text{out}}$ where the $\alpha_i(\cdot)$ may eventually be function of $(n_1,\dots,n_N)$. Because by definition $\sum_{i=1}^{N} q_{\text{out},i} = q_{\text{out}}$, the coefficients $\alpha_i(\cdot)$ verify $\sum_{i=1}^{N} \alpha_i = 1$. If we assume

[2]Note that this comparison is not perfectly fair, as we are not sure that the mentioned authors use $P_s(n)/L$ as the inflow supply.

that the expression of the average trip length $L(t)$ is always defined by equation 4.19 at any time, then the total exit production should be:

$$\sum_{i=1}^{N} L_i q_{\text{out},i}(t) = L(t) q_{\text{out}}(t) \qquad \text{(exit production)} \tag{4.25}$$

Replacing $q_{\text{out},i}$ by $\alpha_i q_{\text{out}}$ in equation 4.25 leads to: $\sum_{i=1}^{N} L_i \alpha_i = L$. Thus our previous demonstration about inflow management in section 4.3.2.1 can also apply here. Knowing that $\sum_{i=1}^{N} \alpha_i = 1$, such a relationship results in: $\forall i \in \{1, \dots, N\}, \alpha_i = n_i/n \cdot L/L_i$.

Therefore it appears that, to be consistent with the reservoir inner dynamics (characterized by the average trip length), the definition of the exit production imposes the exit flows to be all interdependent. Because in our study the supplies $\mu_i(t)$ represent the connection to the "downstream" reservoirs, they may take any values, which in general have a few chances to comply with these interdependence relationships. This is illustrated in an simulation example with two routes in the following. Actually, our analysis shows that there exists only one degree of freedom to restrain the outflow of all routes at the reservoir exit.

In a first approach, we may use the total supply $\mu(t) = \sum_{i=1}^{N} \mu_i(t)$ to limit the total outflow demand. We call this the "total supply" restriction approach, and it gives the following effective outflows:

$$q_{\text{out}}(t) = \min[\mu(t); O(n(t))] \qquad \text{(total effective outflow)} \tag{4.26}$$

$$\forall i \in \{1, \dots, N\}, \qquad q_{\text{out},i}(t) = \frac{n_i(t)}{n(t)} \frac{L(t)}{L_i} q_{\text{out}}(t) \qquad \text{(effective outflow route } i\text{)} \tag{4.27}$$

Note that in free-flow, the expressions of $q_{\text{out},i}(t)$ simplify to $G_i(n_i, n)$. During congestion, this choice ensures that the reservoir exits the maximum flow possible allowed by all downstream reservoirs. However, there is a chance that one or more routes exceed the local exit supply, i.e. there may exist a given $i$ for which $q_{\text{out},i}(t) > \mu_i(t)$ because no constraints are applied locally. This would be critical for the corresponding downstream reservoir (next reservoir in route $i$), as this would mean that an excess of inflow could enter the latter ($\mu_i(t)$ represents the inflow limitation of this downstream reservoir).

For this reason, we propose a second approach called the "most constrained" supply method. To avoid that one route might send excess of flow to the next reservoir, we have to ensure that all exit restrictions are respected: $\forall i \in \{1, \dots, N\}, q_{\text{out},i}(t) \leq \mu_i(t)$. This is possible if we define all the outflows thanks to the most restricted exit $k$:

$$q_{\text{out},k}(t) = \min[\mu_k(t); O_k(n_k, n)] \qquad \text{(most constrained outflow)} \tag{4.28}$$

$$\text{where: } k = \arg\min_{1 \leq i \leq N} \frac{\mu_i}{O_i(n_i, n)}$$

$$\forall i \neq k, \quad q_{\text{out},i}(t) = \frac{n_i(t)}{n_k(t)} \frac{L_k}{L_i} q_{\text{out},k}(t) \qquad \text{(effective outflow route } i\text{)} \tag{4.29}$$

Note that in free-flow, no exit is constrained so that $q_{\text{out},k}(t) = O_k(n_k, n)$. In congestion, with this formulation, the system will adapt to the limitation $\mu_k(t)$ for route $k$, so that at equilibrium we have $G_k(n_k, n) = n_k(t)/L_k V(n(t)) = \mu_k(t)$ (assuming that $\mu_k(t)$ is constant after a given time). Knowing that outflow $k$ is chosen as the most constrained one, i.e. with the highest difference between demand $O_k(n_k, n)$ and supply $\mu_k$ (demand being higher than

supply, see equation 4.28), we have then: $\forall i \neq k, O_i/O_k = n_i/n_k.L_k/L_i \leq \mu_i/\mu_k$, and thus $q_{\text{out},i} \leq \mu_i$ because $q_{\text{out},k} = \mu_k$. It results that all effective outflows $i$ in equation 4.29 will be automatically lower than their respective limitations $\mu_i(t)$. However, while protecting the downstream reservoirs from an excess of flow and thus from a possible gridlock, the consequence of this approach is that the flow in many routes may be actually lower than their respective limitations. This is illustrated in the upcoming section 4.3.5.

### 4.3.4 Implementation in the trip-based model

The management of both inflows and outflows can be easily implemented in the trip-based framework. Inflow restrictions are described following the same principle as in the single reservoir model, we switch to the accumulation-based for each route in congestion. As for the interdependence between the outflows, this is even more simple here: keeping the global order of the vehicles by their arrival times ensures that the reservoir inner dynamics are preserved on each route. In practice, there is a waiting list of users which may have different trip lengths in the reservoir, and because there are all traveling at the same speed, they can be simply ordered by their remaining travel distance.

At the reservoir entry, each route $i$ may restrain its inflow by a supply time $t^{N_i^{\text{in}}}_{\text{entry supply},i}$ for the $N_i^{\text{in}}$th vehicle willing to enter:

$$\forall i \in \{1,\dots,N\}, \qquad t^{N_i^{\text{in}}}_{\text{entry supply},i} = t^{N_i^{\text{in}}-1}_{\text{entry},i} + \frac{1}{I_i(n_i,n)} \qquad \text{(entry supply time)} \qquad (4.30)$$

where $t^{N_i^{\text{in}}-1}_{\text{entry},i}$ is the entry time of the previous vehicle in route $i$, see also Figure 4.3(c). In case only some of the routes are congested, we can also apply the fair merge as described in section 4.3.2.2 to ensure that the total entering production capacity is used. In practice, this is achieved in the event-based resolution scheme by modifying the entry supply times of the congested routes. This modification is done directly in the entry supply function $I_i(n_i,n)$ used in equation 4.30:

$$\forall i, \quad I_i(n_i,n) = \begin{cases} \lambda_i & \text{if } \lambda_i \leq I_i(n_i,n) \\ \dfrac{n_i}{\sum_{q_{\text{in},j} > I_j(n_j,n)} n_j} \dfrac{1}{L_i} \left( P_s(n) - \displaystyle\sum_{q_{\text{in},j} \leq I_j(n_j,n)} L_j q_{\text{in},j} \right) & \text{otherwise} \end{cases} \qquad (4.31)$$

where $P_s(n)$ is the entry production supply defined in equation 4.7.

At the reservoir exit, the vehicles are kept in order inside the reservoir until the next exit is possible. The two outflow management methods differ on the definition of the exit supply time of the first vehicle $N_j^{\text{out}}$ of the waiting list. Here, the $\{N_i^{\text{out}}\}_{1\leq i\leq N}$ represent the numbers of the next vehicles to exit in each route $i$. For the total supply approach, the calculation of the total flow supply $\mu(t)$ as the sum of all $\mu_i(t)$ is translated in a mean time headway estimation:

$$t^{N_j^{\text{out}}}_{\text{exit supply}} = t^{N_k^{\text{out}}-1}_{\text{exit},k} + \frac{1}{\sum_{i=1}^{N} 1/\left( t^{N_i^{\text{out}}}_{\text{exit supply},i} - t^{N_i^{\text{out}}-1}_{\text{exit},i} \right)} \qquad \text{(total supply time)} \qquad (4.32)$$

where $t_{\text{exit supply}}^{N_j^{\text{out}}}$ is the total supply time for the next vehicle to exit (to a given route $j$), $t_{\text{exit},k}^{N_k^{\text{out}}-1}$ is the exit time of the previous vehicle (to another route $k$, which may differ from $j$), and $\left(t_{\text{exit supply},i}^{N_i^{\text{out}}} - t_{\text{exit},i}^{N_i^{\text{out}}-1}\right)$ represents the admissible headway for entering the next reservoir in route $i$.

For the most constrained exit approach, respecting the downstream inflow limitations is ensured by defining the exit supply time as the supply time of the actual route of vehicle $N_j^{\text{out}}$:

$$t_{\text{exit supply}}^{N_j^{\text{out}}} = t_{\text{exit supply},j}^{N_j^{\text{out}}} \qquad \text{(local supply time)} \tag{4.33}$$

where $t_{\text{exit supply},j}^{N_j^{\text{out}}}$ is the local supply time of route $j$, defined by the next reservoir entry limitation.

### 4.3.5 Example of two routes

We present here the above-mentioned results with a simple case of a sudden demand increase in two routes. The network and MFD characteristics are the same as in section 4.2.2. The reservoir configuration is presented in Figure 4.4(a), route 1 has a length of $L_1$ = 2000 m, route 2 of $L_2$ = 1000 m.

#### 4.3.5.1 Illustration of inflow merge when both routes are congested

Figure 4.4(b) presents a scenario of a demand gap on both routes. After 1000 s we have $\lambda_1(t)$ = 1.8 veh/s and $\lambda_2(t)$ = 2 veh/s, whereas the exit supplies are $\mu_1(t) = \mu_2(t)$ = 0.5 veh/s. Note that such extreme and rather unrealistic boundary settings will bring the reservoir state close to gridlock, nevertheless this numerical application is only intended to illustrate the action of the inflow limitation when both routes are highly congested. Figures 4.4(c-g) respectively show the evolution of the average trip length $L(t)$, the accumulations, the mean speed, the inflows and outflows. One can see in Figure 4.4(f) that before congestion reaches the reservoir entry, the total inflow is fully limited by the capacity $P_c/L(t)$ between $t$ = 1000 s and 1200 s while the total accumulation $n$ is still below $n_c$. Once $n \geq n_c$, the limitation becomes $P(n)/L(t)$ which reproduces spillbacks reaching the entries dynamically. Such limitations at entry allow the system to reach a steady state after $t$ = 2000 s, the inflow reaches the same value as outflow on each route, see Figure 4.4(g). The outflow share in steady state will be described in more details latter.

#### 4.3.5.2 Illustration of inflow merge when only one route is congested

A similar demand gap scenario is now studied, but the boundary settings are chosen to create congestion on route 1 only. Here after 1000 s we have $\lambda_1(t)$ = 1.5 veh/s and $\lambda_2(t)$ = 0.4 veh/s, whereas the exit supplies are $\mu_1(t) = \mu_2(t)$ = 0.5 veh/s. Figures 4.5(a1) and (b1) present the evolution of inflows and entering productions respectively, when each route inflow is calculated through equation 4.20 only. One can clearly see that either the total inflow or production capacity is not fully used although inflow on route 1 is significantly limited. Figures 4.5(a2) and (b2) present the same results when the fair merge described

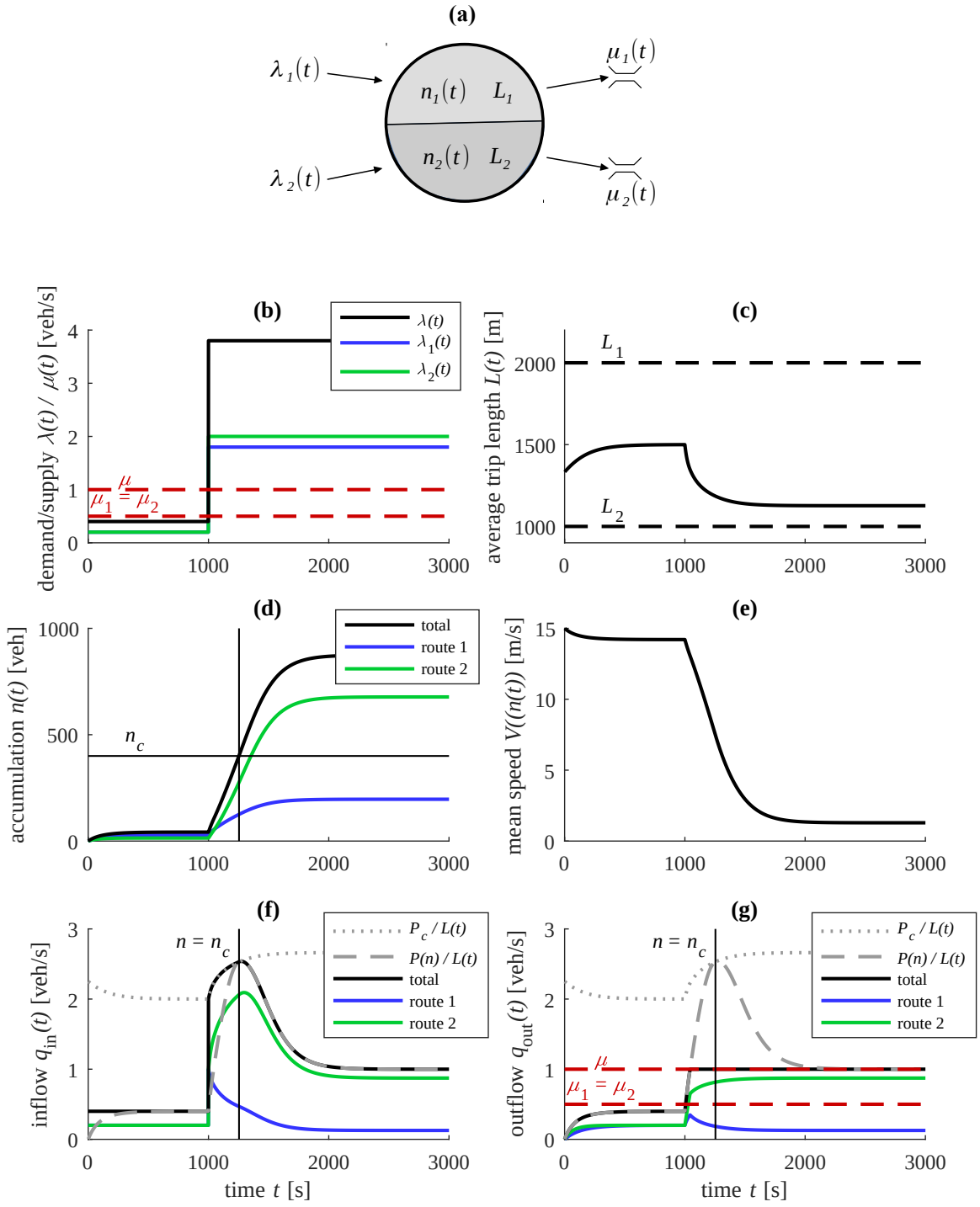



FIG. 4.4 – *(a) Two routes in a reservoir. (b) Case of a demand gap with demand $\lambda_i(t)$ and supply $\mu_i(t)$ profiles, (c) average trip length $L(t)$, (d) accumulation $n_i(t)$ for each route $i$, (e) mean speed $V(n(t))$, (f) inflow $q_{\text{in},i}(t)$ and (g) outflow $q_{\text{out},i}(t)$ for each route $i$*

in section 4.3.2.2 is applied to allocate the remaining flow capacity to route 1. We observe that even though all the flow capacity is used as shown in Figure 4.5(a2), the total entering production exceeds the production capacity, see Figure 4.5(b2). Such an inconsistency

leads us to rather apply the fair merge on production instead, which is presented in Figures 4.5(a3) and (b3). In Figure 4.5(b3), the total entering production is now using the full available capacity, however, we are still missing a small portion of the total inflow capacity, see Figure 4.5(a3). As we show in a simple example in B.1, both inflow and entering production capacity cannot be fully used at the same time, except from very particular situations. Figures 4.5(a4) and (b4) present the same results in the trip-based model with the fair merge on production. A similar evolution is observed, which validates the methods introduced in section 4.3.4 to reproduce our findings from the accumulation-based framework.

#### 4.3.5.3 Illustration of outflow diverge

All these previous results have been performed using the total supply approach. In Figure 4.4(g), we notice that the total outflow corresponds to $\mu(t) = \mu_1(t) + \mu_2(t)$ = 1 veh/s at equilibrium. However it appears clearly that the equilibrium flow on route 2 exceeds its limitation $\mu_2(t)$ = 0.5 veh/s. Consequently, the next reservoir in the considered route may encounter gridlock if such a situation lasts a significant period. As explained previously, this approach may have significant impacts in a multi-reservoir context, which explains why we designed the second method based on the most constrained exit.

Before illustrating the differences in simulation between these two approaches, it is interesting to note that it is almost impossible to identify in advance the most constrained exit, prior to the simulation. While this exit can be determined at each time step or each event during the simulation as shown in sections 4.3.3 and 4.3.4, we cannot predict what the outflow share will be in steady state, and thus find the final critical exit, before simulating the whole congestion propagation in the reservoir. Finding the final outflow distribution would actually require to solve the system equation 4.13 under congested conditions with the first approach. Using equations 4.23 and 4.27 leads to: $\forall i \in \{1, \dots, N\}, q_{\text{in},i}(t) = n_i(t)/L_i V(n(t)) = q_{\text{out},i}(t) = \frac{n_i(t)}{n(t)} \frac{L(t)}{L_i} \mu(t)$, thus $n(t)/L(t)V(n(t)) = \mu(t)$, but this is insufficient to get the final solution. Actually, solving analytically this system seems intractable as the equilibrium state in congestion depends also on the demands $\lambda_i(t)$. This is shown in Figure 4.6(a), where the outflow ratio $q_{\text{out},2}/q_{\text{out},1}$ in steady state is displayed for different simulation runs. We observe that a high demand on a route does not necessarily imply a high outflow on this route. This is because the outflow ratio also depends on the trip lengths, namely a high demand on a long route often results in a low outflow compared to the short route. Actually, four different cases of outflow share can be exhibited out of these results, as presented in Figure 4.6(b):

- Case 1: intuitive, $\lambda_1 < \lambda_2$ gives $q_{\text{out},1} < q_{\text{out},2}$
- Case 2: less intuitive, $\lambda_1 < \lambda_2$ gives $q_{\text{out},1} > q_{\text{out},2}$
- Case 3: less intuitive, $\lambda_1 > \lambda_2$ gives $q_{\text{out},1} < q_{\text{out},2}$
- Case 4: intuitive, $\lambda_1 > \lambda_2$ gives $q_{\text{out},1} > q_{\text{out},2}$

Cases 1 and 4 are intuitive, as the order of flows (according to the demands) is respected in the outflow share. But cases 2 and 3 are less intuitive, as the order of flows is reversed in the outflow share. This phenomenon is due to the differences in trip lengths. Although these graphs can provide some insights about the inner dynamics of such a system, they still depend on specific values for $\lambda_1$ and $L_1$, and could be hardly generalized to three or more routes.

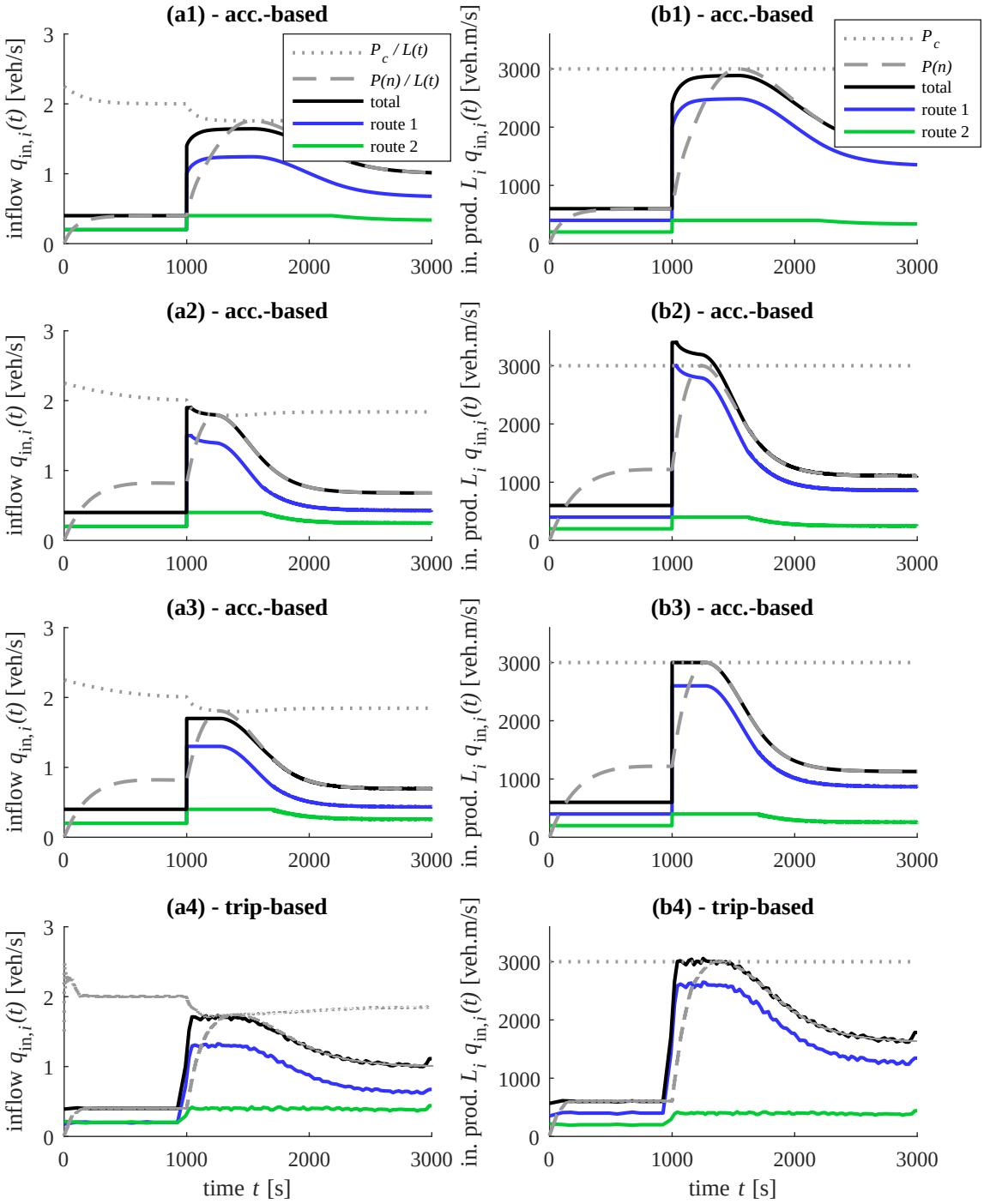



FIG. 4.5 – *(a1) inflow $q_{\text{in},i}(t)$ for each route i calculated without fair merge, (a2) with a fair merge on flows, (a3) with a fair merge on entering production for the accumulation-based and (a4) trip-based model. (b1) entering production $L_i q_{\text{in},i}(t)$ for each route i calculated without fair merge, (b2) with a fair merge on flows, (b3) with a fair merge on entering production for the accumulation-based and (b4) trip-based model*

The differences between the two outflow management approaches are presented in Figures 4.7(a) and (b) for the accumulation-based model, and 4.7(c) and (d) for the trip-based model. The outflow share corresponds to case 4 here, with $L_1$ = 1000 m, $L_2$ = 1800 m, $\lambda_1$ = 1 veh/s, and $\lambda_2$ = 0.9 veh/s. Like the case of one trip length in section 4.2.2, the

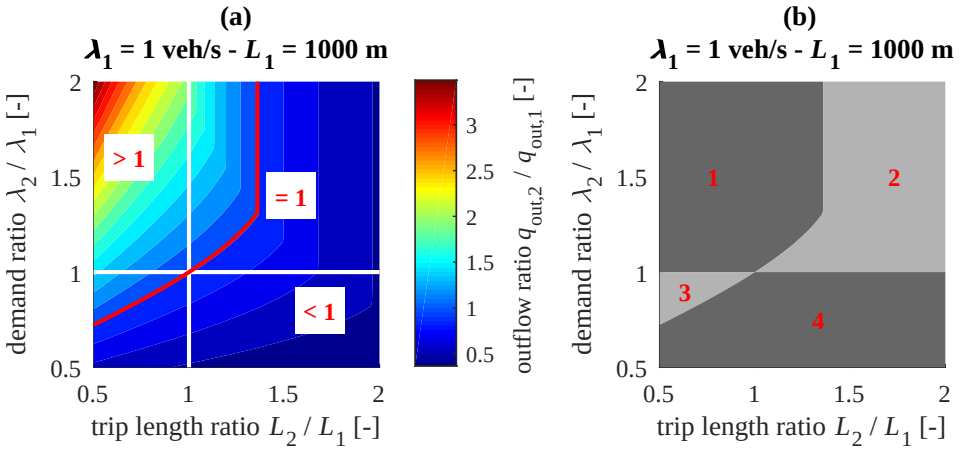



FIG. 4.6 – *(a) Outflow ratio $q_{\text{out},2}/q_{\text{out},1}$ in steady state for different simulations with different trip length $L_2/L_1$ and demand $\lambda_2/\lambda_1$ ratios, (b) four areas in the same graph delimiting different steady state cases*

accumulation- an trip-based models give similar results in congestion for both approaches. Figure 4.7(a) shows the evolution of $(n_1(t), n_2(t))$ in the accumulation plane, where the total trip completion rate function $G(n_1, n_2)$ is also plotted. The red line corresponds to all the equilibrium points verifying $G(n_1, n_2) = \mu(t) = 1$ veh/s. Although the evolution of accumulations is quite similar in both outflow managements, the flow equilibrium is completely different. Figure 4.7(b) shows the evolution of $(q_{\text{in},1}(t), q_{\text{in},2}(t))$ and $(q_{\text{out},1}(t), q_{\text{out},2}(t))$ in the flow plane. Each route exit limitations $\mu_1$, $\mu_2$ are indicated, along with the brown area which represents the total supply. One clearly notices that the only point allowing a maximum global outflow and complying with each local constraint is $\mu_1 = \mu_2 = 0.5$ veh/s. This would be possible if $L_1 = L_2$ and $\lambda_1 = \lambda_2$. But here the first route sends the higher outflow, this is the critical exit in this scenario. With the first approach, its outflow exceeds the downstream limitation, leading the next reservoir to gridlock. Whereas with the second approach, its outflow is set to the right limitation, reducing automatically the outflow of route 2 (the longest) in the same time.

The same conclusions are observed in Figure 4.8, where the outflow share corresponds to case 3, with $L_1 = 1000$ m, $L_2 = 500$ m, $\lambda_1 = 1$ veh/s, and $\lambda_2 = 0.9$ veh/s. It is interesting to notice how the dynamics of the system give priority to route 2 here, whereas this route has a lower inflow demand. Note also that for both accumulation- and trip-based models, the most constrained exit approach does not converge to a point on the red line in the accumulation plane $(n_1, n_2)$. This line actually corresponds to the equilibrium $G(n_1, n_2) = \mu(t) = 1$ veh/s, which is reached with the total supply method only. Whereas with the most constrained approach, the system reaches an equilibrium where the total steady flow is below the total supply $\mu(t)$.

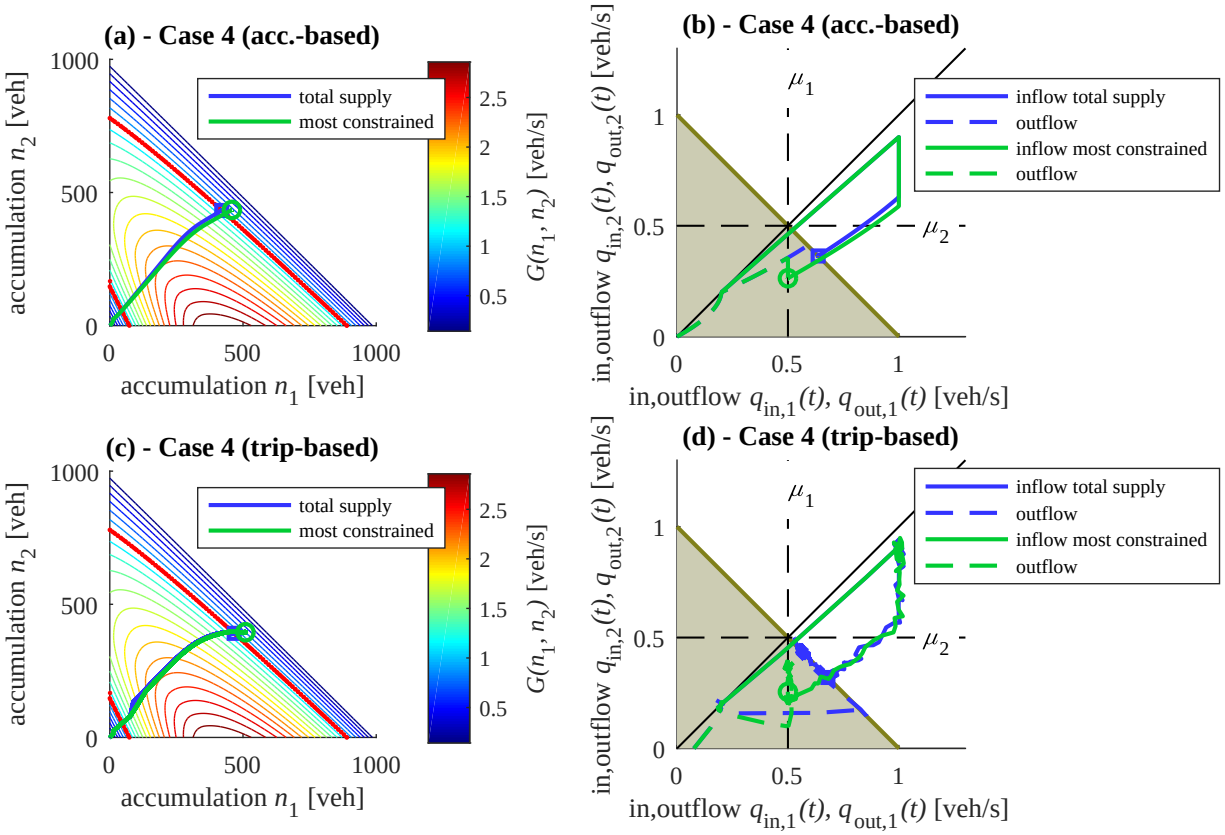



FIG. 4.7 – *Evolution of the system in case 4. (a) Accumulation-based model with the total supply and most constrained supply approaches, evolution of* $(n_1(t), n_2(t))$ *in the accumulation plane, (b) evolution of* $(q_{\text{in},1}(t), q_{\text{in},2}(t))$ *and* $(q_{\text{out},1}(t), q_{\text{out},2}(t))$ *in the flow plane, (c) same results with the trip-based model, evolution of accumulations, (d) inflows and outflows*

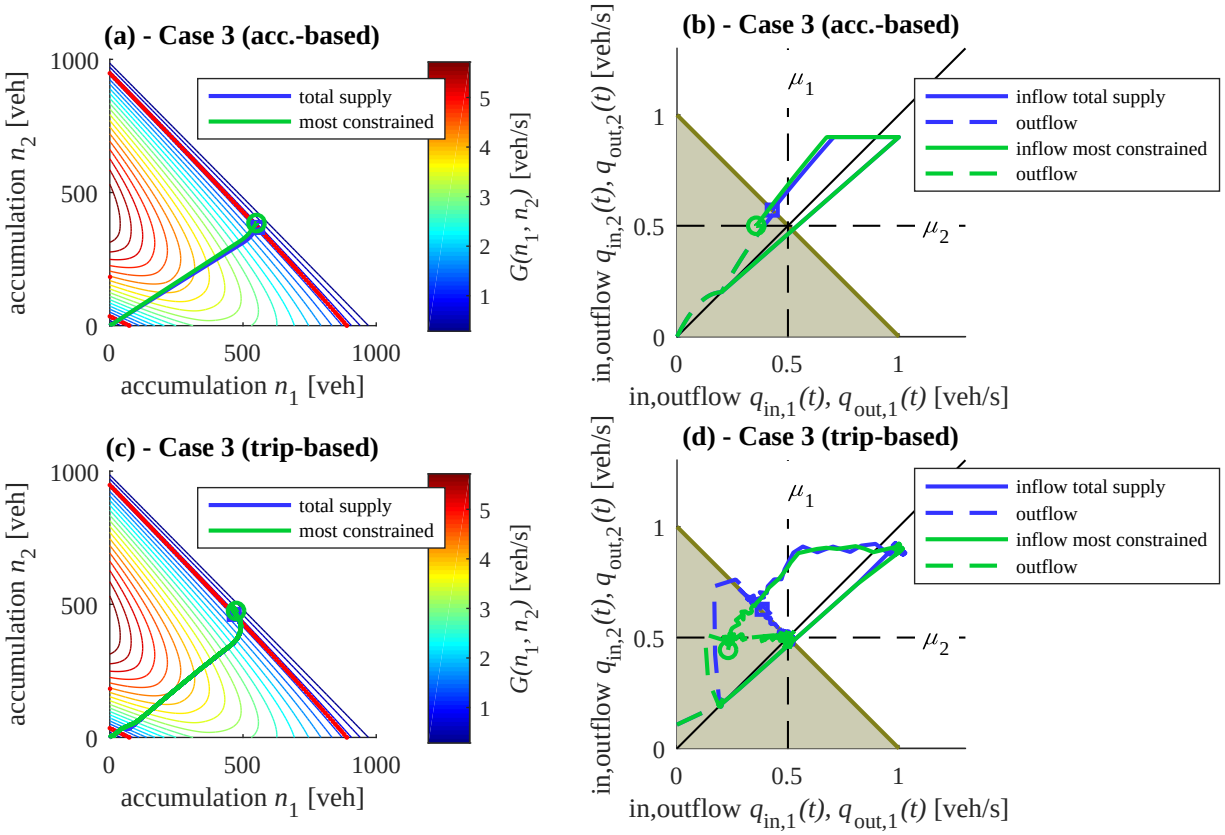



FIG. 4.8 – *Evolution of the system in case 3. (a) Accumulation-based model with the total supply and most constrained supply approaches, evolution of* $(n_1(t), n_2(t))$ *in the accumulation plane, (b) evolution of* $(q_{\text{in},1}(t), q_{\text{in},2}(t))$ *and* $(q_{\text{out},1}(t), q_{\text{out},2}(t))$ *in the flow plane, (c) same results with the trip-based model, evolution of accumulations, (d) inflows and outflows*

## 4.4 Comparisons with existing models

In this section, we present an application of our framework in a simple multi-reservoir environment. We aim at illustrating some phenomena we observed with the previous two-route case, and also compare our framework with two models from the literature.

### 4.4.1 Test case configuration and scenario

The reservoir system configuration is presented in Figure 4.9(a). It basically consists in two macroscopic routes [$R_1$ $R_3$ $R_4$ $R_5$ $R_7$] (route 1) and [$R_2$ $R_3$ $R_4$ $R_6$ $R_8$] (route 2) that share some reservoirs in common. The idea is to see the influence of the modeling choices on the flow exiting from $R_4$ and its consequences on the two downstream reservoirs $R_5$ and $R_6$. The characteristics of the reservoirs are mentioned in table 4.1.

TAB. 4.1 – *Reservoir characteristics, where $L_1$ refers to route 1 and $L_2$ refers to route 2*

| Characteristics | [units] | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ | $R_6$ | $R_7$ | $R_8$ |
|---|---|---|---|---|---|---|---|---|---|
| jam accumulation $n_j$ | [veh] | 1000 | 1000 | 4000 | 1000 | 500 | 500 | 500 | 500 |
| max. production $P_c$ | [veh.m/s] | 3000 | 3000 | 3000 | 3000 | 3000 | 3000 | 600 | 600 |
| free-flow speed $u$ | [m/s] | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| trip length $L_1$ | [m] | 500 | – | 500 | 1000 | 500 | – | 1000 | – |
| trip length $L_2$ | [m] | – | 500 | 500 | 600 | – | 500 | – | 1000 |

The demand scenario is presented in Figure 4.9(b). This scenario together with the reservoir configurations are designed to create congestion on both routes due to the low entry capacity of $R_7$ and $R_8$ (equal to $P_c$, see table 4.1). Thanks to these capacities and the high jam accumulation of $R_3$, the origins $R_1$ and $R_2$ and destinations $R_7$ and $R_8$ are kept free of congestion during the whole simulation period. All queuing vehicles are then stored in $R_3$, $R_4$, $R_5$ and $R_6$.

The two-route case from section 4.3.5 is reproduced in $R_4$ with $L_1$ = 1000 m different from $L_2$ = 600 m. In the following, we will focus on the evolution of the total accumulation in reservoirs $R_4$, $R_5$ and $R_6$ only. Many more things could be described in this system, but the idea is to keep the simulation analysis simple and exhibit some features from the two-route case that have been already explained in details in the previous section.

### 4.4.2 Comparison between the most constrained exit and total supply approaches

Figure 4.9(c) shows the results for the accumulation-based model with the most constrained exit and total supply approaches. Because the demand peak for route 1 arises first, full access is given to this route at the entry of $R_4$ from $t$ = 2000 to 4000 s approximately. During this period, we have a high flow on the longest route and a low flow on the shortest one, which is similar to the situation of section 4.3.5.2. In such a case, we have seen that the total supply approach sends too much flow on the longest route. In the present configuration, this results

in an excess of flow into $R_5$ where spillbacks happen due to the limitation at its exit, i.e. the low entry capacity of $R_7$. Therefore $R_5$ converges to gridlock around $t$ = 3000 s as shown by the evolution of accumulation (green dashed line). On the contrary, the most constrained exit approach (green solid line) limits the flow of route 1 out of $R_4$, so that the situation in $R_5$ evolves very close to gridlock but avoid it as the equilibrium in flow is preserved at each time with this approach. Once accumulation in $R_5$ starts emptying after $t$ = 7000 s, a similar evolution happens in $R_6$ (red solid line). Then, exit of route 2 from $R_4$ becomes the most critical exit in $R_4$, so that gridlock is also avoided in $R_6$, though getting very close to it.

Figure 4.9(d) shows the results for the trip-based model with the most constrained exit and global supply approaches. Not surprisingly, very similar conclusions are observed with this framework too, as the modeling of spillbacks is similar to the accumulation-based model. Both models differ from each other mostly in free-flow conditions. Almost no significant discrepancies are noticed between them here, because our system is quite simple and our test scenario has been designed to exhibit highly congested situations. However, much more differences may be expected in a more realistic case study.

### 4.4.3 Comparison between the most constrained exit approach and two other frameworks from the literature

We then compare the framework of the most constrained exit approach in the accumulation-based model with two simulators from the literature. The first one is developed in Yildirimoglu & Geroliminis (2014) and can account for different trip lengths within a reservoir. The results are shown in Figure 4.9(e). It clearly appears that our considerations about inflow and outflow definition have a significant impact on the simulation results, compared to what is usually done in the literature. To analyze these results, we choose to focus on the most important discrepancy between the two simulations: the congestion of $R_6$ which is predicted in our model but not in the other. This can be mainly explained by the different treatments of the flow entering into $R_4$ and exiting from $R_4$. At $t$ = 3000 s, the sharp increase of accumulation in $R_5$ is due to the limitation at the entry of $R_7$. Because the flow transferring from $R_4$ to $R_5$ becomes limited too, congestion propagates into $R_4$ until a highly oversaturated steady state at $t$ = 4000 s. Up to this time, our simulation and the other one experience quite the same evolution. However, during the propagation of congestion in $R_4$ the outflow from route 2 is not treated the same way in our model. To fulfill the constraint on the total exiting production of $R_4$ as described by equation 4.25, this outflow has to adapt to the dynamics of $R_4$ (the sudden decrease of the mean speed), and is thus also reduced. This interdependency between all the outflows from a same reservoir is precisely discussed in section 4.3.3. Whereas in the model of Yildirimoglu & Geroliminis (2014), this constraint on the total production is not respected, and the two outflows from $R_4$ can therefore behave independently, the outflow on route 1 being restricted, while the one on route 2 can remain a bit higher. It entails that the queue on route 2 in $R_4$ is growing faster in our model because of the lower outflow, which consequently allows more flow to enter (remember that the entry supply is proportional to the ratio $n_i/n$, i.e. the more vehicles on route $i$, the more could enter). The situation is a bit different in the other model as the inflow share is operated through a pro-rata merge, but the smaller queue on route 2 in $R_4$ is also responsible for a lower inflow. In the end, once the exit to $R_5$ (route 1) is no longer the most constrained one after $t$ = 6000 s, the queue on route 2 in $R_4$ can directly exit into $R_6$, which creates congestion in $R_6$

due to the limitation at the entry of $R_8$. Whereas in the other model, the flow into $R_6$ remains low, thus avoiding the creation of congestion. The counterpart is that the vehicles are stored in $R_4$ for a longer period, up to 11,000 s as shown by the evolution of accumulation (blue dashed line).

The second simulator is presented in Knoop & Hoogendoorn (2014) and can distinguish flows between simple routes but assumes the same trip length for all travelers in each reservoir. In this second case, we make some minor modifications in the reservoir settings to ensure a fair comparison: we set $L_1 = L_2 = 1000$ m in $R_4$. We should notice that such a simplification makes flow and production constraints equivalent, as $L_i = L$ for all routes $i$ crossing a given reservoir. The results are shown in Figure 4.9(f). The differences between both models look quite similar as in the previous comparison. But in this case, they are mainly explained by the inflow treatment at the entry of $R_4$. Like in our previous analysis, our model allows more flow to enter $R_4$ on route 2, whereas the other one uses a pro-rata merge at the entry of $R_4$ which results in a different inflow share. Such a situation, as described above, creates a longer queue on route 2 in $R_4$ in our model, which is then responsible of the congestion appearing in $R_6$.

These two comparisons highlight the major difference between our framework and the approaches from other authors, because we focus on the production constraint. In our case, including production limitations both at the entry and exit perimeter of the reservoir provides additional constraints to the definition of inflow and outflow share. Whereas the mentioned authors have more "freedom" to design the allocation of inflows and outflows while only the flow capacity is respected. They usually make pro-rata merges at entry, but one can imagine that other choices like fair merges would also work. We should notice that the case with several trip categories sharing the same trip length, like in Knoop & Hoogendoorn (2014), is a bit special, as flow and production limitations merge into a single constraint. In this case any merge and diverge models would automatically comply with the production restrictions once the ones on flow are verified. Nevertheless, our approach is more robust in this case too, as the extension to multiple trip lengths is straightforward, which is not true for a simulator with another design of merge and diverge than ours.

One still needs to remember that without access to any ground truth, we cannot benchmark the different models over a well-defined reference. However, based on an analysis from the kinematic wave theory, we believe that production constraints must not be ignored. As shown in this study, their inclusion introduces significant changes in the treatment of inflow merge and outflow diverge for one reservoir, and thus different modeling of flow exchanges in multi-reservoir systems.

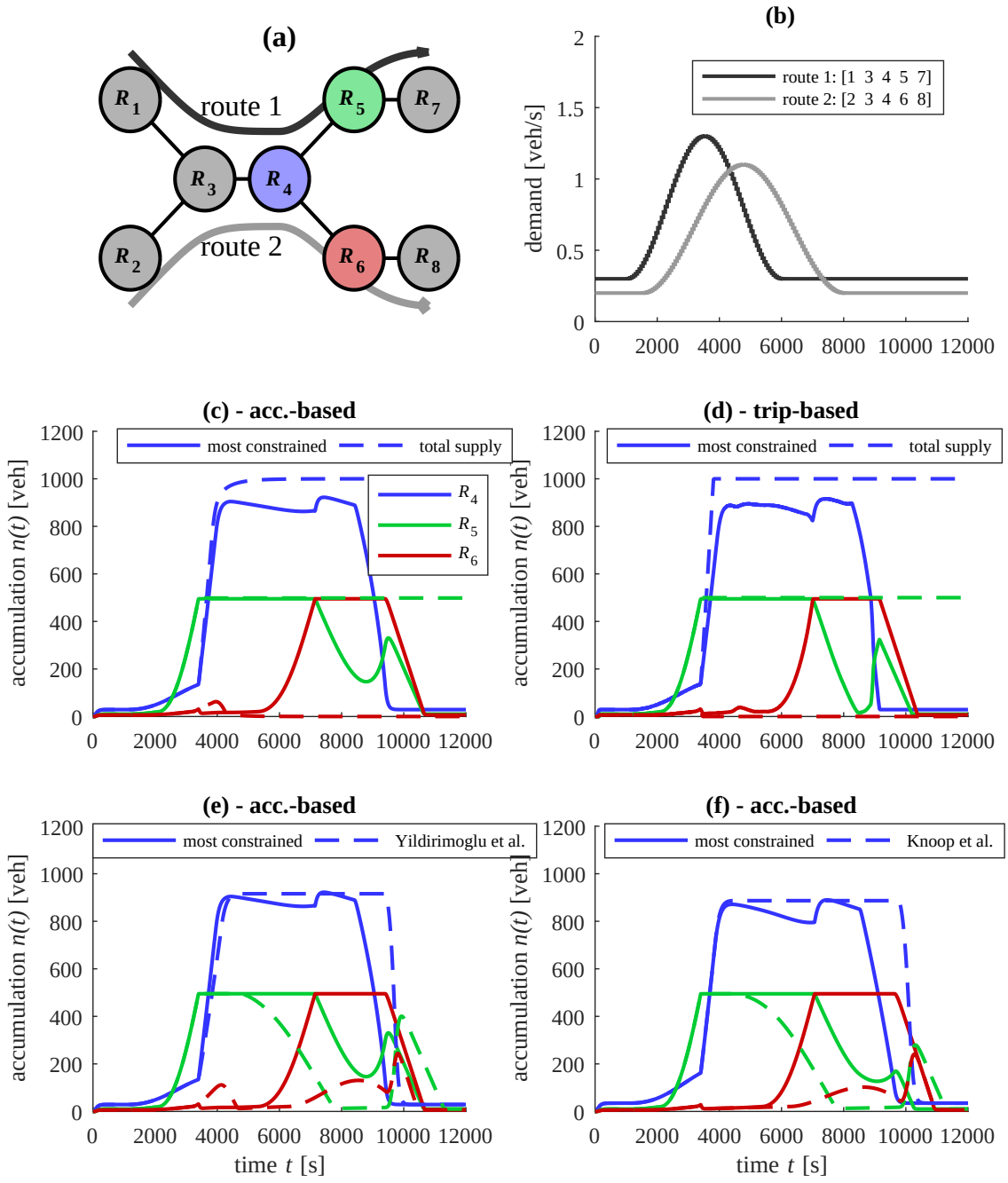



FIG. 4.9 – *(a) Reservoir system configuration, (b) demand profiles for the two routes. (c) Accumulation in reservoirs 4, 5 and 6, comparison between the two outflow management approaches with the accumulation-based model and (d) the trip-based model, (e) comparison between the most constrained supply approach and the model of Yildirimoglu & Geroliminis (2014), (f) comparison between the same approach and the model of Knoop & Hoogendoorn (2014) with identical trip lengths in* $R_4$

## 4.5 Discussion

### 4.5.1 Conclusion about flow exchanges for transfer trips only

This study proposed a general framework to handle flow exchanges in multi-reservoir systems for the accumulation and trip-based models, in particular when congestion spillbacks. We focused on a single reservoir crossed by multiple commodities (e.g. macroscopic routes or trip categories with potential different trip lengths) because it corresponds to the building block of any multi-reservoir simulator. In comparison with existing works from the literature, a major contribution of this chapter is the introduction of production constraints at the reservoir perimeter, defined through the dynamic average trip length. It appears that accounting for such constraints leads to a unique possible design of inflow merge at the reservoir entry, and outflow diverge at the reservoir exit. Although we have developed two approaches to handle outflows, we have illustrated with many examples that only the most constrained exit approach is sustainable in a multi-reservoir model. As shown with a simple multi-reservoir configuration at the end of this chapter, significant differences in simulation results are expected depending on the choice made by the modeler to include production constraints or not. According to basic principles from traffic flow theory, we believe that they must not be ignored, while we still need comparisons with some ground truth (microscopic simulation or real data) to validate our assumptions. The framework we developed (entry flow functions and allocation scheme for downstream capacities) resorts to the accumulation-based formulation, but we demonstrated that it can be easily implemented in the trip-based formulation too by using the same entry supply functions. Not surprisingly, the results are similar for both modeling approaches when congestion is propagating. This constitutes a second important contribution, as the trip-based model has only been studied in undersaturated conditions to the authors' best knowledge.

### 4.5.2 Introduction of internal trips

The last ingredient we can include in our framework is internal trips (starting and ending in the reservoir). Inflows and outflows for such trips should have special treatments as they can start or end anywhere within the reservoir and should not be restricted by perimeter constraints. As such, we assume that internal inflow is unrestricted, i.e. $q_{\text{in},i}(t) = \lambda_i(t)$ for the internal trip category $i$. As for internal outflow (rate at which the users reach their destinations), we consider that it decreases proportionally to the production-MFD in the accumulation-based model, i.e. $q_{\text{out},i}(t) = n_i/n.P(n)/L_i$, or that the vehicles go on driving at the speed-MFD during congestion until they reach destination for the trip-based model. This modeling is intended to capture the limitation of outflow due to internal congestion, as it is traditionally assumed in the literature (see e.g. Daganzo, 2007, Geroliminis & Daganzo, 2007). Nevertheless, if a transfer flow happens to be limited at the reservoir exit, the effective outflow of internal trips will be also impacted, as it still depends on the other outflows through equation 4.29 (most constrained exit approach). Actually, all outflows, internal or transferring, are the result of the reservoir inner dynamics and must then comply with the exit production constraint defined in equation 4.25. A simulation example is presented in Figure 4.10(a) for a simple demand peak scenario with two trips, internal and transferring, in one reservoir. The evolution of the system is shown in Figures 4.10(b), (c) and (d). It

clearly appears that even with the inflow boundary at the perimeter the reservoir can reach oversaturated regimes due to the presence of internal trips. Figure 4.10(d) shows that the total inflow temporarily exceeds the reservoir capacity $P_c/L(t)$. We can also notice a clockwise hysteresis loop for the exit production of internal trips. Note that internal trips may easily lead the reservoir to gridlock when the internal inflow is high, see Mahmassani *et al.* (2013b) for more information on urban gridlock. Although this simple test case certainly needs more investigations, it proves that the integration of internal trips is straightforward in our framework. Note also that other kinds of trips like the ones starting in the reservoir and exiting the reservoir (or the reverse case) simply consist in a combination of internal and transferring trips. In general, a trip starting in the reservoir will not be limited for inflow, and a trip ending in the reservoir will have an outflow decreasing with total accumulation in congestion, just as stated earlier.

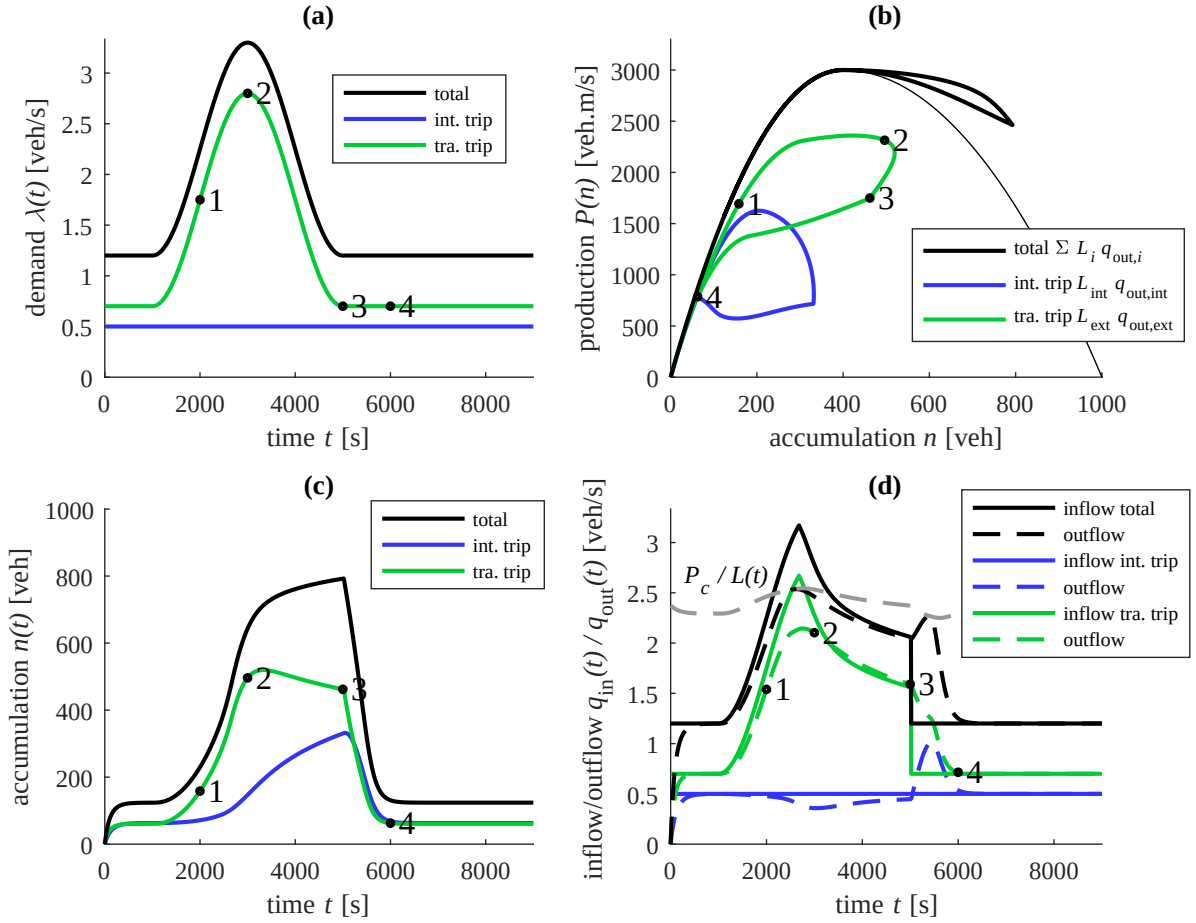



FIG. 4.10 – *Case of a demand peak for internal trips in a single reservoir. (a) demand profile $\lambda(t)$ for internal trips and external trips, (b) evolution of the system in the (accumulation $n_i$, production $P_i = L_i q_{\text{out},i}$) plane for each trip class $i$ ("internal" or "external"), (c) evolution of accumulation and (d) inflow and outflow for each trip class*

### 4.5.3 General conclusion and future work

This study constitutes a significant step forward a deeper understanding on flow exchanges in multi-reservoir systems with multiple regional routes. This chapter focused on theoretical considerations to ensure the framework we developed is consistent with some basic principles from traffic flow theory. The natural next step is to focus on the validation of this framework based on real observations, or by comparing with microscopic simulation results.

Further developments of our MFD-based simulator will include a macroscopic route choice set generator and a DTA module for several applications like routing strategies, search-for-parking or perimeter control. The application of our framework with a control system can be envisioned in two different ways: the first one is the comparison of the control strategy when our approach is used instead of a simpler one; and the second one is the preliminary validation of different control strategies (based on a simple model) by simulating the resulting traffic states with a more detailed multi-reservoir and multi-route framework. Despite the relative complexity of the latter, this second option is considerably less demanding than setting a microscopic simulation in terms of computational effort and demand estimation.

# 5 Comparison between flow exchange modeling approaches and validation versus microscopic simulation

Multi-reservoir systems based on the Macroscopic Fundamental Diagram (MFD) appear attractive for simulating traffic states at large urban scales. Nevertheless, flow limitation due to spillbacks between reservoirs is still not fully understood. In the previous chapter, we analyzed flow exchange properties and formulated new merging and diverging models, different from those previously proposed in the literature. However, both the latter and our approach received very little support from aggregated link-scale data (real or simulated).

The contribution of this chapter is to validate different MFD-based modeling approaches by using microscopic simulation. To this end, we focus on an artificial Manhattan network crossed by two regional paths (West-East and North-South). We then analyze the aggregated accumulation, inflow and outflow of these paths for different congestion scenarios, and compare the results with the outputs predicted by the MFD-based models. These comparisons first allow us to investigate and calibrate the network entry capacity, known as the entry supply function. Then we show that using a demand pro-rata merging scheme in the MFD-based model, as is usually done in the literature, is the best option for determining partial inflows properly. Finally, regarding the network exit, we also show that the outflow diverging scheme is critical to reproduce microsimulation results well. During congestion onset, the widely used approach of a decreasing outflow demand with independent partial outflow treatment provides a good estimation of transient states, but fails to predict reliable network unloading during congestion offset. The modified approach we developed in our previous study takes advantage of setting the reservoir outflow demand to maximum and applying inter-dependency relationships between outflows to overcome this issue.

This chapter is an updated version of the conference paper:

> Mariotte, G., Paipuri, M. & Leclercq, L. (2019). Flow exchanges in multi-trip MFD-based systems: A validation study against microscopic simulation, In *Transportation Research Board 98th Annual Meeting*, Washington DC, USA

## 5.1 Motivations

Since the early works of Daganzo (2007), Geroliminis & Daganzo (2007), using the Macroscopic Fundamental Diagram (MFD) to simulate traffic states at the city scale has attracted increasing interest in the literature. In particular, numerous studies (see e.g., Kouvelas *et al.*, 2017, Sirmatel & Geroliminis, 2017a, Zhong *et al.*, 2017, Yang *et al.*, 2018) have used MFD-based simulation to design promising traffic control frameworks for large-scale networks, where such networks are split into several homogeneous reservoirs (urban areas) with a well-defined MFD. However, there is still a lack in understanding flow exchanges and limitations at the reservoir boundaries in multi-reservoir systems. More precisely, we identify three research questions which need to be investigated: (i) How can the maximum available flow that can enter a reservoir, for both under- and oversaturated conditions be defined dynamically? (ii) How can inflow merging be managed? and (iii) How can outflow diverging be managed when different demand flows are distinguished (by their origins, destinations, or regional paths)?

The literature provides no complete response to any of these questions. The first refers to what is sometimes called the "entry supply function" of the reservoir (also named "receiving capacity" or "boundary capacity"). It is assumed to be a decreasing function of the reservoir accumulation, and represents the total available inflow that can enter the reservoir. Its value is maximum when the accumulation is small (undersaturation), and goes down to zero when the accumulation is high (oversaturation). Its existence has been shown by Geroliminis & Daganzo (2007) through a simulation study. Other authors like Hajiahmadi *et al.* (2013b), Knoop & Hoogendoorn (2014), Lentzakis *et al.* (2016) adopted an entry supply function with a shape based on the Cell Transmission Model (CTM) of Daganzo (1994). The reason behind this approach is that the theoretical formulation of the reservoir dynamics may correspond to cell dynamics in the CTM. Thus, connecting together a sequence of reservoirs would be similar to connecting a sequence of cells. This solution is appealing as we know from kinematic wave theory that such a formulation efficiently reproduces congestion propagation between entities. However, the objection to this approach is that the cell properties cannot be scaled up at the reservoir level. In particular, the maximum inflow capacity and the critical accumulation (at which the inflow capacity starts decreasing) of the entry supply function may be different from the MFD's capacity and critical accumulation. For instance, Geroliminis & Daganzo (2007) showed with their simulation results that the critical accumulation of the supply function is double that of the critical accumulation of the MFD. In Ramezani *et al.* (2015), Sirmatel & Geroliminis (2017a), a simple piecewise linear function was used with these two parameters (the maximum inflow capacity and the critical accumulation). This was also probably the approach taken by Yildirimoglu & Geroliminis (2014), Yildirimoglu *et al.* (2015), although not explicitly mentioned in their works. Interestingly, these authors designed a supply function for each reservoir boundary (thus each function limits inflow from a specific neighboring reservoir), but they did not provide further details about the interactions between these functions. Their implementation was merely justified by technical reasons, i.e. avoiding gridlocks in some scenarios, and the authors acknowledged the fact that their shapes would have negligible influence on their simulation results due to the action of perimeter control or route guidance. This was notably illustrated by Sirmatel & Geroliminis (2017a) who performed a sensitivity analysis on the two function parameters they used, and showed that these parameters had little impact on their study.

More recently, Kim *et al.* (2018) explored flow exchanges between reservoirs with the results from a micro-simulation, and proposed an even simpler shape for this function (linearly decreasing), also used in Zhang *et al.* (2015). Nevertheless, only inflow and outflow shares were tested with network loading scenarios, and the proposed entry supply function was not implemented in an MFD-based simulation.

The second problem about inflow allocation is often solved by using merging rules based on demand pro-rata (see e.g., Geroliminis, 2009, Knoop & Hoogendoorn, 2014, Yildirimoglu & Geroliminis, 2014, Ramezani *et al.*, 2015), or less often fair merging rules (see e.g., Zhong *et al.*, 2018). Ge & Fukuda (2018) developed a unified merge and diverge model similar to Jin & Zhang (2004) to determine turning fractions in their multi-reservoir modeling, which is also based on demand pro-rata rules. On the other hand, we found in chapter 4 that incorporating constraints on production into the merging scheme potentially leads to a different flow allocation. In this case, the inflow share at entry would also depend on the reservoir inner dynamics (i.e. evolution of each partial accumulation) and the different trip lengths inside the reservoir.

Finally, the third question on outflow diverging is rarely investigated in the literature. To the author's best knowledge, almost all studies on MFD-based modeling assume a decreasing reservoir outflow in oversaturated conditions. However, by reproducing the effect of reservoir internal congestion, and through simulation and empirical studies, in chapters 2 and 4 we showed that this assumption may lead to inefficient modeling of congestion propagation between reservoirs. They insisted in particular on the fact that the outflow demand of transferring trips that aim to enter a neighboring reservoir must be maximum in congestion, which is what would be observed on a simple network like an arterial. Recently, Wada *et al.* (2018) derived an analytical method to track congestion patterns and spillbacks in networks by solving an inverse Dynamic User Equilibrium (DUE) problem. Their method provided an analytical formulation of the network throughput, validated with link-level simulation. They showed that the network exit flow remains constant in oversaturated conditions for simple configurations, but decreases for more complex grid networks where users have many different destinations.

In this work, we want to further investigate these three questions by comparing different assumptions in MFD-based approaches with microsimulation outputs. Our case study consists of a regular grid network crossed by two main regional flows (West-East and North-South). On the one hand, network-level traffic states on each regional flow (accumulation, inflow, outflow, production) are estimated by aggregating the link-level outputs from the microsimulation. These states are assumed to represent the ground truth of the grid network traffic dynamics. On the other hand, aggregated traffic states are simulated with an MFD accumulation-based model including multiple trips, so that each regional flow is represented by a partial accumulation, inflow and outflow. Two different representations of traffic dynamics exist in the literature, namely accumulation-based and trip-based models, see chapter 2 for an extensive review. As we will focus on highly congested situations in this study, we use only the accumulation-based approach and not the trip-based one, as the latter was found to behave similarly to the accumulation-based model in such situations. At the network entry, two inflow merging schemes are tested against the microsimulation results: (i) demand pro-rata merging, depending on the ratio of external demands that want to enter the reservoir (commonly used in the literature); and (ii) endogenous merging, depending on the ratio of partial accumulation inside the reservoir (our new approach, see

chapter 4). Also tested at the network exit are two outflow diverging schemes based on (i) a decreasing outflow demand in oversaturation (widely used in the literature), and (ii) based on a maximum outflow demand in oversaturation (our approach). We particularly focus on steady state predictions of network loading scenarios, where congestion is created by fixing exogenous flow limitation at the network exit links. This is the most stressful situation for MFD models as oversaturation is discussed less in the literature. The network MFD is calibrated using stationary network loadings with microsimulation. Our results first allow the calibration of a reliable entry supply function for the network studied, and second, the comparison between the above-mentioned inflow merging and outflow diverging models. Generally, using the demand pro-rata merge at entry with the maximum outflow demand at exit is found to be the best option for reproducing the microsimulation results well.

This chapter takes the following structure: in section 5.2, a review on existing MFD-based modeling approaches is proposed. Then in section 5.3, the simulation case study is presented and the network MFD is estimated. Finally, the microsimulation results are analyzed and compared with the different MFD models introduced earlier, section 5.4 deals with the entry merging problem, while section 5.5 is about the exit diverging problem.

## 5.2 Multiple trips in the single reservoir model: review of existing approaches

In this section, we present the main approaches that have been developed in the literature to describe flow exchanges in a single reservoir with multiple trip categories.

### 5.2.1 General framework

We consider an urban area described by a single reservoir model (Daganzo, 2007, Geroliminis & Daganzo, 2007), i.e. where traffic states are represented by a well-defined production-MFD $P(n)$ (in [veh.m/s]) or speed-MFD $V(n) = P(n)/n$ (in [m/s]), $n(t)$ (in [veh]) being the total accumulation (number of vehicles traveling in the reservoir at time $t$). The production-MFD is defined by: the jam accumulation $n_j$ and the critical accumulation $n_c$ where the production reaches its maximum $P_c = P(n_c)$. We investigate flow exchanges in the framework of multiple trip categories, as presented in Geroliminis (2009, 2015) and further used in Yildirimoglu & Geroliminis (2014), Ramezani *et al.* (2015), Haddad (2015), Zheng & Geroliminis (2016). In this study, a trip category defines a "macroscopic route" (sometimes called "regional path") or simply "route" in the following. It corresponds to the aggregation of multiple individual paths on the real street network that shares certain common characteristics (e.g. similar topology or length, following the same sequence of reservoirs in multi-reservoir systems, etc.). It usually has its own length, which requires considering multiple trip lengths in each reservoir. The reader can refer to Batista *et al.* (2019) for a more detailed description of the methods used to aggregate link-level traveled distances into reservoir-level trip lengths.

Let us assume that the single reservoir considered comprises $N$ routes with lengths $\{L_i\}_{1\leq i\leq N}$ and corresponding accumulations $\{n_i(t)\}_{1\leq i\leq N}$, the total accumulation being $n(t) = \sum_{i=1}^{N} n_i(t)$. The system dynamics are described by the following conservation equa-

tions (Geroliminis, 2015):

$$\forall i \in \{1, \dots, N\}, \qquad \frac{dn_i}{dt} = q_{\text{in},i}(t) - q_{\text{out},i}(t) \tag{5.1}$$

where $q_{\text{in},i}(t)$ and $q_{\text{out},i}(t)$ are route $i$ effective inflow and outflow, respectively. These flow values therefore govern the entire evolution of the system, and are the result of the entry supply and merging and diverging schemes that are presented next. The general configuration of the reservoir exchanging flows with its neighbors is summarized in Figure 5.1.

Depending on the route origin and destination (outside or inside the reservoir), its inflow and outflow treatment may be different. Therefore, to include any possible case, we define four sets of routes:

- $\mathcal{P}_{\text{in}}^{\text{ext}}$ contains all the routes that originate outside the reservoir,
- $\mathcal{P}_{\text{out}}^{\text{ext}}$ contains all the routes that end outside the reservoir,
- $\mathcal{P}_{\text{in}}^{\text{int}}$ contains all the routes that originate inside the reservoir,
- $\mathcal{P}_{\text{out}}^{\text{int}}$ contains all the routes that end inside the reservoir.

Any route $i$ is included in two of the above-mentioned sets.

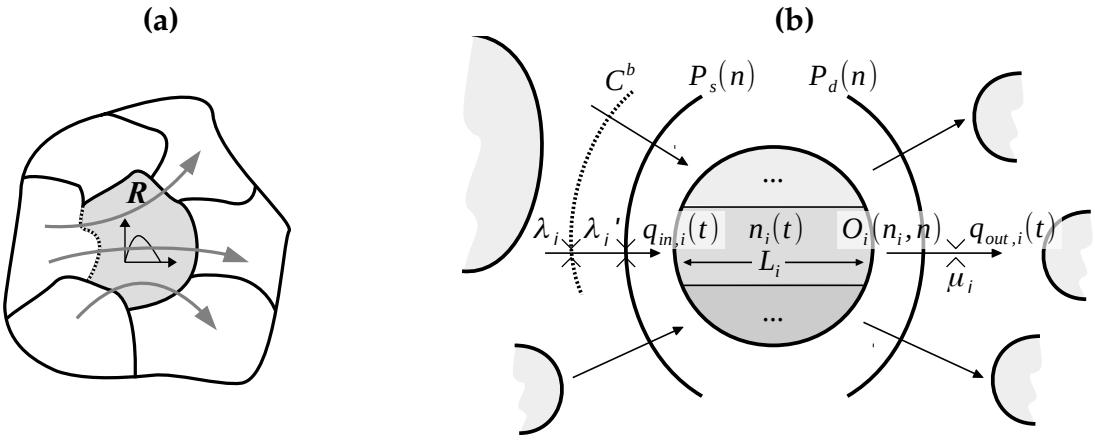



FIG. 5.1 – *Single reservoir configuration crossed by multiple routes, (a) multi-reservoir context where one of the reservoir borders is represented by a dashed line, and (b) detail of flow exchanges on a given route*

### 5.2.2 Reservoir entry supply and inflow merging scheme

The question of inflow limitation only concerns the routes that originate outside the reservoir and thus have to cross the entry boundary (set $\mathcal{P}_{\text{in}}^{\text{ext}}$). The other routes in $\mathcal{P}_{\text{in}}^{\text{int}}$ correspond to vehicles created inside the reservoir that do not have their inflow restricted:

$$\forall i \in \mathcal{P}_{\text{in}}^{\text{int}}, \qquad q_{\text{in},i}(t) = \lambda_i(t) \tag{5.2}$$

where $\lambda_i(t)$ is the inflow demand for route $i$.

#### 5.2.2.1 Total entry supply

At the reservoir entry, the inflow from neighboring reservoirs may be limited by several factors. Two situations must be distinguished depending on the source of limitation. First, in undersaturated conditions, a natural restriction is the sum of the physical capacity of all the links at a specific border of the reservoir (links connected to a specific neighboring reservoir). In this case, it makes sense to design independent and fixed flow capacities per reservoir border, as in Ramezani *et al.* (2015), Sirmatel & Geroliminis (2017a). If we assume that the reservoir considered has $N_B$ borders (i.e. common boundaries with neighboring reservoirs), these capacities $\{C^b\}_{1\leq b\leq N_B}$ are expressed as:

$$\forall b \in \{1,\dots,N_B\}, \qquad C^b = \sum_{\text{link } l\in b} q_l \tag{5.3}$$

where $q_l$ is the flow capacity of link $l$ (in [veh/s], including traffic signal settings of its extremity node intersection).

Second, in oversaturated conditions, the cause of inflow restriction may be seen in two different ways, conveying the same idea but with different implications: (i) the reduction of available space in the reservoir, or (ii) the propagation of congestion waves. The first interpretation (i) accounts for the fact that the higher the number of vehicles traveling in the reservoir, the fewer can enter it. This can be formulated with a simple linearly decreasing function like $C(n) = w_{cs}(n_j - n)$ for $n > n_{cs}$ (the subscript $cs$ is used for the related parameters of the entry supply function to avoid confusion with the MFD parameters). It corresponds to the choice of Ramezani *et al.* (2015), Zhang *et al.* (2015), Sirmatel & Geroliminis (2017a), Kim *et al.* (2018). However, the definition of the rate $w_{cs}$ (in [$s^{-1}$]) at which the inflow reduction applies is critical. In Ramezani *et al.* (2015), Sirmatel & Geroliminis (2017a), this rate depends on the border $b$ considered, it is set as $w_{cs}^b = C^b/(n_j - n_{cs})$. The other interpretation (ii) is related to kinematic wave theory, where inflow reduction mimics congestion spilling back to the entry. In this case, the propagation speed of this information has a physical meaning as it should precisely correspond to the derivative of the congested MFD branch. This interpretation assumes that the MFD can be considered as an extension of the Fundamental Diagram (FD) concept at the network level. That is the viewpoint adopted in chapter 4 and by Hajiahmadi *et al.* (2013b), Knoop & Hoogendoorn (2014), Lentzakis *et al.* (2016). These authors defined inflow limitation with the function $C(n)$ by analogy with the CTM supply function: $C(n) = P_c/L$ for $n \leq n_c$ and $C(n) = P(n)/L$ for $n > n_c$, where $L$ is the average trip length, assumed constant by some authors, or defined dynamically through $L = n \Big/ \sum_{i=1}^{N} \frac{n_i}{L_i}$ (Geroliminis, 2009) by others. The latter formulation ensures that any equilibrium in congestion complies with Little's formula (Little, 1961).

Based on these different approaches, in this study we thus assume a general shape of a decreasing entry supply function, denoted $P_s(n)$, that reproduces spillbacks when the reservoir is oversaturated, as in the second interpretation (ii). It is expressed in production units, because this definition is seen as an intrinsic property of the reservoir, and therefore should be free of any trip length dependency. As in chapter 4, we postulate that internal trip creation may reduce the available entry capacity for transferring routes. Hence, the total entry supply should be modified accordingly: $P_s^{\text{ext}}(n) = P_s(n) - \sum_{i\in\mathcal{P}_{\text{in}}^{\text{int}}} L_i\lambda_i$. This modified entry supply function can be converted into flow units as $P_s^{\text{ext}}(n)/L^{\text{ext}}$ thanks to the dynamic average trip length of routes from outside $L^{\text{ext}} = \sum_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}} n_i \Big/ \sum_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}} \frac{n_i}{L_i}$. The validation and

calibration of this entry supply function will be investigated in section 5.4.1.

#### 5.2.2.2 Merge coefficients

Once we can determine the total available reservoir inflow, an allocation scheme must be designed to calculate partial inflows. For each route $i \in \mathcal{P}_{\text{in}}^{\text{ext}}$, the effective inflow $q_{\text{in},i}(t)$ is the minimum between its corresponding demand $\lambda_i(t)$ from an upstream reservoir and a proportion $\alpha_i(t)$ of the total available inflow. As described in the introduction, to the authors' best knowledge the merge coefficients $\alpha_i$ are almost always assumed to correspond to pro-rata demand (see e.g., Geroliminis, 2009, Knoop & Hoogendoorn, 2014, Yildirimoglu & Geroliminis, 2014, Ramezani *et al.*, 2015):

$$\forall i \in \mathcal{P}_{\text{in}}^{\text{ext}}, \quad \alpha_i(t) = \frac{\lambda_i(t)}{\sum_{j \in \mathcal{P}_{\text{in}}^{\text{ext}}} \lambda_j(t)} \qquad \text{(demand pro-rata coefficients)} \tag{5.4}$$

On the other hand, in chapter 4 we proposed these coefficients as defined endogenously by the reservoir state:

$$\forall i \in \mathcal{P}_{\text{in}}^{\text{ext}}, \quad \alpha_i(t) = \frac{n_i(t)}{\sum_{j \in \mathcal{P}_{\text{in}}^{\text{ext}}} n_j(t)} \qquad \text{(endogenous coefficients)} \tag{5.5}$$

to comply with the two following constraints at the reservoir entry:

$$\sum_{i \in \mathcal{P}_{\text{in}}^{\text{ext}}} q_{\text{in},i}(t) \leq \frac{P_s^{\text{ext}}(n(t))}{L^{\text{ext}}(t)} \qquad \text{(flow constraint)} \tag{5.6}$$

$$\sum_{i \in \mathcal{P}_{\text{in}}^{\text{ext}}} L_i q_{\text{in},i}(t) \leq P_s^{\text{ext}}(n(t)) \qquad \text{(production constraint)} \tag{5.7}$$

The first constraint indicates that the sum of all route inflow restrictions must equal the global reservoir restriction. The second expresses the same idea, but applies to the entering productions instead of inflows. Moreover, unlike previous approaches, these authors split the total entering production into partial entering productions rather than total inflow into partial inflows. They showed that when all the partial inflows are limited, only the choice of endogenous merge coefficients on entering productions can ensure that the total inflow equals $P_s^{\text{ext}}(n)/L^{\text{ext}}$ and that the total entering production equals $P_s^{\text{ext}}(n)$. Note that as long as their sum is one, the choice of any other merge coefficients (including the demand pro-rata coefficients) on inflows ensures by construction that the flow constraint is always satisfied, but not necessarily the production constraint.

#### 5.2.2.3 Merging scheme

Finally, in this study we combine the two sources of inflow limitation for undersaturated and oversaturated situations by designing a two-layer merging scheme, as illustrated in Figure 5.1(b). The first layer applies the border flow restrictions to original demands $\lambda_i$, while the second layer takes the restricted demands $\lambda_i'$ as inputs and applies the reservoir supply $P_s^{\text{ext}}(n)$ to them. For any given merge coefficients, the first layer always consists of a flow merge. On the other hand, the second layer is either a flow merge when using demand

pro-rata coefficients (with available capacity $P_s^{\text{ext}}(n)/L^{\text{ext}}$), or a production merge when using endogenous coefficients (with available capacity $P_s^{\text{ext}}(n)$). This two-layer scheme is described as follows:

$$\forall b \in \{1,\dots,N_B\}, \qquad \{\lambda'_i\}_{i\in\mathcal{P}^b} = \text{Merge}\left(\{\lambda_i\}_{i\in\mathcal{P}^b}, \left\{\frac{\alpha_i}{\sum_{j\in\mathcal{P}^b}\alpha_j}\right\}_{i\in\mathcal{P}^b}, C^b\right) \tag{5.8}$$

$$\begin{cases} \{q_{\text{in},i}\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}} = \text{Merge}\left(\{\lambda'_i\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}}, \{\alpha_i\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}}, \frac{P_s^{\text{ext}}(n)}{L^{\text{ext}}}\right) & \text{for demand pro-rata coeff.} \\ \{L_i q_{\text{in},i}\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}} = \text{Merge}\left(\{L_i\lambda'_i\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}}, \{\alpha_i\}_{i\in\mathcal{P}_{\text{in}}^{\text{ext}}}, P_s^{\text{ext}}(n)\right) & \text{for endogenous coeff.} \end{cases} \tag{5.9}$$

where each set $\mathcal{P}^b$ gathers all the routes crossing the corresponding border $b$. Equation 5.8 corresponds to the first layer merge, and equation 5.9 to that of the second layer. The merge algorithm used here was presented in Leclercq & Becarie (2012) and consists of an extension of the fair merge of Daganzo (1995). For any set of $M$ merging demands $\{\Lambda_i\}_{1\le i\le M}$ with respective merge coefficients $\{\alpha_i\}_{1\le i\le M}$ towards a unique entry with capacity $C$, the resulting effective inflows $\{Q_i\}_{1\le i\le M}$ are calculated as:

$$\forall i \in \{1,\dots,M\}, \quad Q_i = \begin{cases} \Lambda_i & \text{if } \Lambda_i \le \alpha_i C \\ \dfrac{\alpha_i}{\sum_{Q_j>\alpha_j C}\alpha_j}\left(C - \displaystyle\sum_{Q_j\le\alpha_j C} Q_j\right) & \text{otherwise} \end{cases} \quad \text{(fair merge)} \tag{5.10}$$

This merging algorithm ensures that the total available capacity $C$ is always used when only certain inflows are limited while others are not. Its principle is that all demands $\Lambda_i$ below their respective limitation $\alpha_i C$ are served, and then the remaining capacity is shared among the remaining inflows according to merge coefficients. If some of these remaining inflows exceed their respective demand after sharing, then these demands are served and the remaining capacity is adjusted accordingly and shared among the remaining inflows. This process is repeated until all the inflows are served and/or the capacity $C$ is fully used. Note that this algorithm is applied similarly for both productions and flows.

This single reservoir framework also includes a point-queue model for each route to store queuing vehicles at entry when the corresponding demand is not satisfied. Once a queue has formed for a specific route $i$, its demand $\lambda_i$ is set to maximum, equal to its border capacity $C^b$, provided that the queue is not empty.

### 5.2.3 Reservoir exit demand and outflow diverging scheme

The question of outflow diverging concerns all the routes within the reservoir, since it is the result of its internal dynamics. However, the difference between the two sets $\mathcal{P}_{\text{out}}^{\text{ext}}$ and $\mathcal{P}_{\text{out}}^{\text{int}}$ is that each route $i$ ending outside the reservoir may undergo an exogenous outflow limitation $\mu_i(t)$ when crossing the exit boundary (representing inflow restriction to the next reservoir), while the outflow of each route ending inside the reservoir is not exogenously limited:

$$\forall i \in \mathcal{P}_{\text{out}}^{\text{int}}, \qquad \mu_i(t) = +\infty \tag{5.11}$$

#### 5.2.3.1 Exit demand

When considering the routes in $\mathcal{P}^{\text{ext}}_{\text{out}}$, the exit demand is the rate at which users want to cross the reservoir exit boundary to reach a neighboring reservoir. The exit demand is not measurable in practice, but it is an essential concept to define the reservoir outflow in oversaturation. While rarely mentioned in other works, it is implicitly involved as soon as the potential outflow of a reservoir may be limited by the entry capacity of a neighboring reservoir, as in the studies mentioned earlier (e.g., Knoop & Hoogendoorn, 2014, Yildirimoglu & Geroliminis, 2014, Ramezani *et al.*, 2015, Sirmatel & Geroliminis, 2017a). In the accumulation-based model, this rate has been defined through the reservoir production-MFD since the seminal work of Daganzo (2007). Hence, as with the entry supply in the previous section, we define the exit demand function denoted $P_d(n)$ in production units, seen as an intrinsic property of the reservoir and free of any trip length definition. In the literature, the common approach is that $P_d(n)$ simply corresponds to the production-MFD, and thus decreases with $n$ in oversaturated states to mimic the effect of internal congestion:

$$P_d(n) = P(n) \qquad \text{(decreasing exit demand)} \tag{5.12}$$

However, in chapters 2 and 4 we pinpointed a drawback of this formulation for the routes crossing the exit boundary. During the offset of a congestion peak, these authors showed with simple simulation scenarios that oversaturation in the reservoir results in very low demand for outflow, and thus prevents it from proper recovery after the peak. They thus proposed that the exit demand should be maximum for the routes in $\mathcal{P}^{\text{ext}}_{\text{out}}$:

$$P_d(n) = \begin{cases} P(n) & \text{if } n < n_c \\ P_c & \text{otherwise} \end{cases} \qquad \text{(maximum exit demand)} \tag{5.13}$$

#### 5.2.3.2 Diverging scheme

Then, the total exit demand is split between the routes according to Little's formula. The outflow demand of the routes in $\mathcal{P}^{\text{ext}}_{\text{out}}$ is defined as (Geroliminis, 2015):

$$\forall i \in \mathcal{P}^{\text{ext}}_{\text{out}}, \qquad O_i(n_i, n) = \frac{n_i}{n} \frac{P_d(n)}{L_i} \qquad \text{(outflow demand)} \tag{5.14}$$

In chapter 4, we showed that to comply with the accumulation-based main assumption, i.e. that travelers are assumed to have the same speed $V(n)$, a proper outflow model must ensure that Little's formula is valid for all the routes within the reservoir in any steady state situation (under- or oversaturated): the flow on each route $i$ must equal $n_i V(n)/L_i$. This means that all the effective outflows are actually interdependent through the mean speed $V(n)$, or the production $P(n)$.

In the common approach in the literature (decreasing exit demand), this interdependency is ensured by the fact that the exit demand $P_d(n)$ always equals $P(n)$. Thus, the effective outflows can be calculated as:

$$\begin{cases} \forall i \in \mathcal{P}^{\text{ext}}_{\text{out}}, & q_{\text{out},i}(t) = \min[\mu_i(t); O_i(n_i, n)] \\ \forall i \in \mathcal{P}^{\text{int}}_{\text{out}}, & q_{\text{out},i}(t) = \frac{n_i(t)}{n(t)} \frac{P(n)}{L_i} \end{cases} \qquad \text{(for decreasing exit dem.)} \tag{5.15}$$

However, in the approach developed in chapter 4, the fact that we postulated $P_d(n)$ as maximum and constant in oversaturation introduces too many degrees of freedom in the definition of each partial outflow demand $O_i(n_i, n)$. Therefore, these authors used the following constraint on total exit production to ensure interdependency between all the effective outflows:

$$\sum_{i=1}^{N} L_i q_{\text{out},i}(t) = L(t) \sum_{i=1}^{N} q_{\text{out},i}(t) \qquad \text{(exit production)} \tag{5.16}$$

This constraint implies in particular that the definition of the dynamic average trip length $L(t)$ must always be valid at any time. Thus, the authors developed what they called the most constrained exit approach to calculate the effective outflows:

$$q_{\text{out},k}(t) = \min[\mu_k(t); O_k(n_k, n)] \qquad \text{(most constrained outflow)} \tag{5.17}$$

$$\text{where: } k = \arg\min_{1 \le i \le N} \frac{\mu_i}{O_i(n_i, n)}$$

$$\forall i \in \{1, \dots, N\}, i \neq k, \quad q_{\text{out},i}(t) = \frac{n_i(t)}{n_k(t)} \frac{L_k}{L_i} q_{\text{out},k}(t) \qquad \text{(for maximum exit dem.)} \tag{5.18}$$

The latter approach has been designed to comply with the exit production constraint in equation 5.16, while ensuring that none of the partial outflows $q_{\text{out},i}(t)$ exceeds its corresponding exogenous limitation $\mu_i(t)$. Note that these calculations apply to all the routes, i.e. in both $\mathcal{P}_{\text{out}}^{\text{ext}}$ and $\mathcal{P}_{\text{out}}^{\text{int}}$.

## 5.3 Presentation of the validation case study

In order to validate or invalidate the different approaches for the single reservoir model presented in the previous section, the MFD-based simulation results will be compared with heterogeneous microscopic simulation results aggregated at the network level. In the first step, we focus on the two-route case, for which the properties of MFD-based models have been extensively studied in chapter 4. The use of microsimulation on an artificial network is preferred versus real field data, because we need to control link-scale settings (e.g. distribution of link trip lengths to create different average trip lengths, distribution of link paths to ensure quite homogeneous traffic states) to make a good comparison. Moreover, we need a perfect estimation of regional information, such as inflow and outflow per main flow direction, to avoid any bias in the assessment of the MFD-based models. This information is usually very difficult to obtain in real situations. Nevertheless, the microsimulation settings are tuned to provide traffic states that are as realistic as possible. We include a Dynamic Traffic Assignment (DTA) procedure to approximate User Equilibrium (UE) conditions.

### 5.3.1 Network configuration

The network designed for this case study consists of a 5-by-15 Manhattan grid network, as illustrated in Figure 5.2(a). Each link is two-way with two lanes for each way, and 105 m long. Each intersection includes a traffic light with a cycle time $T_c$ of 60 s, green time $T_g$ of 30 s and offset of 0 s. The traffic dynamics on each lane are described by a triangular FD with typical parameter settings for urban traffic conditions: $k_j$ = 0.17 veh/m (jam density),

$w$ = 5.9 m/s (congestion wave speed) and $u$ = 15 m/s (free-flow speed). For each way, each link $l$ has thus the following capacity: $q_l = 2$ lanes $\times\, T_g/T_c \times q_c$ = 0.72 veh/s, where $q_c = k_j/(1/u + 1/w)$ is the maximum flow per lane.

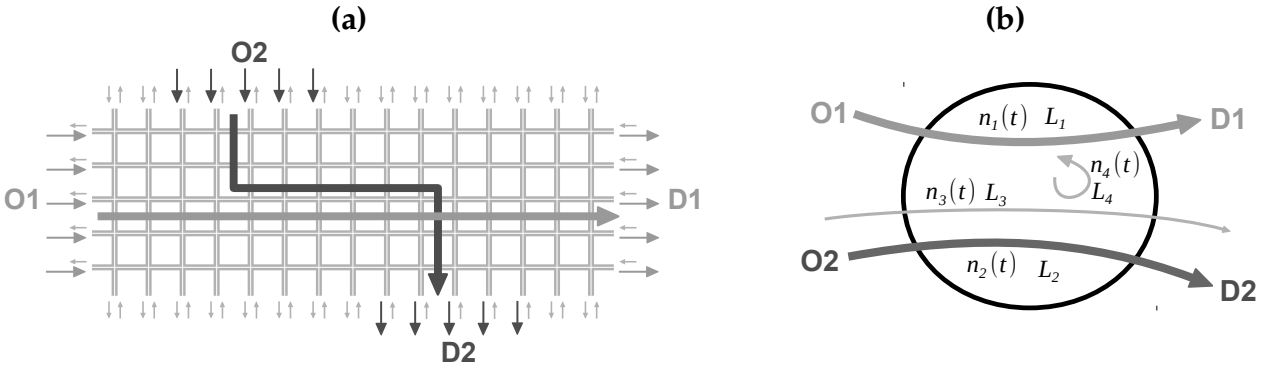



FIG. 5.2 – *(a) Grid network with two major OD pairs and background traffic, and (b) single reservoir modeling with four routes*

### 5.3.2 Simulation settings

Link-level traffic states are simulated by using the microsimulation platform *Symuvia* developed by LICIT (Univ. Lyon, France). This simulator is based on the car-following model of Newell (2002), the lane-changing model of Laval & Leclercq (2008) and further extensions for node merge models and multi-class traffic (Leclercq, 2007b,a, Leclercq & Laval, 2009, Chevallier & Leclercq, 2009). Network loading and unloading are investigated for a simulation duration of 20 hrs. Path flow distributions are updated for every period of 1 hr 40 min to reach UE conditions. The convergence loop uses the classical Method of Successive Averages (MSA), involving the travel times of each path averaged over the last time period (Lu *et al.*, 2009, Ameli *et al.*, 2018). Two main flow directions are defined: O1-D1 aggregates all the trips from the West to the East boundary, and O2-D2 comprises all the trips from 5 entries in the North to five exits in the South boundary, as depicted in Figure 5.2(a).

The demand scenarios for these trips always include a 30 min warm-up period, possibly followed by a high demand surge per origin O1 and/or O2. The total demand for a given macroscopic origin (O1 or O2) is equally distributed among its corresponding entry links. Each origin link flow is sent evenly to the exit links of the corresponding macroscopic destination (D1 for O1, D2 for O2). For a given macroscopic exit (D1 or D2), each exit link thus receives the same portion of flow. Apart from the two main flows O1-D1 and O2-D2, the background traffic is set with a total constant demand of 0.075 veh/s evenly distributed among all the remaining entries and exits. A total constant demand of 0.035 veh/s is also set for the internal trips with the origin and destination links evenly distributed inside the network.

### 5.3.3 MFD-based modeling

In parallel to microsimulation, traffic states are predicted by a multi-trip single reservoir model as described in section 5.2. The reservoir model includes four routes, i.e. 1, 2, 3 and

4, representing link-level trips from O1 to D1, from O2 to D2, and background and internal traffic respectively, see Figure 5.2(b). Based on our review in the previous section, different approaches will be studied, named as follows:

- MFD model 1/1: uses pro-rata merge coefficients at entry, and decreasing exit demand at exit
- MFD model 1/2: uses pro-rata merge coefficients at entry, and maximum exit demand at exit
- MFD model 2/2: uses endogenous merge coefficients at entry, and maximum exit demand at exit

The trip length of each route is estimated as the mean of all individual trips recorded in several microsimulations: $L_1$ = 1850 m, $L_2$ = 1250 m, $L_3$ = 1350 m and $L_4$ = 1330 m. These mean values may change by about 100 m from one simulation to another. The global routing in this network has been designed to obtain a significant difference between the trip lengths $L_1$ and $L_2$ of the two routes 1 and 2. The border capacities of O1 and O2 are calculated as: $C^1 = C^2 = 5 \times q_l$ = 3.6 veh/s.

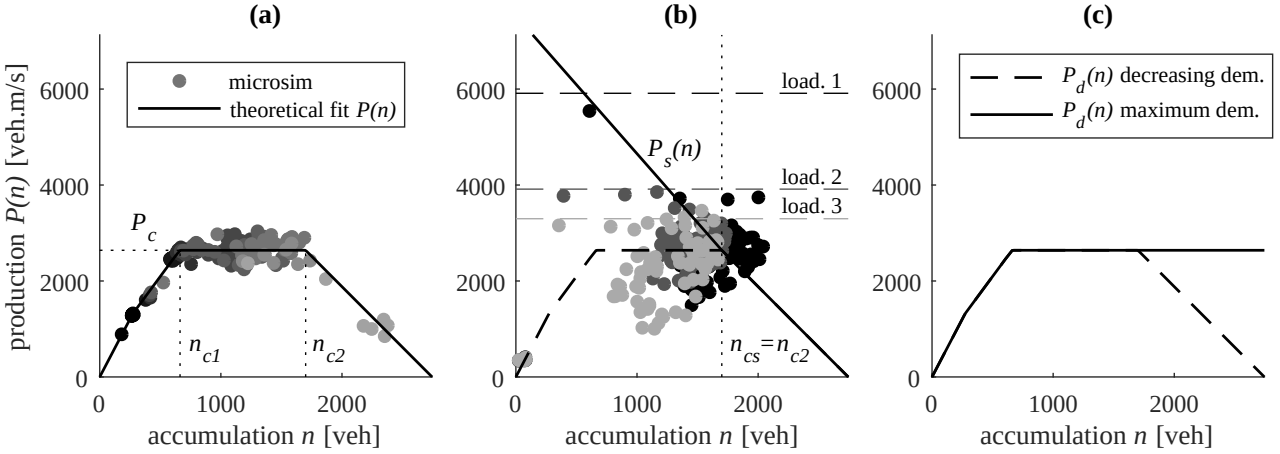



FIG. 5.3 – *(a) Estimation of the production-MFD $P(n)$ of the grid network using different loading scenarios in microsimulation, (b) the calibration of the entry production supply $P_s(n)$ with three constant network loadings indicated by three different shades of gray, and (c) exit production demand $P_d(n)$ for the two diverging approaches*

The production-MFD $P(n)$ of the grid network is estimated with several microsimulations of constant demand loading on all entries. The results of 8 different simulations are presented in Figure 5.3(a), where total production and accumulation are aggregated over 10 min periods. This permits calibrating a piecewise linear MFD with maximum production $P_c$ = 2640 veh.m/s and two critical accumulations $n_{c1}$ = 660 veh, $n_{c2}$ = 1700 veh defining the flat domain in $P(n)$. The points that determine the congested branch of $P(n)$ are obtained by limiting the network outflow exogenously. There are two ways of creating congestion in the network to estimate the congested branch of the MFD. The first is frequently used in other studies and consists in increasing the number of internal trips. Such scenarios favor the creation of local bottlenecks inside the network, thus reducing the average speed of all traveling vehicles. The second method instead focuses on the boundary conditions and exogenously

limits the potential outflow of exit links. In this case bottlenecks are explicitly generated at the network perimeter, from where congestion starts propagating until it finally reaches the entire network. In this study, we adopt this second method because we particularly focus on flow exchanges at the network borders. Thus, we want to calibrate an MFD that mainly describes the behavior of transferring trips. As for the exit demand function, the two models of $P_d(n)$ are directly derived from $P(n)$, as illustrated in Figure 5.3(c).

The first calibration of the entry supply function $P_s(n)$ is shown in Figure 5.3(b). In the first test, a simple way of calibrating $P_s(n)$ consists in loading the network with a total entering production demand $\sum_{i=1}^{4} L_i\lambda_i$ higher than the MFD capacity $P_c$, with each macroscopic demand $\lambda_i$ evenly distributed among all its corresponding entry links. As we know from the MFD calibration that the network cannot sustain a total production higher than $P_c$ at equilibrium, we expect these loading simulations to provide an evolution of the total entering production as follows: from the warm-up level to the demand level $\sum_{i=1}^{4} L_i\lambda_i$, and from the demand level to $P_c$. This latter transient period should provide insight into how the inflows/entering productions adapt dynamically to the network capacity. Moreover, as we assumed that $P_s(n)$ is an intrinsic property of the network, we need to run several loadings to ensure that a unique function is enough to describe the dynamic reduction of the total inflow. To this end, three constant loading scenarios were used to calibrate $P_s(n)$, their respective evolutions in the (accumulation, production) plane are plotted in the same figure with three different shades of gray. While the total entering production demand $\sum_{i=1}^{4} L_i\lambda_i$ of each loading is greater than the MFD capacity $P_c$, we observe in the microsimulation that the traffic states always reach the critical accumulation $n_{c2}$ and stabilize around the capacity $P_c$. Thus we conclude that $n_{cs} = n_{c2}$ is the critical accumulation of $P_s(n)$ in this network configuration, delimiting saturated and oversaturated states in the reservoir. To be consistent with the MFD definition, for $n > n_{cs}$ the entry supply function corresponds to the congested branch of $P(n)$, because this branch was obtained precisely when the inflow equilibrated with the outflow exogenous limitation we set (see the description of the MFD estimation above). On the other hand, the estimation of $P_s(n)$ for $n \leq n_{cs}$ was made using the highest demand loading case in Figure 5.3(b). For this setting, in the microsimulation the entering production decreases along the line of $P_s(n)$ we plotted. This first calibration serves as a baseline for running the MFD simulations. Its relevance is investigated further in the following section on inflow merging.

## 5.4 Comparisons between the entry merging schemes with network loading scenarios

The merging of inflows at the network entry is investigated with three sets of loading scenarios detailed below:

- Entry case 1: simulations with a high demand on route 1 and an increased demand on route 2 at each simulation;
- Entry case 2: simulations with a high demand on route 2 and an increased demand on route 1 at each simulation;
- Entry case 3: simulations with medium demands on routes 1 and 2, and an increased demand on route 4 (internal traffic) at each simulation.

These test cases have been designed to stress the network by setting high demand flows at entry, and thus queues are observed spilling back to the entry links due to interactions between the two routes 1 and 2 in the middle of the network. Each simulation is run for 3 hrs, including a 30 min warm-up period followed by an instantaneous demand increase on both routes at the same time. The network traffic states reach steady states after approximately 1 hr 30 min of simulation.

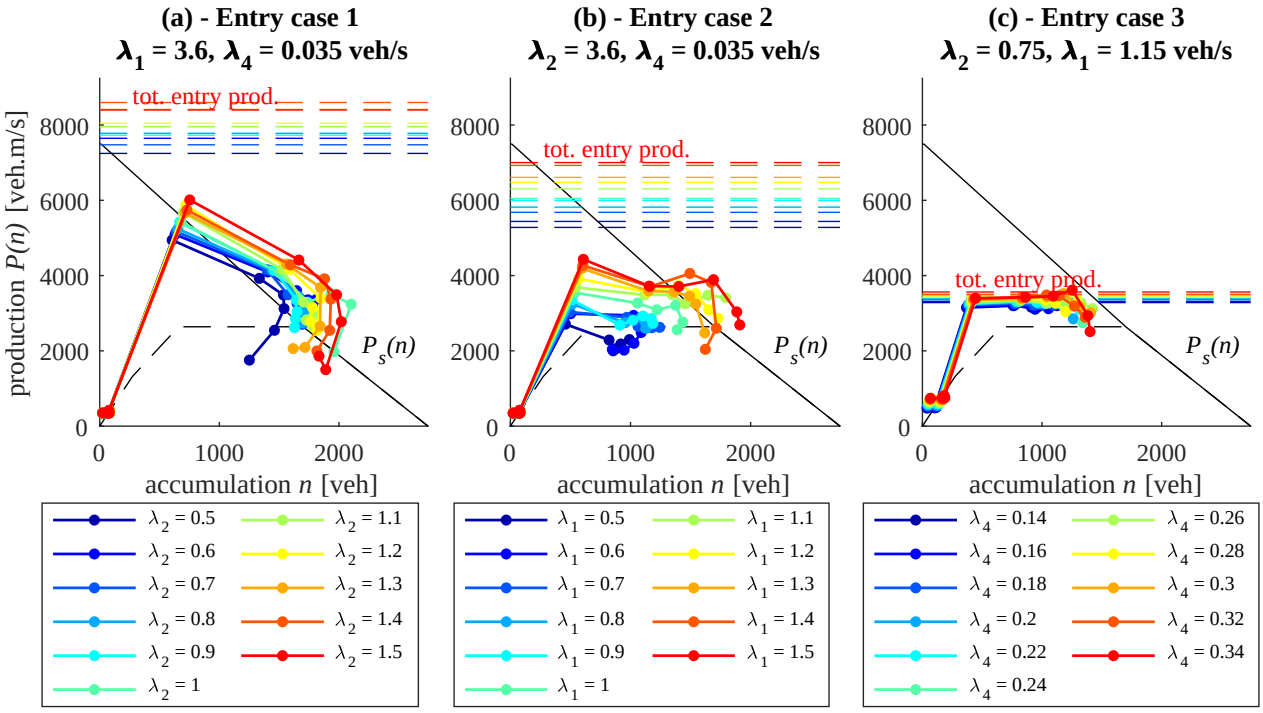



FIG. 5.4 – *Evolution of total entering production* $\sum_{i=1}^{4} L_i q_{\text{in},i}(t)$ *in the (accumulation, production) plane for different network loading cases in microsimulation. The total entering production demand* $\sum_{i=1}^{4} L_i \lambda_i$ *is indicated by the dashed line. Each color corresponds to one simulation. (a) Results of entry case 1 for 11 simulations with increasing demand* $\lambda_2$*, (b) entry case 2 for 11 simulations with increasing demand* $\lambda_1$ *and (c) entry case 3 for 11 simulations with increasing demand* $\lambda_4$

### 5.4.1 Investigation on the network entry supply function

Figure 5.4 shows the evolution of the total entering production $\sum_{i=1}^{4} L_i q_{\text{in},i}(t)$ in microsimulation for the three above-mentioned entry loading cases. For both cases 1 and 2, the demand of one route $i$ is set to its border capacity $C^i = 3.6$ veh/s ($i = 1$ for case 1 and $i = 2$ for case 2) while different simulations are run with an increased demand on the other route in each simulation. We note that the traffic states do not even reach the demand level during the loading. This is due to the aggregation period of 5 min used to calculate the inflows in the microsimulation. They must be higher than the signal cycle time of 1 min to smooth the inflow variations induced by green and red phases. Indeed, with an aggregation period of a few seconds, the entering production can be seen to reach the demand level for less than 1 min, and then rapidly decreas. However, the results are too scattered with such a small aggregation period to identify a clear trend in the transition period of the network loading. This means that the intersections close to the entries are not operating at the maximum

capacity of the entry links (including signal timings), but at a lower capacity due to interactions between turning vehicles. Here, it can be seen that this transition period depends on the demand settings in each simulation. This suggests that a single entry supply function $P_s(n)$ cannot capture the variety of these transition periods. Whereas it might constitute a reliable approximation for cases 1 and 3, we clearly see that the network loadings of the simulations in case 2 are not well described by the shape of $P_s(n)$. Thus, a recalibration of the entry supply function is obviously required for this case if we want the MFD simulation to reproduce this transition period accurately. However, given the fact that in all cases the shape we choose always leads to a reliable steady state near $(n_{c2}, P_c)$, this shape can be considered acceptable if we want to preserve the generality of its application and not go into too much detail for the transient period. This is the modeling choice we adopt for the rest of this study.

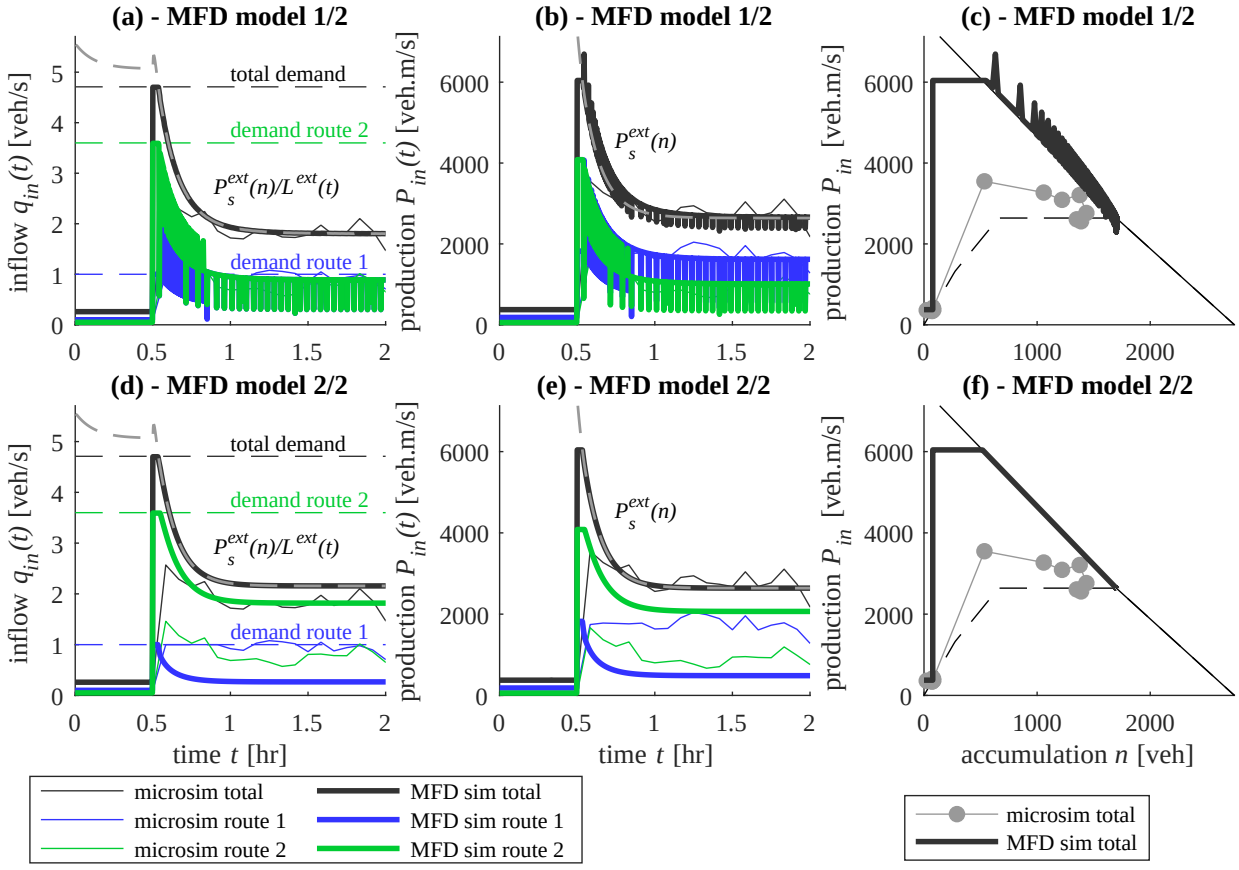



FIG. 5.5 – *Comparison between microsimulation and two MFD models for a network loading case. (a) Evolution of inflows and (b) entering productions vs time and (c) total entering production vs total accumulation for model 1/2. (d) Evolution of inflows and (e) entering productions vs time and (f) total entering production vs total accumulation for model 2/2*

## 5.4.2 Comparisons between the entry merging schemes

We now compare the two merging schemes presented in section 5.2.2 for a given network loading scenario from the entry case 2. We investigate this case first as it gives the most obvious differences between the merging schemes. At equilibrium, we know thanks to our

previous investigations that the total production is always around $P_c$. This will allow properly comparing the share of inflow or entering production between routes 1 and 2 in steady state.

Figures 5.5(a), (b) and (c) show the evolution of inflows, entering productions vs time, and entering productions vs total accumulation, respectively, in both the microsimulation and the MFD model 1/2. Figures 5.5(d), (e) and (f) show the same results but for model 2/2. In this scenario of a sudden demand loading after 30 min, each inflow equals its corresponding demand, so that the first layer of our merge algorithm (applying the border limitations) is not involved here. But it is clear that both the reservoir inflow supply $P_s^{\text{ext}}(n)/L^{\text{ext}}$ for model 1/2 and production supply $P_s^{\text{ext}}(n)$ for model 2/2 limit the demand. In Figures 5.5(c) and (f), clear discrepancies can be seen between the microsimulation and the MFD model during the transient period of the network loading. Note that these discrepancies are also due to the 5 min aggregation period in the microsimulation. However, they only account for around 15 min of simulation (from $t$ = 30 to 45 min), as shown in e.g. Figures 5.5(d) or (e). As the steady state of the total inflow $q_{\text{in}}(t)$ and entering production $P_{\text{in}}(t) = \sum_{i=1}^{4} L_i q_{\text{in},i}(t)$ are well reproduced by both MFD models, we conclude that this entry supply function is sufficient for our needs. We will show with more complex test cases in section 5.5 that the inflows from the microsimulation are well estimated and that the discrepancies observed during the short transient phase of network loadings are negligible.

We observe numerical flow oscillations due to the demand pro-rata merge in model 1/2, see Figures 5.5(a) and (b). This is due to the background traffic of route 3 which also undergoes a restricted inflow, because of its small merge coefficient (proportional to demand). Thus, queuing vehicles are stored at entry, suddenly creating a maximum demand for route 3 which therefore has a higher merge coefficient. The result is that a higher flow portion is temporally allocated to route 3 to empty its small queue, and then to a smaller merge coefficient. This process is thus periodically repeated, entailing the oscillations. This is a well-known problem in merging models, with respect to the invariance principle (basically, here the model is not invariant as it oscillates when demand reaches capacity). It is the consequence of the demand pro-rata rule. This numerical phenomenon is mainly due to the point queue model used to account for the storage of vehicles at the entry of this single reservoir model, but the oscillations would likely disappear in a multi-reservoir context when vehicles are stored in another reservoir.

However, despite this numerical issue, the demand pro-rata merge in model 1/2 is found to better reproduce the inflow or entering production share observed in microsimulation, in comparison to the endogenous merge in model 2/2. This is quite obvious in steady state, where the microsimulation results shows $q^*_{\text{in},1} \approx q^*_{\text{in},2} \approx 1$ veh/s while model 2/2 predicts $q^*_{\text{in},1} \approx 0.1$ veh/s and $q^*_{\text{in},2} \approx 1.9$ veh/s, see Figures 5.5(a) and (d). For model 2/2, the equilibrium inflow share is explained during the transient phase of the network loading. In this test, the maximum demand of route 2 is equal to 3.6 veh/s, which results in a higher increase of accumulation $n_2(t)$ compared to $n_1(t)$. Thus, in this model since the endogenous coefficient assigned to route 2 is $n_2(t)/(n_1(t) + n_2(t))$, the greater the accumulation on this route, the more the flow can enter as this ratio becomes higher. This explains the significant difference between $q_{\text{in},2}(t)$ and $q_{\text{in},1}(t)$ after $t$ = 30 min. Then, the system stabilizes to this share ratio because the same ratio of accumulations is implied in the outflow calculations, which naturally equilibrate with the inflows. On the other hand, for model 1/2, the same trend is observed at the beginning of the loading just after $t$ = 30 min, because the high demand

on route 2 entails a higher allocation ratio for this route. However, as both routes become rapidly limited at entry, a queue forms for both of them and generates a high demand. Both routes are subject to the same demand, because queuing vehicles want to enter as soon as possible regardless of their origin. By default, their maximum entrance rate is fixed to the border capacity $C^i = 3.6$ veh/s as long as the queue for route $i$ is not empty. With $C^1 = C^2$ the two merging coefficients are both equal to 0.5, which explains the identical steady inflows for both routes.

Both models 1/2 and 2/2 are also compared against the microsimulation results for all the simulations of the three network loading cases. Figures 5.6(a), (b) and (c) present the steady state inflow for routes 1 and 2 obtained with the microsimulation and MFD models 1/2 and 2/2, for entry cases 1, 2 and 3, respectively. In each simulation, the steady state inflow is calculated as the mean inflow evolution from 1 hr 30 min to 3 hr. In each plot, one point corresponds to one simulation (one value of steady state inflow per route). As illustrated in Figure 5.6(b), the demand pro-rata merge in model 1/2 clearly better reproduces the inflow share observed in the microsimulation, in comparison to the endogenous merge in model 2/2. This corroborates our first conclusion from Figure 5.5. Model 1/2 also provides a better estimation of the steady state inflows in entry case 1, although a bias can be noticed in the inflow in route 1 between the microsimulation outputs and both MFD models, see Figure 5.6(a). In entry case 3, the difference between both MFD models is less obvious; nevertheless, model 1/2 appears more accurate than model 2/2 for predicting the outputs of the microsimulation, notably for the inflow in route 2, see Figure 5.6(c). In conclusion to this section, despite the bias that may appear in the calibration of $P_s(n)$, the demand pro-rata merging scheme detailed in section 5.2.2 is shown to efficiently reproduce the inflow share observed in the microsimulation in a variety of situations.

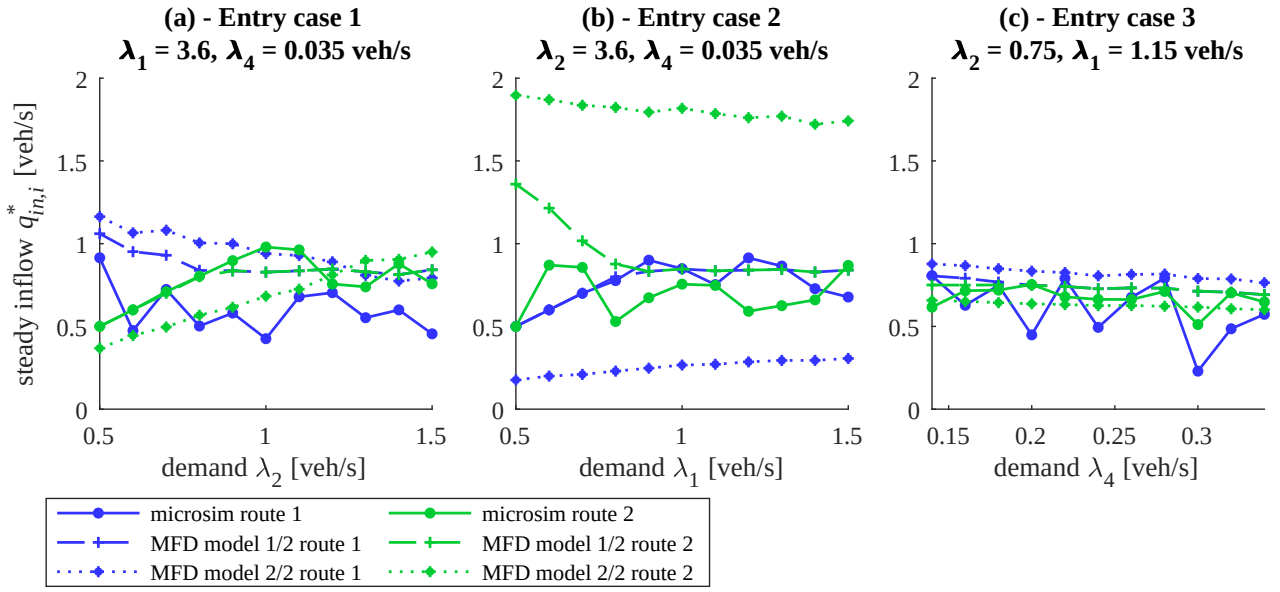



FIG. 5.6 – *Comparison of steady state inflow per route between microsimulation and MFD models 1/2 and 2/2 for different network loading cases. (a) Results of entry case 1 for 11 simulations with increasing demand $\lambda_2$, (b) entry case 2 for 11 simulations with increasing demand $\lambda_1$ and (c) entry case 3 for 11 simulations with increasing demand $\lambda_4$*

## 5.5 Comparisons between the exit diverging schemes with congestion onset-offset scenarios

In this section, we present the comparisons between the microsimulation and MFD models 1/1 and 1/2 to investigate the effect of both exit diverge models: the decreasing outflow demand in oversaturation with the independent treatment of partial outflows (in model 1/1), and the maximum outflow demand in oversaturation with the interdependent treatment of partial outflows (in model 1/2), see section 5.2.3. In both MFD models, we use the demand pro-rata merge at entry, which was found to be the best option in the previous section. The demand scenarios for routes 1 and 2 consist of a 30 min warm-up period (around 0.1 veh/s) followed by a high demand surge of around 1 veh/s per origin O1 and/or O2 (equally distributed among entry links). Congestion is created inside the network by limiting the potential outflow from exit links in D1 and/or D2 below the corresponding origin demand. This supply limitation at the exits is then released at $t_1$ = 6 hr 40 min. Finally, the high demand suddenly falls to its initial level after $t_2$ = 10 hr to observe the full recovery of the network. Three test cases are investigated:

- Exit case 1: homogeneous outflow limitation is applied on the exits of D1 and D2
- Exit case 2: homogeneous outflow limitation is applied on the exits of D1 only
- Exit case 3: homogeneous outflow limitation is applied on the exits of D2 only

Their demand and supply scenarios are presented in Figures 5.7(a), (b) and (c), respectively. Note that these simulation settings are not intended to correspond to any real situations, the high demands and their long durations are designed to stress the network and to observe clear congestion wave propagation until the next steady state. In the microsimulation, accumulation, inflow and outflow are calculated for every 5 min aggregation period. After running the MFD simulations, we filter the inflow oscillations generated by the pro-rata merge coefficients. This helps when presenting the results because only the mean value are of importance for our analysis.

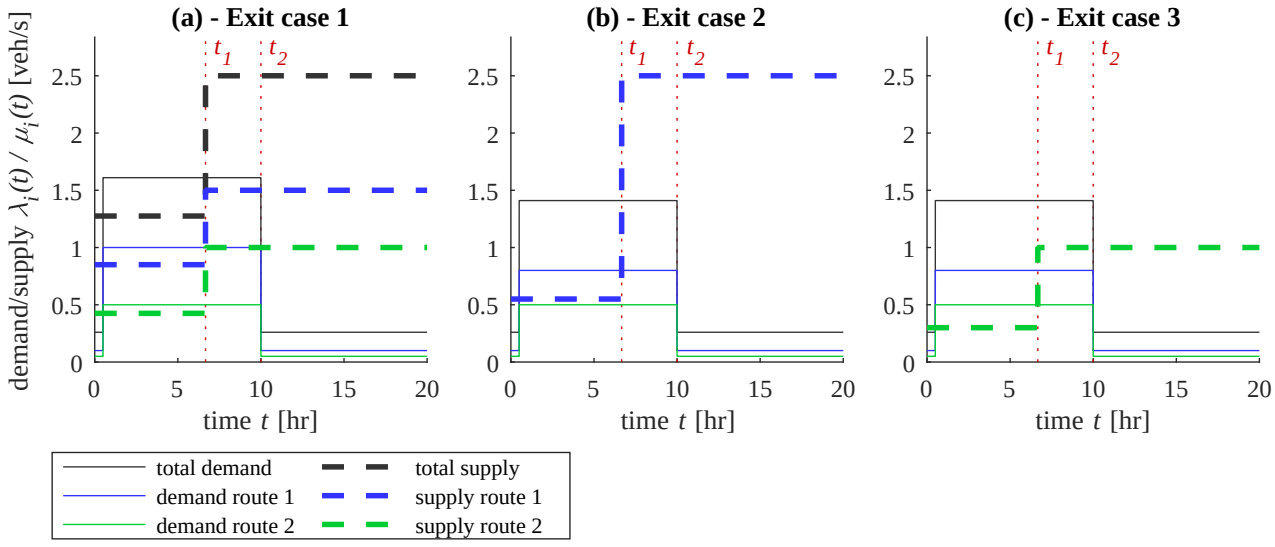



FIG. 5.7 – *(a) demand and supply scenario of exit case 1, (b) exit case 2 and (c) exit case 3*

### 5.5.1 Exit case 1: outflow limitation on both routes

The results for the first case are given in Figure 5.8. The corresponding demand and supply settings are presented in Figure 5.7(a). The evolutions of accumulations, inflows and outflows in microsimulation are compared against the predictions of each MFD model: in Figures 5.8(a), (b) and (c) for model 1/1, and Figures 5.8(d), (e) and (f) for model 1/2. Several interesting observations can be made, as detailed below.

We first focus on the period $[0, t_1]$ which corresponds to the onset of congestion due to the high increase in demand above the outflow limitations for both routes. At the network exit, all the results show that the outflow of route 2 remains equal to its limitation while the outflow of route 1 decreases after $t = 2$ hr. This is due to the interaction between the two routes in the middle of the network: the vehicles traveling longer trips on route 1 are blocked by congestion on route 2, which spills back faster because of shorter distances. This is quite well reproduced in both MFD models.

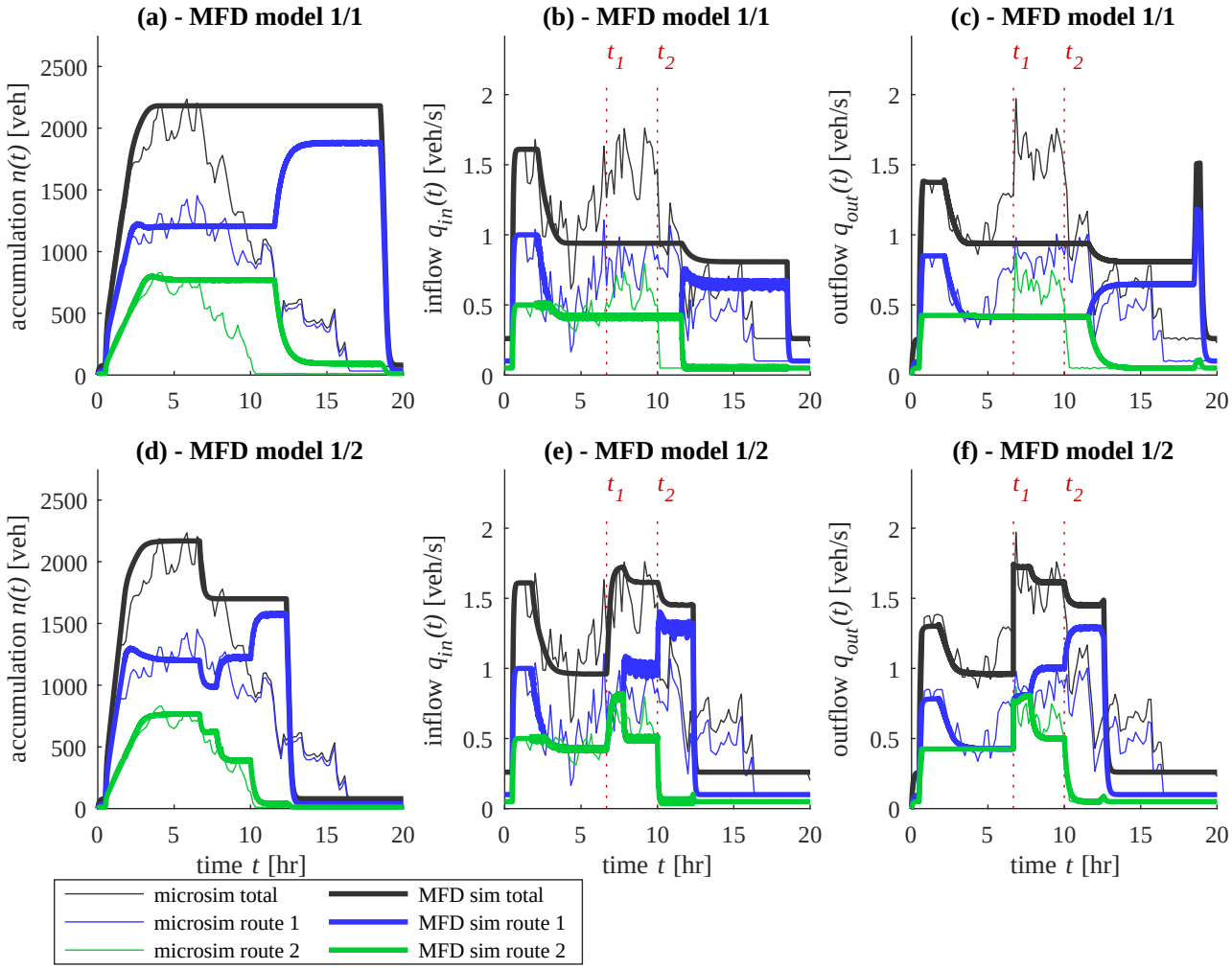



FIG. 5.8 – *Comparison between microsimulation and different MFD models for case 1: exit limitation on both routes. Times $t_1$ and $t_2$ corresponding to supply release and demand decrease, respectively, are indicated. (a) Evolution of accumulation, (b) inflow and (c) outflow with model 1/1, (d) accumulation, (e) inflow and (f) outflow with model 1/2*

We then focus on the second period, from time $t_2$ to the end of the simulation, when all outflow limitations are released. As hypothesized and highlighted in MFD simulation in chapters 2 and 4, it clearly appears that the decreasing exit demand function in model 1/1 prevents the reservoir from recovering after the congestion period. In Figure 5.8(c), it can be seen that the outflows do not react to the release of exit limitations after $t_1$, whereas the microsimulation exhibits a sudden increase of outflow at this time. On the contrary, model 1/2 using the maximum exit demand function adequately reproduces the queue discharge at the network exit after $t_1$. Then, the network can finally recover after $t_2$ once the demands fall to their initial levels. Model 1/1 is still able to recover because of the fall in demand at entry after $t_2$, but the assumption of a low outflow demand results in a much longer congestion period. Congestion vanishes just before $t$ = 20 hr, which is nearly twice as long as in model 1/2. However, this model overestimates the outflow of route 1 after $t_2$, as it predicts a high and fast discharge of the remaining vehicles in the reservoir. This is not the case in the microsimulation, where residual congestion is observed in route 1 until $t$ = 17 hr. The outflow of route 1 is lower and thus the time needed to empty the queues on this route is longer in comparison to the MFD simulation. This phenomenon is the result of internal congestion that can be captured by link-level simulation, but this is hardly reproducible in MFD models. One possible solution would be to adjust the exit demand function $P_d(n)$, for which two opposite approaches have been presented in equations 5.12 and 5.13. Based on this comparative analysis, it appears that the decreasing exit demand model is too pessimistic to reproduce internal congestion. On the other hand, the maximum exit demand model works nicely for short routes like route 2 in this case, but it is too optimistic for longer routes where residual queues need more time to empty. The differences between both routes exhibited in microsimulation are the results of the heterogeneity of traffic states inside the network. Thus, an exit demand function that would work in any situation is impossible to design if such a function is based on a unique MFD $P(n)$, because of the mean speed assumption shared by all vehicles.

The conclusion of this test case comparison is that the best match with microsimulation outputs is obtained with model 1/2, with the maximum exit demand function.

### 5.5.2 Exit case 2: outflow limitation on route 1 only

The results for the second case are given in Figure 5.9. The corresponding demand and supply settings are presented in Figure 5.7(b). The evolutions of accumulations, inflows and outflows in microsimulation are compared against the predictions of each MFD model: in Figures 5.9(a), (b) and (c) for model 1/1, and Figures 5.9(d), (e) and (f) for model 1/2.

In this scenario, similar observations can be made as compared to exit case 1. From $t$ = 0 to $t_1$, both models 1/1 and 1/2 provide on average a reliable estimation of the inflows and outflows of the microsimulation. We note however considerable scatter in these flow values given by the microsimulation. This scatter results in significant variations observed in the evolution of accumulations, in particular for route 1, as shown in Figures 5.9(a) and (d). As in exit case 1 after $t_1$, model 1/1 fails to reproduce the queue release as the outflow of route 1 remains low and equal to its previous exogenous limitation of $\mu_1$ = 0.5 veh/s, applied during [0, $t_1$]. On the other hand, the queue discharge observed in the microsimulation is quite well predicted by model 1/2 after $t_1$. Because of the scatter in the microsimulation outputs, in this case the increase of outflow of both routes after $t_1$ is less obvious than in exit

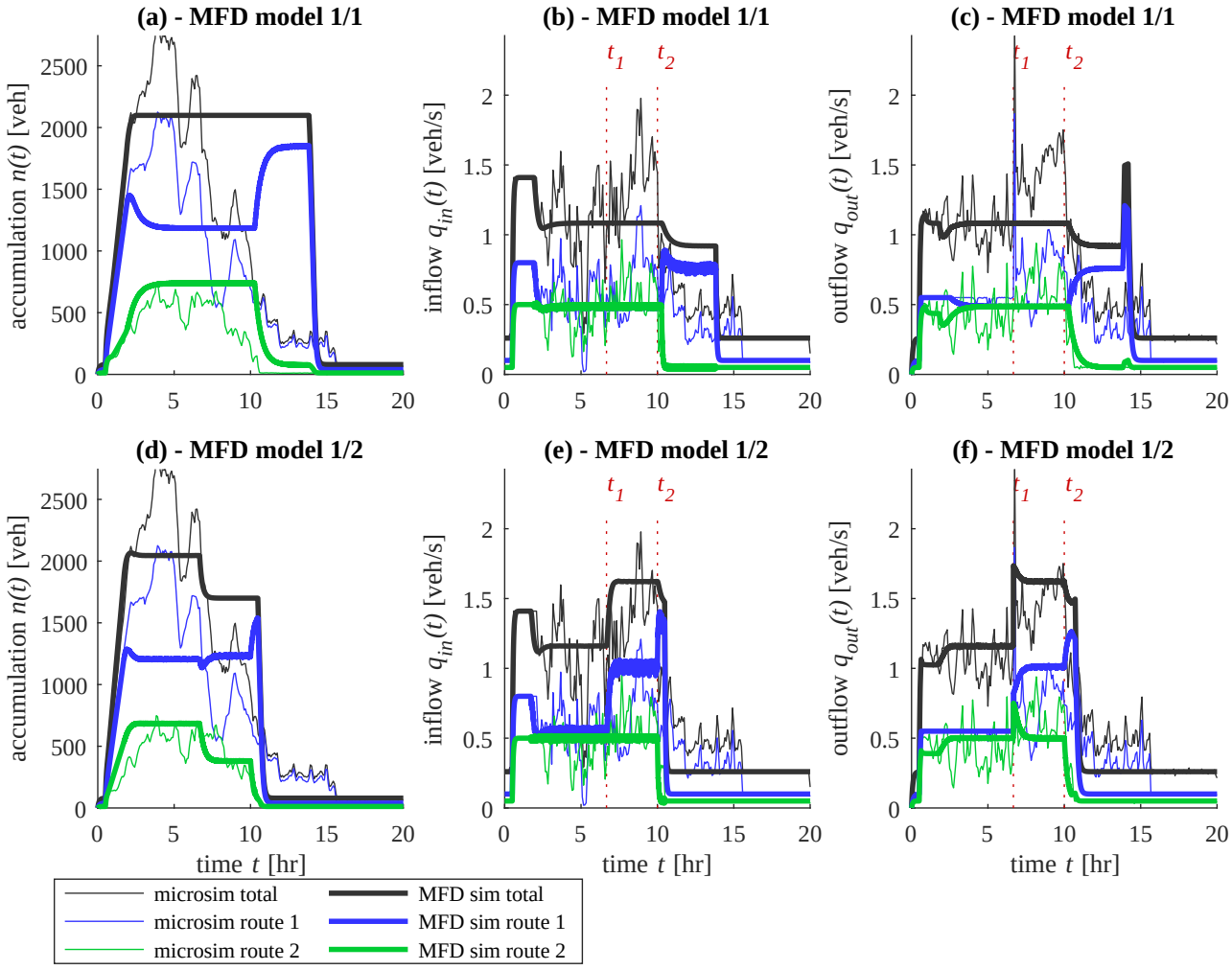



FIG. 5.9 – *Comparison between microsimulation and different MFD models for case 2: exit limitation on route 1. Times $t_1$ and $t_2$ corresponding to supply release and demand decrease, respectively, are indicated. (a) Evolution of accumulation, (b) inflow and (c) outflow with model 1/1, (d) accumulation, (e) inflow and (f) outflow with model 1/2*

case 1. As with the latter case, a residual congestion can be seen in Figures 5.9(c) and (f) in the microsimulation results. Again, none of the MFD models is able to describe it; however, model 1/2 is still the most accurate regarding the whole simulation period. The evolution of traffic states on route 2 is particularly well reproduced with model 1/2, as presented in Figures 5.9(d), (e) and (f).

The conclusions in this test case are hence the same as in exit case 1.

### 5.5.3 Exit case 3: outflow limitation on route 2 only

The results for the third case are given in Figure 5.10. The corresponding demand and supply settings are presented in Figure 5.7(c). The evolutions of accumulations, inflows and outflows in microsimulation are compared against the predictions of each MFD model: in Figures 5.10(a), (b) and (c) for model 1/1, and Figures 5.10(d), (e) and (f) for model 1/2.

First, during the onset of congestion before $t_1$, we can clearly see in Figures 5.10(c) and

(f) the outflow limitation of route 2 and its impact on the outflow of route 1. The latter is significantly reduced, although no exogenous limitation is applied on this route in this test case. Likewise with exit case 1, it is interesting to note that during the first two hours of the simulation, model 1/1 adequately reproduces the high outflow of route 1, while model 1/2 clearly misses this outflow peak. This underestimation is even more significant in this case than in case 1, due to the diverge outflow model. When using the maximum exit demand assumption, in chapter 4 we showed that explicit interdependency relationships must be employed, as presented in equation 5.18. While ensuring consistent steady states after congestion onset, as we can see after $t$ = 2 hr, the drawback is that these relationships are applied instantaneously to route outflows. Consequently, once the outflow of route 2 is limited at the very beginning of the simulation, that of route 1 is instantaneously reduced to accommodate this limitation through the formula in equation 5.18. In this case, the exit of route 2 is that which is most constrained, as described in equation 5.17. The situation is different with the other diverge outflow model in model 1/1. Because each exit limitation is applied independently, the outflow of route 1 can reach the flow level sent by the entry earlier, as can be seen in Figure 5.10(c). But after $t$ = 2 hr, the increase of accumulation in the reservoir due to the onset of congestion causes the mean speed to decrease, and thus reduces the potential outflow of route 1 to finally reach almost the same steady state as in model 1/2. In this approach, the interaction between the route outflows is implicitly described through the speed or production-MFD. Regarding the microsimulation results, it appears that this approach is therefore better for transient evolution during congestion onset.

However, during the offset of congestion after $t_1$, the diverge outflow of model 1/1 is unable to predict the queue discharge on route 2 once its corresponding limitation is released at exit, see Figure 5.10(c). This is similar to what we observed in exit cases 1 and 2. Here also, model 1/2 provides better results for congestion offset, although the total outflow is noticeably overestimated after $t_2$. This model anticipates the unloading of route 1 at the same time as for route 2, due to the interdependency relationships mentioned above. But this does not occur in the microsimulation as this outflow remains low. The simulation duration is not long enough to observe full network recovery, and residual congestion on route 1 is still observed at $t$ = 20 hr. Hence, the exit demand of model 1/2 is again too optimistic for the congestion offset, while the exit demand of model 1/1 is clearly too pessimistic. Though not presented here, model 1/1 is still able to recover due to the decrease in demand at entry after $t_2$. But the assumption of a low outflow demand results in a much longer congestion period. If we run the MFD simulation for a longer time frame, the offset of congestion appears at $t$ = 24 hr, nearly twice as long as in model 1/2.

In conclusion for this test case, although no MFD approach was found to be fully satisfactory to reproduce the microsimulation results due to local congestion (high link-level heterogeneity), the maximum exit demand function of model 1/2 can at least match the evolution of traffic states in route 2 quite well.

## 5.6 Conclusion

In this chapter, we compared different MFD models against microsimulation on an artificial grid network crossed by two regional flows having different mean trip distances. The MFD frameworks we tested include two inflow merging models (demand pro-rata merge with

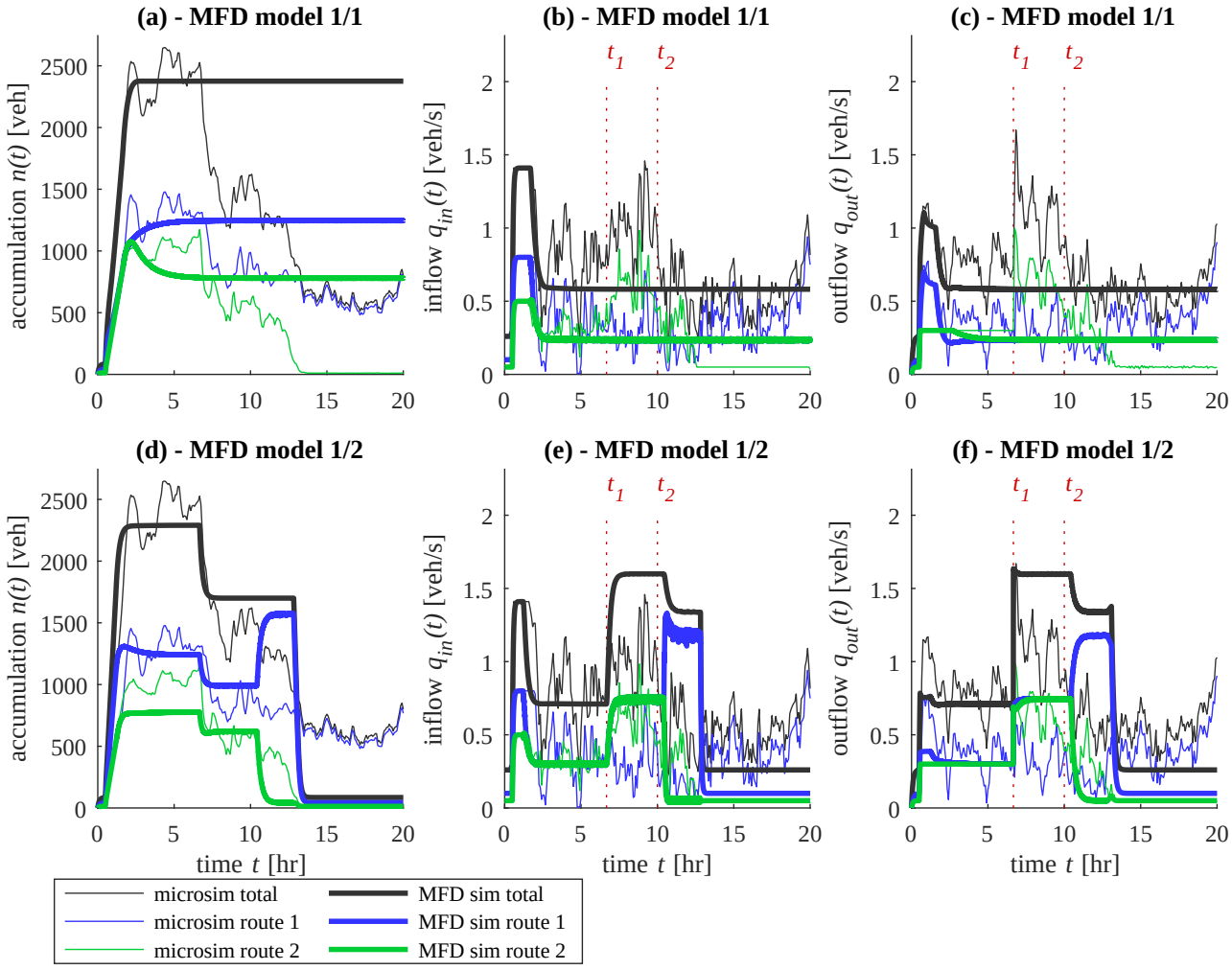



FIG. 5.10 – *Comparison between microsimulation and different MFD models for case 3: exit limitation on route 2. Times $t_1$ and $t_2$ corresponding to supply release and demand decrease, respectively, are indicated. (a) Evolution of accumulation, (b) inflow and (c) outflow with model 1/1, (d) accumulation, (e) inflow and (f) outflow with model 1/2*

flow merge from the literature, and our endogenous merge with production merge), and two diverging outflow models (decreasing outflow demand in oversaturation with independent treatment of partial outflows from the literature, and our maximum outflow demand in oversaturation with interdependent treatment of partial outflows). The first contribution of this chapter was to propose a unified two-layer merging scheme to handle both inflow merging models. Thanks to microsimulation, the reservoir entry supply function was calibrated with several network loadings. A thorough analysis of these loadings with different demand patterns showed that the shape of this function may not be unique, as it depends on the demand configuration for its undersaturated branch. However, this non-uniqueness would be critical only during the transient period of the network loading, because our results also showed that an identical equilibrium point is almost always reached in the production-MFD. As the transition period was found to be quite short in our loading simulation tests, the calibration issue of the entry supply function in undersaturation would not be a major source of error in a general case.

The second contribution was the comparison between the merging and diverging models. At the network entry, three sets of 11 simulations were used to study various situations of network loadings. In all these situations, we found that the demand pro-rata merging model provided a better estimation of flow equilibrium after the congestion onset, in comparison to our endogenous merging model. At the network exit, three test cases of a long scenario with congestion onset and offset were thoroughly investigated. The following observations were made. While adequately reproducing transient states during congestion onset, the approach of a decreasing outflow demand was found to be too pessimistic to predict the unloading of queues once the exit limitations were released. On the other hand, our approach assuming a maximum outflow demand led to better results. Nevertheless, the drawback of the latter framework was that interdependency between all the partial outflows had to be explicitly formulated in the diverging outflow scheme to comply with the reservoir mean speed hypothesis (homogeneity in traffic states). This resulted in faster network recovery in comparison to the microsimulation outputs, where local link-level interactions generated more internal congestion than predicted by the MFD simulation. Interestingly, more internal congestion was found in the case when spillbacks on shorter routes interacted with undersaturated traffic flows on longer routes. Such a scenario is likely to generate more heterogeneous traffic states, and be thus more challenging for MFD-based simulation. Nevertheless, while being rather optimistic during congestion offset, on the whole our approach of the maximum outflow demand proved capable of overcoming the issues encountered with the approach of decreasing outflow demand generally adopted in the literature.

The latter results still need to be confirmed with other network and demand configurations. Our simulation case was designed to create strong internal correlations between the two major flows, which is certainly a limit case compared to reality. However, it provided a better overview of the aggregation of microscopic dynamics and the way MFD models reproduce it. Thus, further work on more complex simulation cases and real field studies is essential to corroborate these results, especially with more OD options on bigger networks.

# Conclusion of part II

In this second part, we investigated flow exchanges in a single reservoir where the general context of multi-reservoir interactions (sending or receiving flows) were represented by exogenous boundary conditions. The extension of the single reservoir model with a unique trip length to multiple trip lengths was also thoroughly analyzed. While the approaches proposed in the literature mainly apply MFD-based simulations in undersaturated situations, we particularly focused on the modeling of congestion propagation within a reservoir, and its effects on inflow restrictions and potential outflows. To this end, we proposed flow and production constraints at the reservoir entry and exit to design new inflow merging and outflow diverging schemes, and ensure their consistency with the reservoir dynamics and homogeneity assumption (i.e. at equilibrium, the accumulation and flow of each trip category should comply with Little's queuing formula, and all trip categories should experience the same speed given by the speed-MFD). In comparison with the existing works in the literature, our inflow restriction formulation consists of a two-layer fair merging scheme, where the first layer accounts for the reservoir physical limitations (link capacities at entry), and the second one reproduces congestion spilling back to the reservoir entry with an entry supply function. On the other hand, our new formulation of outflow in oversaturation consists of a maximum outflow demand function combined with a new outflow diverging scheme. This new formulation is intended to reproduce queue discharge after a congested period due to a temporary flow restrictions at exit (that may represent congestion spilling back from a downstream reservoir). Both entry and exit flow models were implemented in the accumulation-based and trip-based models.

Our new modeling approaches and the ones from the literature were then compared against microscopic simulation outputs, for a grid network crossed by two main sets of trips. At entry, a demand pro-rata merging rule embedded in our two-layer scheme was found satisfactory to reproduce the steady state inflow share observed in the microscopic simulation. Moreover, the study of entering flows allowed us to calibrate a general shape of entry supply function. The latter proved to be accurate enough for steady state estimation. At exit, our new formulation of outflow to describe queue discharge was confirmed by the results from the microsimulation. Besides, the widely used outflow model from the literature (decreasing with accumulation to mimic internal congestion in oversaturation) was found too pessimistic, as it clearly overestimates the congestion period duration in the test cases we studied.

The contributions of this part can be summarized as follows:

- ✓ Proposition of a new framework to handle spillbacks in multi-reservoir systems with multiple trip categories

- ✓ Extension of the event-based resolution method for the trip-based model in oversaturated situations. At the reservoir entry, this consists of the transformation of the accumulation-based merging scheme.

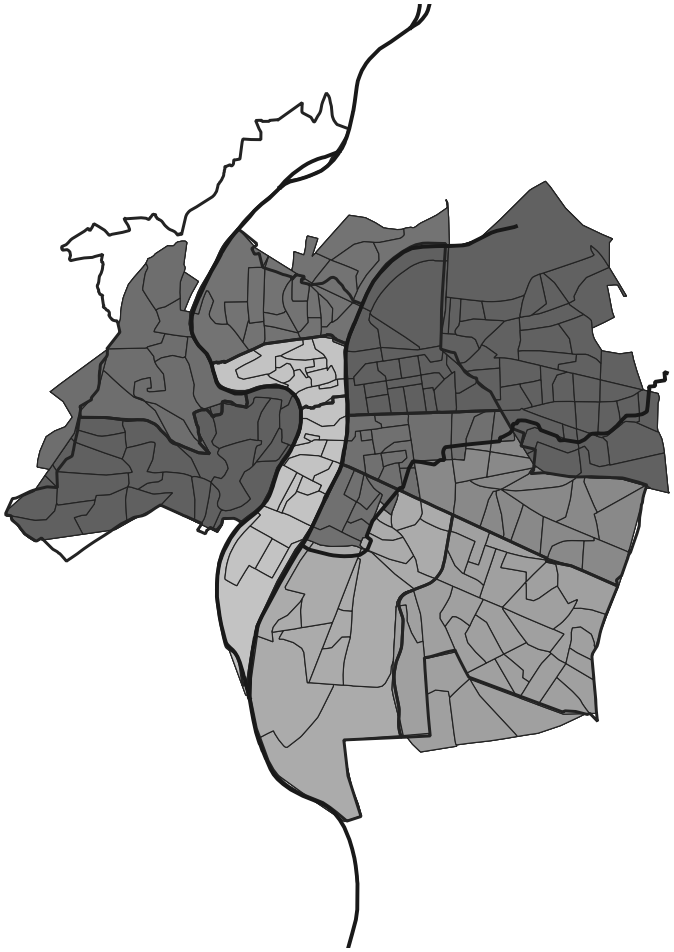
- ✓ Validation of the demand pro-rata merging rule embedded in our two-layer scheme at reservoir entry, and of our new outflow formulation combined with a new outflow diverging scheme against microsimulation results.

The studies conducted in this part were still focused on the traffic dynamics in a single reservoir. Now that we have developed the required component to connect several reservoir together (inflow and outflow models), further applications with multi-reservoir systems should be investigated on larger networks. Moreover, to pursue the effort of our modeling validation, a confrontation between MFD simulation and real data should be envisioned. All these further works are then detailed below:

- Provide an overview of a multi-reservoir simulation for a large-scale network
- Present the implementation of the accumulation-based and trip-based models for multi-reservoir systems
- Calibrate a multi-reservoir simulation and confront its outputs with real data

# Part III

# Developments and applications



# Introduction

The purpose of this part is to introduce the different steps and components required to run a multi-trip and multi-reservoir simulation. It describes the MFD-based simulation platform that has been developed during this PhD, and which constitutes the major practical achievement of the research work presented in the two previous parts. This part includes the presentation of the context and assumptions for the design of a multi-reservoir simulation, the detailed description of the variables and structures used, and the analysis of the pre-processing phase with a discussion about the required input data. A DTA module at the reservoir level, which is the result of a collaboration with another PhD student, is also presented to distribute the flows between the reservoirs. Regarding the modeling of traffic flow, the practical implementation of both the accumulation-based and trip-based frameworks is detailed in the platform solver. This implementation is the result of our previous investigations about flow exchanges for a single reservoir, as studied in the last part.

Finally, a first application to a real city network is carried on. The whole calibration process is presented and discussed, and MFD simulation outputs are compared with real data. Despite the many issues encountered and the relative limitation of the study, this first real test case shows promising results for the multi-reservoir simulation.

# Contents

# 6. Model implementation and first application to a real network

## 6.1 Presentation of the multi-reservoir MFD-based simulation platform

### 6.1.1 Modeling purposes and assumptions

In this section, we introduce the multi-reservoir MFD-based simulation platform that was developed during this PhD thesis. The simulation core of this platform is based on our investigations about flow exchanges in the single reservoir model, as detailed in the previous chapters of the manuscript. This platform is designed to be practice-ready in order to help traffic managers to analyze traffic dynamics at a large scale. It could also help designing or testing new control policies, such as perimeter control or route guidance, to regulate large metropolitan areas. The simulator is a research tool including useful pre-processing and post-processing functions. The platform consists of a standalone source code written in MATLAB™. It will be shared on GitHub in open source when this PhD thesis will be published.

#### General framework and options

The purpose of this MFD-based simulator is to reproduce macroscopic traffic flows in a large urban area (a city for instance) under the multi-reservoir and multi-trip framework. Hence, the studied urban region is divided into $N_r$ reservoirs in which a well-defined MFD is assumed. A set of $N_p$ possible macroscopic routes (sequences of reservoirs) with specific lengths into each reservoir of their respective sequence is also supposed known. Then, considering the macroscopic OD matrix as input (demand profile from each origin reservoir to each destination reservoir), the simulation outputs provide the evolution of aggregated traffic states, such as accumulation, inflow, outflow and travel time in each reservoir per route. The simulator can thus capture undersaturated conditions as well as propagation of congestion on specific routes. A DTA convergence loop to approximate network equilibrium is also embedded in the platform to study user routing.

Several modes and options are available to the modeler for a specific simulation study:

**Utilization mode** The simulator can be used in two ways. The first one is when the simulation is based on a real network. In this case, the network configuration, trip lengths,

MFD, traffic and demand data are supposedly known at the beginning of the simulation. These input data are the result of preliminary studies on the microscopic (link-level) network. The second one is when the simulation is only based on the design of artificial reservoirs encompassing all required information (MFD, trip lengths, demand, etc). This simulation mode is useful to develop benchmark scenarios to test and compare new theoretical modeling approaches.

**Path flow distribution mode** The demand can be defined at two different levels. The first one is the definition for reservoir OD pairs: a demand profile per destination reservoir is set in each origin reservoir. Then a path flow distribution module embedded in the simulator is in charge of distributing these demand profiles among the possible routes. The initial set of routes can be determined in different ways, depending on the initial knowledge of the network one may have. This will be detailed further. The second level is the direct definition of demand profiles on a selection of routes. For this mode, no assignment is proceeded.

**Assignment model** In case of a demand definition per OD, a Dynamic Traffic Assignment (DTA) loop can dynamically calculate the path flow distributions. A fruitful collaboration with PhD student Sergio Batista has allowed to implement different convergence mode: User Equilibrium (UE), Bounded Rational User Equilibrium (BRUE) and Regret Theory User Equilibrium (RTUE).

**Solver model** Both accumulation-based or trip-based traffic modeling approaches can be run. Other specific options about inflow merging and outflow diverging schemes are also available, as investigated in chapter 5.

### Trips starting and ending outside the studied area

The demand flow from an origin reservoir is always supposed to be created inside this reservoir (internal demand creation), and then eventually propagates to the next reservoirs of the corresponding route. The flow reaching its destination reservoir is also always supposed to end in this reservoir. Hence, the modeling of transfer flows outside the studied network perimeter must be handled by adding "ghost" reservoirs to the simulation. These reservoirs represent external sources and destinations. In case of a congestion wave reaching the perimeter, these ghost reservoirs are assumed to store the waiting vehicles, but not to be oversaturated. Therefore, their parameters must be set so that their own dynamics have a negligible influence on the simulation results (very high jam accumulation, critical accumulation and production, high free-flow speed, short trip lengths, compared to the other reservoirs).

### Initial state

By default, the initial simulation state is always an empty network. Although any initial state (initial accumulations) can be set for the accumulation-based model due to its memoryless property, it is difficult to begin a trip-based simulation with traveling vehicles in the network. The trip-based model does not accommodate an initial condition which is not at equilibrium (steady traffic states complying with Little's formula), and even with an initial

steady state condition, it is tricky to assign a remaining travel distance to each vehicle complying with this state. That is why an empty initial state is set by default. Running a short warm-up simulation period then allows to generate any network loading that one might require.

### Variable definitions and notations

The platform conception is based on an object-oriented design. We present here the main structures that characterize a simulation:

- The **Reservoir** structure encompasses all information related to each reservoir ($N_r$ reservoirs, denoted by index $r$):
  - production-MFD $P^r(n^r)$ and speed-MFD $V^r(n^r) = P^r(n^r)/n^r$, $n^r$ being the total accumulation
  - MFD parameters: jam accumulation $n_j^r$, critical accumulation $n_c^r$, maximum production $P_c^r$, free-flow speed $u^r$, and more depending on the MFD shape function
  - entry production supply function $P_s^r(n^r)$ and exit production demand function $P_d^r(n^r)$
  - set of adjacent reservoirs (that share a common border, i.e. directly reachable) $\mathcal{R}^r$
  - maximum boundary capacity per adjacent reservoir (aggregation of physical border limitation due to the real network link connections) $C_{r'}^r(t) = \sum_{l \in \text{Border}(r,r')} q_l$, $r' \in \mathcal{R}^r$, where $q_l$ is the flow capacity of link $l$ (including immediate traffic signal settings)
  - set of routes crossing the reservoir $\mathcal{P}^r$, see below
  - trip lengths per route $L_p^r$, $p \in \mathcal{P}^r$
  - simulation variables: reservoir mean speed $v^r(t)$, accumulation $n_p^r(t)$, inflow $q_{\text{in},p}^r(t)$, outflow $q_{\text{out},p}^r(t)$ per route, cumulative count curves per route $N_{\text{in},p}^r(t)$, $N_{\text{out},p}^r(t)$
- The **Route** structure gathers all data related to each macroscopic route ($N_p$ routes, denoted by index $p$ for "path"):
  - corresponding sequence of reservoirs $R^p = [r_1, \ldots, r_n]$
  - trip lengths in the reservoirs $R^p$
  - path flow coefficient $a^p$, proportion of OD flow, from origin reservoir $r_1$ to destination reservoir $r_n$, assigned to route $p$
  - demand profile $\lambda^p(t)$
  - experienced travel time $T^p(t)$, calculated as the horizontal difference between the origin entering count curve $N_{\text{in},p}^{r_1}(t)$ and destination exiting count curve $N_{\text{out},p}^{r_n}(t)$
- The **Vehicle** structure is only used when running the trip-based model. It includes individual vehicle trajectories in the reservoir network ($N_u$ vehicles, denoted by index $u$ for "user"):
  - creation time $t_{\text{crea}}^u$
  - assigned route $p$
  - sequence of trip lengths experienced in reservoirs $R^p$ of route $p$: $L^u = [L_p^{R^p[1]}, \ldots, L_p^{R^p[n]}]$

  - sequence of entry and exit times in reservoirs $R^p$ of route $p$: $T^u_{\text{in}}$ and $T^u_{\text{out}}$

- The **Simulation** structure contains the required simulation settings and options (denoted by subscript $s$):

  - simulation duration $T_s$ and time step $\delta t_s$ (for the accumulation-based solver)
  - utilization mode (based on a real network or not)
  - path flow distribution mode (demand definition per OD with embedded assignment loop, or per routes, which does not require assignment calculation)
  - solver (accumulation-based or trip-based) and options (inflow merge, outflow diverge models)
  - trip length calculation method (aggregation by either reservoir, destination, origin and destination, or route)

- The **Assignment** structure includes all options related to traffic assignment (denoted by subscript $a$):

  - assignment period $T_a$ for dynamic re-assignment on the routes during the simulation: calculation of constant path flow distributions over $T_a$, and update of these distributions every $T_a$ ($T_a < T_s$ in general, but $T_a$ is set to $T_s$ for a unique assignment period or when no assignment loop is performed)
  - convergence criteria (gap, threshold and tolerance for the number of violations, maximum number of iterations)
  - equilibrium model (UE, BRUE, RTUE)
  - parameters of utility models (number of draws for Monte-Carlo calculations for all equilibriums, bounded rationality and regret aversion parameters for BRUE and RTUE)

The platform is then organized in several modules that load or create these structures and interact with them:

- The **SIMULSETTINGS** module defines the simulation settings and options, via the creation of the **Simulation** and **Assignment** structures.

- The **RESDEF** module loads the **Reservoir** structure with its topology information and properties. The structure is directly created in the module if the simulation is not based on a real network.

- The **DEMDEF** module loads the aggregated demand data. Depending on the assignment mode, the temporal demand profile is either set in each origin reservoir per destination reservoir, or directly in each route.

- The **ROUTECALC** module loads the route choice set and selects some of them according to the simulation options and demand definition.

- The **ASSIGNCALC** module updates the assignment coefficients of the selected routes at the beginning of each assignment period.

- The **MFDSOLVER** module solves traffic dynamics for the current assignment period. It either corresponds to the accumulation-based model, or to the trip-based model where the additional **Vehicle** structure is created.

- The **CONVERGECALC** module tests the convergence criteria at the end of each assignment period, based on the previously simulated traffic states.

More details are given further about the assignment updates, convergence tests and solver algorithms.

### 6.1.2 Simulation input data and pre-processing

Before running a simulation, several information and data must be provided to the platform. These data either come from preliminary studies on the real network at link-scale, or are given as direct inputs in case of the second utilization mode (artificial reservoir network). Figure 6.1 presents the general picture of a simulation for the first utilization mode. The simulation study is divided into four phases: (I) input data collection, (II) pre-processing, (III) simulation loop and (IV) post-processing.

The input data collection (I) involves several data sources, as presented below (see also Figure 6.1):

1. The first essential data input is the real network topology, which is quite easy to access. It consists of link and node connections and link lengths, possibly with additional information such as link speed limit, number of lanes, lane capacity, jam density, and node signal settings. If not available, orders of magnitude can be given for a lane FD characteristics (free-flow speed, capacity and jam density). Together with the link length data, this allows to define upper bounds for the reservoir free-flow speeds and jam accumulations (see also section 1.2.1). The knowledge of the lane number per link is very useful, as it allows to calculate link capacities and aggregate them into reservoir border capacities $C^r_{r'}$. Otherwise a default value may be given. Information on signal timings is useful to refine the calculation of these capacities in accounting for the proximity of signalized intersections. Moreover, it could help estimating the reservoir MFD capacity, according to the method of Laval & Castrillon (2015). The network topology is also an essential ingredient to estimate trip lengths in reservoirs when no real trip data is available (see Batista *et al.*, 2019).

2. The second essential information is the demand profiles. This data can be defined at an aggregated level (urban areas), but which should be equal to or smaller than the future reservoir aggregation level. Having some insights about regional flow distribution would be very useful for the assignment model calibration.

3. The third data source is traffic data. This one may either come from real field measurements, such as loop detector data, probe vehicle data and historical mean speed data, or from simulation studies on the real network in using microsimulation softwares. Like the demand data, traffic information can be aggregated at a similar or lower level than the reservoir level. This data is the best information source to estimate the future reservoir MFDs, and to eventually apply a clustering algorithm to define the reservoirs.

4. The fourth type of information corresponds to vehicle paths (trip data) on the real network. It consists of sequences of links, and may come from real or simulated trajectories. This information is very useful to estimate the possible route choice set and their respective trip lengths in the reservoirs. It may be a part of traffic data, if probe vehicle trajectories are accessible.

Having access to all these data is obviously the best situation, as it ensures a minimum bias when scaling up link-level information to network-level features. Unfortunately, this is not always the case, and sometimes even accessible information sources do not provide sufficient data for a given study (e.g. available traffic data but only in undersaturated conditions).

The pre-processing phase (II) includes the following steps (see also Figure 6.1):

5. The clustering consists in partitioning the network into reservoirs. In the first utilization mode, a reservoir is defined as a set of connected links. The nodes inside the reservoir only belong to this reservoir, while the nodes at its border also belong to the corresponding adjacent reservoirs. On the one hand, the network can be clustered by an algorithm in supervised learning, e.g. $k$-means (Forgy, 1965, Lloyd, 1982) or $k$-NN (Altman, 1992) algorithms, based on link traffic data. This method should ensure that the hypothesis of traffic state homogeneity is satisfied for the created reservoirs. This clustering is thus referred to as "traffic-oriented". However, this method may result in strange area shapes with a lack of connectivity between the links of each reservoir. One might thus prefer methods and algorithms that were especially developed for traffic problems (see Ji & Geroliminis, 2012, Zhou *et al.*, 2012, Saeedmanesh & Geroliminis, 2016, 2018). Insightful discussions about network clustering are also found in Lopez *et al.* (2017). On the other hand, if traffic data is not available, insufficient, or does not give compact shapes with a clustering algorithm, the reservoirs can be designed by hand based on the modeler's experience about the studied network. As the demand data is generally provided on predefined urban areas, these areas may be eventually aggregated to create suitable reservoirs. In this case, the clustering is referred to as "demand-oriented".

6. The MFD estimation is a critical step in the pre-processing phase, as the MFD governs traffic states within its reservoir during the simulation. The MFD estimation based on traffic data is discussed in section 1.2.1. Then any mathematical function form can be used to fit MFD data points. In practice, using a piecewise linear function is convenient, as it can adapt to many cases. In case of missing data, one can only get orders of magnitude of the MFD critical parameters, as detailed earlier for the network data collection.

7. The microscopic trip set (set of link paths) is essential to estimate the mean trip lengths inside the reservoirs, and eventually to get an idea of the most likely macroscopic routes (sequences of reservoirs corresponding to real mobility patterns). It may come directly from trip data if available. Otherwise, another method can be the use of shortest path calculations on the real empty network (only the distances are considered, not the possible reduction of travel time due to traffic). In this last case, a very high number of calculation has to be carried on to cover the traveling possibilities on the

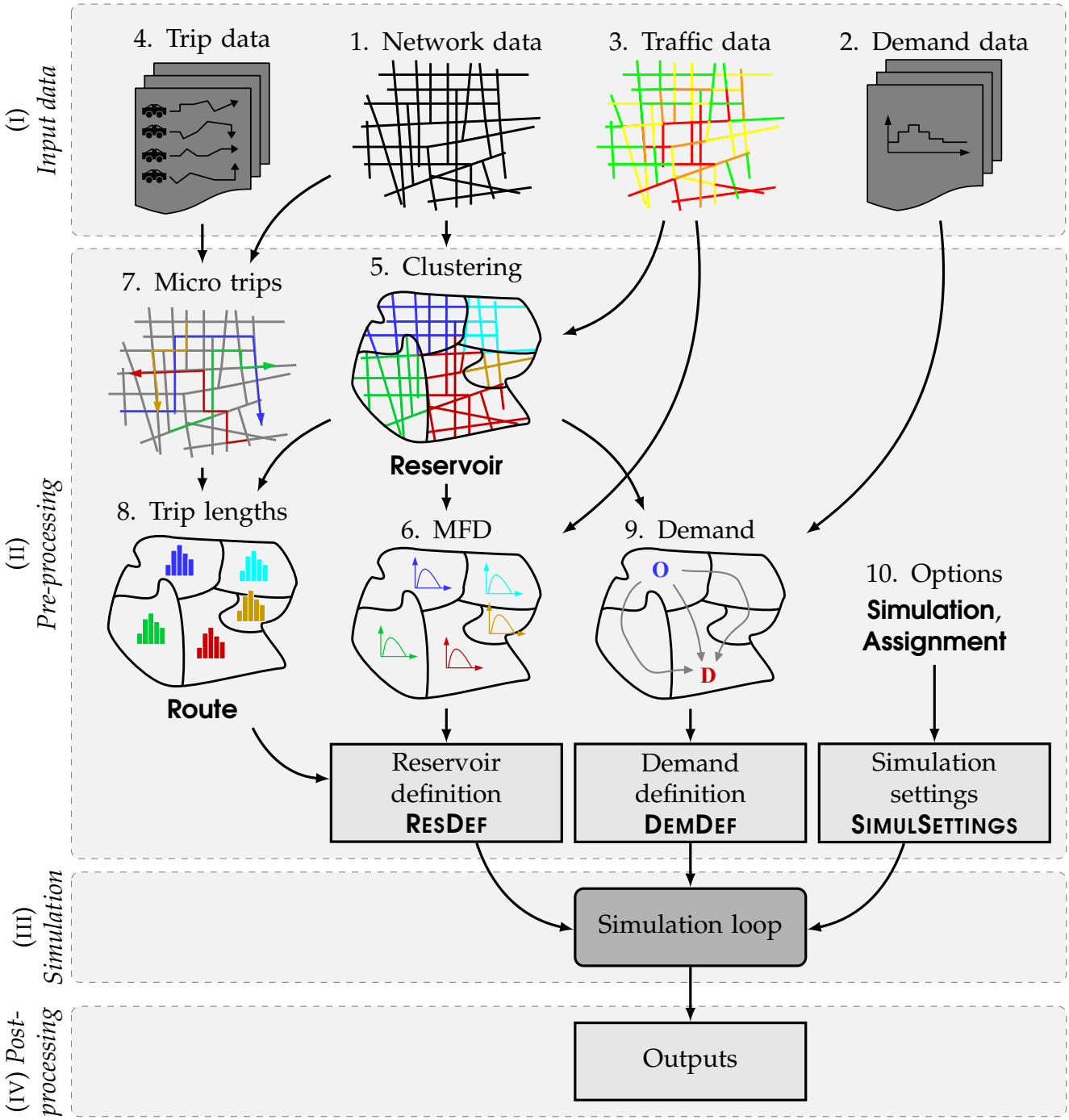



FIG. 6.1 – *General picture of a simulation study with pre-processing phase*

studied network. Each calculation consists of a random draw of an origin and a destination point on the real network, where the shortest path in distance between them is estimated by e.g. Dijkstra's algorithm (Dijkstra, 1959) or A* algorithm (Hart *et al.*, 1968) that achieves better performance on large networks. The drawback of shortest path calculations is that despite a good network coverage (every link has been visited), there is no evidence that the resulting link paths represent the users' mobility on the network studied. Besides, these paths are dependent on the algorithm used. A possi-

ble solution to this issue, proposed in Batista & Leclercq (2018b), Batista *et al.* (2019), is to perform the calculations after the reservoir definition, and force the link paths to follow specific macro routes that are supposed to represent aggregated mobility on the network.

8. The trip length estimation is another critical step in the pre-processing phase, as trip lengths are one of the components of the reservoir trip completion rates. In the first utilization mode, there are four different methods to aggregate link information from the microscopic trip set to the macroscopic route scale. These methods have been proposed and implemented by Batista *et al.* (2019). For a given reservoir, the first method, denoted M1, computes the mean of all the micro trips crossing this reservoir, which results in a unique average trip length whatever the route considered. The second one, M2, distinguishes all the micro trips crossing the reservoir by their destination reservoirs, and computes the mean of each group that share a common destination. The third one, M3, is similar to M2 except that micro trips are distinguished by their origin and destination reservoirs. Finally, the fourth method M4 classifies the micro trips crossing the reservoir by the macro route they follow. This last method is supposed to be the most accurate one, as it accounts for all the topology constraints of each macro route. One must remember that these estimated trip lengths may depend on traffic states if the micro trips are generated via shortest path calculations, as shown by Batista *et al.* (2019). In reality, travelers are indeed likely to adapt their mobility patterns to traffic conditions, a mechanism that may modify mean trip length values, but that cannot be captured by random draw of shortest paths where traffic states are ignored. This issue has been already mentioned in the previous point about the selection of microscopic trip sets. Further research is needed to clarify this concern. In the second utilization mode, the set of macro routes and their corresponding trip lengths in reservoirs are defined manually by the modeler, depending on his/her needs for the simulation studied.

9. The demand aggregation step is aimed at setting the regional OD matrix in the first assignment mode, or directly the demand per route in the second mode. This step is straightforward if the clustering is demand-oriented, as explained earlier. Otherwise if some urban units (at which the demand is defined) are split by the clustering, then the corresponding demand values may be split according to some basic weighting rules (values weighted by the corresponding total link lengths, or socio-demographic information such as number of houses, etc).

10. The simulation settings and options are specific to each study and depend on its purposes. This is detailed in the **Simulation** and **Assignment** structures.

The second utilization mode does not require any real field or simulated data. Hence in this mode, the reservoir network is designed independently. In Figure 6.1, the input data collection phase is thus skipped, as well as the network clustering and microscopic trip set generation. Then, MFD, trip length and demand data are set manually, or by using automated scripts to define regular network configurations. A benchmark example may be the honeycomb city model, as presented in Yildirimoglu *et al.* (2015), Ramezani *et al.* (2015).

Then the simulation phase (III) and post-processing phase (IV) are detailed next.

### 6.1.3 General architecture

#### Main file

We present here the general architecture of the simulation code. A simulation is launched via a main file, see Figure 6.2. In the first assignment mode, two loops are embedded in the simulation process to simulate traffic states with DTA. The detail of these loops is explained further. In the second mode however, the main file successively runs all blocks without any loop.

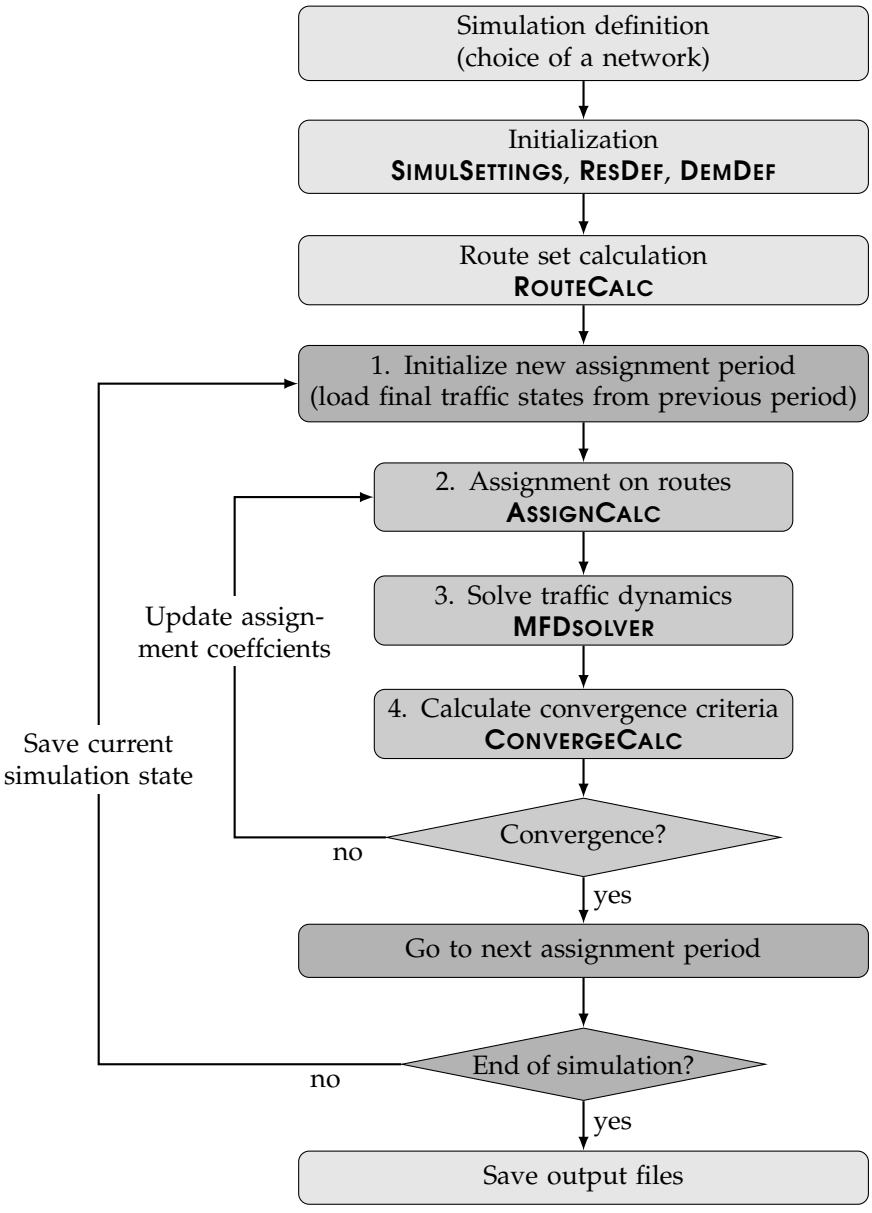



FIG. 6.2 – *Flow chart of the platform main file*

#### Assignment and convergence loops

The **ASSIGNMENT** and **CONVERGECALC** modules are presented in more details here. These modules have been designed during the collaboration with PhD student Sergio Batista.

The DTA process consists of two loops. The first one simply splits the simulation duration in several periods of a duration $T_a$, mentioned in the **Assignment** structure. The second one aims at assigning flows or vehicles on the routes so that the resulting traffic states comply with a given equilibrium specified in **Assignment** for the current period. This optimization problem corresponds to the classical problem of traffic assignment, it is solved via a widely used fixed-point method, known as the Method of Successive Averages (MSA). Hence, the second loop consists of the following steps (see also Figure 6.2):

1. At the beginning of the current period, the final simulation state from the previous period is loaded (basically the **Reservoir**, **Route** and eventually **Vehicle** structures), and it becomes the initial state of the current period. The first period is based on the simulation initial condition (empty network as explained earlier).

2. The assignment coefficients are calculated in the **AssignCalc** module. At each iteration $i_{\text{MSA}}$ of the MSA loop, traffic results from the previous iteration are used to estimate the disutility of each route $p$, i.e. its travel time $T^p(t)$. In classical assignment problems, this travel time is derived from the route cumulative count curves. The situation is different in multi-reservoir systems however, as the correlation between the routes is not only due to fluctuations in the reservoir mean speed but also to fluctuations between trip lengths inside each reservoir. This leads us to calculate the travel time by using both mean speed and trip length data. More details are given in Batista & Leclercq (2018a) who proposed this approach. Then, the new assignment coefficient $a^{p*}_{i_{\text{MSA}}}$ proposed at iteration $i_{\text{MSA}}$ corresponds to the probability of choosing route $p$ for a given OD. The calculation of this probability depends on the equilibrium type. For UE, the choice is given by the minimum of the mean travel time $\langle T^p \rangle_{T_a} = \sum_{r \in R^p} L^r_p / \langle v^r \rangle_{T_a}$ of all routes of a given OD, calculated for the current assignment period. When including uncertainty in UE, Monte-Carlo calculations of possible travel times are performed: the same formula is used, but with distributed values of length and speed. The distribution of trip lengths corresponds to the lengths from the micro trip set, while the distribution of speeds comes from the mean speed evolution during the last simulation iteration $i_{\text{MSA}} - 1$ (of period $T_a$). Then, the route choice for each Monte-Carlo draw is the minimum of the travel times among possible routes of each OD. The average choice among all draws gives each probability $a^{p*}_{i_{\text{MSA}}}$ for the current iteration $i_{\text{MSA}}$ of choosing route $p$ for a given OD. Options are available in the module to select only one source of uncertainty (trip length or mean speed). Other equilibriums like BRUE and RTUE are also included in the module, but not detailed here, see Batista & Leclercq (2018a) for more information. The updated assignment coefficient is finally obtained via the MSA formula (Sheffi, 1985):

$$a^p_{i_{\text{MSA}}} = \frac{1}{i_{\text{MSA}}} a^{p*}_{i_{\text{MSA}}} + \left(1 - \frac{1}{i_{\text{MSA}}}\right) a^p_{i_{\text{MSA}}-1} \tag{6.1}$$

   For the first iteration, the initial mean speed values of the current period are provided.

3. The solver is run for the current period with the updated assignment coefficients that indicate the demand per route. At the end of the period, the mean speed evolution $v^r(t)$ in the reservoirs is saved for the next assignment update.

4. The convergence of the MSA is tested in the **ConvergeCalc** module, where we focus on two criteria. The first one is called the Gap criterion (Sbayti *et al.*, 2007). The Gap corresponds to the total relative difference between the route mean travel time and the minimum travel time among all OD pairs:

$$\text{Gap} = \sum_{OD} \frac{1}{\langle T^{\min} \rangle_{T_a}} \sum_{p \in OD} a^p_{i_{\text{MSA}}} \left( \langle T^p \rangle_{T_a} - \langle T^{\min} \rangle_{T_a} \right) \tag{6.2}$$

where $\langle T^{\min} \rangle_{T_a} = \min_{p \in OD} \langle T^p \rangle_{T_a}$. This criterion is satisfied once the Gap is lower than a given value indicated in **Assignment**. The second criterion is called the number of violations. It corresponds to the number of routes with the difference between old and updated coefficients $a^p_{i_{\text{MSA}}} - a^p_{i_{\text{MSA}}-1}$ greater than a given threshold indicated in **Assignment**. This criterion is satisfied if the proportion of routes in violation is below a given threshold. Then the convergence of the current period is achieved once one or both criteria are satisfied, depending on the options. The procedure goes back to step 2 while the convergence is not achieved.

### 6.1.4 Accumulation-based and trip-based solvers

We present here the accumulation-based and trip-based algorithms used in the **MFDSolver** module. Both solvers take the period initial accumulations and demand on all the routes as input, and return the evolution of the system for the current assignment period (accumulation, mean speed, inflow/outflow or cumulative count curves for each route).

#### Main differences between the accumulation-based and the trip-based solvers

The theoretical formulations of the accumulation-based and trip-based models are presented in chapter 2 for a single reservoir with a unique trip category, and in chapter 4 for a single reservoir in a multi-reservoir environment with multiple trip categories. The differences between both models, in terms of traffic dynamics, are extensively reported in these chapters. Here, we rather describe the general differences regarding the implementation of these models, as explained below:

- About the representation of traffic dynamics, the accumulation-based solver deals with continuous vehicle flow and may eventually handle fractions of vehicles, whereas the trip-based solver keeps track of individual vehicle trajectories and cannot split a single vehicle into several fractions.

- About the general solver algorithm, the accumulation-based model includes a finite difference method with a fixed time step, i.e. the accumulation at the next time step is calculated with the inflow and outflow difference of the previous time step. On the other hand, the trip-based solver is an event-based scheme, as previously detailed in chapter 2

- Concerning the computational complexity, it appears that the accumulation-based solver has a complexity proportional to the number of time steps. Its computational time can be improved by increasing the time step. The number of calculations at each time

step are always the same regardless of each reservoir traffic state. However, the trip-based solver complexity is proportional to the number of traveling vehicles and thus increases during congestion periods where more events are computed during a given time period. In case of very low demand flows, the trip-based solver may eventually have a faster computational time than the accumulation-based one. But very low demand flows are unlikely in large-scale simulation studies, and the trip-based complexity is generally much higher than the accumulation-based one. A simple simulation example is given in the following to illustrate this.

- Regarding the algorithm stability, the accumulation-based solver is deterministic and stable as long as the time step does not exceed the minimum free-flow travel time in any reservoir (general stability condition for a finite difference method). It provides a smooth evolution of flows and accumulations (which may be not very consistent with usual traffic dynamics, see our discussion in chapter 2). Because a vehicle unit cannot be split in the trip-based solver, random draws are required to arbitrate between different vehicles that want to enter or exit a reservoir at the same time. This stochastic component has a few impacts on the system stability in undersaturated conditions, but may be problematic when some reservoirs are oversaturated. In this case, instabilities come from two phenomena: first, there are much more event conflicts than in undersaturation, because all the waiting vehicles in different reservoirs want to enter or exit as soon as possible, and second, the modeling of congestion propagation in the trip-based model may be very sensitive to numerical differences of the order of one vehicle. Actually, the modeling of spillbacks in the trip-based model is an event-based transcription of the flow merging model implemented in the accumulation-based model, where the possibility of splitting fractions of vehicles may be required to ensure the consistency of the merging scheme. When splitting vehicles is not possible as in the trip-based model, numerical errors are thus introduced.

### Accumulation-based solver algorithm

The accumulation-based solver is described by the pseudo-code in algorithm 2. Below are the additional notations used throughout the algorithm:

- In the **Reservoir** structure, $r \in \{1, \dots, N_r\}$:
  - subset $\mathcal{P}^r_{\text{int}}$ of the routes in $\mathcal{P}^r$ that originate inside reservoir $r$
  - subset $\mathcal{P}^r_{\text{ext}}$ of the routes in $\mathcal{P}^r$ that originate outside reservoir $r$
  - modified entry production supply $P^r_{s,\text{ext}}$ accounting for internal trip generation in reservoir $r$
  - average trip length $L^r_{\text{ext}}$ corresponding to the routes originating outside reservoir $r$ only
  - subset $\mathcal{IB}^r_{r'}$ of the routes in $\mathcal{P}^r$ coming from neighboring reservoir $r' \in \mathcal{R}^r$
  - outflow demand $O^r_p$ and outflow supply $\mu^r_p$ per route $p \in \mathcal{P}^r$
  - inflow supply $I^r_p$ and inflow merge coefficient $\alpha^r_p$ per route $p \in \mathcal{P}^r$
- In the **Route** structure, $p \in \{1, \dots, N_p\}$:
  - origin reservoir $R^p[1]$, destination reservoir $R^p[end]$

– previous reservoir $p^-(r)$ and next reservoir $p^+(r)$ for a given reservoir $r \in R^p$

Then the cumulative count curves and travel times are calculated in post-processing:

$$\forall r \in \{1,\dots,N_r\}, p \in \mathcal{P}^r,\ \ N^r_{\text{in},p}(t) = \int_0^t q^r_{\text{in},p}(\tau)d\tau + n^r_p(0) \tag{6.3}$$

$$N^r_{\text{out},p}(t) = \int_0^t q^r_{\text{out},p}(\tau)d\tau \tag{6.4}$$

The *Merge*() algorithm used to calculate inflow supply in reservoirs was proposed in Leclercq & Becarie (2012) and consists in an extension of the fair merge of Daganzo (1995). It ensures that the total available capacity is always fully used when only some of the merging flows are limited. It is described by the pseudo-code in algorithm 1. Note that it can be applied to either flow or production values.

ALG. 1 – *Fair merge with multiple incoming flows*

**Function** *Merge*$(\{\Lambda_i\}_{1\le i\le M}, \{\alpha_i\}_{1\le i\le M}, C)$
**Input:** set of $M$ incoming demand flows (resp. productions) $\{\Lambda_i\}_{1\le i\le M}$ with respective merge coefficients $\{\alpha_i\}_{1\le i\le M}$ towards a unique entry with flow (resp. production) capacity $C$
**Output:** resulting effective inflows (resp. entering productions) $\{Q_i\}_{1\le i\le M}$
**Initialization:**
set of unserved flows: $U = \{1,\dots,M\}$
sum of all coefficients in $U$: $\alpha_U = 1$
total inflow already served: $Q_F = 0$
**while** $U \neq \emptyset$ **do**
set $U' = \emptyset$; $\alpha'_U = 0$; $Q'_F = 0$
**for** $i \in U$ **do**
**if** $\Lambda_i < \alpha_i/\alpha_U(C - Q_F)$ **then**
demand $i$ is served: $Q_i = \Lambda_i$
$Q'_F = Q'_F + Q_i$
**else**
demand $i$ is not served: $Q_i = \alpha_i/\alpha_U(C - Q_F)$
$U' = \{U', i\}$
$\alpha'_U = \alpha'_U + \alpha_i$
**end if**
**end for**
update $U = U'$; $\alpha_U = \alpha'_U$; $Q_F = Q_F + Q'_F$
**if** $\sum_{i=1}^M Q_i = C$ **then**
stop the procedure by setting $U = \emptyset$
**end if**
**end while**
**end function**

ALG. 2 – *Accumulation-based solver*

**Input:** reservoir initial accumulation $n_p^r(t_0)$ per route, route demand profile $\lambda^p(t)$
**Output:** reservoir accumulation $n_p^r(t)$, inflow $q_{\text{in},p}^r(t)$ and outflow $q_{\text{out},p}^r(t)$ per route

**Initialization:** load from previous assignment period $i_{\text{period}} - 1$
initial time: $t_0 = (i_{\text{period}} - 1)T_a$
initial accumulation: $\forall r \in \{1, \dots, N_r\}, p \in \mathcal{P}^r, n_p^r(t_0)$

**for** $t = t_0$ **to** $t_0 + T_a$ **by** $\delta t_s$ **do**
  **for** $r = 1$ **to** $N_r$ **do**
    Outflow demand: $\forall p \in \mathcal{P}^r, O_p^r = \frac{n_p^r(t)}{n^r(t)} \frac{P_d^r(n^r(t))}{L_p^r}$
    Production supply: $P_{s,\text{ext}}^r = P_s^r(n^r(t)) - \sum_{p \in \mathcal{P}_{\text{int}}^r} L_p^r \lambda^p(t)$
    Average trip length: $L_{\text{ext}}^r = \sum_{p \in \mathcal{P}_{\text{ext}}^r} n_p^r(t) / \sum_{p \in \mathcal{P}_{\text{ext}}^r} n_p^r(t) / L_p^r$
  **end for**
  **for** $r = 1$ **to** $N_r$ **do**
    Merging coefficients:
    $$\forall p \in \mathcal{P}_{\text{ext}}^r, \begin{cases} \alpha_p^r = n_p^r(t) / \sum_{k \in \mathcal{P}_{\text{ext}}^r} n_k^r(t) & \text{if endogenous coeff.} \\ \alpha_p^r = O_p^{p^-(r)} / \sum_{k \in \mathcal{P}_{\text{ext}}^r} O_k^{k^-(r)} & \text{if demand pro-rata coeff.} \end{cases}$$
    Border inflow supply:
    $$\forall r' \in \mathcal{R}^r, \{I_p^{r*}\}_{p \in \mathcal{IB}_{r'}^r} = \text{Merge}\left( \{O_p^{p^-(r)}\}_{p \in \mathcal{IB}_{r'}^r}, \left\{ \frac{\alpha_p^r}{\sum_{k \in \mathcal{IB}_{r'}^r} \alpha_k} \right\}_{p \in \mathcal{IB}_{r'}^r}, C_{r'}^r \right)$$
    Reservoir inflow supply:
    $$\begin{cases} \{L_p^r I_p^r\}_{p \in \mathcal{P}_{\text{ext}}^r} = \text{Merge}\left( \{L_p^r I_p^{r*}\}_{p \in \mathcal{P}_{\text{ext}}^r}, \{\alpha_p^r\}_{p \in \mathcal{P}_{\text{ext}}^r}, P_{s,\text{ext}}^r \right) & \text{if endogenous coeff.} \\ \{I_p^r\}_{p \in \mathcal{P}_{\text{ext}}^r} = \text{Merge}\left( \{I_p^{r*}\}_{p \in \mathcal{P}_{\text{ext}}^r}, \{\alpha_p^r\}_{p \in \mathcal{P}_{\text{ext}}^r}, \frac{P_{s,\text{ext}}^r}{L_{\text{ext}}^r} \right) & \text{if demand pro-rata coeff.} \end{cases}$$
  **end for**
  **for** $r = 1$ **to** $N_r$ **do**
    Outflow supply: $\forall p \in \mathcal{P}^r$, **if** $r \neq R^p[end]$ **then** $\mu_p^r = I_p^{p^+(r)}$ **else** $\mu_p^r = +\infty$
    Effective outflow:
    $$\begin{cases} \forall p \in \mathcal{P}^r, q_{\text{out},p}^r(t) = \frac{n_p^r(t)}{n_k^r(t)} \frac{L_k^r}{L_p^r} \min\left[ O_k^r; \mu_k^r \right] & \text{if maximum outflow demand} \\ \quad \text{where: } k = \arg\min_{p \in \mathcal{P}^r} \frac{\mu_p^r}{O_p^r} & \\ q_{\text{out},p}^r(t) = \min\left[ O_p^r; \mu_p^r \right] & \text{if decreasing outflow demand} \end{cases}$$
    Effective inflow: $\forall p \in \mathcal{P}^r, q_{\text{in},p}^r(t) = \begin{cases} \lambda^p(t) & \text{if } r = R^p[1] \\ q_{\text{out},p}^{p^-(r)}(t) & \text{otherwise} \end{cases}$
    Accumulation update: $\forall p \in \mathcal{P}^r, n_p^r(t + \delta t_s) = n_p^r(t) + \delta t_s \left( q_{\text{in},p}^r(t) - q_{\text{out},p}^r(t) \right)$
  **end for**
**end for**

### Trip-based solver algorithm

The trip-based solver is described by the pseudo-code in algorithm 3. Below are the additional notations used throughout the algorithm (other notations from the accumulation-based solver are also used):

- In the **Reservoir** structure, $r \in \{1, \dots, N_r\}$:
  - entry demand time $t^r_{\text{in},d,p}$ and supply time $t^r_{\text{in},s,p}$ per route $p \in \mathcal{P}^r$
  - exit demand time $t^r_{\text{out},d,p}$ and supply time $t^r_{\text{out},s,p}$ per route $p \in \mathcal{P}^r$
  - last entry time $t^r_{\text{last in},p}$ and last exit time $t^r_{\text{last out},p}$ per route $p \in \mathcal{P}^r$
  - estimated inflow demand $q^r_{\text{in},d,p}$ per route $p \in \mathcal{P}^r$
  - possible next entry time $t^r_{\text{in}}$ and next exit time $t^r_{\text{out}}$
  - list of traveling vehicles per route $U^r_p$ sorted by remaining travel distance, $p \in \mathcal{P}^r$
  - traveled distance during the elapsed time $\Delta t$ for all the vehicles in $r$: $TD^r$
- In the **Vehicle** structure, $u \in \{1, \dots, N_u\}$:
  - current traveled distance $TD^u$
  - current traveled time $TT^u$

Then the inflow and outflow are calculated in post-processing:

$$\forall r \in \{1, \dots, N_r\}, p \in \mathcal{P}^r, \quad q^r_{\text{in},p}(t) = \frac{dN^r_{\text{in},p}}{dt} \tag{6.5}$$

$$q^r_{\text{out},p}(t) = \frac{dN^r_{\text{out},p}}{dt} \tag{6.6}$$

In practice, the derivatives are not computed on the direct outputs from the trip-based solver to avoid high oscillations in flow. The cumulative count curves are thus smoothed and filtered over a fixed number of consecutive vehicles before the numerical derivative is applied.

The *MergeTime*() algorithm used to calculate entry supply times in reservoirs corresponds to the adaptation of the function in algorithm 1 used in the accumulation-based solver. It is described by the pseudo-code in algorithm 4.

ALG. 3 – *Trip-based solver*

**Input:** reservoir initial accumulation $n_p^r(t_0)$, cumulative count curves $N_{\text{in},p}^r(t_0)$ and $N_{\text{out},p}^r(t_0)$ per route, route demand profile $\lambda^p(t)$, beginning trajectory of vehicles that already started traveling $L^u$, $T_{\text{in}}^u$ and $T_{\text{out}}^u$

**Output:** reservoir accumulation $n_p^r(t)$, cumulative count curves $N_{\text{in},p}^r(t)$ and $N_{\text{out},p}^r(t)$ per route, vehicle trajectory $L^u$, $T_{\text{in}}^u$ and $T_{\text{out}}^u$

**Initialization:** load from previous assignment period $i_{\text{period}} - 1$
initial time: $t = t_0 = (i_{\text{period}} - 1)T_a$, and elapsed time $\Delta t = 0$
initial accumulation: $\forall r \in \{1, \dots, N_r\}, p \in \mathcal{P}^r, n_p^r(t_0)$
current vehicle: $u_0$ (loaded from previous period or first vehicle if $i_{\text{period}} = 1$)

**while** $t < t_0 + T_a$ **do**
  Set by default:
  $\forall r \in \{1, \dots, N_r\}, p \in \mathcal{P}^r, n_p^r(t) = n_p^r(t - \Delta t), N_{\text{in},p}^r(t) = N_{\text{in},p}^r(t - \Delta t), N_{\text{out},p}^r(t) = N_{\text{out},p}^r(t - \Delta t), v^r(t) = v^r(t - \Delta t)$
  Update traveled distances and times:
  **for** $r = 1$ **to** $N_r$ **do**
    $TD^r = \Delta t.v^r(t)$
    **for** $u \in U^r$ **do**
      $TD^u = TD^u + TD^r$
      $TT^u = TT^u + \Delta t$
    **end for**
  **end for**
  Current vehicle $u_0$ information:
  traveling in reservoir $r_0$, at index $i_0$ on route $p_0$ ($r_0 = R^{p_0}[i_0]$)
  set of reservoirs where accumulation/mean speed changed: $\mathcal{SR} = \emptyset$
  Update entering and/or exiting reservoir information:
  **if** *vehicle* $u_0$ *exits* $r_0$ **then**
    $T_{\text{out}}^{u_0}[i_0] = t$
    $n_{p_0}^{r_0}(t) = n_{p_0}^{r_0}(t) - 1, N_{\text{out},p_0}^{r_0}(t) = N_{\text{out},p_0}^{r_0}(t) + 1, v^{r_0}(t) = V^{r_0}(n^{r_0}(t))$
    $\mathcal{SR} = \{\mathcal{SR}, r_0\}$
    remove $u_0$ from the list $U_{p_0}^{r_0}$
  **end if**
  **if** *vehicle* $u_0$ *enters* $r_0$ **or** *enters next reservoir* $p_0^+(r_0)$ *of the route* $p_0$ **then**
    **if** *vehicle enters* $p_0^+(r_0)$ **then** $r_0 = p^+(r_0), i_0 = i_0 + 1$
    $T_{\text{in}}^{u_0}[i_0] = t$
    $n_{p_0}^{r_0}(t) = n_{p_0}^{r_0}(t) + 1, N_{\text{in},p_0}^{r_0}(t) = N_{\text{in},p_0}^{r_0}(t) + 1, v^{r_0}(t) = V^{r_0}(n^{r_0}(t))$
    $\mathcal{SR} = \{\mathcal{SR}, r_0\}$
    append $u_0$ to the list $U_{p_0}^{r_0}$ and sort the list by remaining travel distance $L^u[i] - TD^u, u \in U_{p_0}^{r_0}$
  **else**
    vehicle $u_0$ completed its trip
    route travel time: $T^{p_0}(t) = T_{\text{out}}^{u_0}[end] - T_{\text{in}}^{u_0}[1]$
  **end if**

Reservoir exit demand times:
**for** $r \in \mathcal{SR}$ **do**
  **for** $p \in \mathcal{P}^r$ **do**
    first vehicle to exit: $u_0 = U_p^r[1]$ at reservoir index $i_0$
    **if** $n^r(t) > n_c^r$ **and** *maximum outflow demand* **then** $TD^{u_0} = L^{u_0}[i_0]$
    $t^r_{\text{out},d,p} = t + (L^{u_0}[i_0] - TD^{u_0})/v^r(t)$
  **end for**
**end for**
Reservoir entry demand times and estimated inflow demand:
**for** $r \in \mathcal{SR}$ **do**
  $\forall p \in \mathcal{P}^r$, **if** $r \neq R^p[1]$ **then** $t^r_{\text{in},d,p} = t^{p^-(r)}_{\text{out},d,p}$ **else** $t^r_{\text{in},d,p} = t^r_{\text{last in},p} + 1/\lambda^p(t)$
  $\forall p \in \mathcal{P}^r$, **if** $r \neq R^p[1]$ **then** $q^r_{\text{in},d,p} = 1/\left(t^{p^-(r)}_{\text{out},d,p} - t^{p^-(r)}_{\text{last out},p}\right)$ **else** $q^r_{\text{in},d,p} = \lambda^p(t)$
**end for**
Reservoir entry supply times:
**for** $r \in \mathcal{SR}$ **do**
  merging coefficients:
  $\forall p \in \mathcal{P}^r_{\text{ext}}, \begin{cases} \alpha^r_p = n^r_p(t)/\sum_{k\in\mathcal{P}^r_{\text{ext}}} n^r_k(t) & \text{if endogenous merge} \\ \alpha^r_p = q^r_{\text{in},d,p}/\sum_{k\in\mathcal{P}^r_{\text{ext}}} q^r_{\text{in},d,k} & \text{if pro-rata merge} \end{cases}$
  modification of entry demand times due to border supply:
  $\forall r' \in \mathcal{R}^r, \{t^{r*}_{\text{in},d,p}\}_{p\in\mathcal{IB}^r_{r'}} =$
  $$\text{MergeTime}\left(\{t^r_{\text{in},d,p}\}_{p\in\mathcal{IB}^r_{r'}}, \left\{\frac{\alpha^r_p}{\sum_{k\in\mathcal{IB}^r_{r'}}\alpha_k}\right\}_{p\in\mathcal{IB}^r_{r'}}, \{1\}_{p\in\mathcal{IB}^r_{r'}}, \{t^r_{\text{last in},p}\}_{p\in\mathcal{IB}^r_{r'}}, C^r_{r'}\right)$$
  $\forall r' \in \mathcal{R}^r, p \in \mathcal{IB}^r_{r'}, t^r_{\text{in},d,p} = \max\left[t^r_{\text{in},d,p}; t^{r*}_{\text{in},d,p}\right]$
  production supply: $P^r_{s,\text{ext}} = P^r_s(n^r(t)) - \sum_{p\in\mathcal{P}^r_{\text{int}}} L^r_p \lambda^p(t)$
  average trip length: $L^r_{\text{ext}} = \sum_{p\in\mathcal{P}^r_{\text{ext}}} n^r_p(t)/\sum_{p\in\mathcal{P}^r_{\text{ext}}} n^r_p(t)/L^r_p$
  entry supply times: $\{t^r_{\text{in},s,p}\}_{p\in\mathcal{P}^r_{\text{ext}}} =$
  $$\begin{cases} \text{MergeTime}\left(\{t^r_{\text{in},d,p}\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{\alpha^r_p\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{L^r_p\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{t^r_{\text{last in},p}\}_{p\in\mathcal{P}^r_{\text{ext}}}, P^r_{s,\text{ext}}\right) \\ \quad \text{if endogenous merge} \\ \text{MergeTime}\left(\{t^r_{\text{in},d,p}\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{\alpha^r_p\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{1\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{t^r_{\text{last in},p}\}_{p\in\mathcal{P}^r_{\text{ext}}}, \frac{P^r_{s,\text{ext}}}{L^r_{\text{ext}}}\right) \\ \quad \text{if pro-rata merge} \end{cases}$$
**end for**
Reservoir exit supply times:
$\forall r \in \{1,\ldots,N_r\}, p \in \mathcal{P}^r$, **if** $r \neq R^p[end]$ **then** $t^r_{\text{out},s,p} = t^{p^+(r)}_{\text{in},s,p}$ **else** $t^r_{\text{out},s,p} = t$

Next possible entry time (vehicle creation):
$\forall r \in \{1, \dots, N_r\}, t^r_{\text{in}} = \min_{p \in \mathcal{P}^r_{\text{int}}} t^r_{\text{in},d,p}$
$t_{\text{in}} = \min_{1 \le r \le N_r} t^r_{\text{in}}$
Next possible exit time:
$\forall r \in \{1, \dots, N_r\}, t^r_{\text{out}} =$

$$\begin{cases} \max[t^r_{\text{out},d,k}; t^r_{\text{out},s,k}] & \text{if maximum outflow demand} \\ \quad \text{where: } k = \arg\min_{p \in \mathcal{P}^r} t^r_{\text{out},d,p} & \\ \min_{p \in \mathcal{P}^r} \left[\max[t^r_{\text{out},d,p}; t^r_{\text{out},s,p}]\right] & \text{if decreasing outflow demand} \end{cases}$$

$t_{\text{out}} = \min_{1 \le r \le N_r} t^r_{\text{out}}$
Next event time:
$t_{\text{event}} = \min[t_{\text{in}}; t_{\text{out}}]$
Save corresponding vehicle $u_0$
Time update:
$\Delta t = t_{\text{event}} - t$
$t = t_{\text{event}}$

**end while**

ALG. 4 – *Fair merge with multiple incoming flows to determine entry supply times*

**Function** *MergeTime*$(\{T_{d,i}\}_{1\leq i\leq M}, \{\alpha_i\}_{1\leq i\leq M}, \{L_i\}_{1\leq i\leq M}, \{T_{\text{last},i}\}_{1\leq i\leq M}, C)$

**Input:** set of $M$ incoming flows with entry demand times $\{T_{d,i}\}_{1\leq i\leq M}$ with respective merge coefficients $\{\alpha_i\}_{1\leq i\leq M}$, trip lengths $\{L_i\}_{1\leq i\leq M}$ and last entry times $\{T_{\text{last},i}\}_{1\leq i\leq M}$ towards a unique entry with flow (resp. production) capacity $C$. If $C$ is in flow units, put by default $\forall i \in \{1,\dots,M\}, L_i = 1$ (no need for the trip length input)

**Output:** resulting entry supply times $\{T_{s,i}\}_{1\leq i\leq M}$

**Initialization:**
set of unserved flows: $U = \{1,\dots,M\}$
sum of all coefficients in $U$: $\alpha_U = 1$
total inflow already served: $Q_F = 0$

**while** $U \neq \emptyset$ **do**
  set $U' = \emptyset$; $\alpha'_U = 0$; $Q'_F = 0$
  **for** $i \in U$ **do**
    Inflow supply for demand $i$: $Q_{s,i} = \alpha_i/\alpha_U(C - Q_F)/L_i$
    Entry supply time for demand $i$: $T_{s,i} = T_{\text{last},i} + 1/Q_{s,i}$
    Effective flow for demand $i$: $Q_i = L_i Q_{s,i}$
    **if** $T_{d,i} \geq T_{s,i}$ **then**
      demand $i$ is served
      $Q'_F = Q'_F + Q_i$
    **else**
      demand $i$ is not served
      $U' = \{U', i\}$
      $\alpha'_U = \alpha'_U + \alpha_i$
    **end if**
  **end for**
  update $U = U'$; $\alpha_U = \alpha'_U$; $Q_F = Q_F + Q'_F$
  **if** $\sum_{i=1}^{M} Q_i = C$ **then**
    stop the procedure by setting: $U = \emptyset$
  **end if**
**end while**
**end function**

### Improving the computational time of the trip-based solver

One of the drawbacks of the trip-based solver is its high complexity compared with the accumulation-based one. Moreover, while the computational time of the latter can be improved by increasing the time step (up to the stability limit), nothing can be done to reduce the number of calculations in the trip-based solver. However, if one is only interested in the evolution of aggregated variables (accumulation, flow, mean speed, etc), and not in the detail of each individual vehicle trajectory, a solution to improve the computational time of this solver exists. Its principle is the following: because the complexity depends on the number of vehicles handled during the simulation, the idea is to reduce the demand to create less vehicles. Obviously, other boundary conditions like supply limitations must also be modified accordingly, as well as each reservoir MFD, to observe the same traffic evolution.

Here, we explain our solution method in more details for a single reservoir with a unique trip category. The demonstration for a multi-reservoir system with multiple trips will follow the same logic but is not presented. This method is based on a transformation of the reservoir conservation equation. Considering a reservoir with a production-MFD $P(n)$, an average trip length $L$, and being loaded by a demand $\lambda(t)$ in undersaturated conditions, the conservation equation governing its traffic dynamics is:

$$\frac{dn}{dt} = \lambda(t) - \frac{P(n(t))}{L} \tag{6.7}$$

where $n(t)$ is the accumulation. In order to reduce the demand, we introduce the scaling factor $\alpha_{\text{scf}}$ in [0, 1]. After multiplying both sides of the previous equation, it comes:

$$\frac{d(\alpha_{\text{scf}}.n)}{dt} = \alpha_{\text{scf}}.\lambda(t) - \frac{\alpha_{\text{scf}}.P(n(t))}{L} \tag{6.8}$$

Then, we can modify the MFD function $P(n)$ to integrate the scaling factor in it. This function can have any shape, but its only requirement is to be concave as suggested by the analytical method for MFD estimation developed by Daganzo & Geroliminis (2008). In practice, this means that a set of linear functions can be a good approximation of any MFD shape:

$$P(n) = \min_{1 \leq i \leq N_b} [w_i.n + P_{c,i}] \tag{6.9}$$

where $N_b$ is the number of MFD branches (linear functions) and $P_i(n) = w_i.n + P_{c,i}$ is the $i$th branch of $P(n)$. Under this approximation, it follows that:

$$\alpha_{\text{scf}}.P(n) = \min_{1 \leq i \leq N_b} [w_i.\alpha_{\text{scf}}.n + \alpha_{\text{scf}}.P_{c,i}] = P_{\alpha_{\text{scf}}}(\alpha_{\text{scf}}.n) \tag{6.10}$$

Thus in modifying the MFD function into $P_{\alpha_{\text{scf}}}(n)$, the conservation equation finally becomes:

$$\frac{d(\alpha_{\text{scf}}.n)}{dt} = \alpha_{\text{scf}}.\lambda(t) - \frac{P_{\alpha_{\text{scf}}}(\alpha_{\text{scf}}.n(t))}{L} \tag{6.11}$$

which can be rewritten in:

$$\frac{dn'}{dt} = \lambda'(t) - \frac{P_{\alpha_{\text{scf}}}(n'(t))}{L} \tag{6.12}$$

Hence, we conclude that solving the initial problem in equation 6.7 is completely equivalent to solve the new problem in equation 6.12, where the demand has been scaled down and the

MFD modified accordingly. Note that the same demonstration applies during a congestion period with an exit limitation $\mu(t)$, where the conservation equation is written as $dn/dt = P_s(n(t))/L - \mu(t)$.

Hence, assuming that the MFD functions are always approximated by piecewise linear functions, the computation improvement for the trip-based solver is described as follows:

1. Apply a scaling factor $\alpha_{\text{scf}}$ to all boundary conditions, i.e. demands and supply limitations, and modify the MFD functions accordingly with the transformation $P_{\alpha_{\text{scf}}}(n)$.

2. Solve the new problem with these scaled boundary conditions and transformed MFD (which is much faster than solving the initial problem because of the demand reduction).

3. Get the results of the initial problem by applying the inverse scaling on the outputs of the new problem: apply the factor $1/\alpha_{\text{scf}}$ to accumulations, inflows and outflows previously calculated with the new problem.

Our previous explanations on a single reservoir do not replace a rigorous mathematical proof of the equivalence of the two problems in any case of multi-reservoir systems with multiple trips. Nevertheless, they bring the idea behind the down-scaling of demand through the transformation of the conservation equation. More investigations are needed to confirm the validity of this method in any case.

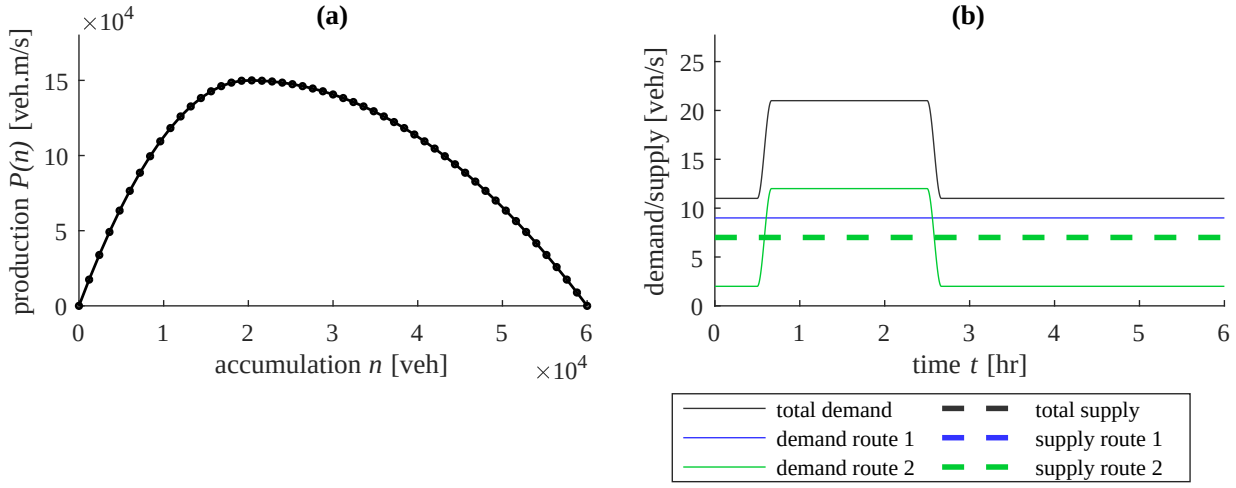



FIG. 6.3 – *Example of a single reservoir settings. (a) Production-MFD defined by multiple branches, (b) demand and supply scenario*

We provide here a simple example on a single reservoir with two trip lengths to illustrate the performance of this method. Two trip-based simulations are performed, the first one corresponds to the initial reservoir settings (initial problem), while the second one uses a scaling factor equal to $\alpha_{\text{scf}} = 0.1$. Figures 6.3(a) and (b) respectively show the production-MFD and the demand/supply scenario considered. This scenario includes a network loading with congestion onset and offset due to the limitation on route 2. The two trip lengths are 4 km and 9 km. Figure 6.4 presents the evolution of accumulation, inflow and outflow for both simulations. On average, it can be seen that the second simulation reproduces the same traffic dynamics as in the initial settings. This second simulation takes 25 s to compute

around 30,600 vehicle trajectories (for a single reservoir, a trajectory corresponds to an entry and exit time), compared to 18 min for the first simulation where 306,000 vehicles have been created (computational time given for a personal laptop with a 1.2 GHz processor and 8 Go memory). This illustrates the use of the scaling factor to greatly reduce the computational burden. As a comparison, the accumulation-based solver only takes 0.5 s to compute 4320 iterations with a time step $\delta t_s = 5$ s.

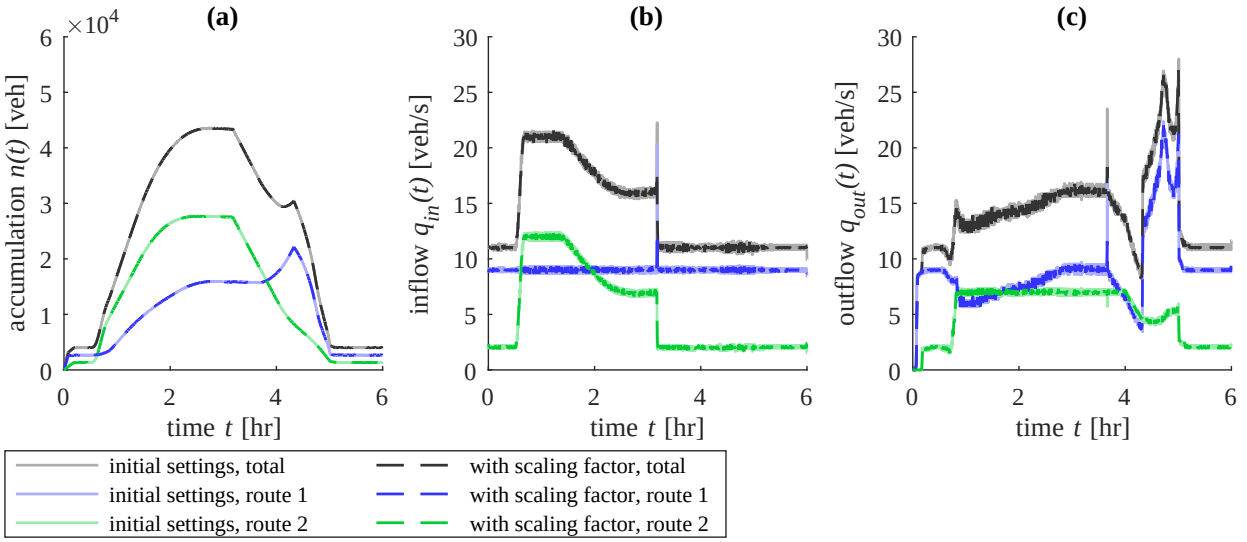



FIG. 6.4 – *Comparison between the initial problem settings and the new results after applying the scaling factor. (a) Evolution of accumulation, (b) inflow and (c) outflow per route*

## 6.2 Simulation of traffic states for the Lyon metropolitan area (France)

We present now the first application of the MFD simulation platform to a real network case: the urban area of Lyon and Villeurbanne (France). This case study is a joint work with master student Victor Boulanger and post-doc Mahendra Paipuri, both involved in the MAGnUM project. It is based on different sources of data processed by other members of the project, master student Yann Lebrech, research engineers Jean Krug and Arthur Burianne.

The aim of this case study is to propose a first validation of the MFD simulation platform on a real large-scale network, by comparing our simulation results with real traffic data. However, due to the specificity of the available data, this validation test is clearly incomplete as it does not imply all the different modules we developed in our simulation platform. The observed traffic states are notably undersaturated, so that the modeling of congestion propagation cannot be tested here. In any case, another purpose of this study is also to illustrate the different steps involved in a MFD simulation, notably the pre-processing phase, as described in the previous section.

### 6.2.1 Case study and input data presentation

#### Network data and configuration

Lyon has the second greatest urban area of France, with more than 2 million inhabitants. The network studied includes the inner city of Lyon and the city of Villeurbanne, which corresponds to the urban area inside the first ring road of Lyon. This network comprises 27,000 links, with an area of 170 km$^2$ and where around 1 million trips are recorded each day. The network configuration with its environment is given in Figure 6.5. This area exchanges traffic with its surroundings via mainly 4 freeways related to 4 origin/destination cities as presented in Figure 6.5: freeway A6 from/to Paris, freeway A42 from/to Geneva, freeway A43 from/to Grenoble and freeway A7 from/to Marseille.

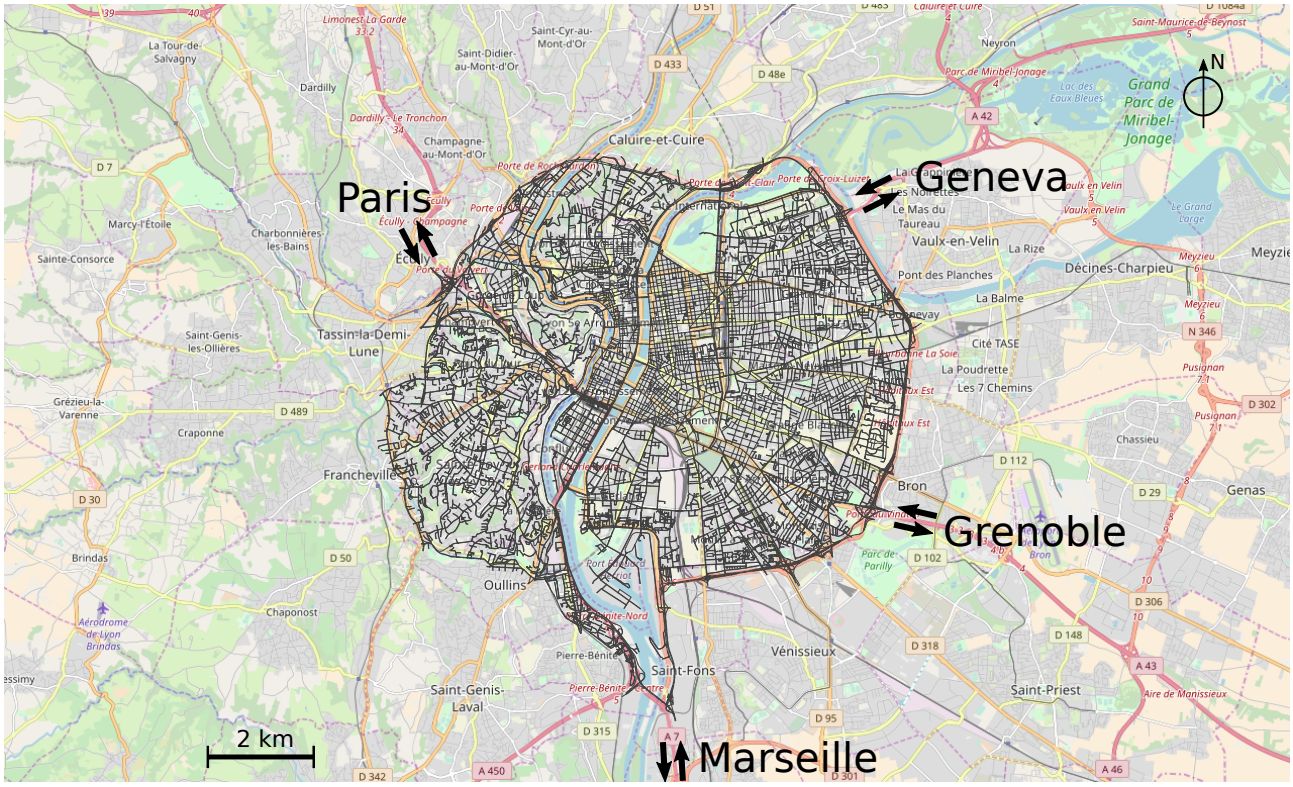



FIG. 6.5 – *The network of Lyon-Villeurbanne with its urban environment (background map credits: © OpenStreetMap)*

The available network data includes the number of lanes and the signal settings at each node with traffic lights. As mentioned in section 6.1.2, this can help to determine some characteristic values for the MFD estimation.

### Demand data

The demand was estimated for a typical weekday in a preliminary study from J. Krug and A. Burianne, research engineers from the MAGnUM project. Their study uses a four-step model based on household trip surveys and socio-demographic data to improve the localization of trip origins and destinations. The estimated OD matrix is defined for each hour of the day at the level of IRIS urban areas, the French partitioning system for demographical data. The spatial extension of an IRIS area may thus vary as its definition is based on a fixed range of inhabitants, workers, etc. In the network studied, each IRIS area comprises around 2000 inhabitants and may extend from 0.5 to 2 km$^2$. The OD matrix also includes the demand for trips from and to outside the perimeter of interest.

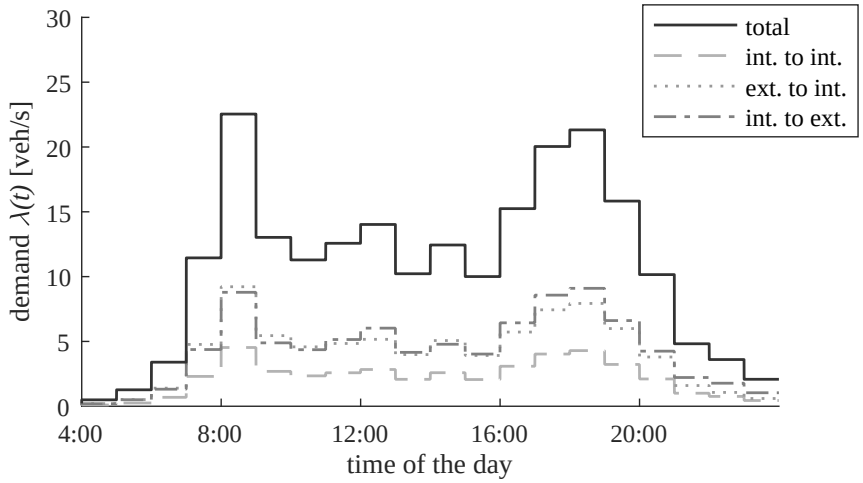



FIG. 6.6 – *Demand profile for a typical weekday split into 3 sets of trips distinguished by their origin and destination*

For this first application example, the network will be clustered into reservoirs by aggregating several neighboring IRIS areas together (demand-oriented clustering). The total demand profile is presented in Figure 6.6. This graph also shows the proportion of the total demand regarding each trip origin and destination location inside or outside the perimeter studied. We distinguish 3 sets of trips: (i) internal origin to internal destination, (ii) external origin to internal destination and (iii) internal origin to external destination. A fourth category of trip actually exists which gathers the trips crossing the whole area (external origin to external destination). Despite being characterized by a high volume of traffic (not shown in the graph), these trips are not considered in our case study because the traffic they generate is mostly located on the ring road where we have a few loop detectors. Thus, this amount of traffic is not well captured in our MFD estimation, so that we decided to neglect this part of the demand to avoid important bias in the following.

### Traffic and trip data

The traffic data was gathered in a previous study from Y. Lebrech to estimate the MFD of the area. This data consists of GPS trajectories of taxi fleets in Lyon, and all the loop detectors available in the area. The data was recorded over several days, three of them were selected for this study. The large amount of taxi trips allows to determine the mean speed of each link with an aggregation period of $\Delta T$ = 18 min. The mean speed $V^r_{\Delta T}$ of a given reservoir $r$ can be thus calculated as:

$$V^r_{\Delta T} = \frac{TTD^r_{\text{taxi}}}{TTT^r_{\text{taxi}}} = \frac{\sum_{\text{taxi}\, i=1}^{N^r_{\text{taxi}}} td^r_i}{\sum_{\text{taxi}\, i=1}^{N^r_{\text{taxi}}} tt^r_i} \tag{6.13}$$

where $TTD^r_{\text{taxi}}$ is the total travel distance and $TTT^r_{\text{taxi}}$ the total travel time in $r$ for the taxis only. There are $N^r_{\text{taxi}}$ taxis circulating in $r$ within the period $\Delta T$, $td^r_i$ is the distance traveled by taxi $i$, and $tt_i$ is the time spent by the same taxi in $r$ during this period. This estimation of the mean speed is considered quite accurate as all the cars in a link experience the same speed on average. However, it is not possible to use the same data to estimate the production or accumulation of reservoir $r$ if we do not know the penetration rate of the taxis among all the circulating vehicles. That is why another source of data is essential to get these values.

The loop data provides a measure of the mean flow of each equipped link with the same aggregation period. Then, the production of reservoir $r$ can be estimated as follows:

$$P^r_{\Delta T} = \frac{L^r_{\text{net}}}{L^r_{\text{equip}}} \left( \sum_{\substack{\text{equipped} \\ \text{link } i \in r}} l_i \langle q_i \rangle_{l_i, \Delta T} \right) \tag{6.14}$$

where $L^r_{\text{equip}} = \sum_{\text{equipped link } i \in r} l_i$ is the total length of equipped links, $L^r_{\text{net}} = \sum_{\text{link } i \in r} l_i$ is the total network length in $r$, $l_i$ is the length of link $i$ (each lane being defined by a separate link) and $\langle q_i \rangle_{l_i, \Delta T}$ is the measured mean flow in link $i$ over $\Delta T$. The ratio of the equipped length over the total length is here to compensate the fact that not all the link mean flows have been measured. The use of this ratio supposes that traffic states are quite homogeneous in $r$. While this is what we expect from a well-defined reservoir, this might not be the case at any time in reality. Hence, we consider that there may be a bias in the estimation of the production.

Finally, the accumulation in reservoir $r$ is derived from the two previous estimations:

$$n^r_{\Delta T} = \frac{P^r_{\Delta T}}{V^r_{\Delta T}} \tag{6.15}$$

The accuracy of the estimation of accumulation is thus the same as the one of the production, i.e. there is the same (unknown) bias in $n^r_{\Delta T}$ and in $P^r_{\Delta T}$.

The taxi data is also used to estimate average distances traveled. The resulting average distances for the three sets of trips defined earlier (inside to inside, outside to inside, inside to outside) are given in Table 6.1. Note that there may be a bias in these values too, as taxi trips are not necessarily representative of other users' trips.

TAB. 6.1 – *Average distances traveled by taxis for the three sets of trips*

| Trip length | [Units] | Value |
|---|---|---|
| Trips from inside to inside $L_{\text{int}>\text{int}}$ | [m] | 4800 |
| Trips from outside to inside $L_{\text{ext}>\text{int}}$ | [m] | 5900 |
| Trips from inside to outside $L_{\text{int}>\text{ext}}$ | [m] | 6000 |

### 6.2.2 Example of a single reservoir simulation

In this first simulation test, we consider the whole area as a single reservoir. Its production-MFD and speed-MFD are estimated with the data of the three days available. With the aggregation period of $\Delta T$ = 18 min, there are 80 points $(n_{\Delta T}, P_{\Delta T})$ and $(n_{\Delta T}, V_{\Delta T})$ per day. The two estimated MFD with the theoretical fit are presented in Figure 6.7. A parabolic fit model is found satisfactory to match both the production-MFD and the speed-MFD (transformed into a linear fit). As the observed traffic states do not reach oversaturation (the production does not decrease with accumulation up to a given point), we only fit the undersaturated part of the MFD. The data points are distinguished by morning and evening data, the morning data corresponds to traffic states from 00:00 to 12:00, while the evening data corresponds

to states from 12:00 to 24:00. We can see that both sets of data follow the same pattern, thus no hysteresis nor bifurcation phenomenon are noticeable. In Figure 6.7, both the accumulation and production data have been modified by a correction factor of 0.8 compared to the original dataset. This correction factor represents the bias of estimation mentioned earlier (see equation 6.14), it has been calibrated thanks to the comparison of accumulation evolution presented next.

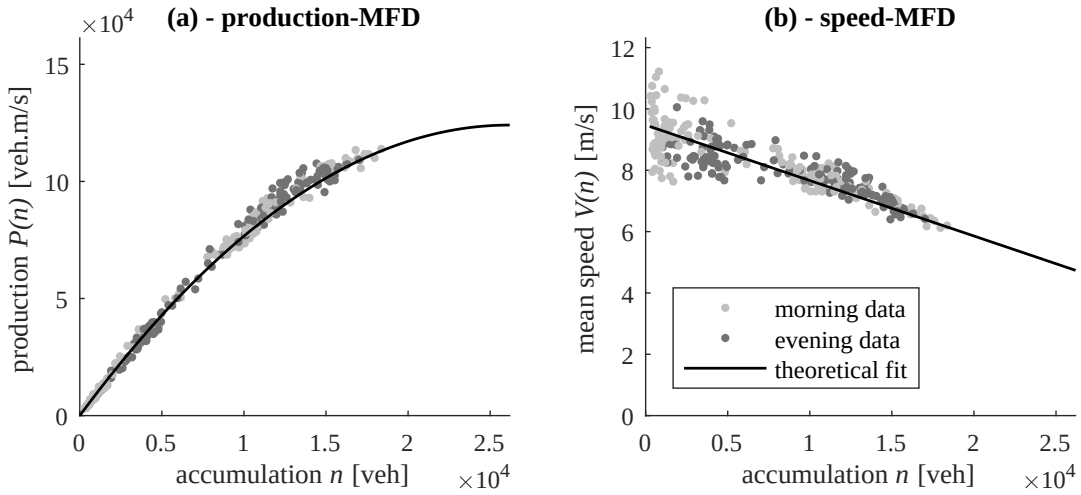



FIG. 6.7 – *MFD of the whole area as a single reservoir. (a) Production-MFD and (b) speed-MFD*

The single reservoir modeling of the area includes three macroscopic routes representing the three sets of trips introduced in the previous section. The trip lengths corresponding to these routes are mentioned in Table 6.1, and the related demand profiles are given in Figure 6.6. The solver used is the accumulation-based model with by default the endogenous inflow merging model at entry and the maximum outflow demand at exit. But because traffic states remain undersaturated, the entry merge and exit diverge models do not matter here.

The resulting total accumulation and mean speed evolution from the MFD simulation is then compared against the 3-day real data. The results are shown in Figure 6.8. The MFD simulation is found to reproduce the shape of the morning peak quite well, as seen in accumulation or mean speed evolution. The original real dataset has been actually adjusted with a correction factor of 0.8 on accumulation to improve the original fit of the morning peak by MFD simulation. Other factors like the trip lengths or the demand levels could have been modified instead, however the corresponding data sources for these factors are considered more reliable than the ones of accumulation/production values, see our discussion in the previous section. Consequently, we assumed that the origin of mismatch between the original data and MFD simulation comes from the estimation bias in accumulation and production, as introduced with equation 6.14. On the other hand, the MFD simulation does not reproduce the shape of the evening peak very well for both accumulation and mean speed evolution. This is probably due to the shape of the demand profiles in Figure 6.6. In undersaturated conditions, the evolution of accumulation follows more or less the shape of the demand profile. As presented in chapter 2, accumulation-based results actually consist of a sequence of exponential functions following the piecewise constant demand evolution. Therefore, when comparing the demand profile in Figure 6.6 and the real accumulation evolution in Figure 6.8(a), we can observe similar shapes for the morning peak but different ones

for the evening peak (the one of the real data is notably wider than the one of the demand profile).

This simulation of a single reservoir has allowed us to identify the first shortcomings in the comparison with real data. For the next simulation test with 5 reservoirs, we will thus apply the same correction factor of 0.8 on accumulation and production data. We can also expect a better match between morning peaks than evening peaks.

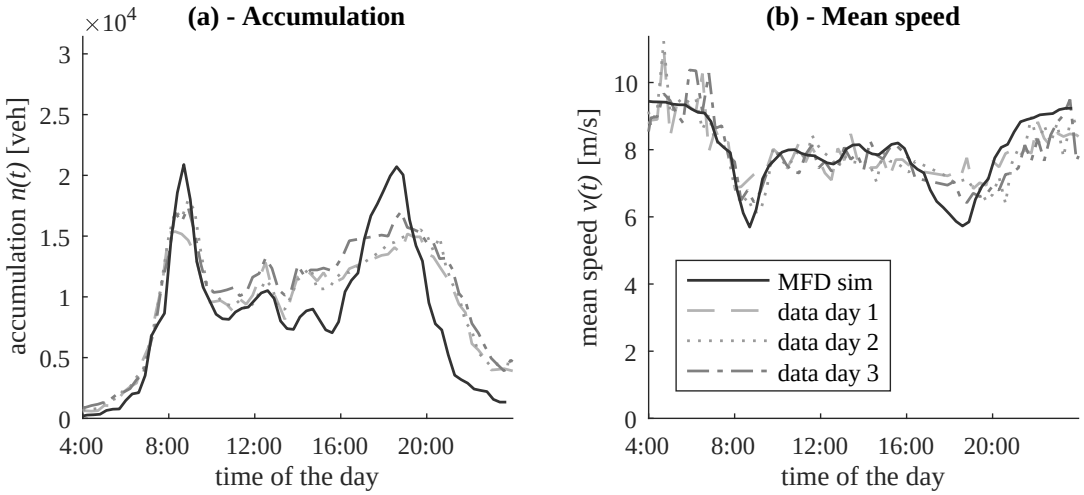



FIG. 6.8 – *Comparison between real data and MFD single reservoir simulation. (a) Evolution of total accumulation and (b) mean speed*

### 6.2.3 Example of a simulation with five reservoirs

In this second test, the studied area is split into 5 reservoirs. The network is clustered by aggregating several adjacent IRIS areas together to keep the demand definition simple (the original OD matrix is estimated at the IRIS area level). The area aggregation is done manually in following a set of practical rules: get a well-defined MFD in each reservoir, get simple and compact reservoir shapes, follow the natural geographical frontiers to define the reservoir borders (e.g. along rivers), do not split two reservoirs along a major road axis (because of the hypothesis of homogeneous traffic states in MFD simulation, we prefer having a concentration of traffic in the middle of a reservoir than at its border in the real network). Our partitioning is presented in Figure 6.9. Reservoir 1 has been first delimited by the two rivers crossing Lyon, the Saône and the Rhône, and then the West and East areas have been split into 2 reservoirs each so that each new reservoir can exhibit a well-defined MFD.

The production-MFD of each reservoir is shown in Figure 6.10. The data points are calculated via the formulas in equation 6.14 and equation 6.15 applied to the 3 days of available data. Likewise the single reservoir modeling, a parabolic fit model is found suitable to match the data in all the reservoirs. Similar observations are also made: undersaturated states only, no hysteresis nor bifurcation. All the MFD are well-defined except the one of reservoir 5 which is quite scattered. This may be explained by the few available data in this reservoir, this part of the city being actually a residential area with few loop detectors. These MFD have been also modified by a correction factor of 0.8 applied on accumulation and production values, as detailed in the previous section.


FIG. 6.9 – *Road network of Lyon-Villeurbanne clustered in 5 reservoirs*

Figure 6.11 presents the histogram of trip lengths in each reservoir for all the routes crossing each reservoir. These trip lengths are based on a set of link paths on the real network, determined with shortest path calculations in distance (no effect of traffic is considered). The resulting trip distances were then aggregated at the reservoir level by using the fourth aggregation method presented in section 6.1.2: for a given macroscopic route $p$, the length $L_p^r$ of the portion of route $p$ crossing reservoir $r$ is calculated as the mean length of all link paths crossing reservoir $r$ and following the sequence of reservoirs defining route $p$. The taxi data was then used to adjust these trip lengths so that the average trip length in each reservoir matches the average distances given by the taxi data.

The OD matrix is calculated by aggregating the demand levels of all the IRIS areas included in each reservoir. The demand from and to outside the perimeter is equally distributed among the four main directions described earlier: Paris, Geneva, Grenoble and Marseille. Each of these directions is represented by a reservoir linked to the area perimeter. Figure 6.12 depicts the general configuration of the five original reservoirs with the added four reservoirs, each of them representing an external origin or destination. These new reservoirs are aimed at sending flow to or receiving flow from the area studied, they are supposed to always remain in quasi free-flow state to have a minimal influence on the simulation. Thus, they are characterized by an artificial parabolic MFD with very high jam accumulation and maximum production. Their trip length is set to 1 km by default, a smaller value could have been chosen but this would imply a small time step for the accumulation-based solver. At a free-flow speed of 15 m/s, such a default trip length introduces a delay of 1 min in the simulation, negligible for a full day time frame. New macroscopic routes are also generated to include the external reservoirs. These reservoirs are simply added to the

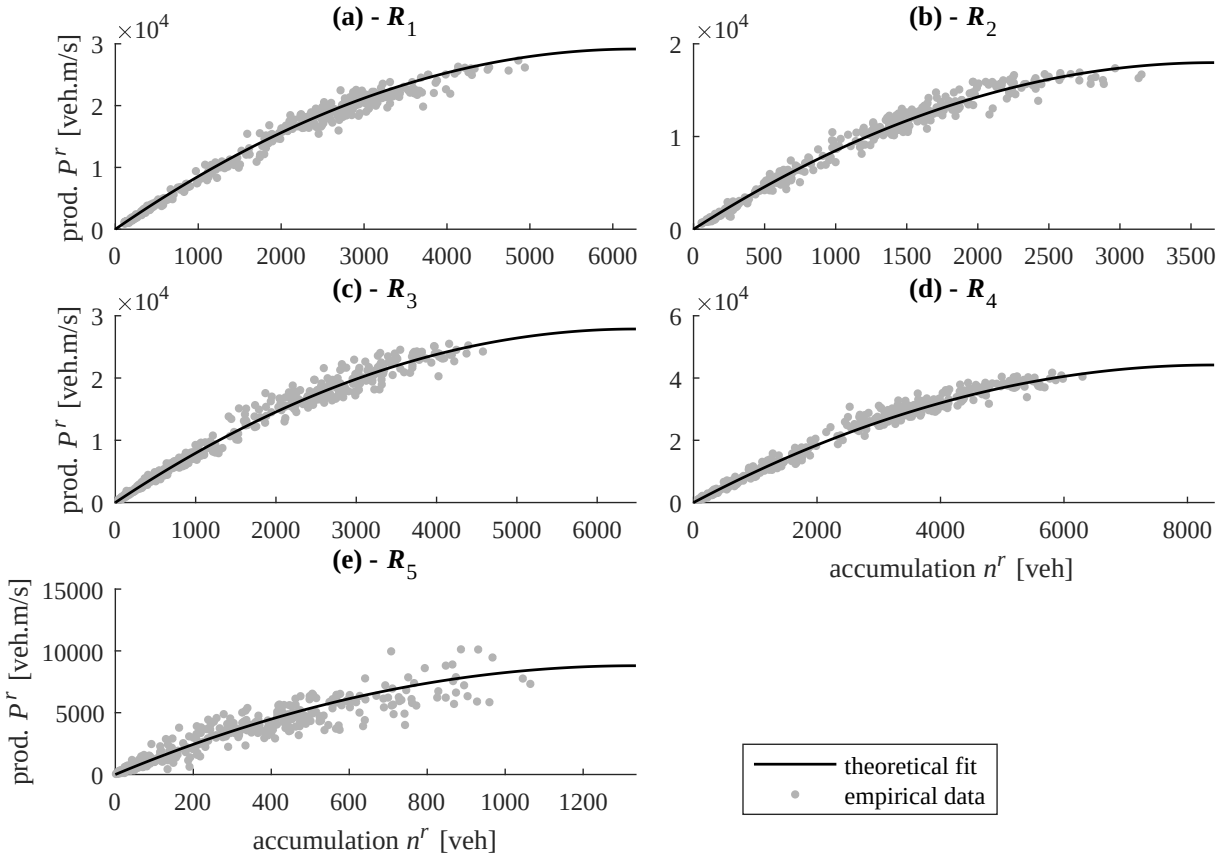



FIG. 6.10 – *Production-MFD $P^r(n^r)$ of each reservoir r. (a) Reservoir $R_1$, (b) $R_2$, (c) $R_3$, (d) $R_4$ and (e) $R_5$*

existing route paths to create new ones, e.g. the route [2 1 3] can generate the new routes [6 2 1 3] and [2 1 3 7]. Like the single reservoir modeling, the trips crossing the perimeter are not considered in this study, e.g. the route [6 2 1 3 7] is not included.

Unlike the single reservoir simulation, this test case requires additional information to understand how the demand is distributed among the possible routes for each reservoir OD pair. This information is critical for reproducing reliable traffic states within the reservoirs, but is unfortunately hardly accessible. In a first test, we assume that users choose their route according to Wardrop's first principle, i.e. that they travel on the route with the shortest travel time to get to their destination (UE conditions). Because traffic states are undersaturated and because the system of reservoirs is quite simple here, UE conditions can be achieved by distributing demand flows on the routes with the shortest free-flow travel time of each OD pair. This first path flow distribution has then been manually adjusted to improve the accuracy of the accumulation prediction by the MFD simulation compared to real data. In total, 136 routes are used in the MFD simulation, represented with arrows in Figure 6.13(a). The thickness of each arrow indicates the assigned demand mean value on the corresponding route (the thicker, the higher the assigned demand). High flow transfers are observed between reservoirs 1, 2, 3 and 4, while lower route flows cross reservoir 5. This is consistent with the lower production estimated in the MFD of reservoir 5, and the fact that this reservoir corresponds to a residential area. Figure 6.13(b) shows the route cumulative demand in each reservoir, i.e. the sum of the demand mean values of the routes crossing

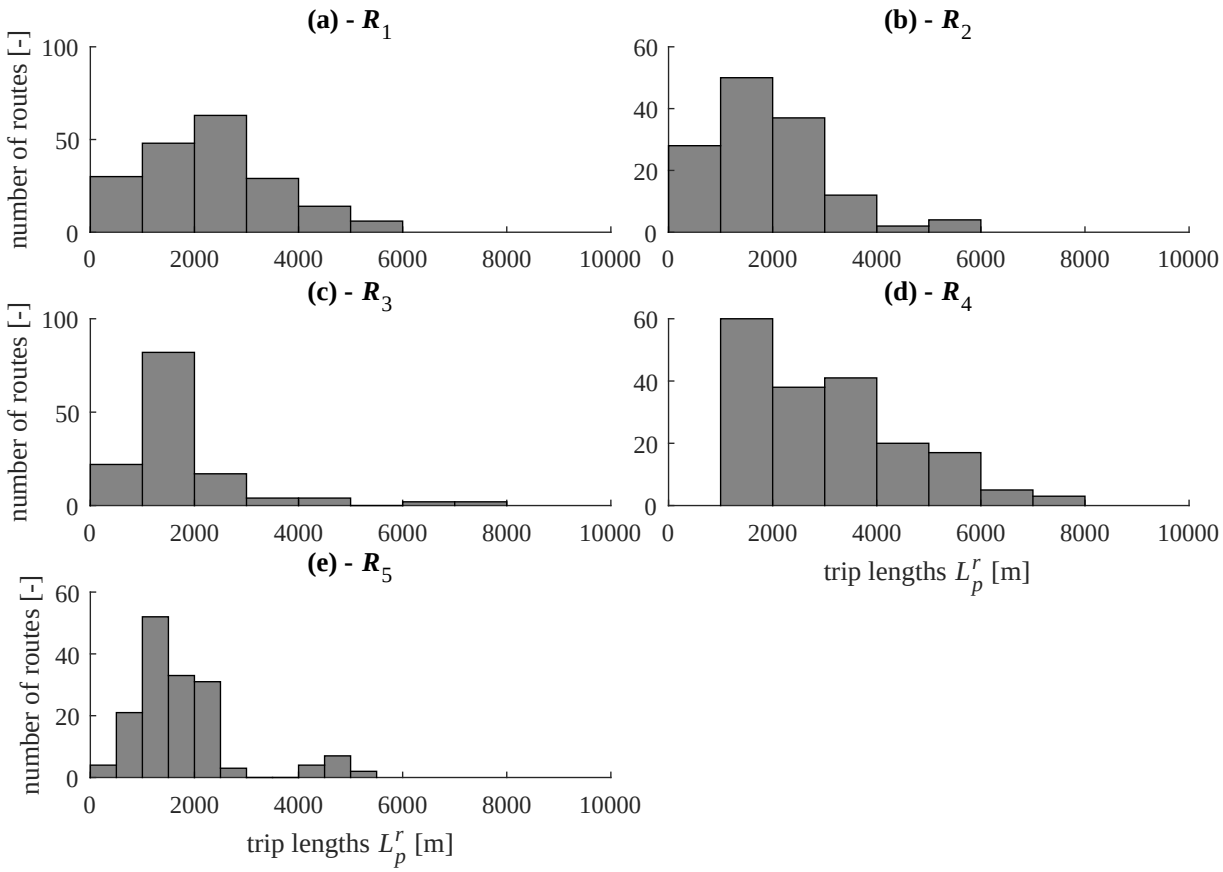



FIG. 6.11 – *Trip lengths $L_p^r$ in each reservoir r for each route p crossing r. (a) Reservoir $R_1$, (b) $R_2$, (c) $R_3$, (d) $R_4$ and (e) $R_5$*

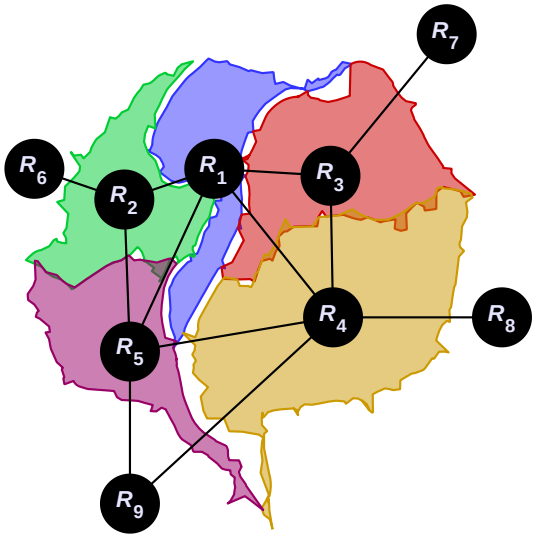



FIG. 6.12 – *Reservoir configuration with external origins/destinations*

each reservoir. Similarly, we notice heavier traffic load on reservoir 1, 2, 3 and 4, and lighter traffic load on reservoir 5. Reservoir 4 which has the widest area is the most loaded one.

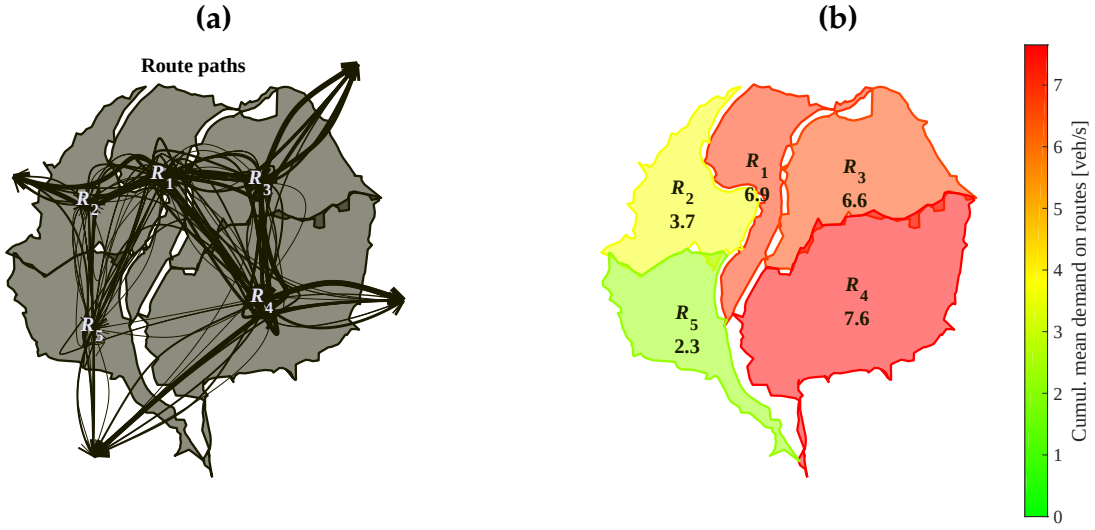



FIG. 6.13 – *Routes configuration for the 5-reservoir simulation. (a) Route paths where their line thickness represents their assigned demand mean value (b) route cumulative demand in each reservoir (in a given reservoir, this corresponds to the sum of the demand values of the routes crossing this reservoir)*

Finally, the comparison between MFD simulation and real data is presented in Figures 6.14 and 6.15, where the evolution of accumulation (respectively mean speed) is plotted for each reservoir. The MFD simulation is run with the accumulation-based solver. Like in the single reservoir simulation, we use by default the endogenous inflow merging model at entry and the maximum outflow demand at exit, although these modules are not involved in undersaturated conditions. While providing a reliable estimation of the accumulation level in reservoir 1, the MFD simulation under-estimates the number of circulating vehicles in other reservoirs 2, 3 and 4, with a relative error of 20-30%, see Figures 6.14(a)-(d). As similarly observed in the single reservoir simulation, in reservoir 1 the morning peak is better reproduced, compared to the evening peak. Because of the scatter in reservoir 5 between the different day datasets, it is hard to evaluate the accuracy of the simulation results in this reservoir, see Figure 6.14(e). Nevertheless, the overall accumulation trend and mean value are consistent with the data. As for the evolution of mean speed in each reservoir, similar observations are made as shown in Figure 6.15. The mean speed in reservoir 1 is well estimated, but the simulation is too optimistic for the other reservoirs, as it over-estimates their mean speed evolution.

In this 5-reservoir simulation test, the causes of the discrepancies between simulation and real data are likely to be multiple. The general under-estimation of accumulation levels may be due to either an under-estimation of the trip lengths in reservoirs 2, 3 and 4, or an under-estimation of the demand crossing these reservoirs. Both reasons are possible, because both the estimation of the trip lengths and the demand may contain some bias. The trip lengths come from shortest path calculations on an empty network (only the distances are considered, not the possible reduction of travel time due to traffic) corrected by taxi trip data. This calculation method is not necessarily representative of the real distances traveled by all vehicles. In the demand profiles, the part of traffic corresponding to trips crossing the area has been removed from our study, as explained in section 6.2.2. This was justified because we assume that these trips are mostly located on the ring road, and are

thus not captured by both sources of data (the loop data and the taxi data). Actually, a small portion of them could take the city streets, which would correspond to the fraction of the accumulation we are missing. The routing of users could be also improved, but this would not solve the problem of accumulation under-estimation, which is general (3 reservoirs out of 5). In any case, additional data is required (either real or simulated in the real network) to better understand the sources of errors in this case study.

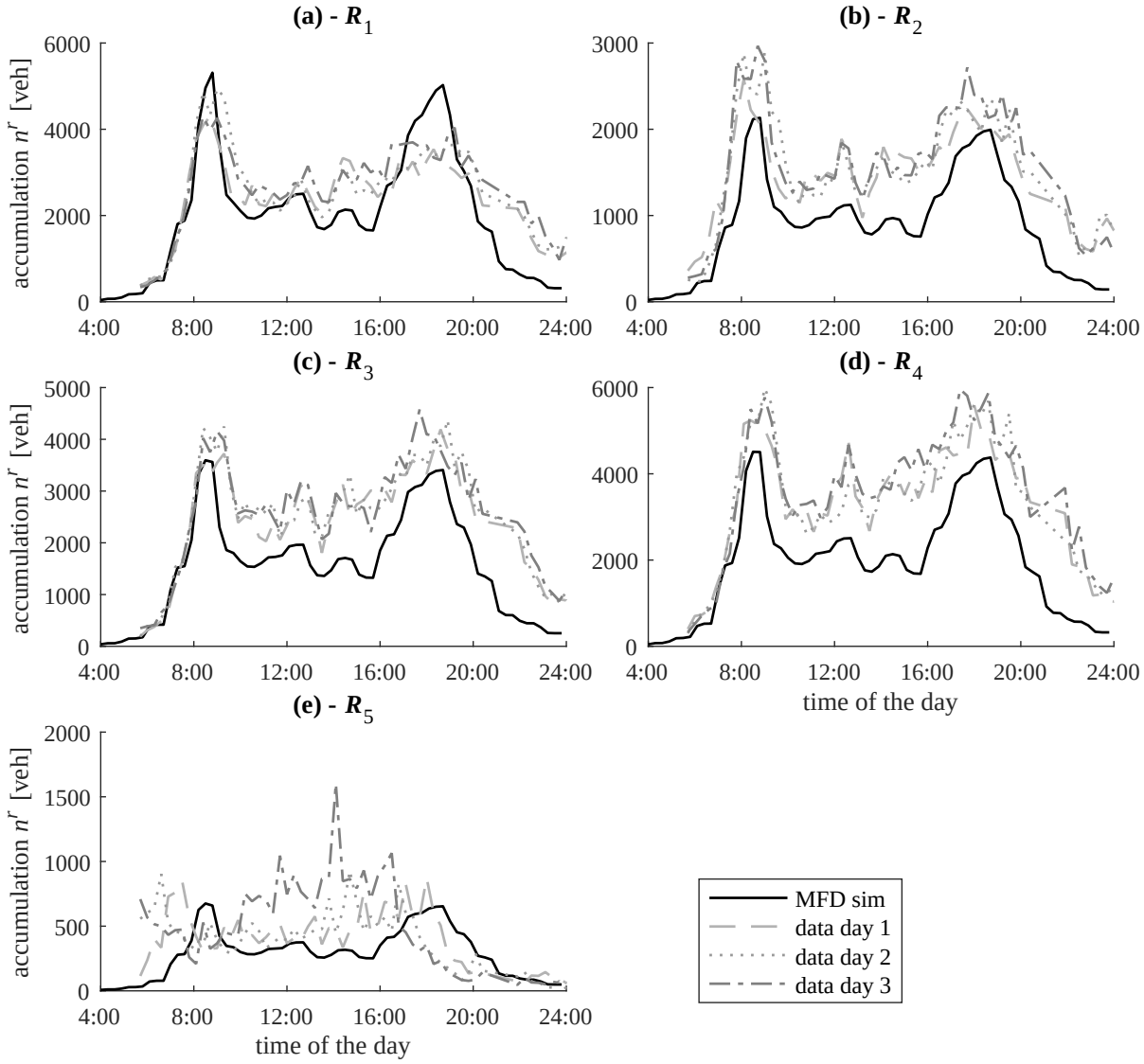



FIG. 6.14 – *Comparison between real data and MFD 5-reservoir simulation. (a) Evolution of accumulation in reservoir $R_1$, (b) $R_2$, (c) $R_3$, (d) $R_4$ and (e) $R_5$*

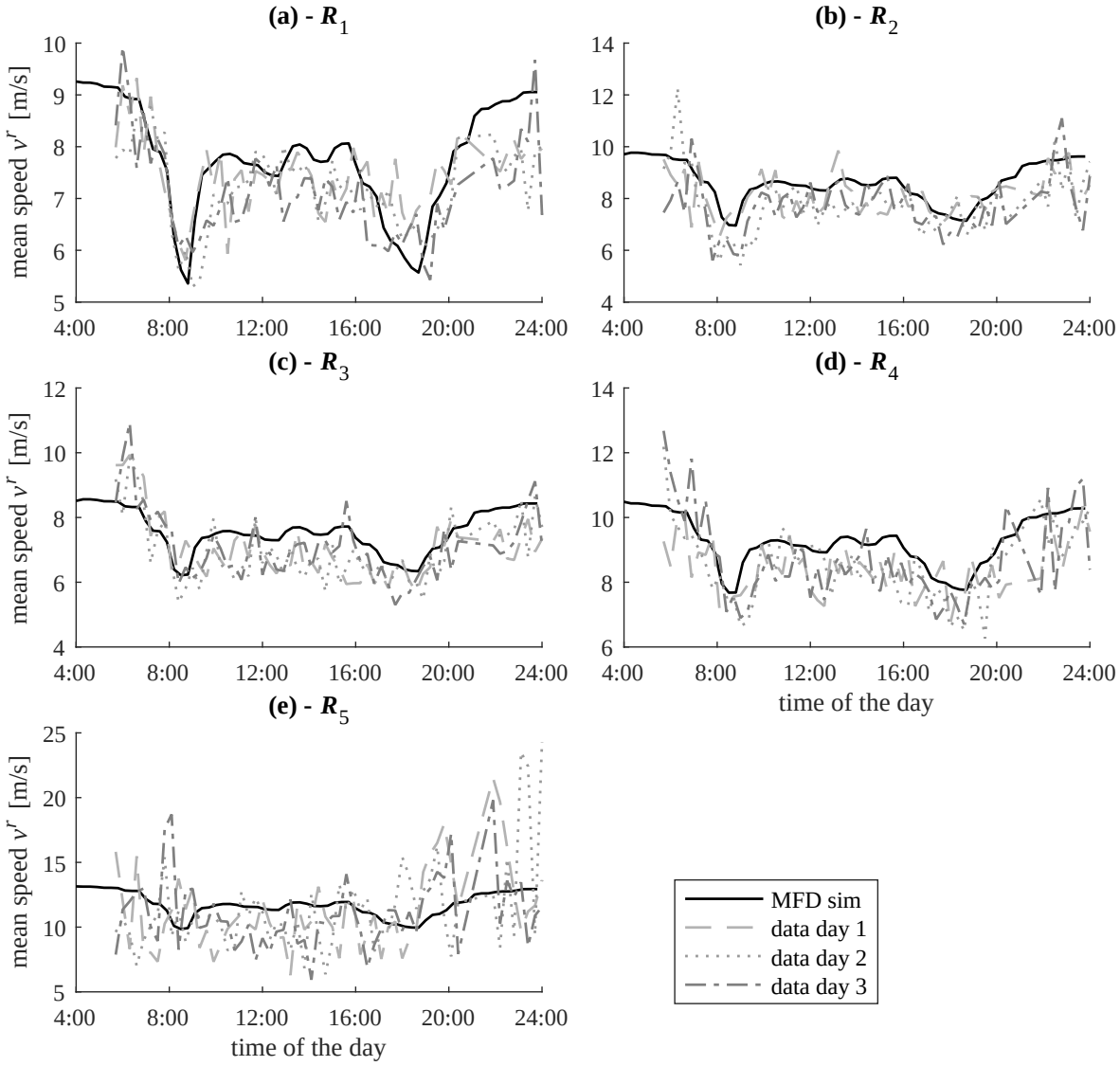



FIG. 6.15 – *Comparison between real data and MFD 5-reservoir simulation. (a) Evolution of mean speed in reservoir $R_1$, (b) $R_2$, (c) $R_3$, (d) $R_4$ and (e) $R_5$*

# Conclusion of part III

In this part, we introduced the multi-trip and multi-reservoir simulation platform developed during this PhD. A complete overview was given about the different assumptions, implementation issues, modeling approaches and options, and the data required to calibrate a simulation. The purpose of this part was to present the practical implementation of our findings about the investigation of flow exchanges in a single reservoir, as detailed in part II. To this end, both the accumulation-based and trip-based solver algorithms were described with all the required simulation variables. The idea of the platform presentation was also to give some insights about the pre-processing of input data. This notably illustrated the collaboration with PhD student S. Batista, who worked on the trip length estimation and the assignment module to calculate path flow distributions on macroscopic routes.

The application of the simulation platform was then tested on a real case, the city of Lyon (France). This consisted in reproducing traffic states for a typical weekday with a 5-reservoir simulation. This application case was limited regarding traffic flow modeling because each reservoir was undersaturated, but it allowed to highlight several issues encountered during the simulation calibration. First, an unknown bias was reported in the estimation of production and accumulation, because both measures come from the same source of loop detector data. The bias is due to the fact that these detectors do not cover all the links in the network. Hence, both variables were estimated as if traffic states were homogeneous in a given reservoir, which might not be always true. We tried to measure this bias with a single reservoir simulation test in assuming that the other parameters were correctly defined. Second, our method to calculate trip lengths in the reservoirs (mean values of shortest distances in the real network for a given reservoir OD pair) was limited because it was probably not representative of users' mobility patterns. Nevertheless, we corrected our estimations thanks to taxi data trajectories. Third, the demand profile we obtained from a preliminary study made by other members of the project might not perfectly correspond to the traffic states we observed in our dataset. And finally, we had no clue on the path flow distributions between many routing possibilities among the 5 reservoirs. The method of trial-and-error was used to get the best fit between the MFD simulation and the real data. Although some of these bias were more or less identified, the main difficulty in this application case was to arbitrate between the different sources of error and try to assess their reliability. At the end, the predictions given by the MFD simulation were quite promising though, knowing the issues faced during the calibration process. Unfortunately, this case lacks oversaturated states to confront the aggregated modeling to real data.

The contributions of this part can be summarized as follows:

- ✓ Detailed presentation of the multi-reservoir platform including the different elements of research from the two previous parts
- ✓ First application of this simulation platform on a real city case. While critical calibration issues were identified, the final results are promising but need to be refined with additional information

# GENERAL CONCLUSION

## Summary and global overview

Modeling the dynamic evolution of traffic states on large-scale urban networks was the purpose of this PhD thesis. Our research has been entirely focused on MFD-based models, for which a rich and prolific literature has been produced over the past decade, as shown in the introduction chapter and the beginning of each following chapter of this manuscript. The MFD itself also received empirical support with data from various cities all around the world, which makes it a promising tool for designing such models. However, despite the growing attention paid to MFD-based modeling, the literature review of our introduction chapter unveiled several theoretical and methodological issues in the way multi-reservoir models are generally implemented. The reason for this situation lies in the fact that most of the time, the use of MFD-based models is included in larger studies dealing with global traffic management methods (e.g., perimeter control, route guidance, search-for-parking). These methods are often promising and take advantage of a much simpler and computationally more efficient formulation of traffic flow problems, but do not aim at providing accurate predictions over a long-term horizon. Only a few studies focused on MFD simulation for traffic state prediction. Moreover, as the development of multi-reservoir models has been generally subordinated to specific applications (like the examples listed above), the domain of use of these models is often limited to what is required for these applications only. In studies about traffic control for instance, the consistency of the model is rarely investigated in highly congested situations because the control module does not let the system go into these situations.

In this context, the objective of our work was two-fold: (i) first, analyze and investigate the properties of the latest and most advanced multi-reservoir models, and (ii) second, extend them to any case of application in proposing new modeling approaches to fill the gap in the literature. For the first objective, our main contribution was the identification and the quantification of theoretical and practical sources of errors introduced by the aggregated representation of traffic states in MFD-based modeling. Although discussed all along the whole manuscript, these questions were mainly in the topics of chapters 2 and 3 (part I). For the second objective, our main contribution was the thorough examination of the dynamic traffic state evolution of a single reservoir in oversaturated regimes. This supported the development of a consistent model of congestion propagation between several reservoirs. The theoretical building of this framework and its confrontation with microscopic simulation

were thus the main questions discussed in chapters 4 and 5 (part II). Parallel to these two objectives, the storyline of this manuscript can be seen as the introduction of gradually refined and complex modeling approaches, from a single reservoir with a unique trip category in chapters 2 and 3, to a single reservoir with multiple trip categories in chapters 4 and 5, and finally the connection between several reservoirs with multiple macroscopic routes in chapter 6.

The evolution of the modeling approaches, from chapter 2 to chapter 5, also aimed at questioning the role of a critical component of MFD simulation: the average trip length. This element is essential as it is the only spatial variable in MFD-based models, and has thus a direct impact on the definition of outflow and travel time. This space variable is hidden in many studies when their approach is grounded on the outflow-MFD and not the production-MFD. Therefore, another important contribution of the present work was the discussion on this component all along the thesis. In that sense, the second part reflected the questions addressed in the first one: the introduction of multiple trip lengths inside a single reservoir not only allowed the proper definition of macroscopic routes for the following development of the multi-reservoir platform, but also permitted to relax the assumption of a unique average trip length that may bring out inaccuracies as notably shown in chapter 3. Interestingly enough, the framework of multiple trip categories moreover proved to be a viable option to reintroduce heterogeneity, at least partially, in the dynamics of the single reservoir model. As described in chapter 5, the MFD model with several trip categories can capture the heterogeneity observed at the network level in microsimulation and due to different boundary conditions applied at two network borders.

Finally, the practical output of our work was the development of the multi-reservoir simulation platform presented in chapter 6. This platform contains one of our notable contributions to the field of MFD simulation, namely the implementation of the trip-based approach for multi-reservoir systems, operating in both under- and oversaturated conditions. From a wider research perspective, the development of this platform also illustrated the different collaborations with other members involved in the MAGnUM project.

## Contributions to our initial research questions

Regarding the list of the research questions that was presented at the beginning of this manuscript (see section 1.3.1), our main contributions are listed below in the same order:

- The refinement of the spatial propagation of information in the single reservoir model was studied with the introduction of the trip-based model in chapter 2. This alternative modeling approach based on the speed-MFD accounts for a delay in the propagation of flow, from downstream to upstream, which results in a better representation of outflow and travel time.

- Concerning the slow-varying hypothesis in the accumulation-based model, it was shown in chapter 2 that this is indeed a limitation for this framework. With simple simulation studies, underestimation of accumulation peak and drop in travel time were the two artifacts found during fast-varying demand scenarios. These artifact thus allowed us to define a validity domain for the accumulation-based model, depending on the amount of error on travel time the modeler is ready to accept.

- About the estimation of the MFD, chapter 3 showed that heterogeneities at the border of a network can trigger very different traffic patterns inside the network that impact both the production-MFD and the average trip length estimation. A direct consequence of this is that recalibration would be needed when boundary conditions drastically change, to avoid significant bias in steady traffic state predictions by the MFD model.

- The assumption of an average trip length for all the travelers inside a single reservoir was first discussed in chapter 3, and the relaxation of this hypothesis together with the implication of multiple trip length modeling on traffic dynamics were studied in chapter 4. In chapter 3, the importance of recalibrating the average trip length was demonstrated on a simple network. In chapter 4, we showed that different categories of macroscopic trips with different trip lengths must still share the same mean speed according to the single reservoir dynamics, and that this assumption implies interdependency between all the outflows of these categories.

- The question of congestion propagation was first introduced in chapter 2, and then further extended to the multiple trip framework in chapters 4 and 5. As mentioned earlier, having multiple trip categories in a reservoir proved to be an efficient option to reproduce heterogeneities at the reservoir perimeter, because spillbacks can be distinguished between different flow directions. However, the shared speed (or production) among all travelers remain a strong assumption. As shown with simulation examples in chapter 5, this makes sense for highly correlated flow directions, but cannot capture heterogeneities inside the network due to local bottlenecks, see chapters 3 and 5.

- The design and calibration of the entry supply function was investigated in chapter 5. It appeared that this function is not unique and should depend on each specific demand pattern for the reservoir considered. Nevertheless, the equilibrium point reached in the MFD plane was found quite unique, so that the exact shape of this function only matters for the transient loading of the reservoir. It was shown that even a rough approximation of its shape provide reliable results in steady state.

- Following the design of the entry supply function, two merging schemes were also studied in chapter 5 to split incoming flows that are exceeding this total entry supply in total. The first merging scheme is a demand pro-rata rule, widely used in the literature, while the second one consists in an endogenous formulation including the ratio of partial accumulations in the reservoir receiving the flows. Comparisons with microsimulation showed that the demand pro-rata rule is the most accurate one.

- At the network exit, a major discussion was started in chapter 2, continued in chapter 4 and finally verified on the comparison with microsimulation in chapter 5. Unlike most studies from the literature, we postulated that the demand for outflow of travelers heading in a neighboring reservoir must be maximum in oversaturation. With this new formulation, we distinguished transfer trips from internal trips (the ones responsible for the decrease of outflow demand in our approach). The consistency of our formulation was illustrated with theoretical simulation examples first, in chapters 2 and 4, and was validated with a comparison against microsimulation then, in chapter 5.

- Comparing the different MFD-based approaches with link-level simulation or even real data was one of the key points in our methodology. Investigations on a very small network were presented in chapter 3, then on a larger one with more complex settings in chapter 5. Despite their own specificity, each of these test cases proved to be useful to observe how accurate MFD-based simulation might be compared to the aggregation of link-level phenomena. An application to a real case was also presented in chapter 6. While limited to undersaturated states, this case allowed to illustrate different issues in the calibration of a multi-reservoir simulation of a city.

## Research perspectives

Different perspectives can be envisioned for the follow-up of this work. We present some of them in this last section:

- First of all, other comparison studies should definitively be carried on to further validate and investigate the different modeling approaches we introduced. Indeed, although we tried to keep them quite general, one major issue with the tests we conducted is that we designed them to create specific congestion patterns. This choice was deliberated and motivated by the fact that we wanted to observe clear differences between two MFD-based approaches for instance. Therefore, to ensure that our results are not too case-specific, other configurations and patterns should be investigated as well. With an artificial grid network, a nice extension of our study in chapter 5 would be to connect explicitly the grid network, represented as a reservoir, to other grid networks representing neighboring reservoirs. Various congestion patterns could be created in distributing trips among these reservoirs. A similar study might be also done on a real network, with synthetic data so as to control the input data and assess the bias that come directly from the aggregated representation of the MFD simulation.

- Following the extension of the single reservoir model to multiple trip categories, a new framework worth investigating could be the hybrid modeling of border nodes for multi-reservoir systems. The idea behind this approach is to discretize the border between two adjacent reservoirs into several connection nodes. These nodes can either correspond to physical link connections (like a bridge or a major arterial), or represent an aggregation of different link connections. The flows in a reservoir are thus distinguished by their entry and exit connection nodes. This idea is extending the concept of macroscopic route, before defined as a succession of reservoirs, and now defined as a succession of connection nodes. This means that two routes may have an identical succession of reservoirs but differ through the connection nodes they used at the reservoir borders. Regarding flow dynamics, this model is identical to the multi-trip representation, as a single reservoir is still crossed by several routes. This framework would allow a more refined representation of the trip lengths inside a reservoir, and the proper treatment of local boundary flow restrictions. This could be useful for the simulation of cities with critical links such as bridges.

- Another hybrid modeling that can be envisioned is the coupling between a multi-reservoir system representing secondary city streets and a microscopic simulation representing major arterials and ring roads. The question of traffic on the ring road was

slightly addressed in the application example of chapter 6, where it appeared clearly that flows on the ring road might be distinguished from flows on city streets (in terms of trip lengths, mean speed, main origin and destination, etc). Such an hybrid modeling is likely to be a promising approach for large-scale cities, however some critical points must be clarified. First, the treatment of flow interfaces between the reservoirs and the major road portions: are all connection nodes being kept? how to eventually aggregate them? Here, the new framework presented previously with connection nodes could be very useful to provide a first answer to these questions. The choice between accumulation-based or trip-based modeling may also depend on the available microsimulator, the use of a CTM would be convenient with an accumulation-based model, while the use of a car-following law would be more suitable for a trip-based model.

- The simulation platform we developed can also be involved in other applications. The problem of cruising-for-parking is one of them, for which a first study has been already carried on. This study was the output of a collaboration with master student Alméria Sénécat, and was reported in a publication authored by Prof. Ludovic Leclercq. Our methodology took advantage of the event-based representation of the trip-based model to assign an extra distance to travel for users that arrived at destination and start looking for a free parking spot (defined as a new event in the resolution scheme). The additional trip length is calculated with a Bernoulli trial on all parking spots, where the probability of success corresponds to the parking occupancy ratio (number of available places over the total number of places). This ratio is thus updated every time a vehicle parks after traveling its previously calculated search distance. This extension of the trip-based, inspired by the work of Cao & Menendez (2015), constitutes an interesting tool to study interactions between traffic and search-for-parking behavior with few modeling and computational efforts.

- Another application is the case of perimeter control and route guidance. Although this topic has been deeply studied in the literature, here the novelty of our approach would lie in the integration of users' behavior and reaction to the control system. This problem was already investigated during a collaboration with post-doc Deepak Ingole for a simple test case of a single reservoir and a freeway. The principle is to implement perimeter control on the single reservoir with several crossing routes, and to give an alternative option for one of the routes: a freeway portion portion with a free-flow travel time a bit higher than the reservoir travel time on the corresponding route in undersaturated conditions. UE discipline is assumed for the behavior of users, so that for this route, all of them choose to travel inside the reservoir (city streets) in normal conditions. But once the perimeter control applies because of some congestion generated in the reservoir, users will wait for entering the reservoir and eventually choose the freeway option to get to their destination. In this framework, UE is ensured by estimating predictive travel times on the cumulative count curves of the reservoir. Such settings allow then to study the overall performance of the system (in terms of total time spent, emissions produced, etc).

# References


**Aboudolas, K. & Geroliminis, N.** (2013). Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transportation Research Part B: Methodological*, 55:265–281.

**Aghamohammadi, R. & Laval, J. A.** (2018). Dynamic traffic assignment using the macroscopic fundamental diagram: a review. In *Transportation Research Board 97th Annual Meeting*, 18-06275. Washington DC.

**Altman, N. S.** (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185.

**Ambuhl, L., Loder, A., Menendez, M. & Axhausen, K. W.** (2017). Empirical macroscopic fundamental diagrams: New insights from loop detector and floating car data. In *Transportation Research Board 96th Annual Meeting*, 17-03331, page 16p. Washington DC.

**Ambuhl, L., Loder, A., Menendez, M. & Axhausen, K. W.** (2018). A case study of zurich's two-layered perimeter control. In *TRA 2018 7th Transport Research Arena*. Vienna, Austria.

**Ambuhl, L. & Menendez, M.** (2016). Data fusion algorithm for macroscopic fundamental diagram estimation. *Transportation Research Part C: Emerging Technologies*, 71:184–197.

**Ameli, M., Lebacque, J.-P. & Leclercq, L.** (2018). Multi-modal multi-class traffic network equilibrium model: a new heuristic optimization approach. *submitted to Transportmetrica B: Transport Dynamics*, in prep.

**Ampountolas, K., Zheng, N. & Geroliminis, N.** (2014). Robust control of bi-modal multi-region urban networks: An lmi optimisation approach. In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pages 489–494.

**Ampountolas, K., Zheng, N. & Geroliminis, N.** (2017). Macroscopic modelling and robust control of bi-modal multi-region urban road networks. *Transportation Research Part B: Methodological*, 104:616–637.

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

**Barcelo, J., Casas, J., García, D. & Perarnau, J.** (2006). A hybrid simulation framework for advanced transportation analysis. *{IFAC} Proceedings Volumes*, 39(12):497–502. 11th {IFAC} Symposium on Control in Transportation Systems.

**Batista, S. F. A. & Leclercq, L.** (2018a). A dynamic traffic assignment framework for mfd multi-regional models. *submitted to Transportation Science*, in prep.

**Batista, S. F. A. & Leclercq, L.** (2018b). Introduction of multi-regional mfd-based models with route choices: the definition of regional paths. In *PLURIS 2018 - 8th Luso-Brazilian*

*Congress for Urban, Regional, Integrated and Sustainable Planning*. Coimbra, Portugal.

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

**Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N. & Mishalani, R.** (2002). *Real Time Simulation of Traffic Demand-Supply Interactions within DynaMIT*, vol. 63, chap. 2, pages 19–36. Springer US, Boston, MA.

**Ben-Akiva, M. E., Gao, S., Wei, Z. & Wen, Y.** (2012). A dynamic traffic assignment model for highly congested urban networks. *Transportation Research Part C: Emerging Technologies*, 24:62–82.

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

**Burghout, W., Koutsopoulos, N., Haris & Andreasson, I.** (2005). Hybrid mesoscopic–microscopic traffic simulation. *Transportation Research Record: Journal of the Transportation Research Board*, 1934:218–225.

**Cao, J. & Menendez, M.** (2015). System dynamics of urban traffic based on its parking-related-states. *Transportation Research Part B: Methodological*, 81(3):718–736.

**Cassidy, M. J., Jang, K. & Daganzo, C. F.** (2011). Macroscopic fundamental diagrams for freeway networks: Theory and observation. *Transportation Research Record: Journal of the Transportation Research Board*, 2260:8–15.

**CEBR** (2014). The future economic and environmental costs of gridlock in 2030. Tech. rep., Centre for Economics and Business Research (CEBR) and INRIX.

**Chevallier, E. & Leclercq, L.** (2009). Do microscopic merging models reproduce the observed priority sharing ratio in congestion? *Transportation Research Part C: Emerging Technologies*, 17(3):328–336.

**Chiabaut, N.** (2015). Evaluation of a multimodal urban arterial: The passenger macroscopic fundamental diagram. *Transportation Research Part B: Methodological*, 81(2):410–420.

**Daganzo, C. F.** (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4):269–287.

**Daganzo, C. F.** (1995). The cell transmission model, part ii: Network traffic. *Transportation Research Part B: Methodological*, 29(2):79–93.

**Daganzo, C. F.** (2007). Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1):49–62.

**Daganzo, C. F.** (2011). On the macroscopic stability of freeway traffic. *Transportation Research Part B: Methodological*, 45(5):782–788.

**Daganzo, C. F., Gayah, V. V. & Gonzales, E. J.** (2011). Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B: Methodological*, 45(1):278–288.

**Daganzo, C. F. & Geroliminis, N.** (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, 42(9):771–781.

**Daganzo, C. F. & Lehe, L. J.** (2015). Distance-dependent congestion pricing for downtown zones. *Transportation Research Part B: Methodological*, 75:89–99.

**de Jong, D., Knoop, V. & Hoogendoorn, S.** (2013). The effect of signal settings on the

macroscopic fundamental diagram and its applicability in traffic signal driven perimeter control strategies. In *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, pages 1010–1015.

**Dijkstra, E. W.** (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271.

**Ding, H., Guo, F., Zheng, X. & Zhang, W.** (2017). Traffic guidance-perimeter control coupled method for the congestion in a macro network. *Transportation Research Part C: Emerging Technologies*, 81:300–316.

**Doig, J. C., Gayah, V. V. & Cassidy, M. J.** (2013). Inhomogeneous flow patterns in undersaturated road networks: Implications for macroscopic fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, 2390:68–75.

**Du, Y., Zhao, C., Zhang, H., Wong, S. C. & Liao, F.** (2017). Modeling park-and-ride services in a multi-commodity discrete/continuum transport system with elastic demand. In *Transportation Research Board 96th Annual Meeting*, 17-02785. Washington DC.

**Edie, L. C.** (1963). Discussion of traffic stream measurements and definitions. In *2nd International Symposium on the Theory of Traffic Flow*, page 139–154. Paris.

**Forgy, E. W.** (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, 21(3):761–777.

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

**Fosgerau, M. & Small, K. A.** (2013). Hypercongestion in downtown metropolis. *Journal of Urban Economics*, 76:122–134.

**Gayah, V. V. & Daganzo, C. F.** (2011). Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability. *Transportation Research Part B: Methodological*, 45(4):643–655.

**Gayah, V. V., Gao, X. & Nagle, A. S.** (2014). On the impacts of locally adaptive signal control on urban network stability and the macroscopic fundamental diagram. *Transportation Research Part B: Methodological*, 70:255–268.

**Gazis, D. C., Herman, R. & Rothery, R. W.** (1961). Nonlinear follow-the-leader models of traffic flow. *Operations Research*, 9(4):545–567.

**Ge, Q. & Fukuda, D.** (2018). A macroscopic dynamic network loading model for multiple-reservoir system. *Transportation Research Part B: Methodological*, in press.

**Geroliminis, N.** (2009). Dynamics of peak hour and effect of parking for congested cities. In *Transportation Research Board 88th Annual Meeting*, 09-1685. Washington DC.

**Geroliminis, N.** (2015). Cruising-for-parking in congested cities with an mfd representation. *Economics of Transportation*, 4(3):156–165.

**Geroliminis, N. & Boyaci, B.** (2012). The effect of variability of urban systems characteristics in the network capacity. *Transportation Research Part B: Methodological*, 46(10):1607–1623.

**Geroliminis, N. & Daganzo, C. F.** (2007). Macroscopic modeling of traffic in cities. In *Transportation Research Board 86th Annual Meeting*, 07-0413. Washington DC.

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

**Geroliminis, N., Haddad, J. & Ramezani, M.** (2013). Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. *Intelligent Transportation Systems, IEEE Transactions on*, 14(1):348–359.

**Geroliminis, N. & Sun, J.** (2011a). Hysteresis phenomena of a macroscopic fundamental

diagram in freeway networks. *Transportation Research Part A: Policy and Practice*, 45(9):966–979.

**Geroliminis, N. & Sun, J.** (2011b). Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research 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 Research Part C: Emerging Technologies*, 42:168–181.

**Gipps, P.** (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, 15(2):105–111.

**Gipps, P.** (1986a). A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological*, 20(5):403–414.

**Gipps, P.** (1986b). Multsim: a model for simulating vehicular traffic on multi-lane arterial roads. *Mathematics and Computers in Simulation*, 28(4):291–295.

**Girault, J.-T., Gayah, V. V., Guler, I. & Menendez, M.** (2016). Exploratory analysis of signal coordination impacts on macroscopic fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, 2560:36–46.

**Godfrey, J. W.** (1969). The mechanism of road network. *Traffic Engineering and Control*, 7(11):323–327.

**Gonzales, E. J., Chavis, C., Li, Y. & Daganzo, C. F.** (2009). Multimodal transport modeling for nairobi, kenya: Insights and recommendations with an evidence-based model. techreport UCB-ITS-VWP-2009-5, UC Berkeley Center for Future Urban Transport.

**Haddad, J.** (2015). Robust constrained control of uncertain macroscopic fundamental diagram networks. *Transportation Research Part C: Emerging Technologies*, 59:323–339.

**Haddad, J.** (2017). Optimal perimeter control synthesis for two urban regions with aggregate boundary queue dynamics. *Transportation Research Part B: Methodological*, 96:1–25.

**Haddad, J. & Geroliminis, N.** (2012). On the stability of traffic perimeter control in two-region urban cities. *Transportation Research Part B: Methodological*, 46(9):1159–1176.

**Haddad, J. & Mirkin, B.** (2016). Adaptive perimeter traffic control of urban road networks based on mfd model with time delays. *International Journal of Robust and Nonlinear Control*, 26:1267–1285.

**Haddad, J. & Mirkin, B.** (2017). Coordinated distributed adaptive perimeter control for large-scale urban road networks. *Transportation Research Part C: Emerging Technologies*, 77:495–515.

**Haddad, J., Ramezani, M. & Geroliminis, N.** (2013). Cooperative traffic control of a mixed network with two urban regions and a freeway. *Transportation Research Part B: Methodological*, 54:17–36.

**Haddad, J. & Shraiber, A.** (2014). Robust perimeter control design for an urban region. *Transportation Research Part B: Methodological*, 68:315–332.

**Haddad, J. & Zheng, Z.** (2017). Adaptive perimeter control with state delays in two urban regions. In *Transportation Research Board 96th Annual Meeting*, 17-02018, page 18p. Washington DC.

**Hajiahmadi, M., Haddad, J., De Schutter, B. & Geroliminis, N.** (2013a). Optimal hybrid macroscopic traffic control for urban regions: Perimeter and switching signal plans controllers. In *Control Conference (ECC), 2013 European*, pages 3500–3505.

**Hajiahmadi, M., Haddad, J., De Schutter, B. & Geroliminis, N.** (2015). Optimal hybrid perimeter and switching plans control for urban traffic networks. *Control Systems Technol-*

*ogy, IEEE Transactions on*, 23(2):464–478.

**Hajiahmadi, M., Knoop, V., De Schutter, B. & Hellendoorn, H.** (2013b). Optimal dynamic route guidance: A model predictive approach using the macroscopic fundamental diagram. In *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, pages 1022–1028.

**Hart, P. E., Nilsson, N. J. & Raphael, B.** (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107.

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

**Jayakrishnan, R., Mahmassani, H. S. & Hu, T.-Y.** (1994). An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies*, 2(3):129–147.

**Ji, Y., Daamen, W., Hoogendoorn, S., Hoogendoorn-Lanser, S. & Qian, X.** (2010). Investigating the shape of the macroscopic fundamental diagram using simulation data. *Transportation Research Record: Journal of the Transportation Research Board*, 2161:40–48.

**Ji, Y. & Geroliminis, N.** (2012). On the spatial partitioning of urban transportation networks. *Transportation Research Part B: Methodological*, 46(10):1639–1656.

**Ji, Y., Luo, J. & Geroliminis, N.** (2014). Empirical observations of congestion propagation and dynamic partitioning with probe data for large scale systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2422(1):1–11.

**Jin, W. & Zhang, H.** (2004). Multicommodity kinematic wave simulation model for network traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, 1883:59–67.

**Joueiai, M., Leclercq, L., Van Lint, H. & Hoogendoorn, S. P.** (2015). Multiscale traffic flow model based on the mesoscopic lighthill–whitham and richards models. *Transportation Research Record: Journal of the Transportation Research Board*, 2491:98–106.

**Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I. & Papageorgiou, M.** (2012). Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological*, 46(10):1393–1403.

**Kim, J. & Mahmassani, H. S.** (2015). Compound gamma representation for modeling travel time variability in a traffic network. *Transportation Research Part B: Methodological*, 80:40–63.

**Kim, S., Tak, S. & Yeo, H.** (2018). Investigating transfer flow between urban networks based on a macroscopic fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(20):75–85.

**Knoop, V. L., De Jong, D. & Hoogendoorn, S. P.** (2014). Influence of road layout on network fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, 1(2421):22–30.

**Knoop, V. L. & Hoogendoorn, S. P.** (2014). Network transmission model: a dynamic traffic model at network level. In *Transportation Research Board 93rd Annual Meeting*, 14-1104. Washington DC.

**Knoop, V. L., Hoogendoorn, S. P. & Van Lint, J. W. C.** (2012). Routing strategies based on macroscopic fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, 2315(1):1–10.

**Knoop, V. L., Van Lint, H. & Hoogendoorn, S. P.** (2015). Traffic dynamics: Its impact on the macroscopic fundamental diagram. *Physica A: Statistical Mechanics and its Applications*,

438:236–250.

**Kouvelas, A., Saeedmanesh, M. & Geroliminis, N.** (2017). Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. *Transportation Research Part B: Methodological*, 96:26–45.

**Lamotte, R. & Geroliminis, N.** (2016). The morning commute in urban areas: Insights from theory and simulation. In *Transportation Research Board 95th Annual Meeting*, 16-2003. Washington DC.

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

**Lamotte, R., Murashkin, M., Kouvelas, A. & Geroliminis, N.** (2018). Dynamic modeling of trip completion rate in urban areas with mfd representations. In *Transportation Research Board 97th Annual Meeting*, 18-06192. Washington DC.

**Laval, J. A. & Castrillon, F.** (2015). Stochastic approximations for the macroscopic fundamental diagram of urban networks. *Transportation Research Part B: Methodological*, 81(3):904–916.

**Laval, J. A. & Chilukuri, B. R.** (2016). Symmetries in the kinematic wave model and a parameter-free representation of traffic flow. *Transportation Research Part B: Methodological*, 89:168–177.

**Laval, J. A. & Leclercq, L.** (2008). Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. *Transportation Research Part B: Methodological*, 42(6):511–522.

**Laval, J. A. & Leclercq, L.** (2013). The hamilton-jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological*, 52:17–30.

**Laval, J. A., Leclercq, L. & Chiabaut, N.** (2018). Minimal parameter formulations of the dynamic user equilibrium using macroscopic urban models: Freeway vs city streets revisited. *Transportation Research Part B: Methodological*, 117:676–686.

**Leclercq, L.** (2007a). Bounded acceleration close to fixed and moving bottlenecks. *Transportation Research Part B: Methodological*, 41(3):309–319.

**Leclercq, L.** (2007b). Hybrid approaches to the solutions of the "lighthill-whitham-richards" model. *Transportation Research Part B: Methodological*, 41(7):701–709.

**Leclercq, L. & Becarie, C.** (2012). Meso lighthill-whitham and richards model designed for network applications. In *Transportation Research Board 91st Annual Meeting*, 12-0387. Washington DC.

**Leclercq, L., Chiabaut, N. & Trinquier, B.** (2014). Macroscopic fundamental diagrams: A cross-comparison of estimation methods. *Transportation Research Part B: Methodological*, 62:1–12.

**Leclercq, L. & Geroliminis, N.** (2013). Estimating mfds in simple networks with route choice. *Transportation Research Part B: Methodological*, 57:468–484.

**Leclercq, L. & Laval, J. A.** (2009). A multiclass car-following rule based on the lwr model. In *Traffic and Granular Flow '07*, pages 151–160. Springer Berlin Heidelberg, Berlin, Heidelberg.

**Leclercq, L., Parzani, C., Knoop, V. L., Amourette, J. & Hoogendoorn, S. P.** (2015). Macroscopic traffic dynamics with heterogeneous route patterns. *Transportation Research Part C: Emerging Technologies*, 55:292–307.

**Leclercq, L., Sénécat, A. & Mariotte, G.** (2017). Dynamic macroscopic simulation of on-

street parking search: A trip-based approach. *Transportation Research Part B: Methodological*, 101:268–282.

**Lentzakis, A. F., Ware, S. I. & Su, R.** (2016). Region-based dynamic forecast routing for autonomous vehicles. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1464–1469.

**Lighthill, M. J. & Whitham, G. B.** (1955). On kinematic waves ii: A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345.

**Little, J. D. C.** (1961). A proof for the queuing formula. *Operations Research*, 9(3):383–387.

**Lloyd, S. P.** (1982). Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2):129–137.

**Loder, A., Ambuhl, L., Menendez, M. & Axhausen, K. W.** (2017). Empirics of multi-modal traffic networks - using the 3d macroscopic fundamental diagram. *Transportation Research Part C: Emerging Technologies*, 82:88–101.

**Loder, A., Ambuhl, L., Menendez, M. & Axhausen, K. W.** (2018). Traffic problems in towns: An empirical analysis with macroscopic fundamental diagrams from cities around the world. In *Transportation Research Board 97th Annual Meeting*, 18-01049. Washington DC.

**Lopez, C., Leclercq, L., Krishnakumari, P., Chiabaut, N. & van Lint, H.** (2017). Revealing the day-to-day regularity of urban congestion patterns with 3d speed maps. *Scientific Reports*, 7(14029):1–11.

**Lowrie, P. R.** (1982). The sydney coordinated adaptive traffic system - principles, methodology, algorithms. In *International Conference on Road Traffic Signalling*, 207, pages 67–70. London, UK.

**Lu, C.-C., Mahmassani, H. S. & Zhou, X.** (2009). Equivalent gap function-based reformulation and solution algorithm for the dynamic user equilibrium problem. *Transportation Research Part B: Methodological*, 43(3):345–364.

**Mahmassani, H., Hou, T. & Saberi, M.** (2013a). Connecting networkwide travel time reliability and the network fundamental diagram of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, 2391:80–91.

**Mahmassani, H., Williams, J. C. & Herman, R.** (1984). Investigation of network-level traffic flow relationships: Some simulation results. *Transportation Research Record: Journal of the Transportation Research Board*, 971:121–130.

**Mahmassani, H. S., Saberi, M. & Zockaie, A.** (2013b). Urban network gridlock: Theory, characteristics, and dynamics. *Transportation Research Part C: Emerging Technologies*, 36:480–497.

**Mahmassani, H. S., Williams, J. C. & Herman, R.** (1987). Performance of urban traffic networks. In *10th International Symposium on Transportation and Traffic Theory*, pages 1–20.

**Mahut, M., Florian, M. & Tremblay, N.** (2003). Space-time queues and dynamic traffic assignment: A model, algorithm and applications. In *Transportation Research Board 82nd Annual Meeting*, 03-1273. Washington DC.

**Mazloumian, A., Geroliminis, N. & Helbing, D.** (2010). The spatial variability of vehicle densities as determinant of urban network capacity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1928):4627–4647.

**Mollier, S., Delle Monache, M. L. & Canudas-de Wit, C.** (2018). A simple example of two-dimensional model for traffic: Discussion about assumptions and numerical methods. In *Transportation Research Board 97th Annual Meeting*, 18-01172. Washington DC.

**Muhlich, N., Gayah, V., Vikash & Menendez, M.** (2015). An examination of mfd hysteresis patterns for hierarchical urban street networks using micro-simulation. In *Transportation Research Board 94th Annual Meeting*, 15-1552. Washington DC.

**Nagle, A. & Gayah, V.** (2013). A method to estimate the macroscopic fundamental diagram using limited mobile probe data. In *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, pages 1987–1992.

**Newell, G. F.** (1993). A simplified theory of kinematic waves in highway traffic, part ii: Queueing at freeway bottlenecks. *Transportation Research Part B: Methodological*, 27(4):289–303.

**Newell, G. F.** (2002). A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological*, 36(3):195–205.

**Ortigosa, J., Menendez, M. & Gayah, V. V.** (2015). Analysis of network exit functions for various urban grid network configurations. *Transportation Research Record: Journal of the Transportation Research Board*, 2491:12–21.

**Ortigosa, J., Menendez, M. & Tapia, H.** (2014). Study on the number and location of measurement points for an mfd perimeter control scheme: a case study of zurich. *EURO Journal on Transportation and Logistics*, 3(3):245–266.

**Ramezani, M., Haddad, J. & Geroliminis, N.** (2015). Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. *Transportation Research Part B: Methodological*, 74:1–19.

**Richards, P. L.** (1956). Shock waves on the highway. *Operations Research*, 4(1):42–51.

**Robertson, D. I. & Bretherton, R. D.** (1991). Optimizing networks of traffic signals in real time—the scoot method. *IEEE Transactions on Vehicular Technology*, 40(1):11–15.

**Saberi, M. & Mahmassani, H.** (2012). Exploring properties of networkwide flow-density relations in a freeway network. *Transportation Research Record: Journal of the Transportation Research Board*, 2315:153–163.

**Saberi, M., Mahmassani, H. S., Hou, T. & Zockaie, A.** (2014). Estimating network fundamental diagram using three-dimensional vehicle trajectories. *Transportation Research Record: Journal of the Transportation Research Board*, 2422(2):12–20.

**Saeedmanesh, M. & Geroliminis, N.** (2015). Empirical observations of mfds and hysteresis loops for multi-region urban networks with stop-line detectors. In *Transportation Research Board 94th Annual Meeting*, 15-2071. Washington DC.

**Saeedmanesh, M. & Geroliminis, N.** (2016). Clustering of heterogeneous networks with directional flows based on "snake" similarities. *Transportation Research Part B: Methodological*, 91:250–269.

**Saeedmanesh, M. & Geroliminis, N.** (2018). Exact formulation of homogeneous and compact-shaped partitioning in large-scale heterogeneous traffic networks. In *Transportation Research Board 97th Annual Meeting*, 18-06566. Washington DC.

**Saumtally, T., Lebacque, J.-P. & Haj-Salem, H.** (2011). Static traffic assignment with side constraints in a dense orthotropic network. *Procedia - Social and Behavioral Sciences*, 20:465–474.

**Sbayti, H., Lu, C.-C. & Mahmassani, H. S.** (2007). Efficient implementation of method of successive averages in simulation-based dynamic traffic assignment models for large-scale network applications. *Transportation Research Record: Journal of the Transportation Research Board*, 2029:22–30.

**Shafiei, S., Gu, Z. & Saberi, M.** (2018). Calibration and validation of a simulation-based dy-

namic traffic assignment model for a large-scale congested network. *Simulation Modelling Practice and Theory*, 86:169–186.

**Sheffi, Y.** (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, chap. Stochastic User Equilibrium, pages 322–331. Prentice-Hall Inc, Englewood Cliffs, New Jersey, USA.

**Shoufeng, L., Jie, W., van Zuylen, H. & Ximin, L.** (2013). Deriving the macroscopic fundamental diagram for an urban area using counted flows and taxi gps. In *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, pages 184–188.

**Sirmatel, I. I. & Geroliminis, N.** (2017a). Economic model predictive control of large-scale urban road networks via perimeter control and regional route guidance. *IEEE Transactions on Intelligent Transportation Systems*, 19(4):1112–1121.

**Sirmatel, I. I. & Geroliminis, N.** (2017b). Integration of perimeter control and route guidance in large-scale urban networks via model predictive control. In *Transportation Research Board 96th Annual Meeting*, 17-04824. Washington DC.

**Smeed, R. J.** (1967). The road capacity of city centers. *Highway Research Record*, 169:22–29.

**Sossoe, K., Lebacque, J., Mokrani, A. & Haj-Salem, H.** (2015). Traffic flow within a two-dimensional continuum anisotropic network. *Transportation Research Procedia*, 10:217–225.

**Sossoe, K. S. & Lebacque, J.-P.** (2018). The dynamic bi-dimensional traffic model. In *Transportation Research Board 97th Annual Meeting*, 18-00901. Washington DC.

**Treiber, M., Hennecke, A. & Helbing, D.** (2000). Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E*, 62:1805–1824.

**Tsekeris, T. & Geroliminis, N.** (2013). City size, network structure and traffic congestion. *Journal of Urban Economics*, 76:1–14.

**Tsubota, T., Bhaskar, A. & Chung, E.** (2014). Macroscopic fundamental diagram for brisbane, australia: Empirical findings on network partitioning and incident detection. *Transportation Research Record: Journal of the Transportation Research Board*, 2421(1):12–21.

**van Wageningen-Kessels, F., van Lint, H., Vuik, K. & Hoogendoorn, S.** (2015). Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics*, 4(4):445–473.

**Wada, K., Satsukawa, K., Smith, M. & Akamatsu, T.** (2018). Network throughput under dynamic user equilibrium: Queue spillback, paradox and traffic control. *Transportation Research Part B: Methodological*, in press.

**Wang, J., He, X., Peeta, S. & Yang, X.** (2017). Perimeter control strategy for traffic events affected region: A feedback approach based on network exit function. In *Transportation Research Board 96th Annual Meeting*, 17-03758. Washington DC.

**Xue, Z., Chiabaut, N. & Leclercq, L.** (2016). Evaluation of the effect of traffic modeling on the control of traffic networks. In *Transportation Research Board 95th Annual Meeting*, 16-2205. Washington DC.

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

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

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

**Zhang, L., Garoni, T. M. & de Gier, J.** (2013). A comparative study of macroscopic fundamental diagrams of arterial road networks governed by adaptive traffic signal systems. *Transportation Research Part B: Methodological*, 49:1–23.

**Zhang, Z., Wolshon, B. & Dixit, V. V.** (2015). Integration of a cell transmission model and macroscopic fundamental diagram: Network aggregation for dynamic traffic models. *Transportation Research Part C: Emerging Technologies*, 55:298–309. Engineering and Applied Sciences Optimization (OPT-i) - Professor Matthew G. Karlaftis Memorial Issue.

**Zhao, T., Li, Z., Huang, B., Mu, B. & Zhang, Y.** (2014). Exploring the influence of traveller information on macroscopic fundamental diagrams. *Intelligent Transport Systems, IET*, 8(1):58–67.

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

**Zhong, R., Chen, C., Huang, Y., Sumalee, A., Lam, W. & Xu, D.** (2017). Robust perimeter control for two urban regions with macroscopic fundamental diagrams: A control-lyapunov function approach. *Transportation Research Procedia*, 23:922–941. Papers Selected for the 22nd International Symposium on Transportation and Traffic Theory Chicago, Illinois, USA, 24-26 July, 2017.

**Zhong, R., Huang, Y., Chen, C., Lam, W., Xu, D. & Sumalee, A.** (2018). Boundary conditions and behavior of the macroscopic fundamental diagram based network traffic dynamics: A control systems perspective. *Transportation Research Part B: Methodological*, 111:327–355.

**Zhou, Z., Lin, S. & Xi, Y.** (2012). A dynamic network partition method for heterogenous urban traffic networks. In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, pages 820–825.

# A. Appendix for chapter 2

## A.1 Trip-based model resolution at higher orders

We use here the notations described in section 2.4.1 and on Figure 2.5. As the vehicle $N$ and the vehicle $N + \delta N$ have the same trip length $L$, the definition in equation 2.8 allows us to write:

$$\int_{t_N^{\text{in}}}^{t_N^{\text{out}}} V(n(t))dt = L = \int_{t_{N+\delta N}^{\text{in}}}^{t_{N+\delta N}^{\text{out}}} V(n(t))dt \tag{A.1}$$

Splitting the integral limits on the left side of this equation in $[t_N^{\text{in}}, t_{N+\delta N}^{\text{in}}] = [t_0, t_0 + \delta N/\lambda(t_0)]$ and $[t_{N+\delta N}^{\text{in}}, t_N^{\text{out}}]$, and these on the right side in $[t_{N+\delta N}^{\text{in}}, t_N^{\text{out}}]$ and $[t_N^{\text{out}}, t_{N+\delta N}^{\text{out}}] = [t_1, t_1 + \delta t]$ simplifies the equation to:

$$\int_{t_0}^{t_0+\delta N/\lambda(t_0)} V(n(t))dt = \int_{t_1}^{t_1+\delta t} V(n(t))dt \tag{A.2}$$

where $V(n)$ is given by the branch $P_i$ of the production-MFD, $V(n) = P_i(n)/n = w_i(1 - \eta_i/n)$ on the interval $[t_0, t_0 + \delta N/\lambda(t_0)]$; and by the branch $P_j$, $V(n) = P_j(n)/n = w_j(1 - \eta_j/n)$ on the interval $[t_1, t_1 + \delta t]$. In order to solve equation A.2 in $\delta t$, we choose to approximate $V(n(t))$ by its Taylor polynomial at a given order. Then the resolution of the integrals should give polynomial expressions in $\delta t$. Let us start with order 2, if $\delta N$ is small enough we have for $t \in [t_0, t_0 + \delta N/\lambda(t_0)]$ the following approximation:

$$V(n(t)) \approx V(n_0) + \frac{\partial V}{\partial n}(n_0) \cdot (n(t) - n_0) + \frac{\partial^2 V}{\partial n^2}(n_0) \cdot \frac{(n(t) - n_0)^2}{2} \tag{A.3}$$

where $n_0 = n(t_0)$. For $\delta t$ small enough, which should be the case if $\delta N$ is small, a similar approximation of $V(n(t))$ can be written for $t \in [t_1, t_1 + \delta t]$ in replacing $t_0$ by $t_1$ and $n_0$ by $n_1 = n(t_1)$. The differences in accumulation $(n(t) - n_0)$ and $(n(t) - n_1)$ can be expressed in assuming a linear variation of accumulation on each interval. With the simplified notations $\lambda_0 = \lambda(t_0)$, $\mu_0 = \mu(t_0)$ and $\lambda_1 = \lambda(t_1)$ we can write:

$$n(t) = \begin{cases} n_0 + (\lambda_0 - \mu_0)(t - t_0) & \text{for } t \in [t_0, t_0 + \delta N/\lambda_0] \\ n_1 + \left(\lambda_1 - \frac{\delta N}{\delta t}\right)(t - t_1) & \text{for } t \in [t_1, t_1 + \delta t] \end{cases} \tag{A.4}$$

which gives the two approximations of $V(n(t))$ on the two intervals:

$$V(n(t)) = \begin{cases} V_0 + \frac{w_i - V_0}{n_0}(\lambda_0 - \mu_0)(t - t_0) - \frac{w_i - V_0}{n_0^2}(\lambda_0 - \mu_0)^2 (t - t_0)^2 & t \in [t_0, t_0 + \delta N/\lambda_0] \\ V_1 + \frac{w_j - V_1}{n_1}\left(\lambda_1 - \frac{\delta N}{\delta t}\right)(t - t_1) - \frac{w_j - V_1}{n_1^2}\left(\lambda_1 - \frac{\delta N}{\delta t}\right)^2 (t - t_1)^2 & t \in [t_1, t_1 + \delta t] \end{cases} \tag{A.5}$$

where $V_0 = V(n_0)$ and $V_1 = V(n_1)$. At order 0 for $V(n(t))$ the latter expressions simplify to the constant terms, and in this case the solution in $\delta t$ after integrating on both intervals is straightforward (order 1 for $\delta t$):

$$\delta t = \frac{V_0}{V_1} \cdot \frac{\delta N}{\lambda_0} \tag{A.6}$$

At order 1 for $V(n(t))$ the approximations of $V(n(t))$ are reduced to the constant terms and the terms in $t$, which gives after integration a second-order polynomial in $\delta t$ to solve (order 2 for $\delta t$):

$$\lambda_1(V_1 - w_j) \cdot \delta t^2 - \left(2n_1 V_1 + (V_1 - w_j)\delta N\right) \cdot \delta t + 2n_1 \left( V_0 \frac{\delta N}{\lambda_0} - \frac{V_0 - w_i}{2n_0} \left(1 - \frac{\mu_0}{\lambda_0}\right) \frac{\delta N^2}{\lambda_0} \right) = 0 \tag{A.7}$$

Note that when $V_0 = w_i$ and $V_1 = w_j$, which is possible only if $w_i = w_j$ is the free-flow speed (i.e. $i = j = 1$, we are on the first branch of the MFD), equation A.7 reduces to equation A.6. In a similar way, order 2 for $V(n(t))$ will provide a third-order polynomial in $\delta t$ which is not mentioned here.

Comparisons between the calculations of $\delta t$ at different orders are presented in Figures A.1(a)-(c). These calculations have been carried on with different values for the accumulation $n_1$ at time $t_1$ (from 30 to 800 veh) and for the demand ratio $\lambda_1/\lambda_0$ (from 0.5 to 2). In each case we have chosen $n_1 - n_0 = 20$ veh and $\lambda_1 = 0.5$ veh/s. Slight differences between the choices of order are only noticed for a high value of $\delta N$ (10 veh). This means for a reasonable value of $\delta N$ (e.g. 0.1 veh), the order 0 (i.e. the order 1 for $\delta t$) is largely sufficient for our resolution scheme.

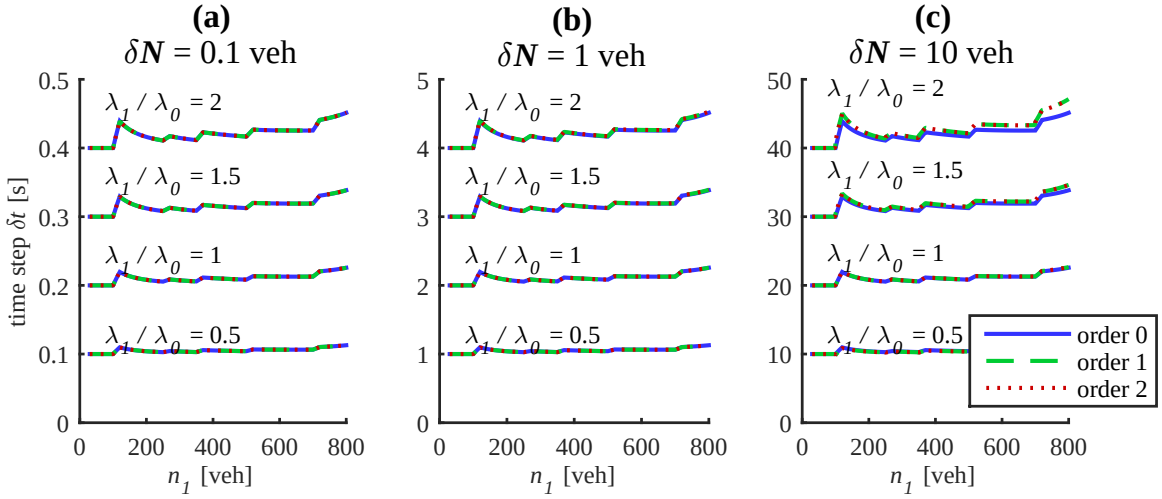



FIG. A.1 – *Values of the time step δt calculated at different orders for $V(n(t))$ and for different accumulation $n_1$ and ratio of demand level $\lambda_1/\lambda_0$. The difference $n_1 - n_0$ is equal to 20 veh in each case and $\lambda_1 = 0.5$ veh/s. (a) Case of a vehicle increment of $\delta N = 0.1$ veh, (b) $\delta N = 1$ veh, and (c) $\delta N = 10$ veh*

## A.2 Numerical investigations for a parabolic MFD with dimensionless parameters

We provide here some insights about a possible generalization of some of our conclusions about the accumulation-based model. For this purpose, we use a parabolic (Greenshield) MFD and the following dimensionless variables:

- $n^* = n/n_j$: the accumulation $n$ to the maximum jam accumulation $n_j$ ratio
- $t^* = t/t_{\text{free}}$: the time to the reservoir free-flow travel time $t_{\text{free}} = L/u$ ratio, where $u$ is the free-flow speed.
- $\lambda^* = \lambda/\mu_c$: the demand to the outflow-MFD capacity $\mu_c = P_c/L$ (maximum outflow) ratio, where $P_c$ is the production-MFD capacity

We use also two dimensionless parameters:

- $\rho = \lambda_0/\mu_c$: the initial demand to the outflow-MFD capacity $\mu_c$ ratio
- $\beta = t_{\text{dem}}/t_{\text{free}}$: the demand relaxation time $t_{\text{dem}}$ to the MFD free-flow travel time $t_{\text{free}}$ ratio

Laval & Castrillon (2015), Laval & Chilukuri (2016) showed that under linear transformation of flow, accumulation, space and time, delays and other performance indicators are invariant. This allows the simplification of any problem to a symmetric MFD shape case. That is why we choose here a simple parabolic MFD, which characteristic parameters are normalized by the transformations mentioned above. Hence the results presented in this appendix are intending to be universally applicable.

Figure A.2 illustrates the influence of the number of steps used to approximate a continuous demand profile. We choose here an exponential function:

$$\lambda(t) = \lambda_0 + (\mu_c - \lambda_0) \cdot \left(1 - \mathrm{e}^{-t/t_{\text{dem}}}\right) \tag{A.8}$$

which in dimensionless form reads:

$$\lambda^*(t^*) = \rho + (1 - \rho) \cdot \left(1 - \mathrm{e}^{-t^*/\beta}\right) \tag{A.9}$$

Here the parameter $\rho$ is chosen equal to 0.1 and $\beta$ equal to 1.5. The solutions provided by our semi-analytical process are compared with a full analytical solution obtained with *Mathematica*® for this kind of demand. The detailed analytical expression can be found in Laval *et al.* (2018). Results show that if the number of demand steps is high enough for our semi-analytical resolution (e.g. more than 40), the solution we get is very close to the exact solution. On the other hand a few branches (around 8) are sufficient to get good approximation of the MFD shape.

Figure A.3 also illustrates the influence of the time step $\delta t$ choice for the full numerical resolution, in comparison with the full analytical one from *Mathematica*® for the same demand case. It is observed that a timestep small enough leads to acceptable approximation of the exact solution too, namely $\delta t$ should be less than 10 % of the free-flow travel time $t_{\text{free}}$.

Figure A.4 presents our analysis in section 2.3.6 about the definition of a validity domain for the accumulation-based model applied to this dimensionless study. The demand profile is a logistic "S-shaped" function that is also characterized by the parameters $\rho$ and $\beta$

in dimensionless form. The demand gap $\Delta\lambda$ is normalized by the initial demand $\lambda_0$, and the travel time drop height $h_0$ and duration $d_0$ are normalized by the free-flow travel time. Hence the dimensionless demand variation rate $Q^*_\lambda$ is obtained from the actual variation rate $Q_\lambda$ with: $Q^*_\lambda = Q_\lambda \cdot t_{\text{free}}/\lambda_0$. If we focus only on the drop height artifact, a demand variation rate $Q^*_\lambda < 0.1$ should ensure a drop height effect below 10 % of the free-flow travel time, according to Figure A.4(c). In fact these charts cannot be used for any demand profile case, as these results are still dependent on $\rho$ (or $\lambda_0$). Here we choose $\rho = 0.1$, but we have verified that the drop height $h_0$ seems to be independent of $\rho$. This work is currently under research.

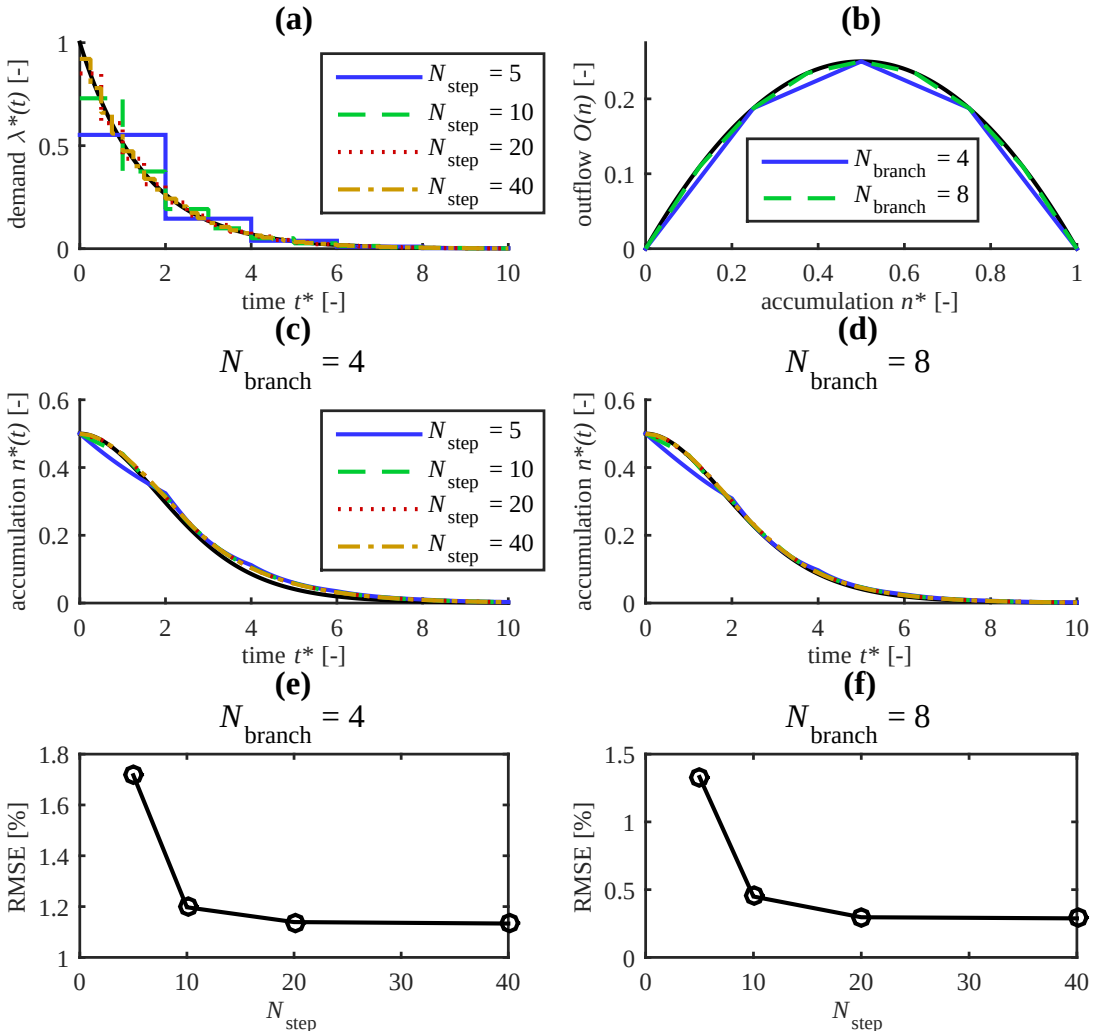



FIG. A.2 – *Convergence test for the semi-analytical method. (a) Exponential demand profile discretization, (b) parabolic MFD discretization, (c) evolution of accumulation for a 4-branch and (d) a 8-branch MFD, (e) root mean square error vs number of demand steps for a 4-branch and (f) a 8-branch MFD. The parameter ρ is chosen equal to 0.1 and β equal to 1.5.*

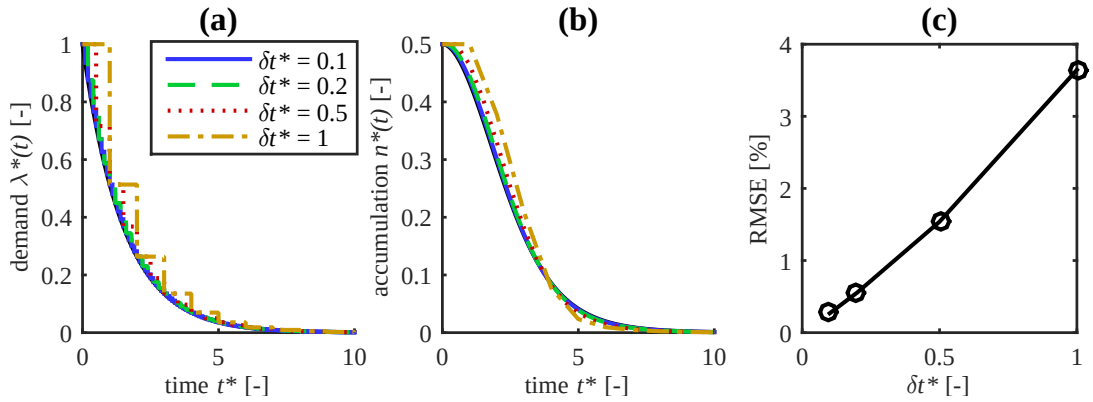



FIG. A.3 – *Convergence test for the full numerical scheme. (a) Exponential demand profile discretization, (b) evolution of accumulation and (c) root mean square error vs timestep*

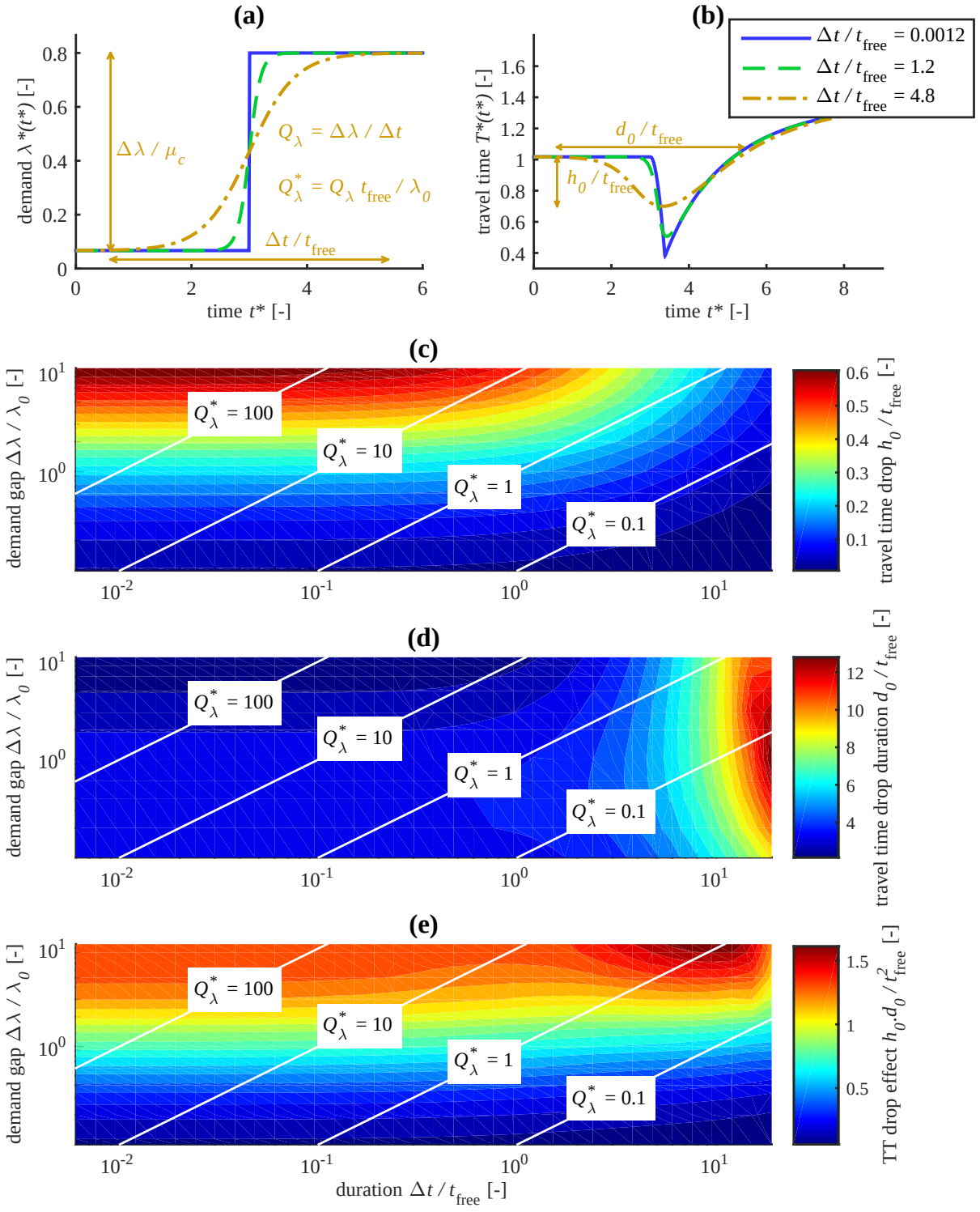



FIG. A.4 – *Influence of the demand variation rate on the accumulation-based model accuracy for travel time. This accuracy is defined regarding the size of the travel time drop artifact. (a) Demand profiles with different variation rates, (b) resulting travel time profiles with the normalized drop height $h_0/t_{\text{free}}$ and drop duration $d_0/t_{\text{free}}$, (c) drop height value in the $(\Delta t^*, \Delta\lambda^*)$ space, (d) drop duration value, and (e) the product height-duration value in the same plane*

# B. Appendix for chapter 4

## B.1 Entering flow and production merge on a simple example

We focus here on the simple case of two routes with lengths $L_1$, $L_2$ to show that flow and production constraints cannot be verified at the same time if one of the route is congested while the other is not. Without loss of generality, let us assume that route 1 is free-flow ($\lambda_1(t) < I_1(n_1(t), n_2(t))$) and that route 2 is congested ($\lambda_2(t) > I_2(n_1(t), n_2(t))$) at a given time $t$, which is then omitted for the sake of simplicity. Let $\alpha_1 = n_1/n$ and $\alpha_2 = n_2/n$ be the merge coefficients, and $L = 1/(\alpha_1/L_1 + \alpha_2/L_2)$ be the average trip length. We denote by $P_s(n)$ the current reservoir entry production capacity, where $n = n_1 + n_2$. We have the following situation:

$$\begin{cases} L_1\lambda_1 < \alpha_1 P_s(n) & \text{(route 1)} \\ L_2\lambda_2 > \alpha_2 P_s(n) & \text{(route 2)} \end{cases} \tag{B.1}$$

which simply corresponds to the above statements, i.e. that route 1 is free-flow whereas route 2 is congested. Applying the fair merge of Daganzo (1995) on production leads to:

$$\begin{cases} L_1 q_{\text{in},1} = L_1\lambda_1 \\ L_1 q_{\text{in},1} + L_2 q_{\text{in},2} = P_s(n) \end{cases} \Leftrightarrow \begin{cases} q_{\text{in},1} = \lambda_1 \\ q_{\text{in},2} = \frac{P_s(n) - L_1\lambda_1}{L_2} \end{cases} \tag{B.2}$$

Then we want to compare the result of this merge with the total flow capacity $P_s(n)/L$. We thus examine the sign of the expression $q_{\text{in},1} + q_{\text{in},2} - P_s(n)/L$. By replacing each inflow by their above expressions, by using the definition of $L$ and the fact that $\alpha_1 + \alpha_2 = 1$, we can show that:

$$q_{\text{in},1} + q_{\text{in},2} - \frac{P_s(n)}{L} = \left(\frac{\alpha_1 P_s(n)}{L_1} - \lambda_1\right)\left(\frac{L_1}{L_2} - 1\right) \tag{B.3}$$

The first factor of the second member is always positive in our case, see equation B.1. It then follows that the total inflow $q_{\text{in},1} + q_{\text{in},2}$ does not exceed the inflow capacity if and only if $L_2 > L_1$, i.e. if the longest route is congested. Moreover, the strict equality, or equivalently the full use of the inflow capacity, is only achieved when $L_1 = L_2$.

Thus this static consideration illustrates that applying a fair merge on production (i) has a few chances to ensure a full use of the inflow capacity as well, and (ii) does not even imply that the system complies with this inflow capacity. While the constraint on flow can

be violated in such a static example, in practice during simulation, the average trip length evolves with the reservoir state so that the inflow supply also adapts dynamically. As the analysis of a simple two-route system seems analytically intractable, we cannot go further with a stronger proof. These investigations remain entirely open to discussion.

FOLIO ADMINISTRATIF

THÈSE SOUTENUE DEVANT L'ÉCOLE NATIONALE DES TRAVAUX PUBLICS DE L'ÉTAT

**NOM :** MARIOTTE **DATE DE SOUTENANCE :** 14/11/2018

(avec précision du nom de jeune garçon/fille le cas échéant)

**Prénoms :** Guilhem

**TITRE :** Modélisation dynamique des grands réseaux de transports

**École doctorale :** ED 162 MEGA (Mécanique, Énergétique, Génie civil, Acoustique)

**NATURE :** Doctorat **Numéro d'ordre :** 2018LYSET010

**Spécialité :** Génie civil

**RÉSUMÉ :** La congestion en milieu urbain est un enjeu majeur que ce soit d'un point vue économique, social ou environnemental. À court et moyen terme, l'utilisation de la simulation dynamique du trafic routier peut permettre d'analyser et de guider des politiques d'optimisation des infrastructures existantes. Aujourd'hui, du fait de la complexité des systèmes de transport, les outils de modélisation classiques sont limités à des échelles géographiques peu étendues (de l'ordre du quartier). À grande échelle, le temps de calcul devient rapidement un facteur limitant tout comme le calibrage et la scénarisation. Néanmoins les dernières décennies ont vu l'apparition d'une nouvelle génération de modèles bien adaptés aux métropoles urbaines. Ceux-ci sont basés sur une relation phénoménologique entre la production de déplacements et le nombre de véhicules dans une zone spatiale d'un réseau routier, appelée Diagramme Fondamental de Zone (*Macroscopic Fundamental Diagram*, MFD). Cette relation, validée empiriquement sur de nombreuses villes, a permis d'étudier différentes méthodes de contrôle du trafic pour une ville entière, mais a été peu utilisée à des fins de prévision de la congestion.

L'objectif de cette thèse est de proposer un premier outil opérationnel de simulation et d'analyse des grands réseaux de métropoles, en utilisant et développant les modèles de trafic basés sur la relation MFD. Cet outil doit posséder un cadre théorique cohérent qui puisse convenir à des applications telles que la prévision d'états de trafic, le développement de nouvelles politiques de contrôle, l'estimation de pollutions liées au trafic, etc. Les contributions de la thèse portent sur deux aspects. Le premier est l'analyse des propriétés mathématiques et physiques des modèles existants, en incluant une formalisation complète de la gestion de plusieurs longueurs de parcours au sein d'une même zone urbaine. En particulier, cette formalisation traite de la distinction des trajets internes à la zone et des problèmes de flux convergents et divergents pour les trajets traversant la zone lorsque la congestion se propage d'une zone à l'autre. Le deuxième aspect est la proposition d'un nouveau modèle basé sur la distance individuelle parcourue à l'intérieur d'une zone urbaine (trip-based). Cette approche permet d'individualiser les usagers (auparavant représentés sous forme de flux continus) et donc de définir plus finement leurs caractéristiques, en vue de coupler leurs déplacements à des modèles d'affectations sur différentes routes. Enfin, des exemples d'application illustrant diverses collaborations sont donnés en dernière partie de la thèse. La simulation du trafic sur l'aire urbaine du Grand Lyon (France) y est présentée, ainsi que de nouveaux modules de modélisation de la recherche de parking ou de contrôle périphérique. Cette thèse est partie intégrante d'un projet européen ERC intitulé MAGnUM : Approche multi-échelle et multimodale de la modélisation du trafic pour une gestion durable de la mobilité urbaine.

**Mots-Clés :** *Modélisation du trafic routier*, *grands réseaux urbains*, *diagramme fondamental de zone (MFD)*, *dynamique de la congestion*, *modèles multi-réservoirs*, *échanges de flux entre zones urbaines*.

**Laboratoire(s) de recherche :** Laboratoire d'Ingénierie Circulation Transports (LICIT)

**Directeur de thèse :** Ludovic LECLERCQ

**Présidente du jury :** Christine SOLNON (INSA, Université de Lyon)

**Composition du jury :**
Nikolas GEROLIMINIS (EPFL, Lausanne), Rapporteur
Monica MENENDEZ (NYU, Abu Dhabi), Rapporteure
Arnaud DE LA FORTELLE (Mines Paris Tech), Examinateur
Serge HOOGENDOORN (TU Delft), Examinateur
Christine SOLNON (INSA, Université de Lyon), Présidente du jury
Ludovic LECLERCQ (ENTPE, IFSTTAR, Université de Lyon), Directeur de thèse