Transportation Research Part B 136 (2020) 62–86

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Calibration and validation of multi-reservoir MFD models: A case study in Lyon

Guilhem Mariotte[a,*], Ludovic Leclercq[a], S.F.A. Batista[a,b], Jean Krug[a], Mahendra Paipuri[a]

[a] Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT, Lyon, F-69518, France
[b] Division of Engineering, New York University Abu Dhabi, United Arab Emirates

ARTICLE INFO

Article history:
Received 24 July 2019
Revised 10 January 2020
Accepted 10 March 2020

Keywords:
Macroscopic Fundamental Diagram
Multi-reservoir systems
Simulation validation
Trip lengths
Path flow distribution
Network traffic
Accumulation-based model

ABSTRACT

Aggregated traffic flow models based on the Macroscopic Fundamental Diagram (MFD), also known as multi-reservoir or multi-region MFD models, have been developed for more than a decade for various applications. While being very appealing for simulating traffic dynamics at a city level, the outputs of these models were rarely confronted with real data measurements. Thus, this paper focuses on calibration and validation of an MFD simulation for a city partitioned into multiple reservoirs. The traffic predictions from the MFD simulation (total accumulation and mean speed) are compared with real data from loop and probe sensors. The questions addressed in this study include the influence of the city partitioning, the MFD and average trip length estimation, and the path flow distribution among reservoirs. This study is carried on the network of Lyon, France, composed by around 27,000 links that extend over an urban area of 80 km^2. Two different partitioning cases are defined and compared, with respectively 5 and 10 reservoirs. Our results notably show that the proper estimation of three elements is critical for accurate traffic state prediction: (i) the total "active" network length of each reservoir, (ii) the regional trip lengths in the reservoirs, and (iii) the path flow distribution at the regional network level. While the network equilibrium found in the 5-reservoir partitioning can be roughly approximated with Wardrop's principle, the 10-reservoir case is more complex and requires to design ad-hoc optimization process to derive regional path flow distributions that fit the data. The global equilibrium found in this latter case turns out to be hardly predictable with any traffic equilibrium principle.

1. Introduction

Traffic modeling based on the Macroscopic Fundamental Diagram (MFD) has become very popular in the traffic flow community. Although bi-dimensional continuous modeling approaches are also developed (Sossoe et al., 2015; Mollier et al., 2019; Aghamohammadi and Laval, 2020), MFD-based models are more often implemented to cope with large-scale traffic

* Corresponding author.
E-mail address: guilhem.mariotte@ifsttar.fr (G. Mariotte).

https://doi.org/10.1016/j.trb.2020.03.006

operations and monitoring. The literature about MFD can be classified into two groups: one investigating the conditions and data for reliable estimations of the MFD itself, and the other developing MFD-based simulation frameworks considering well-defined MFD. The first group often works with real data (Geroliminis and Daganzo, 2008; Lu et al., 2013; Ji et al., 2014; Tsubota et al., 2015; Lu et al., 2018; Dakic and Menendez, 2018), or sometimes simulated data (Nagle and Gayah, 2013; Leclercq et al., 2014; Ambuhl and Menendez, 2016; Du et al., 2016), and always focus on the MFD estimation of a single region (putting aside the questions of network partitioning). The second group mostly uses frameworks with two regions (Haddad and Geroliminis, 2012; Fu et al., 2017; Zhong et al., 2017; Yang et al., 2018; Mohajerpoor et al., 2019), or multiple regions (Aboudolas and Geroliminis, 2013; Ramezani et al., 2015; Kouvelas et al., 2017; Yildirimoglu et al., 2018) with increased complexity and refinement.

However, while providing interesting results in terms of large-scale control strategies and network design, all these MFD-based models were rarely confronted with real data. Actually, it appears quite often that validation is not the primary concern of a lot of studies related to MFD simulation, although it exists notable examples going in this direction. In the literature, these validation elements almost always consist in comparison with simulated data at a link level. Geroliminis and Daganzo (2007) compared a single reservoir simulation with a microsimulation of Downtown San Fransisco, USA. They showed a good match between the two approaches thanks to the homogeneous traffic conditions they observed. Later on, Yildirimoglu and Geroliminis (2014) also presented a good reproduction of microsimulation of a grid network by their multi-reservoir model, but they required regular feedback from the microsimulation to adjust the MFD simulation accordingly. Recently, Saeedmanesh et al. (2019) provided a nice validation of their MFD modeling and control module with microsimulation of the network of Barcelona, Spain. Nevertheless, their framework also uses Extended Kalman Filter to update traffic states in MFD simulation every cycle of 90 s, so that it is hard to conclude which one of the MFD model or the Kalman Filter is responsible for the final result accuracy. We can also highlight two works that especially focus on the validation of the MFD model of a single reservoir, in terms of traffic state prediction. Mariotte and Leclercq (2019b) investigated the effect of changes in the MFD and the average trip length, due to network heterogeneity, on the accuracy of the MFD simulation compared to a link-level mesoscopic model. Paipuri et al. (2019) carried on a microsimulation of the 6^{th} district of Lyon, France, with real demand profiles as input. These studies showed that both the MFD accumulation-based and trip-based models of this network can provide accurate accumulation and mean speed predictions if the MFD is estimated properly. They moreover insisted on the paramount importance of trip length calibration in the MFD-based models.

This study is the first of its kind, as it aims at calibrating and validating a multi-reservoir simulation at the level of a large metropolitan area. We thus want to pursue the effort of the works mentioned above to make a bridge between the investigations of the two groups we presented before. Our case study is the city of Lyon, France, for which the OD matrix, loop detector and probe data (taxi GPS trajectories) were available for February 2011. We choose to use the MFD simulation framework of Mariotte and Leclercq (2019a) who proposed one of the latest advances for multi-reservoir models especially designed for traffic state simulation and prediction. This framework is based on the approach of Yildirimoglu and Geroliminis (2014), who first introduced the concept of “regional” or “macroscopic” routes (i.e. successions of reservoirs) on which OD flows can be distributed. The concept of macroscopic route, later used in Ramezani et al. (2015); Yildirimoglu et al. (2018), also provides a finer description of distances traveled in reservoirs (i.e. instead of a single average trip length, several trip lengths can be defined in each reservoir). Then, Mariotte and Leclercq (2019a) brought the required components (inflow merging and outflow diverging models) to properly reproduce spillbacks between reservoirs when multiple trip lengths are considered. The case of simplified situations with a single average trip length per reservoir (as in, e.g. Kouvelas et al., 2017; Sirmatel and Geroliminis, 2017) will be also examined in this paper. Although both the accumulation-based and trip-based models (for further details, see Mariotte et al., 2017) were developed in Mariotte and Leclercq (2019a), only the accumulation-based one is used here. Our analysis and results show that the proper calibration of multi-reservoir simulation requires much more than an accurate estimation of the MFD. In particular, we highlight the critical role of the total “active” network length, the regional trip lengths and the path flow distribution. We propose and compare different methods to estimate them. What we call the total “active” network length corresponds to the part of the urban network that is extensively used for traffic flow. It obviously contains all major arterial roads, but also a part of secondary roads which is harder to quantify. We will show that this network length is the one to be considered to express the mean circulating flow-MFD in relationship with the production-MFD. We refer to Batista et al. (2019) for the trip length estimation, and we use an optimization framework to derive path flow coefficients and total active network lengths, assuming that traffic conditions are almost stationary at midday. In addition of being the first validation of a large-scale multi-reservoir simulation, this study also contribute to a better understanding of traffic assignment at a regional level. In our last case study, we show that the best traffic equilibrium found to reproduce real data is quite far from Wardrop’s principle. This means that people are not driving on the shortest “regional” routes, while they may still optimize their individual trips on the real network.

This paper is organized as follows. First, the MFD model and its main features are presented in Section 2. Then, the network studied and the different partitioning cases are introduced in Section 3. Issues about the estimation of crucial parameters for the MFD simulation are also discussed here. Finally, the simulation results are showed and compared to the real data in Section 4.

2. The multi-reservoir MFD model

The multi-reservoir model used in this paper is based on the work of Mariotte and Leclercq (2019a), who extended and generalized the approaches found in Geroliminis (2009); Yildirimoglu and Geroliminis (2014) to properly account for congestion propagation between adjacent reservoirs.

2.1. Traffic dynamics

Let assume that an urban area can be partitioned into N_R reservoirs. Each reservoir r has a well-defined production-MFD $P^r(n^r)$ (in [veh.m/s]) or a speed-MFD $V^r(n^r) = P^r(n^r)/n^r$, where n^r is the total accumulation (in [veh]) in reservoir r. A set of N_P regional paths, also called "macroscopic routes" or "routes" in the following, is supposed known. A macroscopic route basically consists of a sequence of reservoirs to get from an origin reservoir to a destination reservoir; it can be seen as an aggregated representation of a mobility pattern gathering a large set of trips from the real network with similar features (e.g. topology, length, succession of reservoirs, etc). The portion of length of route p crossed in reservoir r is written L_p^r. This concept was introduced by Yildirimoglu and Geroliminis (2014), and was then further analyzed in Batista et al. (2019) who provided several methods to estimate the trip lengths L_p^r in the reservoirs by considering the network graph only (topological approach). The demand profile for each route p is written $\lambda_p(t)$ (in [veh/s]). We denote by $\mathcal{P}^r$ the set of routes crossing reservoir r. The state variables are the accumulations $\{n_p^r\}_{1\le p\le N_P, 1\le r\le N_R}$, each $n_p^r(t)$ being the number of vehicles travelling on route p in reservoir r at time t. Their evolution is described by the following conservation equations (Yildirimoglu and Geroliminis, 2014):

$$\forall r \in \{1, \dots, N_R\}, \forall p \in \mathcal{P}^r, \qquad \frac{dn_p^r}{dt} = q_{\text{in},p}^r(t) - q_{\text{out},p}^r(t) \tag{1}$$

where $q_{\text{in},p}^r$ and $q_{\text{out},p}^r$ are respectively the inflow and outflow of route p when crossing reservoir r. The calculation of these quantities for both the accumulation-based and trip-based MFD models, and for both under- and oversaturated conditions was detailed in Mariotte and Leclercq (2019a). These authors proposed different merging and diverging schemes for the inflows and outflows of the routes crossing each reservoir. In this study, we focus on the accumulation-based model only. Moreover, we use the exogenous entry supply model with a demand pro-rata merge for inflows in a reservoir, and the maximum exit demand model for outflows, as developed by the same authors. This choice is motivated by the comparisons between MFD simulation and microsimulation carried on in Mariotte et al. (2019), who showed that these options lead to the best match with the microsimulation outputs. The latter study also introduced a second layer of inflow merging to account for physical flow limitations due to individual link capacities from the real network. This new layer in the calculation of inflows is quite straightforward to implement in the framework of Mariotte and Leclercq (2019a), and will be used in this paper as well with the concept of "macroscopic node" (see details below). As proposed by the same authors, the routes will be distinguished by their types of origin. In reservoir r, the set $\mathcal{P}_{\text{int}}^r$ gathers all the routes originating within r (internal trip creation), and $\mathcal{P}_{\text{ext}}^r$ gathers all the routes entering r from its borders. In practice to ease this distinction in the simulation, the routes are more precisely defined as a succession of "macroscopic nodes". A macroscopic node corresponds to the aggregation of several nodes from the real network; it can be all the inner nodes of a reservoir for an internal origin of trips, or a set of nodes at the reservoir border in case of trips crossing the reservoir boundary. We denote by M_p the sequence of macroscopic nodes defining route p, and by R_p its corresponding sequence of reservoirs. The set $\mathcal{M}^r$ gathers all the macroscopic nodes in r, and $\mathcal{P}^m$ all the routes crossing the macroscopic node m. The set $\mathcal{M}^r$ includes the subset $\mathcal{M}_{\text{int}}^r$ which gathers all the origin macroscopic nodes in r, and the subset $\mathcal{M}_{\text{ext}}^r$ which gathers all the entry border macroscopic nodes (through which the routes in $\mathcal{P}_{\text{ext}}^r$ enter reservoir r). An example of 3 routes is shown in Fig. 1 to illustrate the concept of macroscopic node and their different types. In Fig. 1(b), routes p_2 and p_3 connect the same macroscopic OD pair, i.e. an internal origin in reservoir R_1 to an external exit in R_4. With this definition of the route as a succession of macroscopic nodes, two different routes may correspond to the same sequence of reservoirs, like routes p_1 and p_2. Note that depending on modeling choices, a reservoir can have several external entry or exit macroscopic nodes, and also several macroscopic nodes at a border.

The macroscopic nodes inside a reservoir (of type "origin") are supposed to generate internal traffic without any constraints. The inflow in their corresponding routes is thus equal to the route demand (Mariotte and Leclercq, 2019a):

$$\forall r, \forall p \in \mathcal{P}_{\text{int}}^r, \quad q_{\text{in},p}^r(t) = \lambda_p(t) \qquad \text{(Effective inflow for internal origins)} \tag{2}$$

The macroscopic nodes at a reservoir boundary carry the information of the physical link capacities from the real network. This limit is notably taken into account in the works of Knoop and Hoogendoorn (2014); Kim et al. (2019). Here, in addition to the framework of Mariotte and Leclercq (2019a) and as proposed by Mariotte et al. (2019), a two-layer inflow merge is used for the calculation of inflows in each reservoir. Let denote C_m the aggregated capacity of the macroscopic node m, it is defined as $C_m = \sum_{l\in\mathcal{L}^m} q_l$ where q_l is the flow capacity of link l (including immediate traffic signal settings), and $\mathcal{L}^m$ is the set of all links connected to the nodes defining the macroscopic node m. For the first layer of the inflow merge, the flows $\{I_p^{r*}\}$ of all routes $p \in \mathcal{P}^m$ able to transfer through their respective macroscopic node $m \in \mathcal{M}_{\text{ext}}^r$ are given by

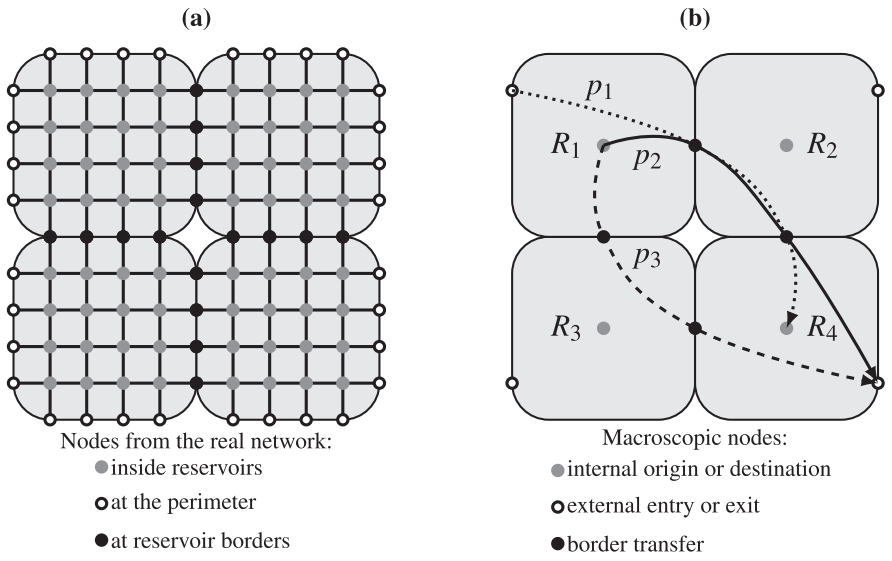


Fig. 1. Illustration of the simulation variables. (a) grid network split into 4 reservoirs, and (b) regional network with its macroscopic nodes and 3 examples of macroscopic routes.

(time t is omitted for the sake of simplicity):

$$\forall r, \forall m \in \mathcal{M}^r_{\text{ext}}, \quad \{I^{r*}_p\}_{p\in\mathcal{P}^m} = \text{Merge}\left(\{\lambda^r_p\}_{p\in\mathcal{P}^m}, \left\{\frac{\alpha^r_p}{\sum_{k\in\mathcal{P}^m}\alpha^r_k}\right\}_{p\in\mathcal{P}^m}, C_m\right) \tag{3}$$

where the *Merge*(·) function is found in Leclercq and 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 in Algorithm 1 in Appendix B, and takes the following arguments: the route inflow demands λ^r_p to enter r, the merging coefficients α^r_p ($p \in \mathcal{P}^m$), and the macroscopic node capacity C_m. The merging coefficients corresponding to a demand pro-rata merging rule are calculated as $\alpha^r_p = \lambda^r_p / \sum_{k\in\mathcal{P}^r_{\text{ext}}} \lambda^r_k$. If the macroscopic node m is an external entry (beginning of the routes), then λ^r_p equals the demand λ_p of route p. Otherwise it is receiving flow from previous reservoirs of the routes, and thus λ^r_p equals the outflow demand $O^{p^-(r)}_p$ from the previous reservoir $p^-(r)$ in route p. The outflow demand of the routes exiting each reservoir will be detailed next. Then, the second layer of the inflow merge treats the potential inflows I^{r*}_p previously calculated as inflow demands into r, and merges them to the reservoir entry supply $P^r_{s,\text{ext}}/L^r_{\text{ext}}$ according to the same function *Merge*():

$$\forall r, \quad \{I^r_p\}_{p\in\mathcal{P}^r_{\text{ext}}} = \text{Merge}\left(\{I^{r*}_p\}_{p\in\mathcal{P}^r_{\text{ext}}}, \{\alpha^r_p\}_{p\in\mathcal{P}^r_{\text{ext}}}, \frac{P^r_{s,\text{ext}}}{L^r_{\text{ext}}}\right) \qquad \text{(Inflow supply)} \tag{4}$$

where $P^r_{s,\text{ext}}$ is the production entry supply for routes crossing the reservoir entry border, equal to $P^r_s(n^r) - \sum_{p\in\mathcal{P}^r_{\text{int}}} L^r_p \lambda_p$ with $P^r_s(n^r)$ being the reservoir entry supply function, and L^r_{ext} is the average trip length for the routes in $\mathcal{P}^r_{\text{ext}}$ entering r from its borders, equal to $\sum_{p\in\mathcal{P}^r_{\text{ext}}} n^r_p / \sum_{p\in\mathcal{P}^r_{\text{ext}}} n^r_p / L^r_p$. The reader can refer to Mariotte and Leclercq (2019a) for more details about this calculation of the entry supply experienced by reservoir inflows.

The effective inflows and outflows of reservoirs are determined after the confrontation of each outflow demand O^r_p to its corresponding inflow supply $I^{p^+(r)}_p$ to enter the next reservoir $p^+(r)$ in route p. When adopting the maximum exit demand model from Mariotte and Leclercq (2019a), the outflow demand is defined as $O^r_p = \frac{n^r_p}{n^r}\frac{P^r(n^r)}{L^r_p}$ if $n^r < n^r_c$, and $O^r_p = \frac{n^r_p}{n^r}\frac{P^r_c}{L^r_p}$ otherwise, where n^r_c and $P^r_c = P^r(n^r_c)$ are respectively the reservoir critical accumulation and maximum production. Then, the effective outflows are calculated according to what the aforementioned authors called the most constrained exit:

$$\forall r, \forall p \in \mathcal{P}^r, \quad q^r_{\text{out},p} = \frac{n^r_p}{n^r_k}\frac{L^r_k}{L^r_p}\min\left[O^r_k; \mu^r_k\right] \qquad \text{(Effective outflow)} \tag{5}$$

where route k is the most constrained one for outflows, i.e. with the minimum supply-to-demand ratio: $k = \arg\min_{p\in\mathcal{P}^r} \frac{\mu^r_p}{O^r_p}$. The route supply μ^r_p when exiting r with route p corresponds to the previously calculated inflow supply $I^{p^+(r)}_p$ to enter the

next reservoir $p^+(r)$ in route p. If r is the destination reservoir of the route, then $\mu_p^r = +\infty$ in case of an internal destination, or $\mu_p^r = C_m$ in case of an external exit, m being the macroscopic node where the route ends.

Finally, the effective inflows are derived from the outflows from the previous reservoirs:

$$\forall r, \forall p \in \mathcal{P}_{\text{ext}}^r, \quad q_{\text{in},p}^r = q_{\text{out},p}^{p^-(r)} \qquad \text{(Effective inflow)} \tag{6}$$

Note that a point queue model replaces the outflow from the previous reservoir when the inflow corresponds to an external entry at the beginning of the route.

The whole accumulation-based multi-reservoir MFD model is described in Algorithm 2 in Appendix B. The cumulative count curves are calculated in post-processing:

$$\forall r \in \{1, \dots, N_R\}, p \in \mathcal{P}^r, \quad N_{\text{in},p}^r(t) = \int_0^t q_{\text{in},p}^r(\tau)d\tau + n_p^r(0) \qquad \text{(Entering cumulative curve)} \tag{7}$$

$$N_{\text{out},p}^r(t) = \int_0^t q_{\text{out},p}^r(\tau)d\tau \qquad \text{(Exiting cumulative curve)} \tag{8}$$

2.2. Traffic assignment and routing

Traffic assignment in multi-reservoir systems was first introduced by Yildirimoglu and Geroliminis (2014), but Batista and Leclercq (2019) showed that network equilibrium at a regional level might have a different meaning compared to what is usually known at the link level. From a theoretical point of view, this question clearly needs to be further investigated. Except from the works mentioned and to our best knowledge, no other studies dealt with this question in much details. In this study however, we only focus on simulation calibration and validation, thus the problem of traffic assignment will be tackled from an operational side. Since we do not know the distribution of flows in the network (this would require significant amounts of floating car data), we need an estimation method to assign OD flows on macroscopic routes. We identify two methods that we can easily apply in our study. The first one relies on a network equilibrium model. We choose to use Wardrop's first principle (Wardrop, 1952) which is widely used in traffic assignment problems, because it is the most common and simple method we have in the literature. It distributes flows to the shortest path options, a problem which is solved with a convergence loop based on the Method of Successive Averages (MSA) to balance travel times in the simulation. Despite its simplicity, this method might not reproduce travelers' mobility at a large scale however (some limitations are already known at a local scale). That is why we introduce a second method that does not rely on any modeling assumption. This second method consists in an optimization problem: we seek to minimize the difference between our MFD simulation predictions and the real observations. As we usually do not have direct observations of path flows, we resort to indirect measurements (i.e. the accumulation or production in each reservoir) to assess the optimal calibration of the path flow distribution. We present these two assignment methods in more details in this section. The first one will be denoted as the "Wardrop flow distribution", while the second one will be referred to as the "optimized flow distribution".

Let assume that the macroscopic or regional OD matrix is known, i.e. a demand profile $\lambda^{OD}(t)$ for each OD pair is known. In our framework, a macroscopic origin is a macroscopic node of type either origin (internal trip generation) or external entry (trips from outside the studied area). Similarly, a macroscopic destination is a macroscopic node of type either destination (trips ending inside the reservoir) or external exit (trips going outside the studied area). For all the routes related to a given OD pair, denoted as the set $\mathcal{P}^{OD}$, the demand λ_p of each route p is defined as:

$$\forall OD, \forall p \in \mathcal{P}^{OD}, \quad \lambda_p(t) = a_p.\lambda^{OD}(t) \qquad \text{(Route demand flow)} \tag{9}$$

where a_p is the path flow coefficient of route p. We have $\sum_{p\in\mathcal{P}^{OD}} a_p = 1$.

For our first assignment method, these coefficients are estimated by using Wardrop's principle which postulates that users travel on shortest paths in time. Over the simulation period, the mean travel time T_p of route p is calculated as the mean horizontal distance between the departure $N_{\text{in},p}^o(t)$ and arrival $N_{\text{out},p}^d(t)$ cumulative curves in the origin o and destination d reservoirs. The set $\mathcal{P}_{\text{min}}^{OD}$ gathers the routes with the minimum travel time: $\forall p \in \mathcal{P}_{\text{min}}^{OD}, T_p = T_{\text{min}}^{OD} = \min_{k\in\mathcal{P}^{OD}} T_k$. Then according to Wardrop's principle we have:

$$\forall OD, \forall p \in \mathcal{P}^{OD}, \quad a_p = \begin{cases} \frac{1}{|\mathcal{P}_{\text{min}}^{OD}|} & \text{if } p \in \mathcal{P}_{\text{min}}^{OD} \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Wardrop flow distribution)} \tag{10}$$

where the notation $|\cdot|$ is used for the number of elements in a given set. Because the travel times are unknown at the beginning of the simulation, the well-known fixed-point method of the MSA is applied to converge towards the situation described in Eq. 10:

- step $i_{\text{MSA}} = 1$. Initialize route travel time to free-flow travel time: $T_p = \sum_{r\in R_p} V^r(0)/L_p^r$, calculate $a_p^{i_{\text{MSA}}}$ according to Eq. 10.
- step i_{MSA}. Run the simulation and calculate the new (simulated) travel times T_p with the departure $N_{\text{in},p}^o(t)$ and arrival $N_{\text{out},p}^d(t)$ curves: $T_p = \text{mean}(t_p(t))$ where the experienced travel time $t_p(t)$ at t follows the relationship $N_{\text{in},p}^o(t - t_p(t)) = N_{\text{out},p}^d(t)$. Then, calculate $a_p^{i_{\text{MSA}}*}$ according to Eq. 10 and apply the MSA formula (Sheffi, 1985):

$$a_p^{i_{\text{MSA}}} = \frac{1}{i_{\text{MSA}}} a_p^{i_{\text{MSA}}*} + \left(1 - \frac{1}{i_{\text{MSA}}}\right) a_p^{i_{\text{MSA}}-1} \tag{11}$$

- Stop the loop when convergence is reached and set $a_p = a_p^{i_{\text{MSA}}}$. The convergence can be checked with the value of the Gap (see below), and/or the comparison between $a_p^{i_{\text{MSA}}}$ and $a_p^{i_{\text{MSA}}-1}$ which should stabilize after enough iterations

The Gap corresponds to the total relative difference between the route mean travel time and the minimum travel time among all OD pairs (Sbayti et al., 2007):

$$\text{Gap} = \sum_{OD} \text{Gap}^{OD} = \sum_{OD} \frac{1}{T_{\min}^{OD}} \sum_{p \in \mathcal{P}^{OD}} a_p \left(T_p - T_{\min}^{OD}\right) \tag{12}$$

This value can be used in any situation to assess how far a network is from Wardrop equilibrium. For a given OD pair, the higher its Gap^{OD}, the more users are traveling on non-shortest options. The Gap is equal to zero for a perfect Wardrop equilibrium as all users are traveling on shortest options.

For our second assignment method, the following optimization problem is solved:

$$\min \sum_{r=1}^{N_R} \left(P_{\text{sim}}^r - P_{\text{obs}}^r\right)^2, \quad \text{where: } P_{\text{sim}}^r = \sum_{p \in \mathcal{P}^r} L_p^r . a_p . \lambda^{OD}(t_0) \qquad \text{(Optimized flow distribution)} \tag{13}$$

s.t.

$$\forall p, \quad 0 \le a_p \le 1 \tag{14}$$

$$\forall OD, \quad \sum_{p \in \mathcal{P}^{OD}} a_p = 1 \tag{15}$$

$$\forall r, \forall m \in \mathcal{M}^r, \quad \sum_{p \in \mathcal{P}^m} a_p . \lambda^{OD}(t_0) \le C_m \tag{16}$$

where P_{sim}^r is the simulated total production in reservoir r and P_{obs}^r the observed production (from real data). The last constraint in equation 16 stipulates that the distribution of flows cannot violate each macroscopic node capacity. This problem is solved with a Particle Swarm algorithm. In practice, a simulation for a large city is likely to involve thousands of routes as detailed in the next sections, so that calculating P_{sim}^r with the traffic dynamics presented in Section 2.1 at each step of the optimization solver would be virtually impossible. Therefore, we look for a time t_0 of the day when traffic can be considered in a steady and undersaturated state. In that case, the total production in each reservoir r can be roughly estimated as the instantaneous cumulative effect of the demand of all the routes crossing r: $P_{\text{sim}}^r = \sum_{p \in \mathcal{P}^r} L_p^r . \lambda_p(t_0) = \sum_{p \in \mathcal{P}^r} L_p^r . a_p . \lambda^{OD}(t_0)$, because we have $q_{\text{in}}^r(t_0) = q_{\text{out}}^r(t_0) = \sum_{p \in \mathcal{P}^r} \lambda_p(t_0)$ in steady state conditions. A good candidate for such conditions is midday time where traffic states look stable and undersaturated, see Fig. 3(b). We then further assume that the estimated path flow distribution is constant over the day, i.e. valid for any time $t \neq t_0$. This is a strong assumption, but as soon as traffic conditions fluctuate or the network becomes congested, there is no simple relationship between the demand and the mean flow inside reservoirs.

3. The network studied with different partitioning cases

We present now the first application of the MFD model introduced in the previous section on a real field: the urban area of Lyon, France. Our available data include the network layout, the city regional OD matrix, taxi GPS trajectories and loop detector data recorded in February 2011.

3.1. Network data and configuration

Lyon has the second greatest urban area of France, with more than 2 million inhabitants. The network studied includes the cities of Lyon, Villeurbanne, Sainte-Foy-lès-Lyon and La Mulatière (in total almost 700,000 inhabitants), which corresponds to the urban area inside the first ring road of Lyon. This network comprises 27,000 links, with an area of 80 km^2 and where around 1 million trips are done each day. The network configuration with its environment is given in Fig. 2. This area exchanges traffic with its surroundings via mainly 4 freeways related to 4 origin/destination cities as presented in Fig. 2. The available network data includes the number of lanes and the signal settings at each node with traffic lights.

3.2. Demand estimation

The demand was estimated for a typical weekday in a preliminary study, which uses a four-step model based on household trip surveys and socio-demographic data. The final demand profile was validated with historical data used by Lyon authorities. The estimated OD matrix is defined for each hour of the day at the level of IRIS urban zones, the French partitioning system for demographic data, see Fig. 3(a). The spatial extension of an IRIS zone may vary as its definition is based

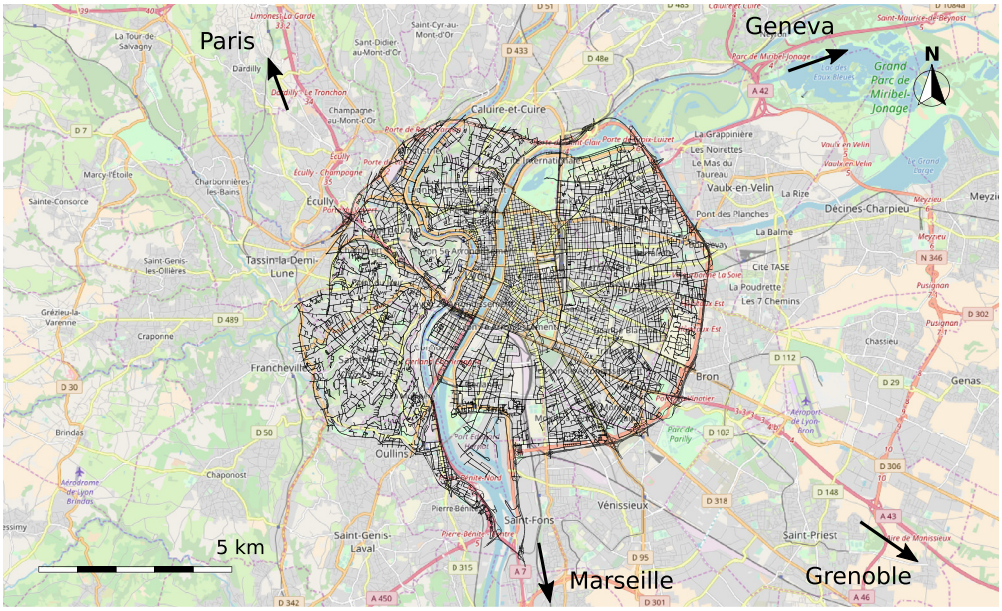


Fig. 2. The network of Lyon inside the first ring road with its urban environment (background map: © OpenStreetMap).

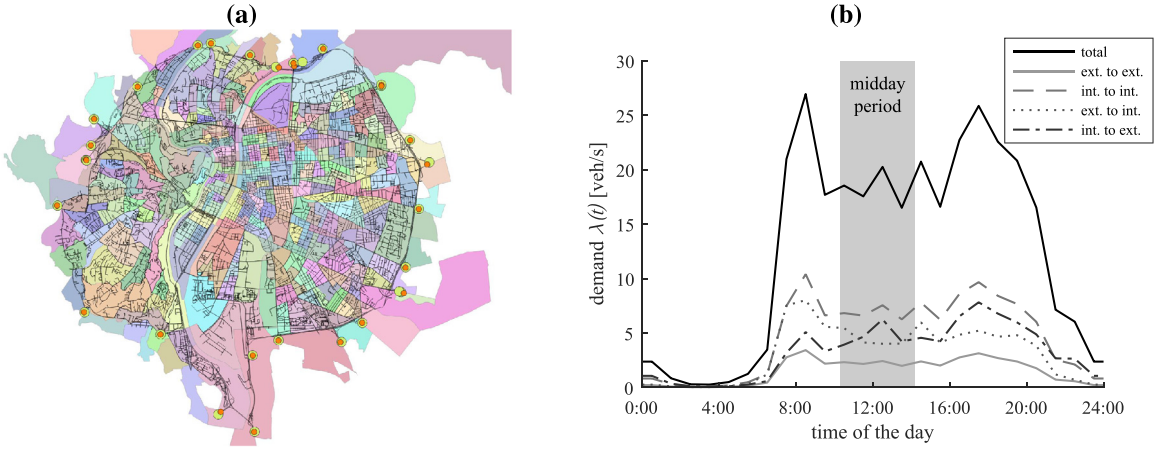


Fig. 3. (a) Demographic partitioning of the Lyon network in 285 IRIS zones with the main entry and exit node connections at the outside perimeter. (b) Smoothed demand profile for a typical weekday split into 4 sets of trips distinguished by their major origin and destination.

on a fixed range of inhabitants, workers, etc. In the network studied, each IRIS zone 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. The total demand profile is presented in Fig. 3(b), where the midday period is represented by the gray area. This graph also shows the proportion of the total demand regarding each route origin and destination location inside or outside the perimeter studied. Based on this demand data at the IRIS scale, the OD matrix used for the simulation is obtained by summing demand profiles from the IRIS OD matrix when they correspond to a given couple of origin-destination macroscopic nodes.

3.3. Two different network partitioning cases

Two partitions are investigated in this study. The first one splits the city into 5 reservoirs and is presented in Fig. 4(a), whereas the second one consists of 10 reservoirs as shown in Fig. 4(b). In both cases, the network is partitioned into reservoirs by aggregating several neighboring IRIS zones together to easily aggregate demand levels. These zones are small enough to get a large variety of reservoir shapes. Both partitioning cases were obtained by a trial-and-error process that successively accounts for the following constraints: (i) keep natural borders from the network topology when they exist, (ii) ensure compact reservoir shapes, (iii) avoid reservoir borders too close to major arterial roads and crossroads, and (iv)

Fig. 4. (a) Road network of Lyon partitioned in 5 reservoirs, and (b) 10 reservoirs.

ensure a good definition of low-scattered MFD. In practice, a first partitioning is built according to the first three topological constraints, and the good definition of MFD is checked with traffic data. Then the reservoir shapes may be slightly modified, keeping the topological constraints in mind, until the MFD exhibit low enough scatter. Point (iii) is here to ensure that the macroscopic routes are representative of general mobility patterns at the city scale. Having a major arterial right at a reservoir border is likely to generate routes with an alternate sequence of reservoirs (e.g., $[R_1, R_2, R_1, R_2, \dots]$) which cannot be easily handled in MFD modeling. For the first constraint, two rivers cross the city and were thus kept as natural borders between reservoirs. In Fig. 4(a), the two rivers are surrounding reservoir R_1, while in Fig. 4(b), they define the borders of reservoirs R_8 and R_{10}. Then the partitioning of the remaining reservoirs was done by complying with all the other constraints. A partitioning algorithm using historical traffic data could have been applied instead, as the ones developed in Ji and Geroliminis (2012); Zhou et al. (2012); Saeedmanesh and Geroliminis (2016, 2017); Lopez et al. (2017). However these works mostly focused on the homogeneity of traffic states within the reservoirs, regardless the relevance of their shapes to the network general topology and mobility patterns. Recently, Saeedmanesh and Geroliminis (2018) started addressing this question as they presented a trade-off between traffic homogeneity and compactness of partitions. This research direction is critical to us to define a proper partitioning in MFD-based simulation, that is why we adopted the aforementioned constraints in our study.

3.4. Trip length estimation

The trip lengths form a critical component in MFD traffic dynamics as detailed in Section 2.1. As highlighted by Yildirimoglu and Geroliminis (2014); Leclercq et al. (2015); Mariotte and Leclercq (2019b), a wrong estimation of these parameters can have a significant impact on the final simulation results. Actually, Batista et al. (2019) showed that the question of trip lengths in reservoirs is also intricately related to the set of macroscopic routes that will be considered in the simulation. In this work, we therefore use the method developed by the same authors to both generate the set of feasible macroscopic routes and their respective lengths in reservoirs. It consists in shortest path calculations in the real network (only the distances are considered, not the possible increase of travel time due to traffic) to cover all possible traveling options in the studied area. This is obtained by randomly selecting a large number of origin and destination points and then calculating the shortest path in distance between them. The drawback of this method is that despite a good network coverage (every link is visited), there is no guaranty that the resulting paths represent user real paths in the city. At the end, the set of link-level trips found in the real network is matched with the definition of reservoirs and macroscopic nodes, which provide a set of macroscopic routes. Then, the trip lengths $\{L_p^r\}_{r \in R_p}$ of a given route p are computed as the mean of the lengths crossed in each reservoir by the link paths following the same sequence of reservoirs (or, more precisely, the same sequence of macroscopic nodes).

One must remember that these estimated trip lengths may depend on traffic states, as discussed in 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. On the other hand, one can argue that this method only provides trip length values as an intrinsic property of the topology of the network. A sufficiently high number of draws should ensure that at least a reliable trend is captured in the reservoirs.

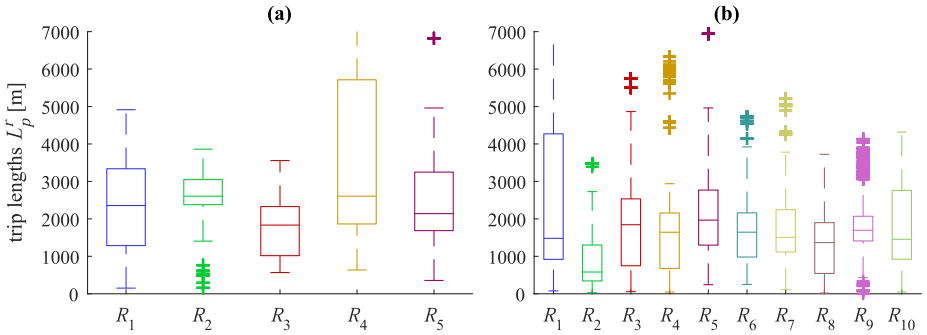


Fig. 5. Distribution of trip lengths L_p^r in each reservoir r for each route p crossing r. (a) Distributions for the 5-reservoir case, and (b) the 10-reservoir case.

Then, the effect of users adaptation to traffic states would be independent from the trip lengths, because already accounted at the regional level with the question of path flow distribution among the macroscopic routes. Although further research is needed to clarify this concern, we choose here to adopt the method of shortest path calculations for its robustness and practical aspects. We should also acknowledge that we do not have data to monitor user daily trips, thus no practical way to estimate trip lengths except from the network graph. Furthermore, this estimation method has the following advantages. First, the set of link paths obtained by such calculations is purely topological, thus free from a possible bias inherent of the sampling of any real dataset. Second, this link path set is calculated one for all, and can be used for any partitioning on the same network to compute the trip lengths then. Our preliminary results confirm the reliability of these trip length estimations, as presented latter in Section 4.1.

Figs. 5 (a) and (b) show the distributions of trip lengths we estimated for the 5-reservoir and 10-reservoir cases, respectively. These distributions depict the full route set available. The actual trip lengths involved in the simulation of traffic dynamics will come from the selection of the most relevant routes provided by the assignment procedure. As already mentioned in Batista et al. (2019), note the great variety of lengths inside each reservoir. This underlines the interest of tracking the specific lengths crossed by the macroscopic routes, instead of assuming a single average trip length per reservoir as it is implicitly done in a lot of studies which use outflow-MFD (see e.g., Haddad and Zheng, 2018; Mohajerpoor et al., 2019; Kim et al., 2019).

3.5. MFD estimation

The estimation of the production-MFD for each reservoir is based on two datasets. The first one consists of GPS trajectories from a taxi company in Lyon (around 5000 trips recorded per day), and the second one gathers 330 loop detectors operated by Lyon authorities. The data was recorded over several months in 2011, February 2011 was selected for this study because less missing data was observed. Among all days, only weekdays from 01/02/2011 to 18/02/2011 are kept in our analysis, because weekends usually have different traffic patterns as shown by Ambuhl et al. (2019); Shim et al. (2019). The end of the month is also removed from our study because it is school winter holidays. Vehicles counts are recorded for 6 min intervals in the loop data, so that we chose an aggregation period of $\Delta T = 18$ min for the MFD estimation. As suggested by Leclercq et al. (2014); Ambuhl and Menendez (2016), the combined use of probe and loop data should ensure a reliable estimation of the MFD. According to these authors, the mean speed can be derived from the probe data only, and the accumulation can be approximated with the probe data corrected by the loop data. Actually, the probe data can only give the accumulation of probe vehicles with enough accuracy, and the loop data is needed to estimate their penetration rate so that the total accumulation can be estimated properly. However, we do not adopt this method here, because we expect the penetration rate of taxis to be very low given the dataset we collected (5000 trips recorded among 1 million in a day, thus a penetration rate of less than 1%). Instead, we choose to estimate the production thanks to the mean flow measurements given by the loop data, and then deduce the accumulation by dividing the production by the mean speed. The drawback of this method is that we face an unknown parameter that is usually disregarded in the literature: the network total length to consider to scale up the observed production (on links equipped with loops) to the full network production.

Thanks to the probe (taxi) data, the mean speed $V_{\Delta T}^r$ of a given reservoir r can be calculated as:

$$V_{\Delta T}^r = \frac{TTD_{\text{taxi}}^r}{TTT_{\text{taxi}}^r} = \frac{\sum_{\text{taxi } i=1}^{N_{\text{taxi}}^r} td_i^r}{\sum_{\text{taxi } i=1}^{N_{\text{taxi}}^r} tt_i^r} \qquad \text{(Space-mean speed estimation)} \tag{17}$$

where TTD_{taxi}^r is the total travel distance and TTT_{taxi}^r the total travel time in r for taxis only. N_{taxi}^r is the number of taxis circulating in r within the period ΔT, td_i^r 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, also used by Leclercq et al. (2014); Lu et al. (2018), is considered quite accurate as all the cars in a link experience the same speed on average.

The loop data provides a measure of the mean flow of each equipped link with the same aggregation period. Then, the production of the equipped network in reservoir r is written:

$$P^r_{\Delta T,\text{equip}} = \sum_{i \in \mathcal{L}^{\mathbf{r}}_{\mathbf{equip}}} l_i \langle q_i \rangle_{l_i,\Delta T} \qquad \text{(Equipped production estimation)} \tag{18}$$

where $\mathcal{L}^{\mathbf{r}}_{\mathbf{equip}}$ is the set of alllinks equipped with loops in r (each lane being defined by a separate link), l_i is the length of link i and $\langle q_i \rangle_{l_i,\Delta T}$ is the measured mean flow in link i over l_i and ΔT. Like Lu et al. (2018), if we assume that traffic states are homogeneous in reservoir r (which is the basic hypothesis for MFD-based simulation), the mean circulating flow observed in the equipped network of r must equal the mean flow in the full network of r. Because the mean flow is defined as the production divided by the total length of the corresponding network, the production $P^r_{\Delta T}$ of the full network in r can be estimated as follows:

$$P^r_{\Delta T} = \frac{L^r_{\text{tot}}}{L^r_{\text{equip}}} P^r_{\Delta T,\text{equip}} = \frac{1}{\Gamma^r} P^r_{\Delta T,\text{equip}} \qquad \text{(Production estimation)} \tag{19}$$

where $L^r_{\text{equip}} = \sum_{i \in \mathcal{L}^{\mathbf{r}}_{\mathbf{equip}}} l_i$ is the total length of the equipped network in r and $L^r_{\text{tot}} = \sum_{i \in \mathcal{L}^r} l_i$ is the total network length in r. We call "equipped length factor" the ratio of the equipped length over the total length, denoted as Γ^r. The estimation of this parameter is almost never thoroughly discussed in the literature, because the calculation of total network lengths seems unambiguous. But whereas there is no bias on the total length of the equipped network L^r_{equip} (we know exactly which links are equipped), the relevant total network length L^r_{tot} to consider has no obvious definition when looking at the city network. Based on the MFD homogeneity assumption, L^r_{tot} should correspond to the network that carries the circulating flow with ideally homogeneous loading. Should we include all the secondary roads and residential streets in the calculation? Maybe not, because unlike arterials, these roads usually only carry local traffic for trip endings. We call the network length L^r_{tot} that should indeed be considered in the total production calculation the total "active" network length. It definitely comprises all major arterials but certainly also a part of secondary roads, which is hard to estimate from the city map only. Thus, we can express L^r_{tot} as a fraction of the total maximum network length $L^r_{\text{tot,max}}$, i.e. including all the lanes for cars from the road network in r: $L^r_{\text{tot}} = \gamma^r . L^r_{\text{tot,max}}$. The fraction γ^r is called the "total length factor" ($\gamma^r < 1$). It follows that the equipped length factor can be written as:

$$\Gamma^r = \frac{L^r_{\text{equip}}}{L^r_{\text{tot}}} = \frac{L^r_{\text{equip}}}{\gamma^r L^r_{\text{tot,max}}} = \frac{1}{\gamma^r} \Gamma^r_{\text{min}} \qquad \text{(Equipped length factor)} \tag{20}$$

where Γ^r_{min} is the minimum equipped length factor (also referred to as the loop sensor coverage), obtained with the total maximum network length $L^r_{\text{tot,max}}$.

Finally, the accumulation in reservoir r is derived from the two previous estimations:

$$n^r_{\Delta T} = \frac{P^r_{\Delta T}}{V^r_{\Delta T}} \qquad \text{(Accumulation estimation)} \tag{21}$$

The accuracy of the estimation of accumulation is the same as the one of the production, i.e. there is the same (unknown) bias in $n^r_{\Delta T}$ and in $P^r_{\Delta T}$, expressed through the equipped length factor Γ^r. Methods are discussed to evaluate this bias in the next section.

4. Results and comparison with real data

The confrontation of the MFD simulation with the real data is carried on as follows: the first four days of data (01-04/02/2011) are used to calibrate the MFD, and the remaining days (excluding weekends) are used to validate the simulation results. Note that we have a single demand profile corresponding to a typical weekday. So, we will compare the simulation results for this typical day with different network observations over several days (07–11/02/2011 and 14–18/02/2011). The comparisons are made considering the total accumulation and mean speed in each reservoir.

4.1. The single reservoir case

In this first simulation test, we consider the whole area as a single reservoir. This test case has the advantage of eluding the question of path flow distributions, and thus allows a direct estimation of the city-wide average trip length L and equipped length factor Γ. The average trip length could be calculated as the mean of all route lengths found in the initial route set. But we know that all the routes will not be used in the simulation of 5 or 10 reservoirs, thus we prefer estimating L as the mean of route lengths weighted by their mean demand. Because we have no clue on the relevant path flow distribution to use in a multi-reservoir simulation for now, we assume that each OD demand goes on its shortest route in distance. Consequently, the average trip length is calculated as:

$$L = \frac{\sum_{OD} \overline{\lambda^{OD}} . \min_{p \in \mathcal{P}^{OD}} \left(\sum_{r \in R_p} L^r_p \right)}{\sum_{OD} \overline{\lambda^{OD}}} \tag{22}$$

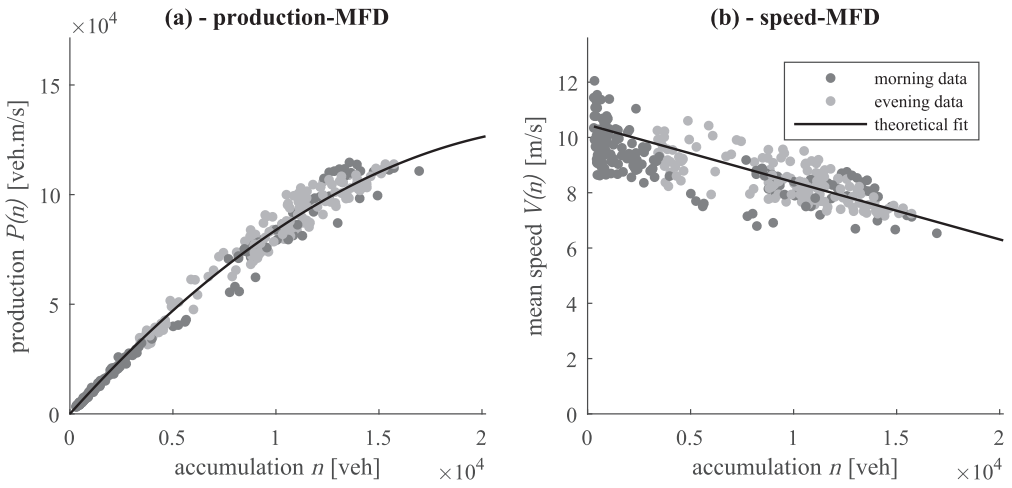


Fig. 6. MFD of the whole area as a single reservoir. (a) Production-MFD and (b) speed-MFD.

where $\overline{\lambda^{OD}}$ is the mean of the demand profile over the full day. We find $L = 4300$ m with the OD matrix and the trip lengths of the 5-reservoir partitioning. This value is actually corroborated by other sources of data like travel surveys operated by Lyon authorities. The computation of the mean distance traveled by the taxis also gives a similar value, although taxi trip distances may not be representative for personal cars. The hypothesis behind Eq. 22 of a simplistic path flow distribution will be discussed in the next section about the 5-reservoir simulation. By averaging the total demand during midday (from 10:30 am to 02:30 pm, represented by the gray area in Fig. 3(b)), we get: $\lambda^{\text{midday}} = 18.1$ veh/s. At the same time, the production of the equipped network calculated with Eq. 18 gives: $P^{\text{midday}}_{\text{equip}} = 4850$ veh.m/s. Providing that undersaturated and slow-varying traffic conditions are experienced at this time, if we assume that the total demand equals the total inflow for the city, we can express the total production at midday as:

$$L.\lambda^{\text{midday}} = \frac{1}{\Gamma} P^{\text{midday}}_{\text{equip}} \tag{23}$$

This also comes from the application of the queuing formula of Little (1961) on the whole city. We deduce that the city-wide equipped length factor Γ equals 0.062, and that the total length factor γ equals 0.35, as $\Gamma_{\min} = L_{\text{equip}}/L_{\text{tot,max}} = 0.022$ for the whole city. Note that the equipped length factor is supposed to be responsible of all the discrepancy between both estimations of the total production. This seems reasonable, because both the demand and the trip length were validated thanks to Lyon authorities' data. A bias could eventually be observed on L, but improving our confidence in this parameter must involve more sources of data (e.g. a larger set of GPS trajectories) that are not available to us for now. Thus the equipped length factor of 0.062 is used for the MFD estimation and the total accumulation chronicles of the remaining days. The result of a low γ is appealing, because it means that only 35 % of the full network carry significant flow values. While this should be investigated for other cities as well, a corollary to this is that the microsimulation of large-scale networks could be drastically simplified by removing a large part of the secondary road network (at least for an analysis of aggregated traffic states, not for studying local phenomena).

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 Fig. 6. 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.

The single reservoir modeling of the area includes a unique macroscopic route of length L. The solver used is the accumulation-based model. The resulting total accumulation and mean speed evolution from the MFD simulation is then compared against the real data (07–11/02/2011 and 14–18/02/2011). The results are shown in Fig. 7. The MFD simulation is found to reproduce the evolution of traffic states quite well, as seen in accumulation or mean speed profiles. We notice an underestimation of the accumulation during the night however, which might be due to an inaccurate calibration of the demand profile at this time. The average trip length or the equipped length factor could also be different during this period, but in any case, the period of interest (working hours) for traffic monitoring is well predicted enough. This simulation of a single reservoir provides a first validation for the MFD modeling, and notably allows the calibration of the city-wide equipped length factor.

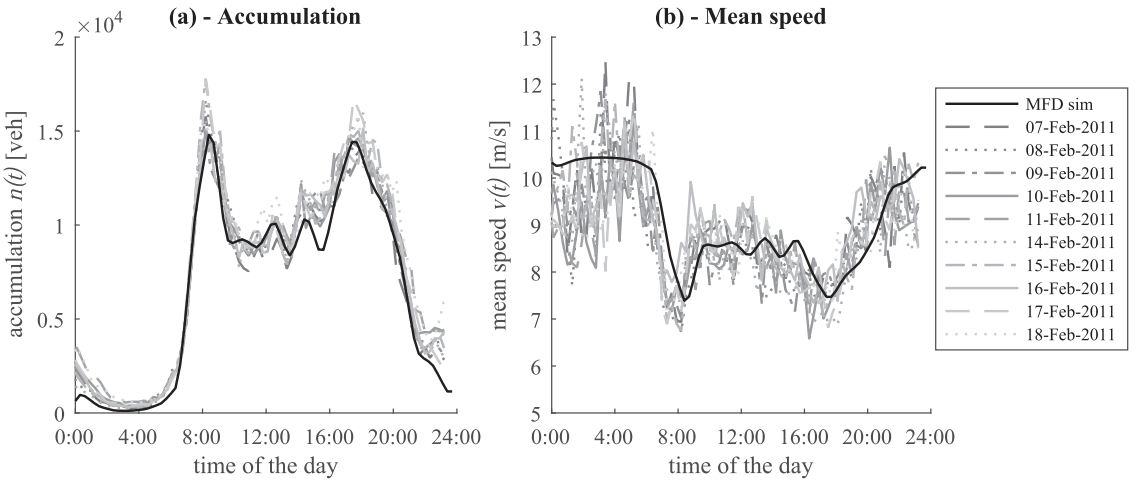


Fig. 7. Comparison between real data and MFD single reservoir simulation. (a) Evolution of total accumulation and (b) mean speed.

4.2. The 5-reservoir case

In this second test case, the city is split into 5 reservoirs as illustrated in Fig. 4(a). The main difference compared to the single reservoir case is that now each reservoir is loaded according to a path flow distribution which is a priori unknown. The consequence is that the prior estimation of the equipped length factors Γ^r in each reservoir r is no longer possible, because the method to calibrate them requires the total demand loading in each reservoir. But this loading actually depends on the (unknown) path flow distribution. As presented in Section 2.2, we will employ two methods to derive a rough estimation of this distribution. The first one is based on Wardrop's principle with a MSA convergence loop. This loop requires the network characteristics (i.e. supply) do not change over its iterations, which in particular means that the equipped length factors must be set one for all before running the simulation. Therefore, we adopt the following methodology. Assuming that the final path flow coefficients will not be very different from the first MSA iteration, we load the reservoirs instantaneously with a static assignment (which also supposes that free-flow travel times, as well as spillbacks between reservoirs, are neglected). We can then estimate the equipped length factors by comparing the total production given by the static assignment and the value given by real data in each reservoir. The values we obtain are detailed in Table 1. We can see that the equipped length factors and total length factors are similar to the city-wide estimations of $\Gamma = 0.062$ and $\gamma = 0.35$. We can check this by recalculating the city-wide equipped length factor by aggregating the values from the 5 reservoirs: $\Gamma \approx \sum_{r=1}^{N_R} P_{\text{equip}}^{\text{midday},r} / \left(\sum_{r=1}^{N_R} P_{\text{equip}}^{\text{midday},r} / \Gamma^r \right) = 0.066$.

The second method to determine the path flow distribution consists in optimizing the path flow coefficients a_p for each route p so that the total load in each reservoir match the real data. As introduced in Section 2.2, the match to real data is checked for a stationary phase of the simulation (i.e. at midday). But unlike what was presented in the latter section, the optimization problem is slightly modified here to include new decision variables, i.e. the equipped length factors. Hence we

Table 1
Equipped length factors, total length factors and accumulation RMSE for the 5-reservoir simulation.

Simulation	Parameter	Notation	R_1	R_2	R_3	R_4	R_5
(All)	Min equipped length factor	Γ^r_{min}	0.022	0.024	0.034	0.022	0.002
Wardrop	Equipped length factor	Γ^r	0.075	0.055	0.095	0.073	0.003
	Total length factor	γ^r	0.29	0.43	0.35	0.31	0.58
	mean RMSE [veh]	Δn^r_{rmse}	220	280	274	380	486
Optimization	Equipped length factor	Γ^r	0.054	0.052	0.094	0.076	0.004
	Total length factor	γ^r	0.4	0.45	0.36	0.29	0.43
	mean RMSE [veh]	Δn^r_{rmse}	284	265	278	354	366

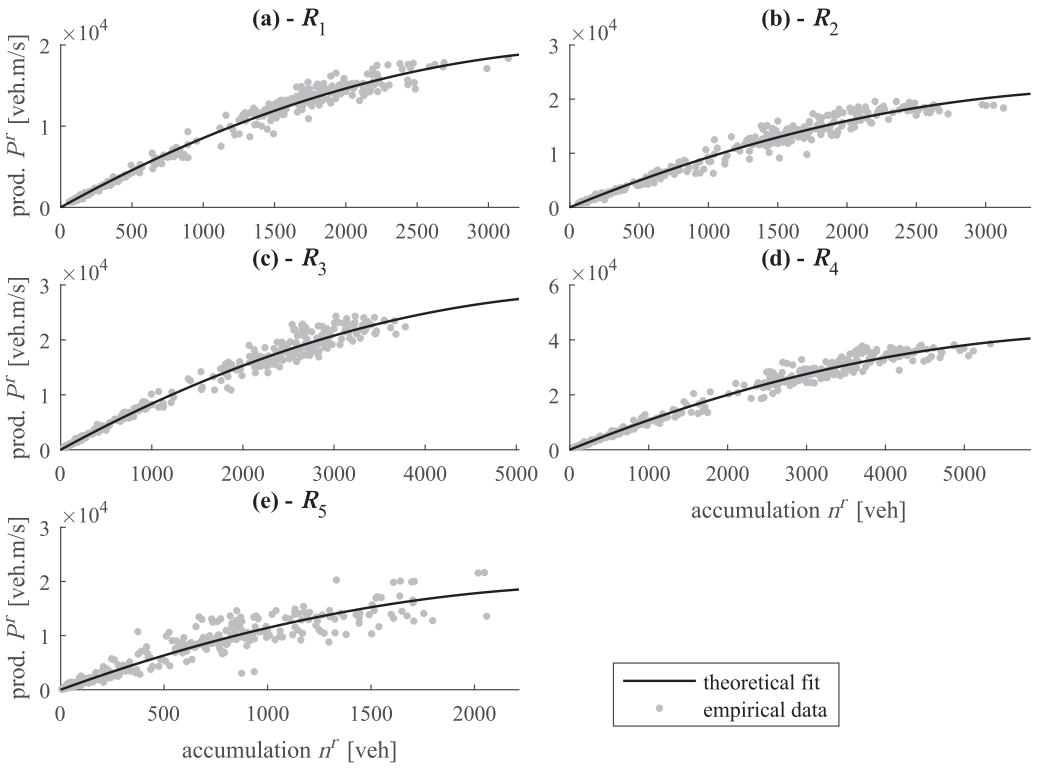


Fig. 8. Production-MFD $P^r(n^r)$ of each reservoir r for the 5-reservoir partitioning. (a) Reservoir R_1, (b) R_2, (c) R_3, (d) R_4 and (e) R_5.

solve:

$$\min \sum_{r=1}^{N_R} \left(P^r_{\text{sim}} - P^{\text{midday},r}_{\text{equip}} \right)^2, \quad \text{where: } P^r_{\text{sim}} = \Gamma^r \sum_{p \in \mathcal{P}^r} L^r_p . a_p . \lambda^{OD}(t_{\text{midday}}) \qquad \text{(Optimized flow distribution)} \tag{24}$$

s.t.

$$\forall p, \quad 0 \leq a_p \leq 1 \tag{25}$$

$$\forall OD, \quad \sum_{p \in \mathcal{P}^{OD}} a_p = 1 \tag{26}$$

$$\forall r, \forall m \in \mathcal{M}^r, \quad \sum_{p \in \mathcal{P}^m} a_p . \lambda^{OD}(t_{\text{midday}}) \leq C_m \tag{27}$$

$$\left| \frac{\sum_{r=1}^{N_R} P^{\text{midday},r}_{\text{equip}}}{\sum_{r=1}^{N_R} P^{\text{midday},r}_{\text{equip}} / \Gamma^r} - \Gamma \right| < \epsilon \tag{28}$$

where Γ is the city-wide equipped length factor, equal to 0.062, and ϵ is a tolerance margin. The productions $P^{\text{midday},r}_{\text{equip}}$ of the equipped network are estimated with the first days of data (01–04/02/2011). This is a non-linear and certainly non-convex problem. Fortunately, we are not looking for the global solutions but for an admissible solution with reasonable fit with the data. This flow optimization method allows us to estimate a set of equipped length factors, different from the first method based on Wardrop assignment, as shown in Table 1. Both methods give similar trends however. The special case of reservoir R_5 is worth noticing, as it exhibits a much lower equipped length factor compared to the ones of R_1 to R_4. This is explained by a very scattered data due to a low coverage of this reservoir by both loop detectors and taxi trips. The ratio of equipped network length over the total network length is indeed very low compared to other reservoirs, because this part of the Lyon network corresponds to a residential area with low traffic volumes, less signalized intersections and thus less loop detectors.

The production-MFD of each reservoir is shown in Fig. 8. The data points are calculated via the formulas in Eq. 19 and Eq. 21 applied to the first 4 days used for MFD calibration (01–04/02/2011), with the equipped length factors obtained with

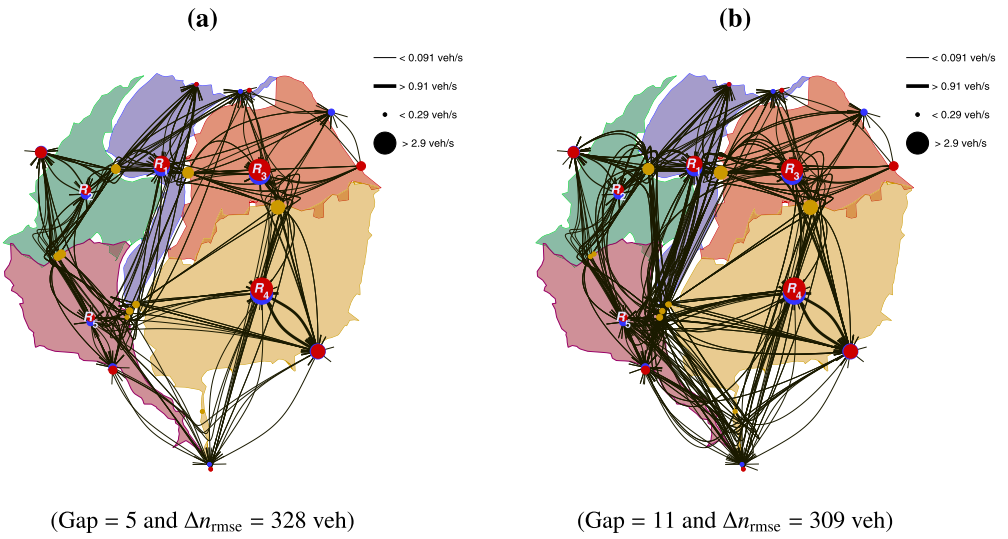


Fig. 9. Routes configuration for the MFD 5-reservoir simulation. (a) Route paths and macroscopic nodes with corresponding flows for the Wardrop flow distribution, and (b) the optimized flow distribution.

the Wardrop flow distribution. 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 R_5 which is quite scattered, for the reasons reported before.

Fig. 9 depicts the path flow distribution for both estimation methods (Wadrop and optimized flow distribution). The macroscopic nodes are represented with dots of different colors at symbolic locations (inside or at reservoir borders): blue symbolizes origin or entry, red destination or exit, and yellow border transfer. The size of the dots is proportional to the effective flow (averaged over a full day) being either generated, received or transferred through the node depending on its type. The macroscopic routes are then represented with lines connecting these macroscopic nodes. The line thickness is proportional to the demand assigned on the route. In total, there are 308 routes available for all the 138 OD pairs, with 3 routes per OD at maximum. In the first simulation with Wardrop flow distribution, only 158 routes are used. The Gap equals 5 (as calculated with Eq. 12), and there are only 3 % of OD pairs that have a Gap^{OD} above 0.2. The higher Gap^{OD} of a few OD pairs is mostly explained by a penalty we introduced for the routes crossing the macroscopic nodes at the South of R_1. The reason for this penalty is detailed here. As illustrated in Figs. 4(a) and 9(a), the South of R_1 corresponds to the only points of transfer between R_1 and R_5 on the one hand, and between R_1 and R_4 on the other hand, respectively by two bridges over the two rivers crossing the city. Thus, the macroscopic nodes there physically correspond to these bridge connections, and significantly restrict the potential flow transfers between these reservoirs due to their low capacity (around 2 veh/s, to be compared with potential transfers of 10 veh/s). Consequently, the South of R_1 is a critical bottleneck in the simulation, which is likely to bring reservoirs to gridlock if this limitation is not properly handled in the path flow distribution. In practice, the MSA loop is not flexible enough to massively reroute users when they face such a bottleneck, given the high number of routes in the system. A penalty has therefore been put on all the routes crossing the South of R_1, which for instance prioritizes the route $[R_2, R_1, R_3, R_4]$ over $[R_2, R_1, R_4]$ to get from R_2 to R_4. However, these penalties introduce a bias in Wardrop flow distribution, as they may favor non-shortest routes. This explains the higher Gap^{OD} of the corresponding OD pairs, although the total Gap remains low and consistent with a Wardrop equilibrium.

In the second simulation with the optimized flow distribution, 255 routes are used. This means that OD demand flows are more distributed over the possible routes, as compared with the Wardrop flow distribution case. This can be seen in Fig. 9(b) where more connection lines are visible between macroscopic nodes. In this simulation, no penalty on routes crossing the South of R_1 is needed, because the capacity constraint of the macroscopic nodes is already included in the optimization problem, see Eq. 27. But here, the equilibrium found is not really complying with Wardrop's principle: the Gap equals 11, and 19 % of OD pairs have a Gap^{OD} above 0.2.

Finally, the comparison of accumulation evolution between MFD simulation and real data is presented in Fig. 10 for the Wardrop flow distribution case. Overall, a good match is observed between the MFD model and the data, except from R_5 where a lot of scatter has been already noticed in this reservoir. The prediction accuracy is not perfect, as we can see an underestimation of the morning peak in R_1 and R_2, whereas it is overestimated in R_4. More precisely, these comparison results can be analyzed with the Root Mean Square Error (RMSE) of the difference between the accumulation predicted by the MFD model and the one observed in the real data over a given period. In each reservoir r, the RMSE for the data of day d from t_1 to t_2 is expressed as:

$$\Delta n^r_{\text{rmse},d} = \sqrt{\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \left(n^r_{\text{sim}}(t) - n^r_{\text{day }d}(t)\right)^2 dt} \qquad \text{(Accumulation RMSE)} \tag{29}$$

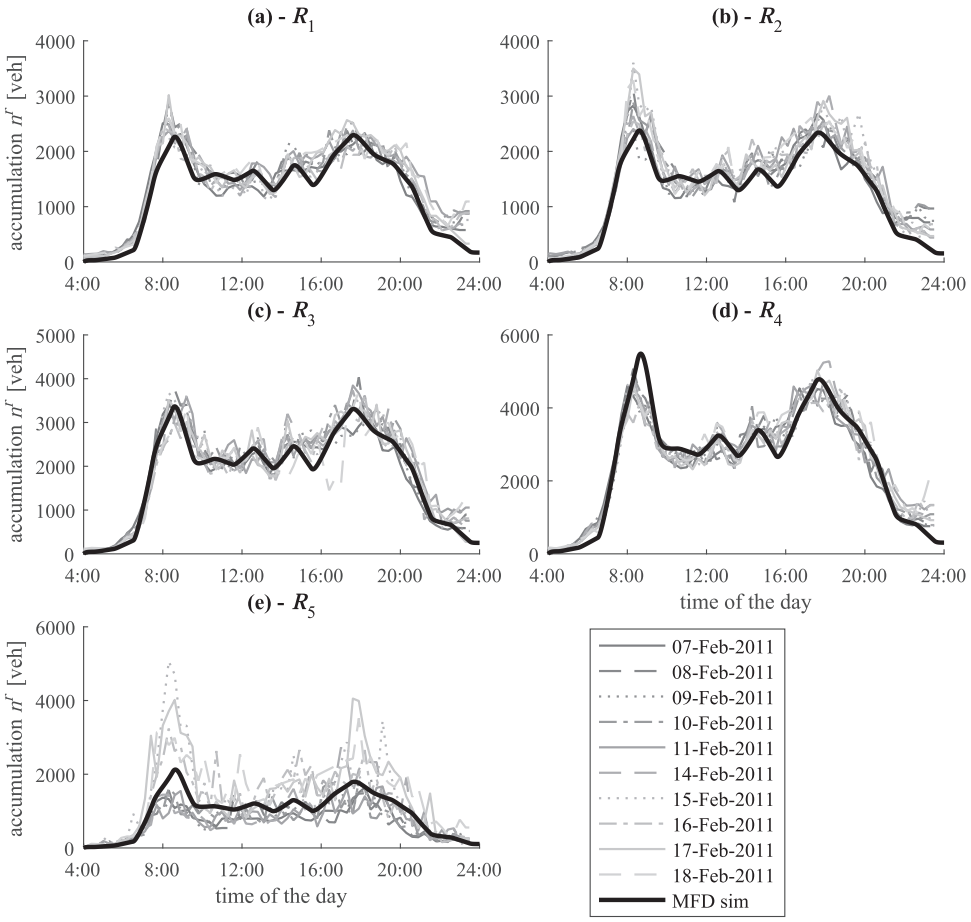


Fig. 10. Comparison between real data and MFD 5-reservoir simulation with the Wardrop flow distribution. (a) Evolution of accumulation in reservoir R_1, (b) R_2, (c) R_3, (d) R_4 and (e) R_5.

where $n^r_{\text{sim}}(t)$ is the accumulation predicted by the MFD simulation and $n^r_{\text{day } d}(t)$ the one observed on day d. Fig. 12(a) displays the RMSE in each reservoir of all days from $t_1 = 4{:}00$ am to $t_2 = 11{:}00$ pm. The differences in prediction accuracy are more visible here, as we notice a lower RMSE for R_1, R_2 and R_3. The scatter in R_5 is also illustrated with various RMSE across the days. Nevertheless, these errors in accumulation remains quite low (below 400 veh) compared to daytime values (around 3000 veh on average). The RMSE of each reservoir averaged over all days is summarized in Table 1, denoted as $\Delta n^r_{\text{rmse}} = 1/10 \sum_{d=1}^{10} \Delta n^r_{\text{rmse},d}$ (calculated for the 10 days of data, 07–11/02/2011 and 14–18/02/2011). The average over all reservoirs gives the following global RMSE: $\Delta n_{\text{rmse}} = 328$ veh.

This simulation case can also validate the hypothesis made to calculate the city-wide average trip length for the single reservoir case, see Eq. 22. Now that we have a reliable set of path flow coefficients $\{a_p\}_{1 \le p \le N_p}$ that provide accurate traffic state predictions, we can calculate the city-wide average trip length L as:

$$L = \frac{\sum_{p \in \mathcal{P}^{OD}} a_p . \overline{\lambda^{OD}} \sum_{r \in R_p} L^r_p}{\sum_{OD} \overline{\lambda^{OD}}} \tag{30}$$

We find $L = 4400$ m, which is close to the value of 4300 m adopted in Section 4.1. Overall, these results show that Wardrop equilibrium can be used at a regional scale to approximate regional path flow distributions for this 5-reservoir partitioning, providing the fact that critical transfer flow limitations are properly handled (the South of R_1 here).

Fig. 11 presents the comparison of accumulation evolution between MFD simulation and real data for the optimized flow distribution case. Here again, a good match is observed between the MFD model, displayed in thick solid black line, and the data. The same morning peak underestimation in R_1 and R_2, and overestimation in R_4 are noticed, but note that the absolute values of accumulation are not the same because the equipped length factors are different in both simulation cases, see Table 1. The good results obtained show that our flow optimization method works properly to derive suitable path flow coefficients and equipped length factors. We should recall that the problem of the estimation of path flow coefficients is under-determined, that is, there are not enough equations to determine the path flow distribution without ambiguity.

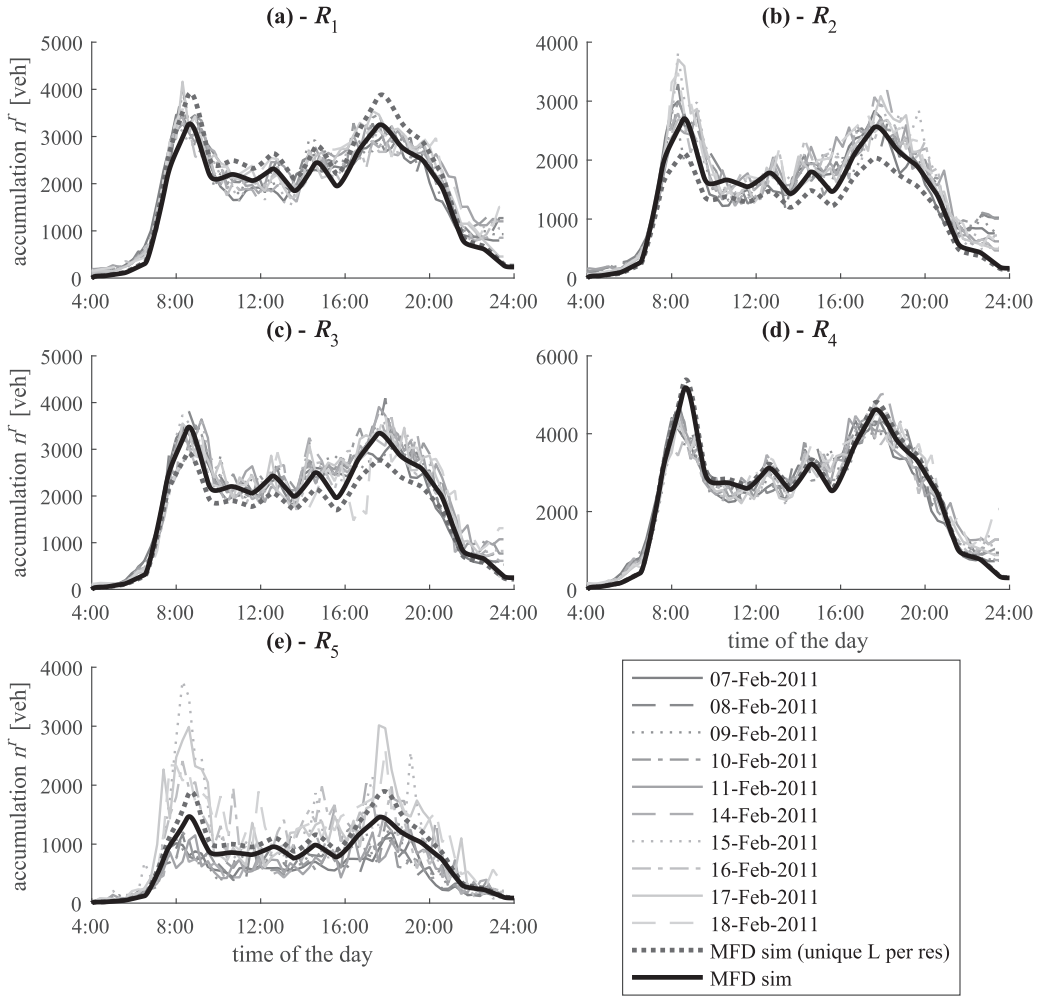


Fig. 11. Comparison between real data and MFD 5-reservoir simulation with the optimized flow distribution. (a) Evolution of accumulation in reservoir R_1, (b) R_2, (c) R_3, (d) R_4 and (e) R_5.

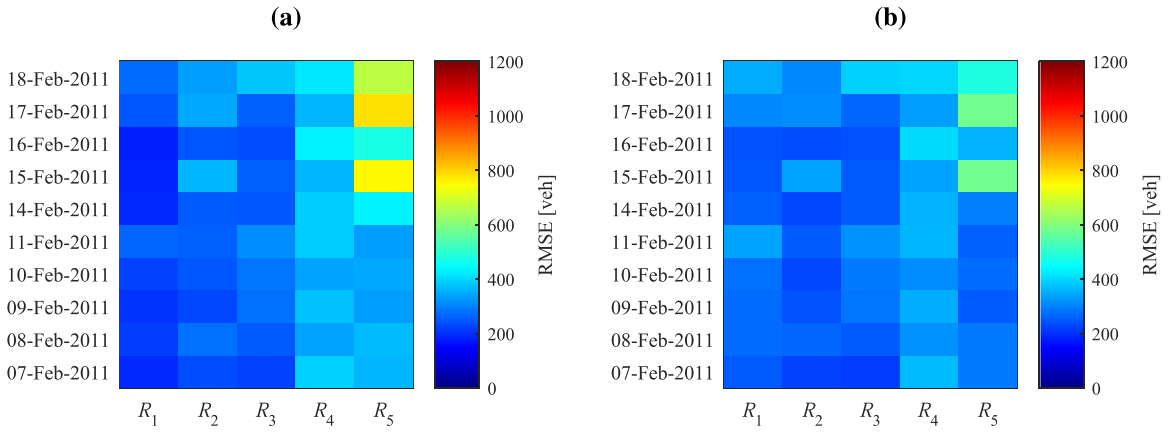


Fig. 12. Difference in accumulation (RMSE $\Delta n^r_{\text{rmse},d}$ for reservoir r and day d) between MFD 5-reservoir simulation and real data. (a) MFD simulation with the Wardrop flow distribution, and (b) the optimized flow distribution.

Therefore, the solution presented here is not unique, and the results of the Wardrop flow distribution case are just an example of another possible combination of path flow coefficients and equipped length factors that works fine as well. While the levels of accumulation are similar between the two simulation case for R_2, R_3 and R_4, note however the difference of around 1000 veh between them for R_1. Like the first simulation with Wardrop flow distribution, the RMSE in each reservoir for each day is presented in Fig. 12(b). Slightly better results are obtained with the optimized flow distribution, as it can also be seen in Table 1 showing the RMSE of each reservoir averaged over all days. The global RMSE (average over all reservoirs and all days) equals $\Delta n_{\text{rmse}} = 309$ veh, a bit lower than the RMSE of 328 veh obtained with the Wardrop flow distribution.

Another MFD model is also compared versus the real data. This one consists in a simplification of the set of trip lengths, it assumes that there is only a single average trip length L^r per reservoir, i.e. we have: $\forall r, \forall p \in \mathcal{P}^r, L_p^r = L^r$. This assumption is notably adopted in every study based on the outflow-MFD instead of the production-MFD (as in e.g., Kouvelas et al., 2017; Sirmatel and Geroliminis, 2017). The simulation results are displayed in thick dotted black line in Fig. 11 for this MFD modeling approach. We can see that this trip length simplification leads to inaccurate traffic predictions in R_1, R_2 and R_3. With this MFD model, the global RMSE is 429 veh (averaged over all reservoirs and days), which is significantly higher than what was obtained with the previous MFD model. As already pointed out by Yildirimoglu and Geroliminis (2014); Leclercq et al. (2015); Mariotte and Leclercq (2019b); Paipuri et al. (2019), this underlines the critical role of trip lengths in MFD modeling.

4.3. The 10-reservoir case

In this third and last test case, the city is split into 10 reservoirs as depicted in Fig. 4(b). In total, there are 992 routes available for 385 OD pairs, with 3 routes per OD at maximum. Because of the high number of routes, finding a suitable path flow distribution is more complex with this partitioning configuration. Unlike the 5-reservoir case, the first method based on

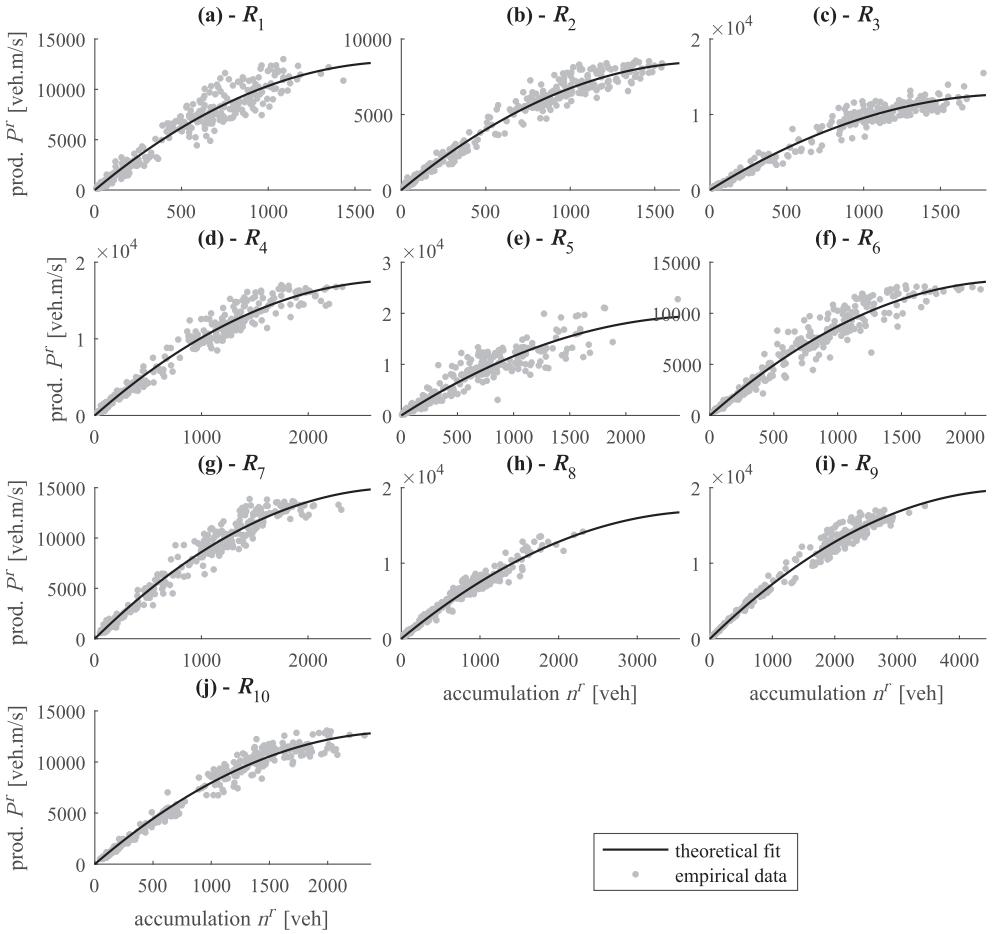
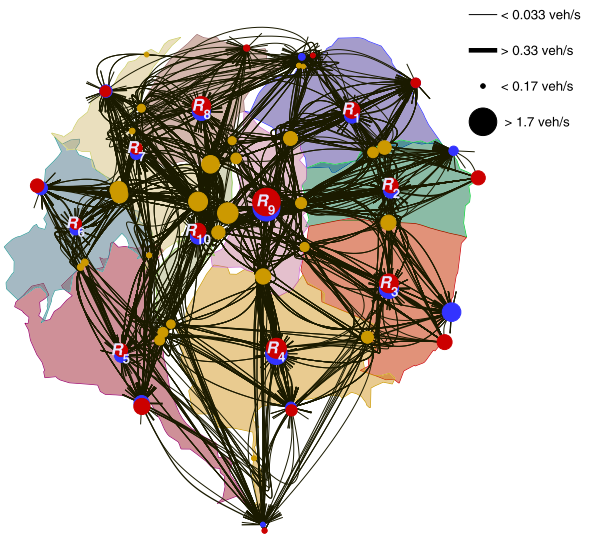


Fig. 13. Production-MFD $P^r(n^r)$ of each reservoir r for the 10-reservoir partitioning. (a) Reservoir R_1, (b) R_2, (c) R_3, (d) R_4, (e) R_5, (f) R_6, (g) R_7, (h) R_8, (i) R_9 and (j) R_{10}.

Table 2
Equipped length factors and total length factors for the 10-reservoir simulation.

Parameter	Notation	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
Min equipped length factor	$\Gamma^r_{\min}$	0.017	0.017	0.014	0.025	0.002	0.011	0.037	0.012	0.072	0.043
Equipped length factor	Γ^r	0.062	0.055	0.052	0.087	0.003	0.025	0.063	0.026	0.11	0.072
Total length factor	γ^r	0.27	0.31	0.28	0.29	0.62	0.43	0.59	0.45	0.65	0.59
mean RMSE [veh]	Δn^r_{rmse}	239	240	235	289	625	390	411	314	428	439



(Gap = 64 and $\Delta n_{\text{rmse}} = 361$ veh)

Fig. 14. Routes configuration for the MFD 10-reservoir simulation, route paths and macroscopic nodes with corresponding flows for the optimized flow distribution.

Wardrop's principle to determine path flows is no longer applicable, unless unreasonable assumptions are made as detailed here. In this case, the implementation of Wardrop's principle always leads to a global gridlock of the MFD simulation due to overloading of some reservoirs, despite the MSA loop which tries to balance the flows. An eventual remedy to this is to tune the equipped length factor of each reservoir Γ^r down to a low value, so that the capacity of the reservoirs is artificially increased, see Eq. 19. Our tests showed that we can run a simulation without gridlock if we have $\Gamma^r \approx 0.02$ in reservoirs on average. This means that each reservoir capacity is approximately increased by a factor 3 compared to the situation of the single reservoir or the 5-reservoir modeling where the equipped length factors were found to be around 0.06. However, adopting low Γ^r factors very different from the city-wide equipped length factor Γ is hardly justifiable, because that would mean that the 10-reservoir simulation is not consistent with the MFD of the whole city. Note that the city-wide MFD has been estimated with $\Gamma = 0.062$ in Section 4.1 with a simple but reliable hypothesis of steady state conditions at midday, and without assuming any kind of path flow distribution. This is why we consider this estimation of Γ as the most reliable one. If having $\Gamma^r \approx 0.02$ in reservoirs is the only solution to run a simulation with the Wardrop flow distribution that does not converge to gridlock, then this estimation of the path flow distribution is likely to be inaccurate.

Consequently, we choose to rely on our second optimization method (i.e. the optimized flow distribution) that approximates both the path flow distribution and the equipped length factors. This method is detailed in Section 4.2, see Eq. 24. It uses the first 4 days of data (01–04/02/2011) for estimating $P^{\text{midday},r}_{\text{equip}}$. Its application leads to reliable equipped length factors Γ^r and total length factors γ^r, as given in Table 2. We can then recalculate the city-wide equipped length factor by aggregating the values from the 10 reservoirs: $\Gamma \approx \sum_{r=1}^{N_R} P^{\text{midday},r}_{\text{equip}} / \left(\sum_{r=1}^{N_R} P^{\text{midday},r}_{\text{equip}} / \Gamma^r \right) = 0.059$, which is consistent with our reference estimation of 0.062 obtained with the single reservoir case.

The production-MFD of the reservoirs are presented in Fig. 13. As previously, the MFD are calibrated in using the first days of data (01–04/02/2011). Parabolic functions are found to fit the data quite well. As in the 5-reservoir case, we notice some scatter in R_5 which also corresponds to the residential area in the South-West part of the city. This scatter is the reason for the much lower equipped length factor $\Gamma^5 = 0.003$ of R_5 as compared to other reservoirs, see Table 2.

The path flow distribution is illustrated in Fig. 14, where each line indicates a route. Each route connects a sequence of macroscopic nodes, displayed as dots in the figure. Each line thickness is proportional to the flow assigned to the corresponding route, and each dot size is proportional to the flow transferring through the macroscopic node. Almost all the

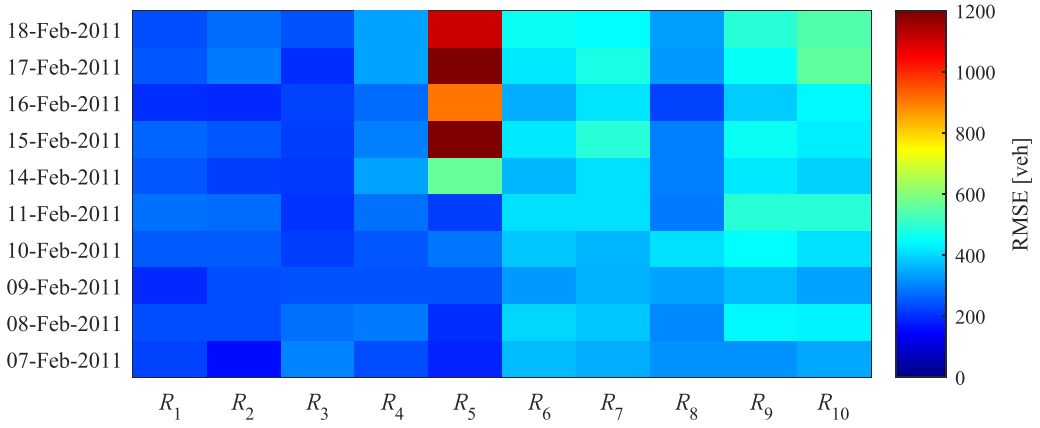


Fig. 15. Difference in accumulation (RMSE) between MFD 10-reservoir simulation and real data.

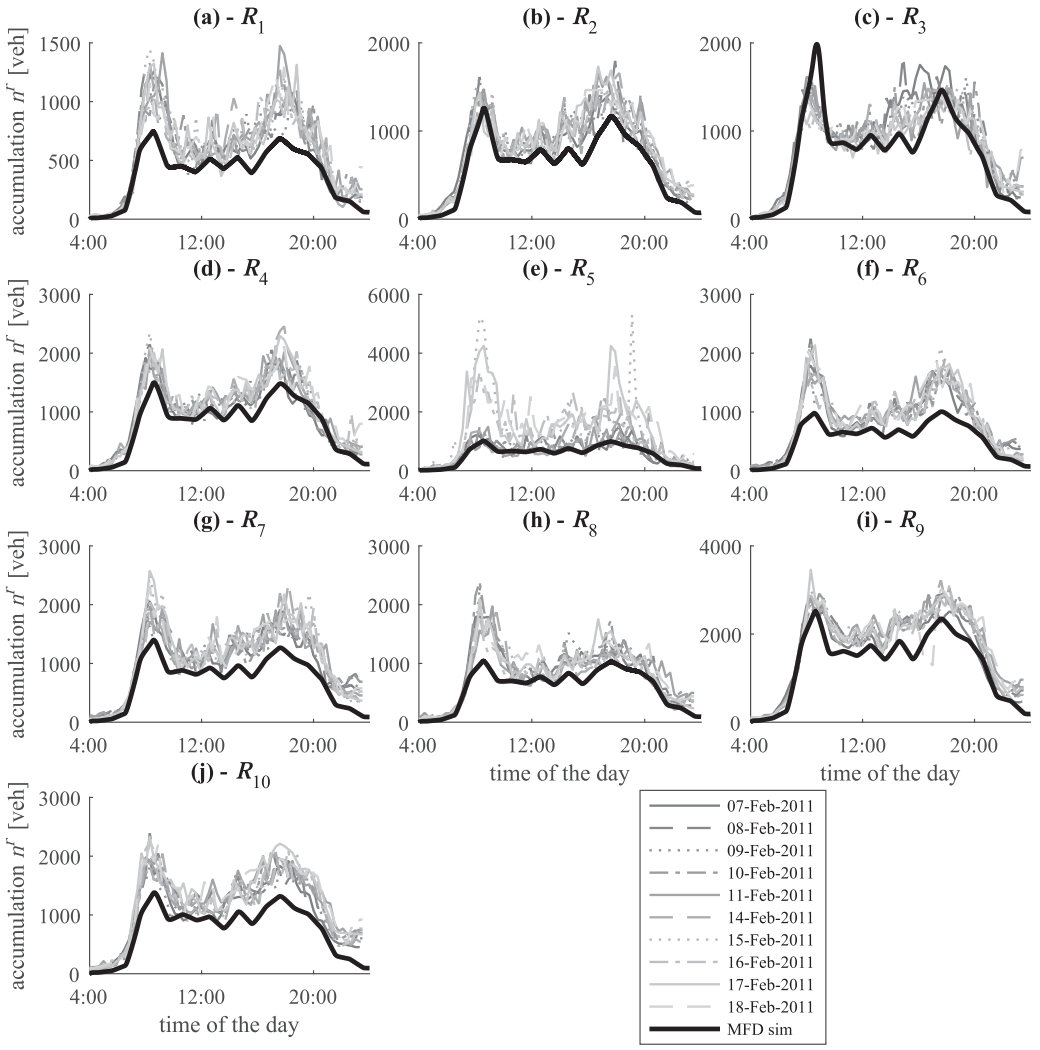


Fig. 16. Comparison between real data and MFD 10-reservoir simulation with the optimized flow distribution. (a) Evolution of accumulation in reservoir R_1, (b) R_2, (c) R_3, (d) R_4, (e) R_5, (f) R_6, (g) R_7, (h) R_8, (i) R_9 and (j) R_{10}.

available routes are used, a total of 909 among the 992 available. This underlines a large distribution of demand flows among all possible options. The equilibrium found is far from Wardrop's principle, as the Gap equals 64 and 34 % of OD pairs have a Gap^{OD} above 0.2.

The comparison between the MFD simulation and the real data is shown in Fig. 16 with the evolution of the total accumulation in reservoirs. As in the previous sections, the observations from the real field correspond to the following days of the month (07–11/02/2011 and 14–18/02/2011). The match to the real data is not as good as in the 5-reservoir case, as we see that the MFD simulation mostly underestimates the accumulation actual evolution. The error of prediction is also reported in Fig. 15 through the RMSE of accumulation $\Delta n^r_{\text{rmse},d}$ calculated as in Eq. 29. This figure reveals higher errors in the West part and center of the city, in R_5, R_6, R_7, R_9 and R_{10}. The high discrepancies in R_5 are explained by the scatter of the data due to a low coverage of loop and probe data in this area. The slightly higher errors in the other above-mentioned reservoirs are likely to be the result of an inaccurate path flow distribution or a bad estimation of the equipped length factors. The values of RMSE averaged over all days of data Δn^r_{rmse} are also mentioned in Table 2. The method for the optimized flow distribution probably reaches its limits, because it provides a path flow distribution valid at midday only, and under the assumption of steady state and undersaturated conditions. Moreover, the estimated path flow coefficients are supposed constant for the whole day, and in particular during peak hours. Nevertheless, the global RMSE Δn_{rmse} equals 361 veh, which is still acceptable, albeit a bit higher than the values found with the 5-reservoir case.

5. Conclusion and discussion

In this paper, we have carried on one of the first validation and calibration of a MFD multi-reservoir simulation. Our test case was the city of Lyon, for which two partitioning cases have been studied: one with 5 reservoirs, and the other with 10 reservoirs. A real dataset from February 2011 with loop detectors and (taxi) probe vehicles has been used to calibrate the MFD model and test its prediction accuracy by comparing the evolution of total accumulation in reservoirs. The MFD model framework is based on the previous theoretical study of Mariotte and Leclercq (2019a). We have notably shown that its prediction accuracy greatly depends on three components: (i) the total active network length that carries the main circulating flow, (ii) the average trip lengths of the macroscopic routes crossing the reservoirs, and (iii) the path flow distribution of the regional demand on these routes.

The first component is crucial for the proper scaling of the total production in a reservoir based on observations from a small set of links (the part of the network equipped with loops). Thanks to steady state considerations at midday, it was found that only 35 % of the total network length (including all secondary roads) must be considered as the total active network length (representing main traffic loading) at the whole city scale.

Regarding the second component, the method developed by Batista et al. (2019) was found satisfactory to estimate the trip lengths of the routes based on shortest path computations in the city map only. In particular, we showed that keeping multiple trip lengths per reservoir led to better results than averaging all these lengths into one single trip length in each reservoir.

Finally, we pointed out that the third component, the regional path flow distribution, is very critical and non trivial to determine. In this study, we proposed two methods that could roughly estimate it. The first one is based on Wardrop's principle with a MSA convergence loop, while the second one uses an optimization of flows at midday to match the loading of reservoirs to the real observations. The first method is a common and easily tractable approach, but may be too naive in many situations. On the contrary, the second method should always ensure a consistent path flow distribution by construction (it consists of an optimization problem in which any constraint can be embedded), but is hardly tractable and may lead to many acceptable solutions as the problem is often under-determined (high number of routing options for a low number of indirect observations). The comparisons between the MFD model predictions and the real data showed that the first method was suitable in the 5-reservoir case, but not in the 10-reservoir case. This is actually explained by the lower number of reservoirs in the first case. Here, spatially close origins and destinations are likely to be included in reservoir internal trips, thus limiting the number of macroscopic routes, and more distant origins and destinations have often an obvious shortest reservoir sequence option. This is no longer true in the 10-reservoir case, where a higher number of reservoirs offers more path choices which are not obvious to discriminate. For this case, only the second method based on flow optimization was appropriate. The high value of the Gap found with this method tends to confirm that the network equilibrium is actually far from Wardrop's principle in the 10-reservoir case. More investigations on traffic assignment at aggregated levels are required to confidently draw more general conclusions. This would require specific data about the user paths at the whole city scale.

Despite these valuable insights, this study also has some limitations that we have previously acknowledged. One of the main assumptions is to consider that the discrepancy observed at the city scale between the total demand loading and the production measurements comes from the definition of the total active network length only, i.e. from the scaling-up of the equipped network measurements with the equipped length factor. But due to the low coverage by loop sensors (less than 3 % of the total network is equipped), the real observations could be not representative of the full network in case of heterogeneous traffic states in the reservoirs. In our case, we have argued that the homogeneity of traffic states is confirmed by the good definition of the estimated MFD. Moreover, our probe dataset is also limited, first because the penetration rate of taxis is low (which again raises the question of representativeness), and second because taxis may experience slightly higher speeds compared to normal cars as they can sometimes drive on dedicated lanes. This could introduce a bias in

mean speed (and thus accumulation with our calculations) and trip length estimations. Here also, we have argued that the orders of magnitude found are consistent with other estimations from Lyon authorities. In any case, the calibration of the multi-reservoir MFD model that we presented in this paper should be definitely improved with larger sets of loop and probe data. Our conclusions on the total active network length and the path flow distribution should also be confronted with similar studies of other cities. Nevertheless, we hope this work can contribute to the further development of MFD simulation as a tool to analyze traffic states at a city scale. This could help understanding better mobility patterns and dynamics at such a scale, together with designing control applications and regulation policies such as perimeter control or route guidance.

CRediT authorship contribution statement

Guilhem Mariotte: Conceptualization, Methodology, Validation, Writing - original draft. **Ludovic Leclercq:** Supervision, Conceptualization, Methodology, Writing - review & editing. **S.F.A. Batista:** Methodology, Writing - review & editing. **Jean Krug:** Data curation. **Mahendra Paipuri:** Methodology, Validation.

Acknowledgements

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

Appendix A. Notation glossary

Table A1
Notations related to each route p.

Notation	Definition [units]
a_p	path flow coefficient [-]
$\lambda_p(t)$	$= a_p.\lambda^{OD}(t)$, demand profile [veh/s]
T_p	average travel time, calculated as the mean horizontal distance between entering and exiting curves [s]
R_p	sequence of reservoirs corresponding to route p
M_p	sequence of macroscopic nodes corresponding to route p
$p^-(r)$	previous reservoir from reservoir r in the sequence R_p
$p^+(r)$	next reservoir from reservoir r in the sequence R_p

Table A2
Notations related to each macroscopic node m.

Notation	Definition [units]
C_m	capacity, maximum flow that can transfer through the macroscopic node m [veh/s]
$\mathcal{L}^m$	set of network links connected to nodes that define the macroscopic node m

Table A3
Notations related to each macroscopic Origin-Destination pair OD.

Notation	Definition [units]
$\lambda^{OD}(t)$	demand profile [veh/s]
$T^{OD}_{\min}$	$= \min_{p\in\mathcal{P}^{OD}} T_p$, minimum average travel time for the OD pair [s]
Gap^{OD}	Gap of the OD pair
$\mathcal{P}^{OD}$	set of routes corresponding to the OD pair
$\mathcal{P}^{OD}_{\min}$	$\subset \mathcal{P}^{OD}$, set of routes having the minimum travel time $T^{OD}_{\min}$

Appendix B. Accumulation-based solver algorithm

The accumulation-based solver is described by the pseudo-code in Algorithm 2. It takes the initial accumulations and demand on all the routes as input, and return the evolution of the system (accumulation, mean speed, inflow/outflow or cumulative count curves for each route).

The *Merge*() algorithm used to calculate inflow supply in reservoirs was proposed in Leclercq and 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.

Algorithm 1: Fair merge with multiple incoming flows.

Function *Merge*$\left(\{\Lambda_i\}_{1\leq i\leq M}, \{\alpha_i\}_{1\leq i\leq M}, C\right)$

Input: set of M incoming demand flows (resp. productions) $\{\Lambda_i\}_{1\leq i\leq M}$ with respective merge coefficients $\{\alpha_i\}_{1\leq i\leq M}$ towards a unique entry with flow (resp. production) capacity C

Output: resulting effective inflows (resp. entering productions) $\{Q_i\}_{1\leq i\leq M}$

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**

    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

Table A4
Notations related to each reservoir r.

Notation	Definition [units]
$n^r(t)$	total accumulation [veh]
$P^r(n^r)$	production-MFD [veh.m/s]
$V^r(n^r)$	$= P^r(n^r)/n^r$, space-mean speed-MFD [m/s]
n^r_j	jam (maximum) accumulation [veh]
n^r_c	critical accumulation [veh]
P^r_c	$= P^r(n^r_c)$, maximum (capacity) production [veh.m/s]
u^r	$= dP^r/dn^r(0)$, free-flow speed [m/s]
$L^r(t)$	dynamic average trip length [m]
$L^r_{\text{tot,max}}$	total maximum network length (including all lanes of all roads in reservoir r) [m]
L^r_{tot}	$< L^r_{\text{tot,max}}$, total active network length (part of the network that carries the main circulating flow) [m]
L^r_{equip}	$< L^r_{\text{tot}}$, total equipped network length (part of the network equipped with loop detectors) [m]
γ^r	$= L^r_{\text{tot}}/L^r_{\text{tot,max}}$, total length factor [-]
Γ^r	$= L^r_{\text{equip}}/L^r_{\text{tot}}$, equipped length factor [-]
Γ^r_{min}	$= L^r_{\text{equip}}/L^r_{\text{tot,max}}$, loop sensor coverage, minimum equipped length factor [-]
$\mathcal{P}^r$	$= \mathcal{P}^r_{\text{int}} \oplus \mathcal{P}^r_{\text{ext}}$, set of routes crossing or originating in reservoir r
$\mathcal{P}^r_{\text{int}}$	$\subset \mathcal{P}^r$, set of routes originating in reservoir r
$\mathcal{P}^r_{\text{ext}}$	$\subset \mathcal{P}^r$, set of routes entering reservoir r from its borders
$\mathcal{M}^r$	set of macroscopic nodes in reservoir r
$\mathcal{M}^r_{\text{int}}$	$\subset \mathcal{M}^r$, set of origin macroscopic nodes in reservoir r (at which the routes in $\mathcal{P}^r_{\text{int}}$ start)
$\mathcal{M}^r_{\text{ext}}$	$\subset \mathcal{M}^r$, set of entry border macroscopic nodes in reservoir r (through which the routes in $\mathcal{P}^r_{\text{ext}}$ enter)
$\mathcal{L}^r$	set of network links (defined with one lane per link) in reservoir r
$\mathcal{L}^r_{\text{equip}}$	set of network links equipped with loop sensors in reservoir r
L^r_p	trip length of route p crossing reservoir r [m]
$n^r_p(t)$	accumulation per route p [veh]
$q^r_{\text{in},p}(t)$	effective inflow per route p [veh/s]
$q^r_{\text{out},p}(t)$	effective outflow per route p [veh/s]
$N^r_{\text{in},p}(t)$	entering cumulative count curve per route p [veh]
$N^r_{\text{out},p}(t)$	exiting cumulative count curve per route p [veh]
$P^r_{s,\text{ext}}(n^r)$	modified entry production supply accounting for internal trip generation in reservoir r [veh.m/s]
$L^r_{\text{ext}}(t)$	average trip length corresponding to the routes originating outside reservoir r only [m]
$\lambda^r_p(t)$	inflow demand from route p to enter reservoir r [veh/s]
$I^r_p(t)$	inflow supply to route p to enter reservoir r [veh/s]
$\mu^r_p(t)$	outflow supply to route p to exit reservoir r [veh/s]
$O^r_p(t)$	outflow demand from route p to exit reservoir r [veh/s]

Algorithm 2: Accumulation-based solver.

Input: reservoir initial accumulation $n_p^r(t_0)$ per route, route demand profile $\lambda_p(t)$, simulation duration T_s and timestep δt_s

Output: reservoir accumulation $n_p^r(t)$, inflow $q_{\text{in},p}^r(t)$ and outflow $q_{\text{out},p}^r(t)$ per route

for $t = t_0$ **to** $t_0 + T_s$ **by** δt_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: $\forall m \in \mathcal{M}_{\text{int}}^r, \{I_p^{r*}\}_{p \in \mathcal{P}^m} = \text{Merge}\left(\{\lambda_p^r\}_{p \in \mathcal{P}^m}, \left\{ \frac{\alpha_p^r}{\sum_{k \in \mathcal{P}^m} \alpha_k^r} \right\}_{p \in \mathcal{P}^m}, C_m \right)$

    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: $\begin{cases} \forall p \in \mathcal{P}_{\text{int}}^r, q_{\text{in},p}^r(t) = \lambda_p(t) \\ \forall p \in \mathcal{P}_{\text{ext}}^r, q_{\text{in},p}^r(t) = q_{\text{out},p}^{p^-(r)}(t) \end{cases}$

    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

References

Aboudolas, K., Geroliminis, N., 2013. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. Transportation Research Part B: Methodological 55, 265–281. doi:10.1016/j.trb.2013.07.003. URL: www.sciencedirect.com/science/article/pii/S0191261513001185

Aghamohammadi, R., Laval, J.A., 2020. A continuum model for cities based on the macroscopic fundamental diagram: A semi-lagrangian solution method. Transportation Research Part B: Methodological 132, 101–116. doi:10.1016/j.trb.2019.04.011. URL: www.sciencedirect.com/science/article/pii/S0191261518311573

Ambuhl, L., Loder, A., Leclercq, L., Axhausen, K.W., Menendez, M., 2019. Identifying reproducible macroscopic traffic patterns in a year-long data set. In: Transportation Research Board 98th Annual Meeting. Washington DC.

Ambuhl, L., Menendez, M., 2016. Data fusion algorithm for macroscopic fundamental diagram estimation. Transportation Research Part C: Emerging Technologies 71, 184–197. doi:10.1016/j.trc.2016.07.013. URL: www.sciencedirect.com/science/article/pii/S0968090X16301267

Batista, S.F.A., Leclercq, L., 2019. Regional dynamic traffic assignment framework for macroscopic fundamental diagram multi-regions models. Transportation Science 53 (6), 1563–1590. doi:10.1287/trsc.2019.0921.

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. doi:10.1016/j.trb.2019.02.009. URL: www.sciencedirect.com/science/article/pii/S0191261518311603

Daganzo, C.F., 1995. The cell transmission model, part ii: Network traffic. Transportation Research Part B: Methodological 29 (2), 79–93. doi:10.1016/0191-2615(94)00022-R. URL: www.sciencedirect.com/science/article/pii/019126159400022R

Dakic, I., Menendez, M., 2018. On the use of lagrangian observations from public transport and probe vehicles to estimate car space-mean speeds in bi-modal urban networks. Transportation Research Part C: Emerging Technologies 91, 317–334. doi:10.1016/j.trc.2018.04.004. URL: www.sciencedirect.com/science/article/pii/S0968090X18304418

Du, J., Rakha, H., Gayah, V.V., 2016. Deriving macroscopic fundamental diagrams from probe data: Issues and proposed solutions. Transportation Research Part C: Emerging Technologies 66, 136–149. doi:10.1016/j.trc.2015.08.015. Advanced Network Traffic Management: From dynamic state estimation to traffic control

Fu, H., Liu, N., Hu, G., 2017. Hierarchical perimeter control with guaranteed stability for dynamically coupled heterogeneous urban traffic. Transportation Research Part C: Emerging Technologies 83, 18–38. doi:10.1016/j.trc.2017.07.007. URL: www.sciencedirect.com/science/article/pii/S0968090X17301882

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

Geroliminis, N., Daganzo, C.F., 2007. Macroscopic modeling of traffic in cities. In: Transportation Research Board 86th Annual Meeting. 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. URL: www.sciencedirect.com/science/article/B6V99-4S6GRGM-1/2/dd42cae455ac898c88ae053d14330edb
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. doi:10.1016/j.trb.2012.04.004. URL: www.sciencedirect.com/science/article/pii/S0191261512000641
Haddad, J., Zheng, Z., 2018. Adaptive perimeter control for multi-region accumulation-based models with state delays. Transportation Research Part B: Methodological in press. doi:10.1016/j.trb.2018.05.019. URL: www.sciencedirect.com/science/article/pii/S0191261518303096
Ji, Y., Geroliminis, N., 2012. On the spatial partitioning of urban transportation networks. Transportation Research Part B: Methodological 46 (10), 1639–1656. doi:10.1016/j.trb.2012.08.005. URL: www.sciencedirect.com/science/article/pii/S0191261512001099
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. doi:10.3141/2422-01.
Kim, S., Tak, S., Lee, D., Yeo, H., 2019. Distributed model predictive approach for large-scale road network perimeter control. Transportation Research Record: Journal of the Transportation Research Board 2673 (5), 515–527. doi:10.1177/0361198119838521.
Knoop, V.L., Hoogendoorn, S.P., 2014. Network transmission model: a dynamic traffic model at network level. In: Transportation Research Board 93rd Annual Meeting. Washington DC.
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. doi:10.1016/j.trb.2016.10.011. URL: www.sciencedirect.com/science/article/pii/S019126151630710X
Leclercq, L., Becarie, C., 2012. Meso lighthill-whitham and richards model designed for network applications. In: Transportation Research Board 91st Annual Meeting. 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. doi:10.1016/j.trb.2014.01.007. URL: www.sciencedirect.com/science/article/pii/S0191261514000174
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. doi:10.1016/j.trc.2015.05.006. URL: www.sciencedirect.com/science/article/pii/S0968090X15001783
Little, J.D.C., 1961. A proof for the queuing formula. Oper Res 9 (3), 383–387.
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. Sci Rep 7 (14029), 1–11. doi:10.1038/s41598-017-14237-8.
Lu, S., Knoop, V.L., Keyvan-Ekbatani, M., 2018. Using taxi gps data for macroscopic traffic monitoring in large scale urban networks: calibration and mfd derivation. Transportation Research Procedia 34, 243–250. doi:10.1016/j.trpro.2018.11.038. International Symposium of Transport Simulation (ISTS18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS18) Emerging Transport Technologies for Next Generation Mobility, URL: www.sciencedirect.com/science/article/pii/S2352146518303272
Lu, S., Wang, J., van Zuylen, H., Liu, X., 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, pp. 184–188. doi:10.1109/ITSC.2013.6728231.
Mariotte, G., Leclercq, L., 2019. Flow exchanges in multi-reservoir systems with spillbacks. Transportation Research Part B: Methodological 122, 327–349. doi:10.1016/j.trb.2019.02.014. URL: www.sciencedirect.com/science/article/pii/S019126151731175X
Mariotte, G., Leclercq, L., 2019. Heterogeneous perimeter flow distributions and mfd-based traffic simulation. Transportmetrica B: Transport Dynamics 7 (1), 1378–1401. doi:10.1080/21680566.2019.1627954. URL: https://www.tandfonline.com/doi/abs/10.1080/21680566.2019.1627954
Mariotte, G., Leclercq, L., Laval, J.A., 2017. Macroscopic urban dynamics: Analytical and numerical comparisons of existing models. Transportation Research Part B: Methodological 101, 245–267. doi:10.1016/j.trb.2017.04.002. URL: www.sciencedirect.com/science/article/pii/S0191261516307846
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.
Mohajerpoor, R., Saberi, M., Vu, H.L., Garoni, T.M., Ramezani, M., 2019. H-infinity robust perimeter flow control in urban networks with partial information feedback. Transportation Research Part B: Methodological in press. doi:10.1016/j.trb.2019.03.010. URL: www.sciencedirect.com/science/article/pii/S0191261518308609
Mollier, S., Monache, M.L.D., de Wit, C.C., Seibold, B., 2019. Two-dimensional macroscopic model for large scale traffic networks. Transportation Research Part B: Methodological 122, 309–326. doi:10.1016/j.trb.2019.02.016. URL: www.sciencedirect.com/science/article/pii/S0191261518305435
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, pp. 1987–1992. doi:10.1109/ITSC.2013.6728521.
Paipuri, M., Leclercq, L., Krug, J., 2019. Validation of macroscopic fundamental diagrams-based models with microscopic simulations on real networks: importance of production hysteresis and trip lengths estimation. Transportation Research Record: Journal of the Transportation Research Board 2673 (5), 478–492. doi:10.1177/0361198119839340.
Ramezani, M., Haddad, J., Geroliminis, N., 2015. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. Transportation Research Part B: Methodological 74, 1–19. doi:10.1016/j.trb.2014.12.010. URL: www.sciencedirect.com/science/article/pii/S0191261515000028
Saeedmanesh, M., Geroliminis, N., 2016. Clustering of heterogeneous networks with directional flows based on “snake” similarities. Transportation Research Part B: Methodological 91, 250–269. doi:10.1016/j.trb.2016.05.008. URL: www.sciencedirect.com/science/article/pii/S0191261515302605
Saeedmanesh, M., Geroliminis, N., 2017. Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. Transportation Research Part B: Methodological 105, 193–211. doi:10.1016/j.trb.2017.08.021. Supplement C, URL: www.sciencedirect.com/science/article/pii/S0191261517304411
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.
Saeedmanesh, M., Kouvelas, A., Geroliminis, N., 2019. A real-time state estimation approach for multi-region mfd traffic systems based on extended kalman filter. In: Transportation Research Board 98th Annual Meeting. 19–02756. Washington DC.
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. doi:10.3141/2029-03.
Sheffi, Y., 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall Inc, Englewood Cliffs, New Jersey, USA, pp. 322–331. Stochastic User Equilibrium
Shim, J., Yeo, J., Lee, S., Hamdar, S.H., Jang, K., 2019. Empirical evaluation of influential factors on bifurcation in macroscopic fundamental diagrams. Transportation Research Part C: Emerging Technologies 102, 509–520. doi:10.1016/j.trc.2019.03.005. URL: www.sciencedirect.com/science/article/pii/S0968090X18304042
Sirmatel, I.I., Geroliminis, N., 2017. Economic model predictive control of large-scale urban road networks via perimeter control and regional route guidance. IEEE Trans. Intell. Transp. Syst. 19 (4), 1112–1121. doi:10.1109/TITS.2017.2716541.
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. doi:10.1016/j.trpro.2015.09.071. URL: www.sciencedirect.com/science/article/pii/S2352146515002586
Tsubota, T., Bhaskar, A., Nantes, A., Chung, E., Gayah, V.V., 2015. Comparative analysis of traffic state estimation cumulative counts-based and trajectory-based methods. Transportation Research Record: Journal of the Transportation Research Board 2491, 43–52. doi:10.3141/2491-05.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. In: Road Engineering Division Meeting, London, pp. 325–362.
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. doi:10.1016/j.trc.2017.08.014. URL: www.sciencedirect.com/science/article/pii/S0968090X17302243

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

Yildirimoglu, M., Sirmatel, I.I., Geroliminis, N., 2018. Hierarchical control of heterogeneous large-scale urban road networks via path assignment and regional route guidance. Transportation Research Part B: Methodological 118, 106–123. doi:10.1016/j.trb.2018.10.007. URL: www.sciencedirect.com/science/article/pii/S0191261518301152

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. doi:10.1016/j.trpro.2017.05.051. Papers Selected for the 22nd International Symposium on Transportation and Traffic Theory Chicago, Illinois, USA, 24-26 July, 2017. URL: www.sciencedirect.com/science/article/pii/S2352146517303290

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, pp. 820–825. doi:10.1109/ITSC.2012.6338712.