Transportation Research Part B 122 (2019) 192–217

Contents lists available at ScienceDirect

Transportation Research Part B

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

Estimation of regional trip length distributions for the calibration of the aggregated network traffic models

S.F.A. Batista[a,b], Ludovic Leclercq[a,*], Nikolas Geroliminis[b]

[a] University of Lyon, ENTPE, IFSTTAR, LICIT, Lyon F-69518, France
[b] School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ARTICLE INFO

Article history:
Received 3 December 2018
Revised 15 February 2019
Accepted 19 February 2019
Available online 26 February 2019

Keywords:
Trip lengths
Regional network
Regional path
Macroscopic fundamental diagram
Trip-based macroscopic fundamental diagram model

ABSTRACT

The calibration of trip lengths is an important challenge for multi-regional MFD-based applications, as it can influence the dynamics of regional densities (and speeds). Existing research has not paid significant attention in the topic, giving opportunities for answering some fundamental questions. In this paper, we propose an original methodology to explicitly determine trip length distributions based on a subset of trips in the city network and its partitioning. Since the full description of all realized trips is difficult to estimate for a large city network, we propose to define a set of virtual trips corresponding to a full coverage of potential local origin-destination pairs and shortest-paths in distance. We investigate how different levels of information can influence the accuracy of the multi-regional dynamic MFD-based models, through the estimated trip length distributions. This information ranges from regional trip lengths without any information of origin, destination, previous or next regions up to the specific regional path associated to each trip.

We test this methodology on a simulated environment with a significant amount of real information for a network with 757 links that corresponds to the 6th district of Lyon. We first provide guidance on how to properly define the set of virtual trips. We show that a single average trip length is not representative of all possible trip lengths estimated by the most detailed levels of information for one region. We also highlight that the first level of information assigns similar trip lengths for regional paths that are composed by the same regions, but in a reverse ordered sequence. Nevertheless, this is not the case for the most detailed levels of information since they capture the directed links city network topology. We propose a procedure to update the trip lengths when an update of the regional Origin-Destination matrix changes, without a need for a full recalculation. We also investigate the influence of the trip length tuning on the traffic dynamics of the regional network. The traffic states are modeled by a trip-based MFD model. We show that the setting of the trip lengths influences the traffic dynamics in the regions. Moreover, the traffic conditions predicted by the two most detailed levels of information are close.

* Corresponding author.
E-mail address: ludovic.leclercq@entpe.fr (L. Leclercq).

https://doi.org/10.1016/j.trb.2019.02.009

1. Introduction

A network level aggregated traffic model was early introduced by Godfrey (1969) and later revisited by Herman and Prigogine (1979) and Mahmassani et al. (1984). But only after the works of Daganzo (2007) and Geroliminis and Daganzo (2008) did this kind of traffic models attract more the attention from the scientific community. In these aggregated traffic models, the city network (Fig. 1 (a)) is partitioned into regions (Fig. 1 (b)), where the traffic conditions are approximately homogeneous (Ji and Geroliminis, 2012). The city network encompasses all the connected links that correspond to the existing roads. Let A be the set of directed links and N be the set of nodes that define the city network. To partition a network in a number of regions, it is possible to apply the clustering techniques discussed in the literature (Saeedmanesh and Geroliminis, 2016; 2017; Lopez et al., 2017; Casadei et al., 2018). Let X be the set of the regions that define the regional network (Fig. 1 (c)). The system dynamics inside each region is governed by a conservation equation, where the outflow is given by a well-defined relation between the mean flow and the accumulation, called the Macroscopic Fundamental Diagram (MFD) (Daganzo, 2007). Using traffic data from the city of Yokohama (Japan), Geroliminis and Daganzo (2008) provided ground truth evidence of the existence of the MFD. The existence and properties of this relationship have also been confirmed by other authors (Geroliminis and Sun, 2011; Ambühl and Menendez, 2016; Loder et al., 2017).

The mathematical formulation of the MFD was introduced by Daganzo (2007) for a single region r. The traffic dynamics is governed by a state equation that associates the accumulation of vehicles ($n_r(t)$) with the balance between the inflow ($Q_{in,r}(t)$) and the outflow ($Q_{out,r}(t)$):

$$\frac{dn_r}{dt} = Q_{in,r}(t) - Q_{out,r}(t), t > 0 \tag{1}$$

Depending on the assumption made on $Q_{out,r}(t)$, two models can be distinguished in the literature: the accumulation-based model (Daganzo, 2007; Geroliminis and Daganzo, 2008); and the trip-based model (Arnott, 2013; Fosgerau, 2015; Lamotte and Geroliminis, 2016; Mariotte et al., 2017; Leclercq et al., 2017; Mariotte and Leclercq, 2018). A key ingredient for implementing these models is the definition of the trip lengths L, i.e., the travel distance of vehicles in each region. The simplest assumption is to consider a constant average trip length for all vehicles traveling in the same region (see e.g., Daganzo, 2007). Up to today, the applications of the MFD-based models have mostly been for testing different control algorithms or design perimeter control strategies. The first perimeter flow control strategies were designed for single region cities (see e.g., Daganzo, 2007; Keyvan-Ekbatani et al., 2012; Ekbatani et al., 2013), where a constant average trip length is assumed. Several other authors considered the outflow MFD application in theoretical studies on a 2-region network (see e.g., Geroliminis et al., 2013; Sirmatel and Geroliminis, 2018; Haddad, 2017; Zhong et al., 2017; Yang et al., 2018), where the travel distances were equal for all the vehicles in the same region. Other perimeter control studies have been applied to real city networks, in detailed simulation environments. For example, Aboudolas and Geroliminis (2013) designed a perimeter control strategy where the San Francisco (USA) downtown network was partitioned into three zones. The MFD model was implemented considering a similar trip length value for all vehicles traveling in the same region. Kouvelas et al. (2017) implemented an outflow MFD-based model for a perimeter control application in part of the Barcelona (Spain) city network, divided into four regions, by controlling only inter-transfers among regions and not the inflow of the outer regions of the network (as happened in Aboudolas and Geroliminis (2013)). As with the previous studies, a similar trip length per region associated with the outflow MFD assumption, is considered. Compared to the previous studies, Ramezani et al. (2015) and Sirmatel and Geroliminis (2018) considered variable mean trip lengths in the implementation of an MFD model for perimeter control purposes. The average trip lengths were estimated dynamically for a toy network, based on the exchange flows between adjacent regions and the accumulation of vehicles. Moreover, the trip lengths are kept constant per Origin-Destination regional pair in each region. Route guidance coupled with an MFD-based models has been recently addressed in Yildirimoglu et al. (2015) and Yildirimoglu et al. (2018), with similar approaches to determine the trip lengths.

Little attention has been paid in the literature to the proper setting of trip lengths for MFD-based applications. In general, a similar distance for all vehicles traveling inside the same region is considered. This is a rough approximation, clearly

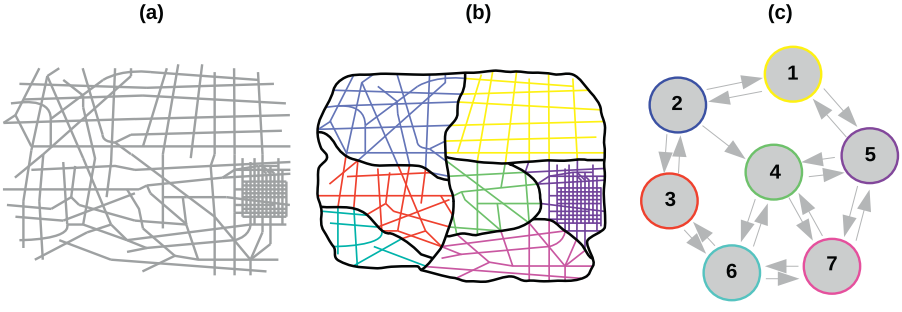

Fig. 1. (a) City network. (b) Partition of the city network. (c) Regional network.

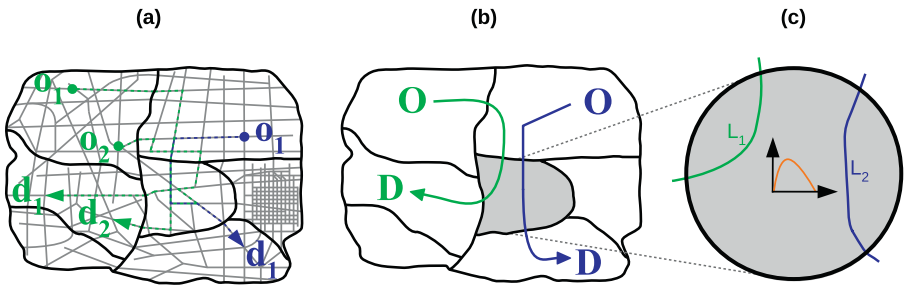

Fig. 2. (a) Partition of the city network with three trips represented. (b) Regional paths that correspond to the blue and green trips. (c) Zoom in the gray region, where the trip lengths distributions L_1 and L_2 of the green and blue regional paths are represented, respectively. A scheme of an MFD diagram is also represented. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

influencing the modeled traffic states even when the trip lengths are distinguished considering different regional paths, e.g., per regional OD pair inside the regions, the trip lengths are assumed to be given the MFD model. In reality, the proper setting and calibration of trip lengths for the application of MFD-based models is a complex task. To better highlight this, consider as an example the partitioned city network in Fig. 2 (a), where three trips are represented. These trips represent vehicle trajectories in the city network. As can be observed, these trips will cross a different sequence of regions from their origin to their destination nodes, defining different regional paths (Yildirimoglu and Geroliminis, 2014). A regional path is the ordered sequence of regions crossed by the trips from their Origin (O) to their Destination (D) regions. Generically, a regional path p can be defined as:

$$p = (p_1, \ldots, p_m, \ldots, p_R), \forall m = 1, \ldots, R \wedge m \in X \tag{2}$$

where R is the number of regions that define p; p_1 and p_R are the Origin (O) and Destination (D) regions, respectively. Batista and Leclercq (2018) distinguishes between regional paths that cross different regions and internal paths defined by trips that only travel inside the same region.

The associated regional paths to the blue and green trips are represented in Fig. 2 (b). Let us now focus on the green trips and the region highlighted in gray in Fig. 2 (b). As can be seen in Fig. 2 (a), the green trips have different travel distances inside this gray region. The latter are also different from the travel distance of the blue trip inside this same region. The green and blue regional paths are therefore characterized by different trip length distributions L_1 and L_2 (Fig. 2 (c)), respectively. As far as we know, Yildirimoglu and Geroliminis (2014) is one of the first efforts that proposes a methodology to set up different travel distances inside the same region, for MFD-based model applications. The authors discussed a methodology where the time-varying average trip lengths are implicitly estimated. Nevertheless, this approach does not allow taking into account explicitly estimated distributions of trip lengths. As far as we know, no study has been performed as yet to describe a methodological framework based on trip patterns and city network partitioning, aimed at explicitly defining distributions of trip lengths for the calibration of dynamic MFD models. The purpose of this paper is to fill this gap. We propose a methodology to explicitly estimate distributions of trip lengths based on a set of trips in the city network and its partitioning, in order to calibrate an MFD-based model. The true trip patterns in a city network are unknown and change over time. The question now is how to calibrate the MFD model when the true set of trips is not available. To this end, we construct a set of virtual trips by randomly sampling origin-destination (od) nodes in the city network and calculating the shortest-path in distance for all od pairs. This process is performed only one time before running the MFD model. Let Γ be the set of virtual trips and N_{od} be the total number of od pairs that are uniformly sampled in the city network. Once the set of virtual trips has been obtained, the question regarding the calculation of trip length distributions is how should we differentiate the trips inside each region? To tackle this question, we propose a methodology in this paper, that considers different levels of information from the regional network. These different levels include:

- *Level 1:* no information on the Origin and Destination regions of the trips.
- *Level 2:* the next region to be traveled by the trips.
- *Level 3:* the previous and next regions traveled by the trips.
- *Level 4:* the specific regional path associated with each trip.

The mathematical formulation of the proposed methodology in this paper is introduced in Section 2. We test this methodology on the 6th district of the Lyon network (France), which is divided into eight regions. In Section 3, we first investigate how the set Γ should be properly calibrated in terms of N_{od}. We then analyze the distributions of trip lengths for each region, considering the four levels of information. Changes in the origins and destinations of the trips can have a significant influence on the trip lengths (Leclercq et al., 2015). Therefore, we introduce a procedure to update the trip lengths, when the OD flows are modified inside the regional OD matrix, without the need to re-sample the set of virtual trips in the city network. The updated OD matrix has the same zoning for origins and destinations but a different number of trips between OD pairs. In Section 3, we consider the static case when the OD matrix is constant over time. In Section 4, we investigate how the trip lengths estimated on the basis of the four levels of information influence the traffic dynamics

in the regions. To do this, we consider the trip-based MFD model as described in Mariotte et al. (2017). We consider two scenarios. In the first scenario, we assume that we have the full information on the trip patterns in the city network. So, we can not only determine the time-dependent evolution of the OD matrix but also the trip distributions for all regional paths. In the second scenario, we still know the time-dependent OD matrix but we compute the trip distributions based on the Deterministic User Equilibrium. In Section 5, we outline the conclusions of this paper.

2. Methodological framework: trip length distributions for the calibration of MFD-based models

One of the core variables for the calibration of an MFD-based model is the setting of the trip lengths, i.e. the distance that vehicles have to travel in a region. In Section 2.1, we describe the role of trip lengths in the formulation of the accumulation-based (Daganzo, 2007; Geroliminis and Daganzo, 2008) and trip-based (Arnott, 2013; Fosgerau, 2015; Lamotte and Geroliminis, 2016; Mariotte et al., 2017; Leclercq et al., 2017; Mariotte and Leclercq, 2018) MFD models. In Section 2.2, we introduce the mathematical basis of the methodological framework proposed in this paper to estimate the distributions of trip lengths. In Table 1 we summarize the notations of all the symbols and variables used in this paper.

2.1. The role of trip lengths in the formulation of MFD-based models

The traffic states inside a region r are modeled through the temporal evolution of a state variable called accumulation $n_r(t)$ (see Eq. (1)). In the literature, two MFD-based models can be distinguished: the accumulation-based model (Daganzo, 2007; Geroliminis and Daganzo, 2008); and the trip-based model (Arnott, 2013; Fosgerau, 2015; Lamotte and Geroliminis, 2016; Mariotte et al., 2017; Leclercq et al., 2017; Mariotte and Leclercq, 2018).

In the accumulation-based model, the outflow function $Q_{out,r}(t)$ (Daganzo, 2007) is defined as the ratio between the production-MFD $P_r(n_r(t))$ and the average trip length $\bar{L}_r$:

$$Q_{out,r} \approx \frac{P_r(n_r(t))}{\bar{L}_r} \tag{3}$$

where $\bar{L}_r$ is the distance that vehicles have to travel in region r to complete their trips.

This outflow function (see Eq. (3)) is then embedded in the dynamic equation that describes the evolution of traffic states in the system. For a single region, the state is the accumulation $n_r(t)$ of the region. For multi-regions the state representation is more complicated as Origin-Destination flows matter. The state representation for multi-regional systems can have different degrees of complexity. The most parsimonious models utilize states of type n_{ij}, where i is the current region and j the final one. Then the same trip length can be assumed for all vehicles crossing a region i independently of where they are coming from or what is the regional path towards their destination (Geroliminis et al., 2013). More complicated applications (Yildirimoglu and Geroliminis, 2014; Geroliminis, 2015; Mariotte and Leclercq, 2018) consider different regional paths crossing the same region r are assigned different trip lengths. Let $\bar{L}_{rp}$ be the average trip length of regional path p inside region r. The outflow function $Q_{out,r}(t)$ is rewritten as:

$$Q_{out,rp} \approx \frac{n_{rp}(t)}{n_r(t)} \frac{P_r(n_r(t))}{\bar{L}_{rp}} \tag{4}$$

where $n_{rp}(t)$ is the accumulation of vehicles traveling on regional path p and inside region r; and $n_r(t)$ is the total accumulation of vehicles in region r. For the simplest models, Eq. (4) simplifies to $Q_{out,ij} \approx \frac{n_{ij}(t)}{n_i(t)} \frac{P_r(n_r(t))}{\bar{L}_i}$.

In the trip-based model (Arnott, 2013; Fosgerau, 2015; Lamotte and Geroliminis, 2016; Mariotte et al., 2017; Leclercq et al., 2017; Mariotte and Leclercq, 2018), the MFD dynamics is centered on the vehicle trip length L_{ir}:

$$L_{ir} = \int_{t_{entry}}^{t_{exit}} v_r(n(s))ds \tag{5}$$

where t_{entry} and t_{exit} are the entry and exit times of the vehicle i in region r, respectively; and $v_r(n(s))$ is the speed-MFD. The travel time of vehicle i in region r is $t_{travel} = t_{exit} - t_{entry}$. We assume that all vehicles traveling on the same regional path p and inside region r have the same travel distance. Moreover, we assume that L_{ir} is time-invariant and that both t_{travel} and $n_r(t)$ are continuous functions and differentiable at t, based on Eq. (5) and under fluid conditions, it is possible to gather the following relation for the outflow $Q_{out,r}(t)$ (Arnott, 2013; Mariotte et al., 2017; Mariotte and Leclercq, 2018):

$$Q_{out,r}(t) = Q_{in,r}(t - t_{travel}) \frac{v_r(n_r(t))}{v_r(n_r(t - t_{travel}))} \tag{6}$$

The setting of trip lengths for the calibration of an MFD model cannot be neglected. In an MFD framework, the vehicles speed is assumed to be homogeneous. Thus, the trip lengths dictate the time that vehicles have to spend to complete their trips inside a region r. Longer trip lengths mean that vehicles spend more time to complete their trips inside the region, increasing the accumulation n_r and thus the modeled traffic states. In the next section, we introduce the methodology to estimate the distance to be traveled by the vehicles inside each region, that will depend on the level of information that we consider from the regional network as stated in the introduction. The influence of the trip lengths tuning on the traffic dynamics modeled by a trip-based MFD model, will be investigated in more detail in Section 4.

Table 1
Nomenclature used in this paper.

City and regional networks:	
o	Origin node.
d	Destination node.
O	Origin region.
D	Destination region.
W	Set of regional OD pairs.
p	Regional path.
R	Number of regions that define regional path p.
X	Set of regions that define the regional network.
r	Region.
k	Virtual trip in the city network.
N	Set of nodes of the city network.
A	Set of directed links of the city network.
N_{nodes}	Total number of nodes in the city network.
N_{links}	Total number of links in the city network.
Ω^{OD}	Regional choice set of regional OD pair.
MFD models:	
t	Time instant.
T	Simulation period.
$n_r(t)$	Accumulation of vehicles in region r.
$Q_{in,r}(t)$	Inflow function of region r.
$Q_{out,r}(t)$	Outflow function of region r.
$P_r(n_r(t))$	Production MFD.
$v_r(n(s))$	Speed-MFD.
n_{ij}	Accumulation of vehicles that travels on region i and goes to j.
$\overline{L}_r$	Average trip length inside region r.
i	Vehicle.
t_{entry}	Entry time of vehicle i in region r.
t_{exit}	Exit time of vehicle i of region r.
t_{travel}	Travel time of vehicle i in region r.
$\overline{L}_{rp}$	Average trip length of regional path p inside region r.
Set of virtual trips:	
Γ	Set of virtual trips.
N_{od}	Number of virtual trips listed in Γ.
$N_{used}^{nodes}(N_{od})$	Total number of nodes visited in the city network.
$N_{used}^{links}(N_{od})$	Total number of links visited in the city network.
Ψ	Set of regional paths defined by the virtual trips listed in Γ.
$\overline{L}_p^{M_x}$	Average trip length of regional path p considering the level x (M_x) of information.
Z and Z^*	Regional OD matrices.
Level 1:	
L_r	Distribution of trip lengths of region r.
$\overline{L}_r$	Average trip length of region r.
l_{rk}	Length of trip k in region r.
δ_{rk}	Binary variable that equals 1 if trip k travels on region r, or 0 otherwise.
Level 2:	
j	Next adjacent region to r.
L_{rj}	Distribution of trip lengths to go from region r to region j.
$\overline{L}_{rj}$	Average trip length to go from region r to j.
δ_{rjk}	Binary variable that equals 1 if trip k travels in region r and goes to region j, or 0 otherwise.
Λ	Set of adjacent regions to r.
γ	Binary variable that equals 1 if p is composed of 2 or more regions, or 0 otherwise.
Level 3: internal trips	
L_r	Distribution of trip lengths of region r.
$\overline{L}_r$	Average trip length of region r.
η_{rk}	Binary variable that equals 1 if k is an internal trip of region r, or 0 otherwise.
Level 3: crossing trips	
h	Previous origin region.
L_{hrj}	Distribution of trip lengths of region r, for trips coming from region h and going to region j.
$\overline{L}_{hrj}$	Average trip length to go from region h to j and crossing region r.
δ_{hrjk}	Binary variable that equals 1 if trip k comes from region h and goes to j by crossing region r, or 0 otherwise.
γ	Binary variable that equals 1 if p is composed 3 or more regions, or 0 otherwise.

(*continued on next page*)

Table 1 (*continued*)

Level 4:	
L_r^p	Distribution of trip lengths for p on region r.
$\bar{L}_r^p$	Average trip length of p on region r.
δ_{rk}^p	Binary variable that equals 1 if trip k travels in region r and both define regional path p, or 0 otherwise.
Other variables:	
ρ	Relative differences of the average trip lengths calculated for the $i-$th level of information, between the sets of $N_{od} = 400$ and 5000 and the reference set of $N_{od} = 10,000$.
δ_{lr}	Binary variable that equals 1 if link l lies inside region r.
δ_{nr}	Binary variable that equals 1 if node n lies inside region r.
β_{ip}	Relative differences between the average trip lengths calculated by the $j-$th level of information ($j = 1, 2, 3$) and the reference level of information M_4.
α_r^{OD}	Scaling factor.
$\hat{L}$	Estimated trip length based on the updated OD matrix.
θ	Relative differences between $\hat{L}$ and $\bar{L}$.
ζ_s	Descent step of the Method of Successive Averages.
s	Descent step iteration of the Method of Successive Averages.
$N_{\max}$	Total number of descent step iterations.
Gap	Relative gap between the regional paths travel time and the travel time at the network equilibrium.
tol	Tolerance for the *Gap* criterion.

2.2. Methodological framework for the estimation of trip length distributions

The methodology proposed in this paper to estimate the trip length distributions, it utilizes a representative set of trips in the city network as well as the definition of its partitioning. To partition the city network, we can apply one of the different methods described in the literature (Saeedmanesh and Geroliminis, 2016; 2017; Lopez et al., 2017; Casadei et al., 2018). To gather the set of trips, the ideal scenario would be to use real trips in the city network. However, the full set of the latter is usually unknown and difficult to estimate. To bypass this obstacle, we construct a set of virtual trips in the city network, by uniformly sampling N_{od} od pairs and calculating a path that connects each of them. In this paper, we consider the shortest-paths in distance estimated by the Dijkstra algorithm. We further note that other approaches can be used to estimate the virtual trips. To name a few examples, we can consider the link elimination (Azevedo et al., 1993) or link penalty (de la Barra et al., 1993) approaches, the branch-and-bound algorithm (Prato and Bekhor, 2006), the doubly stochastic approach (Nielsen, 2000) or the implementation of the Metropolis–Hasting algorithm, as proposed by Flötteröd and Bierlaire (2013). Exploring the influence of the local trips calculation on the generation of the virtual trips set is beyond the scope of this study.

Once the set of virtual trips Γ is defined, to estimate the distributions of trip lengths, we differentiate the trips inside each region r according to four levels of information from the regional network. These levels are:

- *Level 1 (M_1): no information on the regional Origin and Destination of the trips.* This level follows the original idea of Daganzo (2007) for MFD-based models applications and consists of a single average trip length ($\bar{L}_r$) for all vehicles traveling in region r. All trips that cross region r are considered, independently of the previous and the next regions that these trips travel, to estimate an average trip length $\bar{L}_r$ as:

$$\bar{L}_r = \frac{\sum_k l_{rk}}{\sum_k \delta_{rk}}, \forall k \in \Gamma \tag{7}$$

  where l_{rk} is the length of the part of the trip k that occurs in region r; and δ_{rk} is a binary variable that equals 1 if trip k travels on region r, or 0 otherwise.
  The distribution of trip lengths of region r is:

$$L_r = \{\delta_{rk} l_{rk}\}, \forall k \in \Gamma \tag{8}$$

  In Fig. 3 (a) we show the application of this aggregation level for the green region. This region is highlighted and four trips are shown: their origin nodes are represented by black circles and the destination nodes by arrows. Each trip is identified by an identification number. Three of these trips (i.e., trips 1, 2 and 4) simply cross the green region, while one has its origin node inside the green region (i.e., trip 3). To estimate the distribution of trip lengths for the green region, we aggregate the length of the part of these trips that occur inside the green region. These lengths are represented by the solid black lines in Fig. 3 (a). Note that, although the internal trips (i.e. that have origin and destination nodes inside the green region) are not shown, they are nonetheless considered.
- *Level 2 (M_2): next region to be traveled by the trips.* The idea is to consider all the trips that cross region r and that go to the same next adjacent region, independently of their previous adjacent region. We filter the trips on Γ that cross region r and that go to the same next adjacent region j. This estimation is consistent with the dynamics of state accumulations

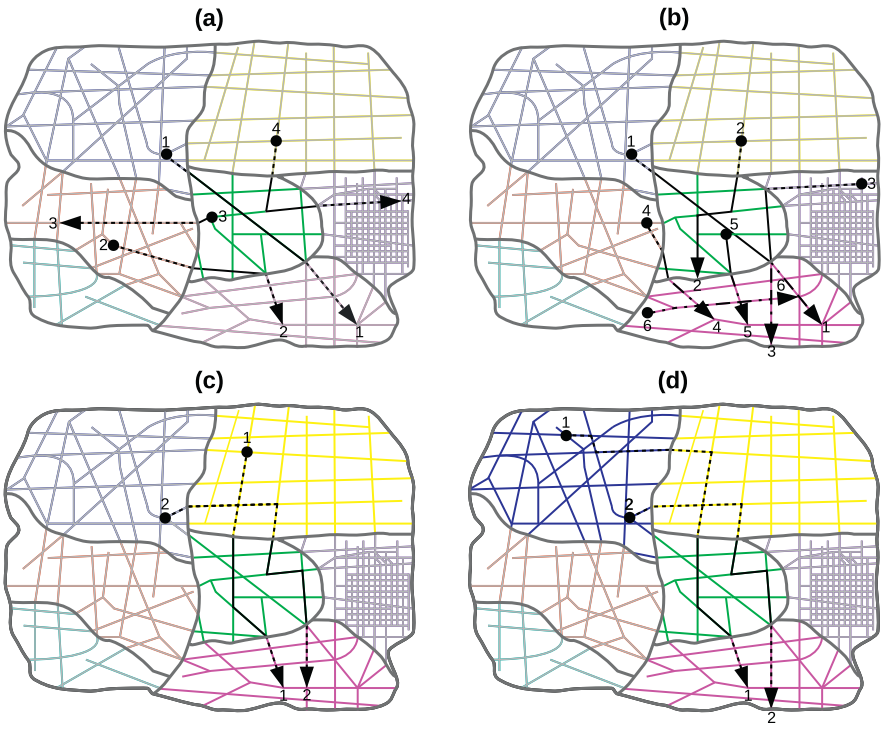

Fig. 3. Example of application of the methodology to estimate the distributions of trip lengths, considering the four levels of information. Levels 1 to 4 are represented in the sub figures (a) to (d), respectively.

that are utilized in different efforts for perimeter control of multi-regional networks (see for example Ramezani et al., 2015). For these works it is not important to know from where vehicles started their trips, but only what is the next region towards the destination.

The average trip length $\bar{L}_{rj}$ to go from region r to j is estimated as:

$$\bar{L}_{rj} = \frac{\sum_k \delta_{rjk} l_{rk}}{\sum_k \delta_{rjk}}, \forall k \in \Gamma \wedge (\forall j \in \Lambda \vee j = r) \tag{9}$$

where δ_{rjk} is a binary variable that equals 1 if trip k travels in region r and goes to region j; and Λ is the set of adjacent regions to r, or 0 otherwise. Note that $j = r$ represents the case of a destination region of the trips.

The distribution of trip lengths of region r that go to j is:

$$L_{rj} = \{\delta_{rjk} l_{rk}\}, \forall k \in \Gamma \tag{10}$$

In Fig. 3 (b) we show the application of this aggregation level. Five examples of trips are shown for the green region. All of these trips have the same adjacent destination region (the pink one). Trip 2 has a destination node on the border between the green and the pink regions. The length of these trips that occur inside the green region are aggregated. This is represented by the solid black lines. Now let us consider the pink region shown in Fig. 3 (b), which is a common destination region of the trips represented in it. Trip 6 represents an example of an internal trip. In this case, the lengths of trips 1 to 6 that occur inside the pink region are aggregated to define the distribution of trip lengths for this region.

- *Level 3 (M_3): the previous and the next regions traveled by the trips.* The idea is to consider all the trips that cross region r and that come from the same previous adjacent region h and go to the same next adjacent region j. The reasoning behind this calculation is that when partitioned, regions have asymmetric and uneven shapes and vehicles that come from different previous regions might have significantly different trip lengths in the current region of consideration. In this case, we must consider two distinct types of trips:
  - The first type corresponds to the internal trips of region r, i.e. internal paths. In this case, the average trip length $\bar{L}_r$ is:

$$\bar{L}_r = \frac{\sum_k \eta_{rk} l_{rk}}{\sum_k \eta_{rk}}, \forall k \in \Gamma \tag{11}$$

where l_{rk} is the length of trip k; and η_{rk} is a binary variable that equals 1 if trip k is an internal trip of region r, or 0 otherwise.

The distribution of trip lengths of region r is:

$$L_r = \{\eta_{rk} l_{rk}\}, \forall k \in \Gamma \tag{12}$$

– The second type corresponds to trips that cross region r. In this case, all trips traveling in region r, coming from the same previous adjacent region h and going to the next adjacent region j are aggregated to define the average trip length $\bar{l}_{hrj}$, to go from region h to j and crossing region r:

$$\bar{l}_{hrj} = \frac{\sum_k \delta_{hrjk} l_{rk}}{\sum_k \delta_{hrjk}}, \forall k \in \Gamma \wedge \forall (h, j) \in \Lambda \wedge h \neq j \tag{13}$$

where δ_{hrjk} is a binary variable that equals 1 if trip k comes from the previous region h and goes to the next region j, by crossing region r, or 0 otherwise.
The distribution of the trip lengths of region r, for trips coming from region h and going to region j is:

$$L_{hij} = \{\delta_{hijk} l_{ik}\}, \forall k \in \Gamma \tag{14}$$

In Fig. 3 (c), we show an example of an application of this case of Level 3 for intermediate regions. To do this, we consider the two trips that come from the yellow region (the equivalent to region h), cross the green region (the equivalent to the region r), and go to the pink region (the equivalent to region j). The lengths of the part of these two trips inside the green region are aggregated together to estimate the distribution of trip lengths L_{hrj}. This is represented by the solid black lines.

- *Level 4 (M_4): the related regional path p defined by each trip.* This is the most general case requiring a significant number of information. That is, we only consider the trips that cross region r, following the same specific sequence hij of regions and that define the same regional path p. Thus, all the trips that cross region r and define the same regional path p, are considered to estimate the average trip length of p in region r ($\bar{L}_r^p$):

$$\bar{L}_r^p = \frac{\sum_k \delta_{rk}^p l_{rk}}{\sum_k \delta_{rk}^p}, \forall k \in \Gamma \tag{15}$$

where δ_{rk}^p is a binary variable that equals 1 if trip k travels in region r and is associated with the regional path p, or 0 otherwise.
The distribution of the trip lengths of the regional path p inside region r is:

$$L_r^p = \{\delta_{rk}^p l_{rk}\}, \forall k \in \Gamma \tag{16}$$

In Fig. 3 (d), we show the application of this aggregation method. Two examples of trips that define the same regional path, defined by the sequence of the blue-yellow-green-pink regions, are shown. To determine L_r^p for the green region, we aggregate the length of the part of these trips inside this region. This is represented by the solid black lines.

Level 1 is the more generic one, where all the trips that cross the same region are considered for the aggregation. Level 4 is the most refined in terms of the level of information considered. It allows defining a specific distribution of trip lengths for a regional path that crosses a given region, based on the specific sub-set of trips that define this regional path. Thus, we consider Level 4 as the reference throughout this paper, even if it might be the most challenging to compute with real data. We reinforce the idea that our goal is to assess the influence of the trip lengths tuning for MFD-based model applications. In the latter case, it is M_1 which is usually considered in the absence of more information on a proper calibration framework. The purpose of this paper is to set up this framework and to investigate the influence on the traffic dynamics in the regions, when the trip lengths are estimated between M_1, M_2 and M_3 and the reference M_4.

The average trip length of regional path p can be directly estimated considering any of the previously discussed methods as follows:

1. *Level 1* (M_1):

$$\bar{L}_p^{M_1} = \sum_{m=1}^{R} \bar{L}_{p_m} \tag{17}$$

2. *Level 2* (M_2):

$$\bar{L}_p^{M_2} = \gamma \sum_{m=1}^{R-1} \sum_{l=m+1}^{R} \bar{L}_{p_m p_l} + \bar{L}_{p_R p_R} \tag{18}$$

where γ is a binary variable that equals 1 if $R \geq 2$, or 0 otherwise.

3. *Level 3* (M_3):

$$\bar{L}_p^{M_3} = \begin{cases} \bar{L}_{p_R} & \text{if } R = 1 \\ \bar{L}_{p_1 p_2} + \gamma \sum_{m=2}^{R-1} \bar{L}_{p_{m-1} p_m p_{m+1}} + \bar{L}_{p_{R-1} p_R} & \text{if } R \geq 2 \end{cases} \tag{19}$$

where γ is a binary variable that equals 1 if $R \geq 3$, or 0 otherwise. Note that: $\bar{L}_{p_R}$ is estimated according to Eq. (11); $\bar{L}_{p_1 p_2}$ and $\bar{L}_{p_{R-1} p_R}$ are estimated according to Eq. (9); and $\bar{L}_{p_{m-1} p_m p_{m+1}}$ is estimated according to Eq. (13).

4. *Level 4* (M_4):

$$\bar{L}_p^{M_4} = \sum_{m=p_1}^{p_R} \bar{L}_m^p \tag{20}$$

Although the four levels of information might give similar estimations of L_p (Eqs. (17)–(20)) for the mean value when the shape of the regions is symmetrical and the distribution of demand and congestion is homogeneous, this might not be the case with the variance. For accumulation-based MFD models only the average trip lengths matter. However, for the trip-based MFD models, the variance of the trip length distributions plays a significant role in the outflow. Lamotte et al. (2018) shows that trip length distributions can easily be implemented in this framework.

The trip lengths depend on factors such as the definition of N_{od}, the regional OD matrix (Leclercq et al., 2015) and the traffic conditions (i.e. congestion patterns). In the next sections, we investigate the influence of the first two factors on the trip lengths. The influence of congestion together with a methodology to estimate time-varying trip length distributions will be investigated in future research.

3. Analysis of the trip length distributions obtained by the four levels of information

In this section, we analyze the differences between the trip lengths estimated through the four levels of information ($M1$ to $M4$). The test network is the 6^{th} district of the Lyon network (France), divided into eight regions. The test network and respective MFD functions are introduced in Section 3.1. In Section 3.2, we investigate the importance of calibrating N_{od} well to obtain the set of virtual trips and estimate the trip length distributions. Based on this analysis, we choose an appropriate set of virtual trips and in Section 3.3 we analyze the regional trip length distributions estimated by the four levels of information. In Section 3.4, we investigate the influence of the regional OD matrix on the trip lengths. We propose a procedure that allows updating the regional trip lengths when the number of trips between the different OD pairs change, without the need to re-sample the set of virtual trips Γ. This is very useful to run different MFD-based model applications for different demand scenarios.

3.1. Network definition

The city network considered for testing the methodology proposed in Section 2.2 is the 6th district of the Lyon network (France) shown in Fig. 4. It is composed of 757 links and 431 nodes. The city network is partitioned into eight regions, for which we fitted the MFD functions (Fig. 4) considering the dimensions of each one. We assume a bi-parabolic shape for the MFD (one parabola for the increasing and one for the decreasing part of the MFD, with a first derivative equal to zero for the critical accumulation that maximizes production).

3.2. Definition of N_{od} for the calculation of trip lengths

In this section, we analyze the importance of properly setting the value of N_{od} to calibrate the set of virtual trips Γ. We also investigate its influence on the calculation of the trip lengths, considering the four levels of information (i.e. M_1 to M_4).

The proper setting of N_{od} is essential to ensure that the full city network has been covered by the set Γ. A low value of N_{od} does not guarantee that all the links and nodes of the city network have been visited. To better elucidate this, we execute 100 trials of Γ for different values of N_{od}: 200; 400; 600; 800; 1000; 5000; and 10,000. The *od* pairs are uniformly sampled in the city network for all N_{od} trials. For these 100 trials and for each N_{od} value, we estimate the fraction of links and nodes that have been visited by the sets of virtual trips. We define two criteria that estimate the percentage of the city network nodes ($N_{cov}^{nodes}(N_{od})$) and links ($N_{cov}^{links}(N_{od})$) visited:

$$N_{cov}^{nodes}(N_{od}) = \sum_{i=1}^{N_{nodes}} \frac{N_{used}^{nodes}(N_{od})}{N_{nodes}} \tag{21}$$

$$N_{cov}^{links}(N_{od}) = \sum_{i=1}^{N_{links}} \frac{N_{used}^{links}(N_{od})}{N_{links}} \tag{22}$$

where $N_{used}^{nodes}(N_{od})$ and $N_{used}^{links}(N_{od})$ are the total number of nodes and links covered in the city network for one trial of a fixed N_{od}, respectively; and N_{nodes} and N_{links} represent the total number of nodes and links of the city network, respectively. In this case, $N_{nodes} = 431$ and $N_{links} = 757$.

For each value of N_{od} previously stated and based on the 100 trials, we estimate the average and standard deviations for $N_{used}^{nodes}(N_{od})$ (Eq. (21)) and $N_{used}^{links}(N_{od})$ (Eq. (22)). These results are shown in Fig. 5. When $N_{od} = 200$, we cover around 90% of the nodes and 78% of the links of the city network. As N_{od} increases, both $N_{used}^{nodes}(N_{od}) \to 1$ and $N_{used}^{links}(N_{od}) \to 1$ with low standard deviations, showing that the fraction of nodes and links traveled by the virtual trips listed in Γ are very close to the total number of nodes and links in the city network. Moreover, the standard deviation is close to 0 for $N_{od} \geq 5000$. This highlights that the fraction of nodes and links visited by the 100 trials for a fixed N_{od} value are equal. Nevertheless, what

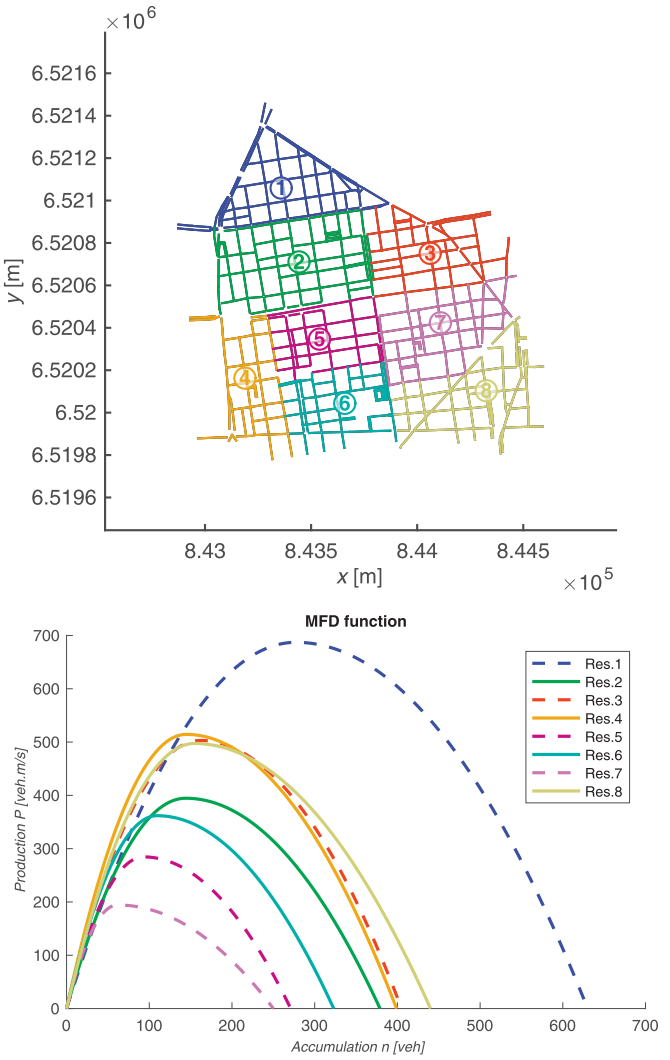

Fig. 4. *Top:* Lyon 6th district network divided into 8 regions. *Bottom:* MFD function of each region.

is the influence of good city network coverage for the calculation of the trip length distributions? To answer this question, we analyze the relative differences ρ between average trip lengths estimated through all four levels of information (M_1 to M_4). We choose three sets of $N_{od} = 400, 5000$ and 10,000. The last value of N_{od} is considered as the reference since all nodes ($N_{used}^{nodes}(N_{od}) = 1$) have been visited by the virtual set of trips (see Fig. 5). We also consider $N_{od} = 400$ where $N_{used}^{nodes}(N_{od}) \sim 0.95$ and $N_{od} = 5000$ where $N_{used}^{nodes}(N_{od}) \to 1$. These values guarantee that we compare a case with a situation similar to the reference (i.e. $N_{od} = 5000$) and another that is not (i.e. $N_{od} = 400$). This analysis provides better understand of the link between the results shown in Fig. 5 and the calculation of the trip lengths. We estimate the relative differences ρ, between the average trip lengths estimated from the sets of $N_{od} = 400$ and 5000 and the reference set of $N_{od} = 10,000$, as:

$$\rho = \frac{\bar{L}_{M_i}^{x} - \bar{L}_{M_i}^{ref}}{\bar{L}_{M_i}^{ref}}, \forall i = 1, \ldots, 4 \wedge x = 400, 5000 \tag{23}$$

where M_i refers to ith level of information (i.e. M_1 to M_4) and x refers to the $N_{od} = 400, 5000$.

The distributions of the relative differences ρ for all four levels of information are summarized in the boxplots shown in Fig. 6 (a) for $N_{od} = 400$ and in Fig. 6 (b) for $N_{od} = 5000$. The horizontal red lines represent the average of the ρ distributions while the red crosses represent the outliers. Levels 3 (M_3) and 4 (M_4) are split into the four possible cases discussed in Section 2.2, i.e. for the calculation of the trip length distributions for the Origin (O), Intermediate (I) and Destination (D) regions of the regional paths as well as for the internal paths (referred to as int. in Fig. 6). We observe that the variances of the ρ distributions are lower for $N_{od} = 5000$ compared to $N_{od} = 400$. The average of the ρ distributions are also basically

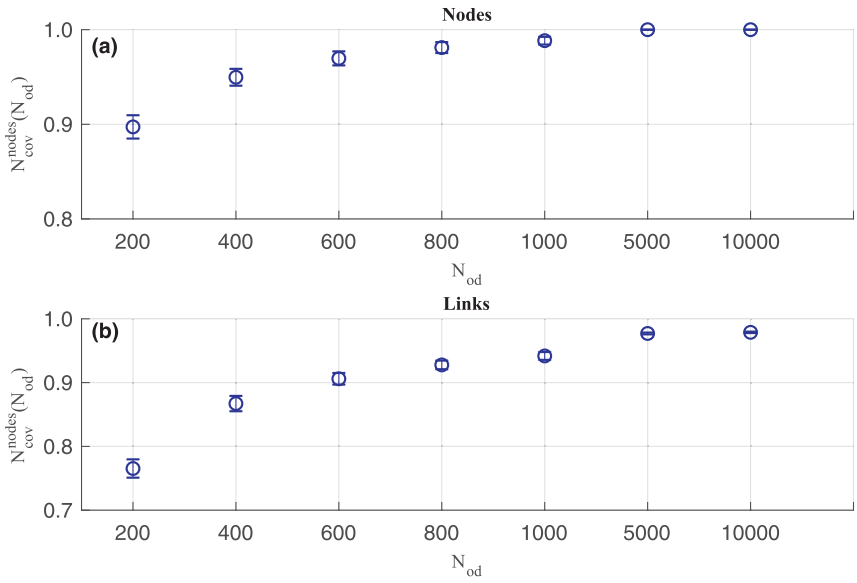

Fig. 5. Average and standard deviation for $N_{used}^{nodes}(N_{od})$ (a) and $N_{used}^{links}(N_{od})$ (b) estimated based on the 100 trials and for each N_{od} value considered. The blue dots represent the average values and the vertical bars represent the standard deviations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

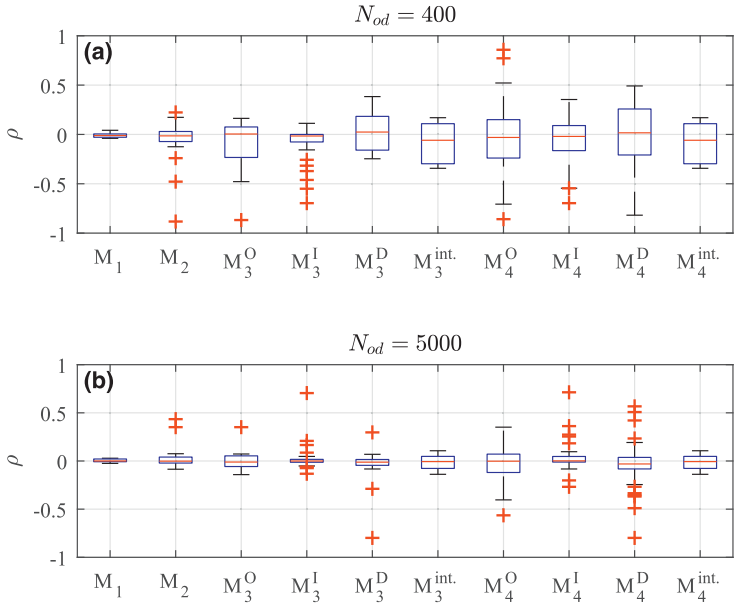

Fig. 6. Distributions of the relative differences ρ for the four levels of information (M_1 to M_4) and both $N_{od} = 400, 5000$ compared to the reference $N_{od} = 10,000$. Levels M_3 and M_4 have been divided into the cases for Origin (O), intermediate (I) and destination (D) of the regional paths as well as the internal paths (int.).

0 for $N_{od} = 5000$. This means that the average trip lengths are very similar for sets of virtual trips associated with distinct values of N_{od}, when $N_{used}^{nodes}(N_{od}) = 1$. This means that a node has been visited by the virtual trips at least one time. For this specific case, this occurs for $N_{od} \geq 5000$. We now analyze the results shown in Fig. 6 in more detail. Level 1 (M_1) is the most robust regarding the definition of N_{od}, i.e. where the relative differences of the average trip lengths for both $N_{od} = 400, 5000$ and the reference value $N_{od} = 10,000$ are the lowest. Indeed, for M_1, we consider all the virtual trips that travel in each region to estimate the trip length distributions. It is thus more probable to obtain a statistical number of virtual trips to estimate the distributions of trip lengths for M_1 than with the other three levels of information M_2 to M_4. For $N_{od} = 400$ compared to the reference $N_{od} = 10,000$, the largest relative differences are $\sim$10 m for regions 4 and 5, meaning that $\rho \sim 4\%$. The numbers of virtual trips used to define the distributions of trip lengths for regions 4 and 5 are 95 and 162, respectively.

Table 2
Average trip lengths [in meters] for the outliers of the ρ distributions shown in Fig. 6, for M_2 and the three $N_{od} = 400, 5000, 10,000$ values. The number of trips that define the associated trip length distributions are also listed.

Region		$N_{od} = 400$		$N_{od} = 5000$		$N_{od} = 10,000$	
r	j	$\bar{L}_{rj}$	$\Sigma_{k\in\Gamma}\delta_{rjk}$	$\bar{L}_{rj}$	$\Sigma_{k\in\Gamma}\delta_{rjk}$	$\bar{L}_{rj}$	$\Sigma_{k\in\Gamma}\delta_{rjk}$
3	7	259	8	∼	∼	341	233
5	4	19	1	232	14	161	19
7	6	220	36	∼	∼	180	1010
8	6	171	1	443	1	327	8

Table 3
Average trip lengths [in meters] for the outliers of the ρ distributions shown in Fig. 6, for M_3. The number of trips that define the associated trip length distributions are also listed.

Region			$N_{od} = 400$		$N_{od} = 10,000$		ρ
h	r	j	$\bar{L}_{hrj}$	$\Sigma_{k\in\Gamma}\delta_{hrjk}$	$\bar{L}_{hrj}$	$\Sigma_{k\in\Gamma}\delta_{hrjk}$	
1	2	3	103	1	339	95	−0.70
4	5	6	144	1	321	50	−0.55
4	6	5	185	1	344	42	−0.46

These numbers are higher than the number of links and nodes of the city network that lie inside these two regions. Let δ_{kr} be a binary variable equal to 1 if virtual trip k travels on region r, or 0 otherwise; and δ_{lr} and δ_{nr} be binary variables that equal 1 if link l or node k lie inside region r, respectively, or 0 otherwise. Thus $\frac{\sum_{k\in\Gamma}\delta_{kr}}{\sum_{n\in N}\delta_{nr}} \sim 1.2$ and $\frac{\sum_{k\in\Gamma}\delta_{kr}}{\sum_{l\in A}\delta_{lr}} \sim 1.7$ for region $r = 4$; and $\frac{\sum_{k\in\Gamma}\delta_{kr}}{\sum_{n\in N}\delta_{nr}} \sim 2.5$ and $\frac{\sum_{k\in\Gamma}\delta_{kr}}{\sum_{l\in A}\delta_{lr}} \sim 3.9$ for region $r = 5$.

The other three levels of information M_2 to M_4 require a larger virtual trip set size, N_{od} for a good estimation of the trip lengths due to their nature. For level M_2, we observe the presence of outliers in Fig. 6 for both $N_{od} = 400, 5000$ values. Let us first analyze the case of $N_{od} = 400$ to understand the presence of these four outliers. These outliers occur for the average trip lengths $\bar{L}_{37}$ (i.e. for the average trip length to go from region $r = 3$ to the adjacent region $j = 7$ - see Eq. (9)), $\bar{L}_{54}$, $\bar{L}_{86}$ and $\bar{L}_{76}$. These average trip lengths are listed in Table 2 for the three N_{od} values, as well as the number of virtual trips associated with these trip length distributions. Note that, as defined in Section 2.2 for M_2, $\Sigma_{k\in\Gamma}\delta_{rjk}$ defines the total number of trips that travel in region r and go the same adjacent region j. As can be seen, to define the trip length distribution when going from region $r = 3$ to $j = 7$ there are only 8 virtual trips for $N_{od} = 400$, while there are 233 virtual trips for the reference $N_{od} = 10,000$. The average trip length $\bar{L}_{76}$ represents the lowest outlier in Fig. 6, where $\rho \sim 0.2$. In this case, the number of associated virtual trips is 36, which is statistically more significant. For the other two trip length distributions (i.e. to go from regions $r = 5$ and $r = 8$ to $j = 4$ and $j = 6$, respectively) the number of associated virtual trips is low, despite its increase verified for the three N_{od} values. Indeed, the entry and exit links of the nodes located on the borders of the city network partitioning might not allow traveling in both directions between two adjacent regions. This is because of the presence of one-way links, which occurs in most of the 6th district of the Lyon network. Fig. 4 shows that there are 5 nodes located on the border between regions 4 and 5, while 4 nodes located on the border between regions 8 and 6. Nevertheless, only two nodes allow traveling from regions 5 to 4 and only one node allows traveling from region 8 to 6. With the increase of N_{od}, we increase the probability of having a larger number of longer virtual trips that cross the borders of the city network partitioning at these nodes. Thus, for very specific trip length distributions, the number of connection nodes between adjacent regions also plays a role. This becomes more significant as we increase the level of information for M_3 and M_4. This might further reduce the total number of virtual trips used to define the trip length distributions.

We now focus our analysis on M_3 and M_4. As can be seen in the boxplots shown in Fig. 6, the distributions of ρ are narrower for $N_{od} = 5000$ compared to $N_{od} = 400$. The increase of N_{od}, increases the probability that a larger number of virtual trips are considered to estimate the distributions of trip lengths. However, several outliers can be seen, especially for the intermediate and destination regions. Let us first analyze the outliers of the ρ distributions for M_3 and $N_{od} = 400$. The average trip lengths $\bar{L}_{hrj}$ (see Eq. (13)) for these three outliers are listed in Table 3. Note that as defined in Section 2.2 for M_3, $\Sigma_{k\in\Gamma}\delta_{hrjk}$ defines the total number of virtual trips k associated with the trip length distribution L_{hrj}. As shown in Table 3, these outliers occur because for $N_{od} = 400$ there is only one virtual trip to define the trip length distribution. However, the number of virtual trips become much more significant for the reference $N_{od} = 10,000$. We also analyze the largest outlier of the ρ distribution for the intermediate regions and for $N_{od} = 5000$. This outlier occurs for the trip length distribution L_{645}, for virtual trips traveling on region $r = 4$, coming from region $h = 6$ and going to region $j = 5$. This sequence of regions is not covered by the virtual trip for $N_{od} = 400$. The average trip lengths $\bar{L}_{645}$ are 179 m for $N_{od} = 5000$ and 105 m for $N_{od} = 10,000$. These values are estimated from a set of 3 virtual trips for $N_{od} = 5000$ and 2 virtual trips for $N_{od} = 10,000$. The lower number of virtual trips associated with these sets of trip lengths for larger N_{od} values is related with the topology

Table 4
Number of virtual trips between each regional OD pair, for the 6th Lyon district network.

		Destination region							
		1	2	3	4	5	6	7	8
Origin region	1	165	258	135	153	96	133	125	171
	2	215	298	164	176	147	202	184	282
	3	160	188	114	135	97	103	128	185
	4	156	247	105	104	96	106	155	175
	5	108	139	69	93	43	99	86	138
	6	140	159	102	67	85	65	89	162
	7	147	227	115	81	75	79	135	152
	8	199	272	141	132	103	149	151	175

of the city network as well as the definition of its partitioning. These five virtual trips have an Origin in region 8 and Destination in region 5, defining the regional path $p = \{87645\}$. In the set of $N_{od} = 5000$ there are a total of 42 virtual trips connecting this regional OD pair, while in the set of $N_{od} = 10,000$ there are 94 virtual trips. Out of the total numbers of virtual trips mentioned previously, there are $\sim$90% that define the regional path $p = \{875\}$. Travel is therefore more likely on the regional path $p = \{875\}$ rather than on $p = \{87645\}$, for the regional OD pair 8-5 (see also Fig. 4).

Fig. 6 shows that the most refined level of information M_4 is that most sensitive to the calibration of the virtual trip sets Γ. It also shows several outliers in the ρ distributions that come from the calculation of trip lengths based on regional paths that are not statistically significant. When the regional paths are not statistically significant, the bias induced from the sampling of origin and destination nodes plays an important role in the calculation of the trip length distributions for the Origin and Destination regions. Nevertheless, this bias is considerably reduced for distributions of trip lengths estimated on the basis of statistically significant regional paths. To better elucidate the latter, we analyze several examples. We first consider the regional path $p = \{357\}$ which is not statistically significant. This regional path is defined by 3 virtual trips for $N_{od} = 5000$ and by 4 virtual trips for $N_{od} = 10,000$. We estimate the relative differences ρ for the three regions. The ρ values are -0.41 for the Origin region, 0.73 for the intermediate region and -0.35 for the Destination region. We now analyze the regional path $p = \{578\}$ which is statistically significant. This regional path is defined by 37 virtual trips for $N_{od} = 5000$ and by 66 virtual trips for $N_{od} = 10,000$. The ρ values are 0.02 for the Origin region, -0.03 for the intermediate region and -0.05 for the Destination region. We observe that for the regional path $p = \{578\}$, the average trip lengths are very close for the three regions. However, this does not occur for the regional path $p = \{357\}$.

In this section, we show that the criterion $N_{cov}^{nodes}(N_{od})$ provides good guidance for ensuring a proper definition of the virtual trip sets Γ for the calculation of the trip length distributions. We also carried out a detailed analysis of the influence of the N_{od} setting on the calculation of the trip lengths and identified that the presence of outliers in the ρ distributions, especially for the more detailed levels of information (i.e. M_3 and M_4), originate from trip length sets that are not statistically significant. As mentioned previously, these outliers originate from the scaling of the virtual trips to the aggregated level, taking into account the definition of the city network partitioning. One solution to avoid this problem is to first analyze the set of regional paths defined by the virtual trips set Γ. Let Ψ be this set of regional paths. First, it is possible to filter the non-statistically significant regional paths of this set Ψ, i.e. filter the virtual trips from the set Γ that define these regional paths. Then, based on these virtual trips, the distributions of trip lengths can be reestimated according to the four levels of information M_1 to M_4. By applying this procedure, we reduce the number of outliers on the ρ distributions to 2 for M_3^I and M_4^I and to 3 for M_4^D.

Based on the analysis done in this section, we select the $N_{od} = 10,000$ set of virtual trips Γ as the reference for the next two Sections 3.3 and 3.4. The latter yields $\frac{N_{od}}{N_{links}} \sim 13.2$. The virtual trips are uniformly sampled in the city network and define the regional OD matrix listed in Table 4. Let Z be this regional OD matrix. These virtual trips define a total of 205 regional paths.

3.3. Analysis of the trip length distributions

In this section, we start by analyzing the differences between the averages of the trip length distributions estimated by the different levels of information M_1 to M_4. We consider M_4 as the reference distribution as stated in Section 2.2. The average of the trip length distributions of the other three levels of information, M_1 to M_3, are analyzed against the reference level M_4. We then analyze the differences of the trip length distributions estimated by the four levels of information for a representative regional path $p = \{124\}$.

We investigate the differences between the average trip lengths estimated by the levels of information M_1 to M_3 compared to the reference level M_4. We do this analysis at the region level. First, we focus our analysis at the region level. Fig. 7 depicts the average trip lengths estimated by the four levels of aggregation for the eight regions of the 6th district Lyon network (see Fig. 4). The blue dots represent the average trip lengths estimated for each region by M_1 (see Eq. (7)). The green triangles represent the average trip lengths estimated for M_2 (see Eq. (9)) and for the full set of Λ as well as the

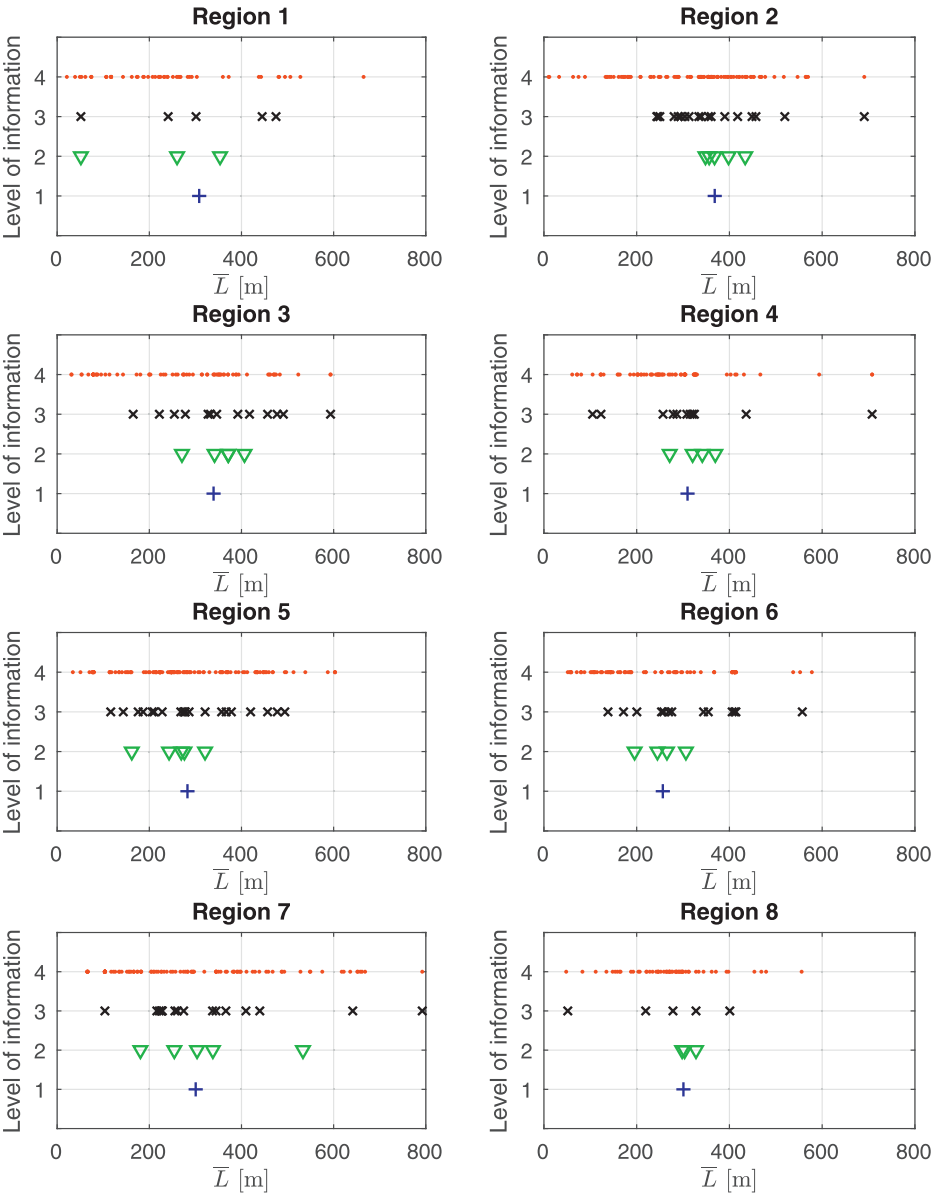

Fig. 7. Average regional trip lengths $\overline{L}$ [in meters] estimated through the four levels of information M_1 to M_4, for the eight regions. The blue dots refer to M_1 (see Eq. (7)). The green triangles refer to M_2 (see Eq. (9)). The black crosses refer to M_3 (see Eqs. (11) and 13). The red dots refer to M_4 (see Eq. (15)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

internal trips of each specific region. The black crosses represent the average trip lengths estimated for M_3. This includes the average trip lengths for the Origin, intermediate and Destination (see Eq. (11)) regions of the regional paths as well as the internal paths (see Eq. (13)). The red dots represent the average trip lengths estimated for M_4 (see Eq. (15)). Each of the red dots represented in Fig. 7 for each region corresponds to a different regional path that crosses this specific region. The results shown in Fig. 7 show that considering only one average trip length for each region (i.e. M_1) is not representative of all the possible average trip lengths of the regional paths that cross each region (i.e. M_4). The level of information increases the diversity of the possible average trip lengths in each region.

We discuss these differences in more detail for region 5, which has 4 adjacent regions $\Lambda = \{2, 4, 6, 7\}$. As can be seen in Fig. 4, there are five regions adjacent to region 5. However, the city network node that lies on the border between regions 3 and 5 only allows traveling from region 5 to 3. Note that similar conclusions can be made for the other regions as well. The first level of information M_1 provides an average trip length $\overline{L} = 273$ meters (see Eq. (7)). For M_2, we have five average trip lengths that correspond to the adjacent regions to region 5 and to the internal path $p = \{5\}$. The estimated trip lengths range from $\overline{L}_{54} = 203$ meters to $\overline{L}_{57} = 321$ meters. As can be seen M_2 already provides a heterogeneity of average trip lengths that

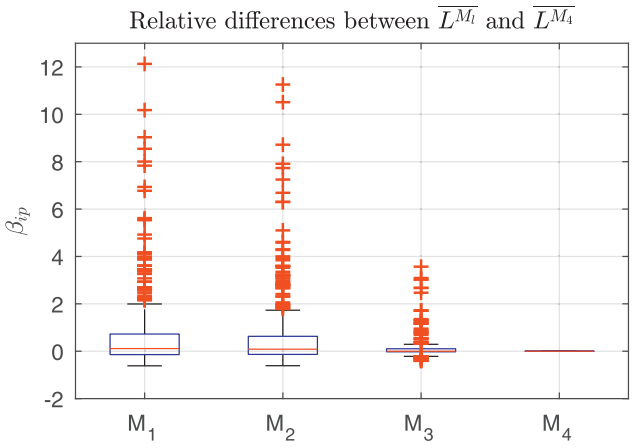

Fig. 8. Distribution of the relative differences β, between method $l = 1, 2, 3, 4$ (from the left to the right) and the reference method 4.

M_1 is unable to deliver. In the case of M_3, we observe a total of 20 black crosses. They correspond to the origin, intermediate and destination regions of the regional paths as well as the internal path $p = \{5\}$. The set of 20 trip lengths given by M_3 has an average of 296 m and standard deviation 111 m. As shown, by increasing the level of information, we increase the range of possible average trip lengths compared to M_2 and, in particular, to M_1. This is because we consider in more detail all the possible trips that might occur inside a region. When we consider the level of information M_4, the range of possible average trip lengths increases even more. A total of 124 regional paths out of 205 cross region 5. The average trip lengths of these 124 regional paths are represented by the red dots in Fig. 7. The average of these regional paths inside region 5 can range from a few meters up to $\sim$600 m. For example, the regional paths $p = \{5687\}$ and $p = \{645\}$ are among the shortest. Whilst, the regional paths $p = \{45768\}$ and $p = \{245768\}$ are longest ones that cross region 5. As can be seen, the more detailed level of information M_4 allows distinguishing between regional paths that just cross a region by a few meters from others that have much longer trip lengths.

These four levels of information M_1 to M_4 provide regional trip lengths for each regional path crossing each region of the regional network. Therefore, it is relevant to estimate the relative differences between the regional trip lengths estimated through the levels of information M_1 to M_3 and the reference level M_4, when assigning regional trip lengths to regional paths (p) inside the regions i. This enhances a more aggregated comparison between the different levels of information than in Fig. 7. The relative differences β_{ip} are estimated as:

$$\beta_{ip} = \frac{\bar{L}_i^{M_j} - \bar{L}_i^{M_4}}{\bar{L}_i^{M_4}} \delta_i^p, \forall j = 1, 2, 3 \wedge \forall p \in \Psi \tag{24}$$

where δ_i^p is a binary variable that equals 1 if regional path p crosses region i, or 0 otherwise. $\bar{L}_i^{M_4}$ is the average regional trip length estimated through Eq. (15) (i.e. level of information M_4). For level of information M_1, $\bar{L}_i^{M_j}$ is estimated through Eq. (7). For M_2, $\bar{L}_i^{M_j}$ is estimated through Eq. (9), where $i = p_m$ is the current region and $j = p_{m+1}$ is the next region adjacent to i on the sequence of regional path p. For M_3, $\bar{L}_i^{M_j}$ is estimated through Eq. (13), where $i = p_m$ is the current region, $h = p_{m-1}$ is the previous region and $j = p_{m+1}$ is the next region adjacent to i on the sequence of regional path p.

The estimated relative differences β_{ip} are illustrated in Fig. 8. As shown, the relative differences β_{ip} decrease as we increase the level of information, i.e. from M_1 to M_3. This clearly shows how far the average trip lengths provided by the level of information M_1 are from all the possible trip lengths of the regional paths crossing the different regions. The setting of the trip lengths is crucial for the MFD-based models. Indeed, as shown by Leclercq et al. (2015), the hypothesis of considering an average trip length for all the vehicles traveling inside the same region for MFD-based models applications is a strong limitation. Disaggregating the definition of trip lengths by the different regional paths that cross each region is clearly a better approach. This is because it allows gathering short trip lengths for certain regional paths that just cross a given region. However, one can also assign longer trip lengths for other regional paths. This has a non-negligible influence on the modeled traffic states, a point that will be discussed in Section 4.

We now analyze the differences between the trip length distributions estimated by the four levels of information M_1 to M_4 for the regional path $p = \{124\}$. The distributions of trip lengths for each region of this regional path and for each of the four levels of information are presented in Fig. 9. The black horizontal dashed lines represent the average of each trip length distribution inside each region. As can be observed, the variance of the trip length distributions decreases as we increase the level of information considered. The diversity of virtual trips, i.e. the number of virtual trips that is considered to gather the distributions of trip lengths decreases from M_1 to M_4. In particular, this clearly influences the shape of the trip length distribution for the intermediate region 2, as observed in Fig. 9 for M_3 and M_4. Increasing the level of information to define

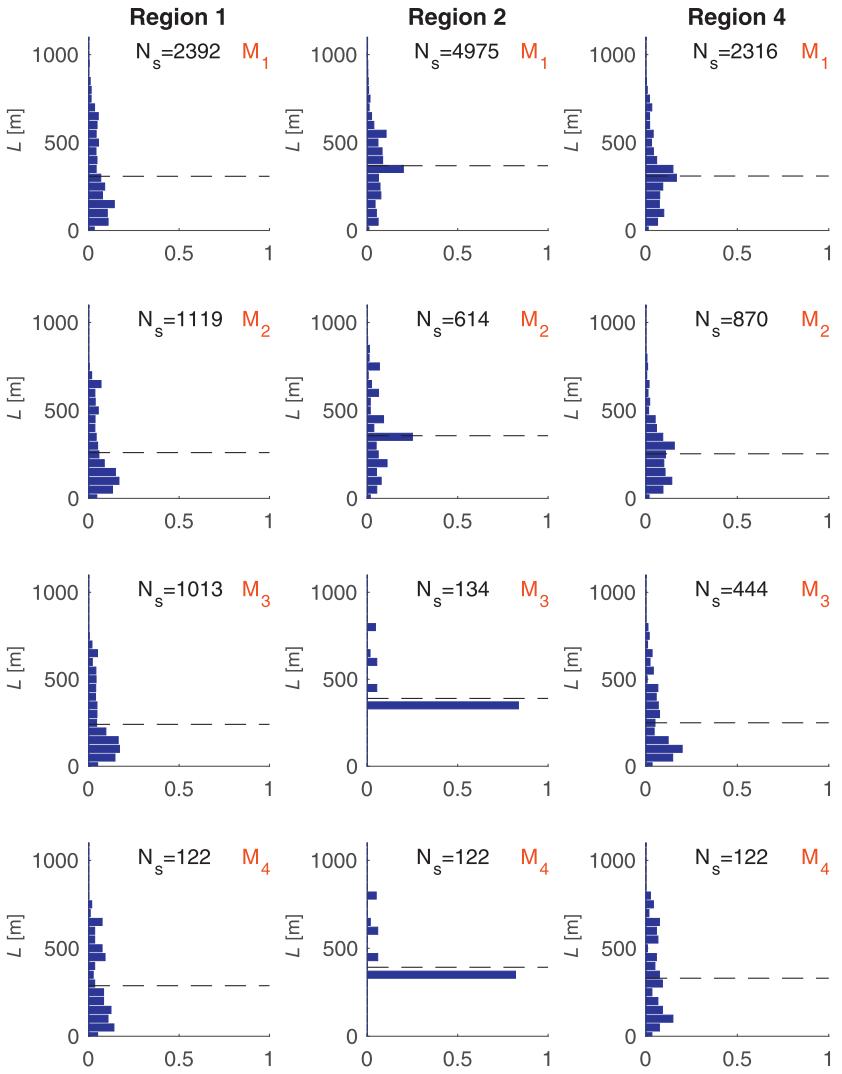

Fig. 9. Regional trip length distributions estimated through the four levels of information M_1 to M_4 for each region of the regional path $p = \{124\}$. The total number of trips considered for each distribution is identified at the top of each subplot. Each row of the subplots represents the results for each level of information, while each column of subplots represents each region of the regional path. The horizontal dashed lines represent the average of the trip length distributions.

the trip length distributions for intermediate regions gives priority to the virtual trips that specifically travel this intermediate sequence. This explains the peak observed for the trip length distributions for the intermediate region 2 gathered by M_3 and M_4. Indeed, the fraction of the city network that lies inside region 2 is similar to a grid network (see Fig. 4). The links of the city network that lie inside region 2 are mostly one directional. This increases the probability that the virtual trips traveling on this specific sequence of intermediate regions, use a specific sequence of links of the directional links of the city network more often. Whereas the trip length distributions for region 2 gathered from M_3 and M_4 are similar, this might not be the case for all intermediate regions. Note that, for M_3 we consider all virtual trips that cross the specific sequence of regions 4-2-1. For M_4, we only consider the virtual trips that define the same regional path p. The trip length distributions for the intermediate regions are estimated based on the length of each trip inside these regions. As shown in Fig. 9, there are 122 virtual trips that define the regional path $p = \{124\}$ and that are then used to estimate the trip length distribution for region 2. For M_3 there are 12 additional virtual trips that define other regional paths that are considered to define the trip length distribution for region 2.

We also elucidate that sequence $p = \{124\}$ defines a regional path different from the sequence $p = \{421\}$. In Table 5, we list the average trip lengths for each region of these two regional paths, estimated through the four levels of information. M_1 attributes similar average trip lengths for the three regions of these two regional paths. Due to its nature, we consider all the virtual trips that cross each of these regions to define the trip length distributions. Nevertheless, as we increase the level of information M_2 to M_4, we assign different average trip lengths for the three regions of each regional path. This

Table 5
Average trip lengths for the regional paths $p = \{124\}$ (left) and $p = \{421\}$ (right) estimated by the four levels of information for each region.

Region	Level of information				Region	Level of information			
	M_1	M_2	M_3	M_4		M_1	M_2	M_3	M_4
1	312	277	253	314	4	311	362	419	439
2	364	353	390	393	2	365	396	355	355
4	311	268	243	296	1	312	345	247	285

Table 6
Regional OD matrix Z^*.

		Destination region							
		1	2	3	4	5	6	7	8
Origin region	1	135	231	81	933	173	881	142	838
	2	153	198	140	208	119	285	137	377
	3	137	153	75	95	65	91	90	133
	4	102	156	83	84	77	72	94	106
	5	76	119	64	58	38	61	65	99
	6	96	112	68	46	42	37	77	104
	7	105	137	78	72	53	43	80	102
	8	147	157	107	81	73	106	114	142

has a clear influence on the traffic states modeled by an MFD-based model. A longer average trip length inside a region means that a vehicle has to spend more time to complete its trip inside the region. Consider two vehicles departing at the same time, but traveling on different regional paths $p = \{124\}$ and $p = \{421\}$. Furthermore, assume that we have fluid traffic conditions in the network. According to the average trip lengths assigned by M_1, these vehicles will complete their trips at the same time. M_2 to M_4 assign different average trip lengths for these vehicles to travel inside the same region, leading to different travel times for vehicles traveling in the same region and along different regional paths.

In this section, we emphasize the importance of considering higher levels of information from the regional network to define the trip length distributions. First, we highlight that considering an average trip length for all vehicles traveling inside the same region and on different regional paths (i.e. M_1) is not representative of all the possible average trip lengths of all the regional paths that cross this region (i.e. M_4). Second, we show that the level of information considered also plays a role in the shape of the trip length distributions. Moreover, we show another limitation of M_1 compared to the other three levels of information. We show that for two regional paths that are defined by the same regions, but in reverse order, the assigned average trip lengths for each region is similar. However, the other three levels of information M_2 to M_4 assign different average trip lengths. Because of their nature, they are able to capture important information from the possible directions of the city network links. In Section 4, we discuss the importance of trip lengths setting for MFD-based models.

3.4. Influence of the OD matrix on the trip lengths

The trip length distributions are clearly influenced by the regional OD matrix (Leclercq et al., 2015). In fact, the regional OD matrix defines the distribution of od pairs in the city network, which may lead to very different sets of virtual trips. This happens mostly when considering the relative quantities that go from one region to another. In this section, we introduce a methodology to update the trip lengths for when the number of trips between OD pairs is modified. This methodology is only applicable for changes of the OD matrix caused by differences in the trip patterns and not for updating trip length distributions influenced by the spatial distribution of the city network od pairs. As discussed previously, this is the case of the trip length distributions for the Origin and Destination regions estimated by the level of information M_3. Nevertheless, we introduce an estimation procedure for these two cases, to cover all three possibilities for M_3 dissected in Section 2.2. We consider an updated regional OD matrix Z^*, also composed of $N_{od} = 10,000$ virtual trips. Out of these 10,000 virtual trips, a total of 1000 vehicles travel specifically on each of the following regional OD pairs: 1-4; 1-6; and 1-8. The remaining 7000 virtual trips are uniformly distributed over the city network. For the updated OD matrix, we re-estimate the set of virtual trips Γ and the distributions of trip lengths according to the four levels of information M_1 to M_4. The updated OD matrix Z^* is listed in Table 6.

The methodology applied to update the regional trip lengths is based on the regional OD matrix (Z) defined by the set of virtual trips Γ as well as the trip lengths that are estimated and the updated regional OD matrix Z^*. One solution is to estimate a new set of virtual trips based on Z^* and update the trip length distributions accordingly. Nevertheless, the computational cost to do so largely increases for large city networks. We propose an alternative methodology to circumvent the computational power required to update the trip length distributions. The general idea is to balance the weight of virtual trips to adjust their frequency to the updated regional OD matrix. This procedure is very useful as it updates the regional trip lengths considering the current set of trips Γ when the regional OD matrices change over time, or when another trip-

based MFD model application has to be run considering a regional OD matrix different from the original one. This procedure is described below.

The trip lengths for the i–th region and for the first level of information M_1, are estimated as:

$$\hat{L}_i = \frac{\sum_r \sum_k l_{ik}^{OD} \alpha_r^{OD}}{\sum_O \sum_D \sum_r \sum_k \delta_{ik} \alpha_r^{OD}}, \forall k \in \Gamma \tag{25}$$

where $\hat{L}_i$ is the estimated average regional trip length for region i; l_{ik}^{OD} is the length of virtual trip k on region i that has the regional OD pair; δ_{ik} is a binary variable that equals 1 if virtual trip k crosses region i, or 0 otherwise. α_r^{OD} is a scaling factor that depends on the demand of regional OD pair of Z weighted by the demand of the same regional OD pair of Z^*, that crosses region i:

$$\alpha_r^{OD} = \begin{cases} \frac{n^{OD}}{n_*^{OD}}, \text{if } n_*^{OD} \geq 1 \\ 1, \text{if } n_*^{OD} = 0 \end{cases} \tag{26}$$

where n^{OD} is the demand going from regional origin O to regional destination D of matrix Z; and n_*^{OD} is similar but for the updated regional OD matrix Z^*. Such calculations are simple and allow updating trip lengths as a function of dynamic variations of the regional OD matrix.

For the level of information M_2, the trip length from region i to the next adjacent region j is estimated as:

$$\hat{L}_{ij} = \frac{\sum_O \sum_D \sum_r \sum_k \delta_{ijk} l_{ik}^{OD} \alpha_r^{OD}}{\sum_O \sum_D \sum_r \sum_k \delta_{ijk} \alpha_r^{OD}}, \forall k \in \Gamma \tag{27}$$

where $\hat{L}_{ij}$ is the estimated regional trip length from the current region i to the adjacent region j; δ_{ijk} is a binary variable that equals 1 if virtual trip k, crosses region i and goes to the adjacent j, or 0 otherwise.

For the level of information M_3, we have to distinguish three possible cases for the estimation of the regional trip lengths:

- the case where the regional path is composed of only one region ($R = 1$). The average regional lengths are estimated as:

$$\hat{L}_i = \frac{\sum_O \sum_D \sum_r \sum_k \delta_{ik} l_{ik}^{OD} \alpha_r^{OD}}{\sum_O \sum_D \sum_r \sum_k \delta_{ik} \alpha_r^{OD}}, \forall k \in \Gamma \tag{28}$$

  where δ_{ik} is a binary variable that equals 1 if virtual trip k is an trip internal to region i, or 0 otherwise.
- the case where the regional path is composed by two or more regions ($R \geq 2$) and for the Origin and Destination regions. The average regional lengths are estimated as:

$$\hat{L}_{ij} = \frac{\sum_O \sum_D \sum_r \sum_k \delta_{ijk} l_{ik}^{OD} \alpha_r^{OD}}{\sum_O \sum_D \sum_r \sum_k \delta_{ijk} \alpha_r^{OD}}, \text{if } ((i = p_1, j = p_2) \vee (i = p_{R-1}, j = p_R)) \wedge \forall k \in \Gamma \tag{29}$$

  where δ_{ijk} is a binary variable that equals 1 if virtual trip k travels in region i and goes to j (for the Origin regions) or if virtual trip k travels in region j coming from region i (for the Destination regions), or 0 otherwise.
- the case where the regional path is composed of three or more regions ($R \geq 3$) and for the intermediate regions. The average regional lengths are estimated as:

$$\hat{L}_{hij} = \frac{\sum_O \sum_D \sum_r \sum_k \delta_{hijk} l_{ik}^{OD} \alpha_r^{OD}}{\sum_O \sum_D \sum_r \sum_k \delta_{hijk} \alpha_r^{OD}}, \text{if } (h = p_{m-1}, i = p_m, j = p_{m+1}) \wedge \forall m = 2, \ldots, R-1 \forall k \in \Gamma \wedge \forall (h, j) \in \Lambda \wedge h \neq j \tag{30}$$

  where δ_{hijk} is a binary variable that equals 1 if virtual trip k, with regional OD pair, crosses the specific sequence of $h - i - j$ regions, or 0 otherwise.

Note that the trip lengths for the intermediate regions estimated for M_3 and M_4 are not influenced by changes in the OD matrix. For this, one needs to ensure that the sample size used to calculate the trip lengths are statistically significant. Nevertheless, changes in the spatial distribution of the city network od pairs have a non-negligible influence on the calculation of the trip length distributions for the Origin and Destination regions.

Based on Eqs. (25) and (30) we estimate that: $\hat{L} \approx \bar{L}$. We estimate average regional trip length for all eight regions based on Z and Z^* and define θ as the relative difference between $\hat{L}$ and $\bar{L}$ as:

$$\theta = \frac{\hat{L} - \bar{L}}{\bar{L}} \tag{31}$$

where $\hat{L}$ is the estimated average trip length through Eqs. (25)–(30); and $\bar{L}$ is the average trip length estimated based on the new set of virtual trips of the Z^* OD matrix. Fig. 10 shows the distributions of the relative differences θ for the levels of information M_1 and M_2 as well as for the intermediate regions case of level of information M_3. As shown, the methodology proposed to update the trip lengths yields good estimation results. The averages of the θ distributions, represented by the horizontal red lines are close to 0 and the largest differences are of the order of $\sim$0.1. As discussed previously, the trip

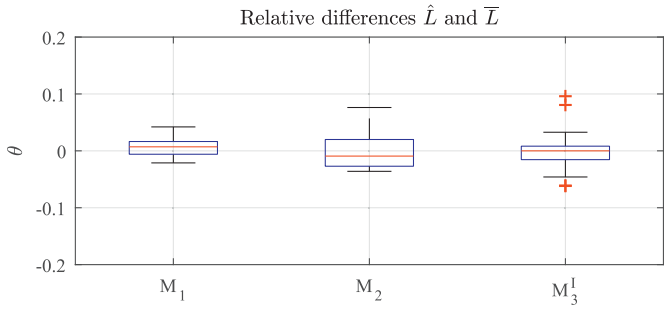

Fig. 10. Relative differences θ between the estimated trip lengths $\hat{L}$ and the average trip lengths $\overline{L}$ for the levels of information M_1 to M_3.

lengths for the Origin and Destination regions estimated by M_3 are influenced by the spatial distribution of the city network od pairs and are not included in Fig. 10. For these cases, the average of the θ distributions are also close to 0. Nevertheless, the largest differences are of the order of ~ 0.3.

This methodology is simple to implement, provides good estimations of the trip lengths, and is computationally inexpensive.

4. Impacts of trip-lengths estimation methods on dynamic models

In this section, we investigate the importance of properly tuning the trip lengths for multi-regional MFD-based model applications. We focus our analysis on the trip-based MFD model and consider two scenarios. We start by introducing these two scenarios in Section 4.1. In the first scenario (Section 4.2), we consider that we have the full information on the trip patterns in the city network, while, in the second scenario (Section 4.3), we do not know the real trip patterns in the city network and seek the Deterministic User Equilibrium. For both scenarios, we investigate how the tuning of the trip lengths influences the predicted traffic states.

4.1. Trip-based MFD model settings

For the trip-based MFD traffic model application, we consider the network and MFD functions shown in Fig. 4. The implementation of the trip-based model follows the procedure described in Mariotte et al. (2017). For all the tests in this section, we consider a total simulation period of $T = 8000$ s. Moreover, we consider three regional OD pairs: 1-6; 3-4; and 8-4. For these OD pairs, we consider a maximum of two regional paths and we define the following regional choice sets: $\Omega^{16} = \{1256; 12576\}$; $\Omega^{34} = \{324; 3524\}$; and $\Omega^{84} = \{87524; 8764\}$. The regional paths are directly gathered from the set of virtual trips Γ (Batista and Leclercq, 2018), considering the most significant ones for each OD pair, i.e. the regional paths associated with the largest number of virtual trips.

As previously stated, in this section we investigate the influence of trip length calibration in the modeled traffic states by a trip-based MFD-based model (Mariotte et al., 2017) in two distinct scenarios:

- *Scenario 1*: We assume that the full information on the trip patterns in the city network is known. Consequently, the regional path assignment is known. Therefore, the total travel distance of all vehicles inside the same region r is the same for all levels of information M_1 to M_4. For this test scenario, we consider the demand curves dissected in Fig. 11 (a). We estimate the set of virtual trips such that the total number of vehicles traveling on each OD pair corresponds to those defined by the demand curves. There are 1980 vehicle trips traveling on regional path $p = \{1256\}$, 220 on regional path $p = \{12576\}$, 1900 on regional path $p = \{324\}$ and 2400 on regional path $p = \{8764\}$. The trip lengths for these four regional paths are estimated based on this set of virtual trips and for the four levels of information M_1 to M_4. The average trip lengths are listed in Table 7. The modeled traffic states are analyzed in Section 4.2.
- *Scenario 2*: We assume that we do not know the full trip patterns in the city network. For this test scenario, we consider the demand curves as depicted in Fig. 11 (b). For these demand profiles, the corresponding regional network equilibrium is unknown. Therefore, we assume that vehicle drivers are perfectly rational and seek to minimize their own regional path travel times. The latter correspond to the first Wardrop principle (Wardrop, 1952) and defines the Deterministic User Equilibrium (DUE). To solve the DUE of the regional network, we apply of the Method of Successive Averages (MSA) (Sheffi, 1985; Sbayti et al., 2007). This method consists of a descent step iteration procedure to solve a fixed point problem. The convergence properties of the MSA are guaranteed by the good tuning of the descent step ζ_s, where s is the descent iteration step. In this paper, we assume that $\alpha_s = \frac{1}{s}$. The iterative descent procedure works as follows. At each descent step s, the vehicles' drivers choose the regional path with the minimal travel time based on an all-or-nothing procedure. The trip-based MFD model provides travel times estimations for every given path flow distribution at each descent step. The regional network equilibrium is achieved when no vehicle driver can decrease their own travel times by unilaterally changing regional paths. We consider two convergence criteria for the MSA. We consider the relative *Gap*

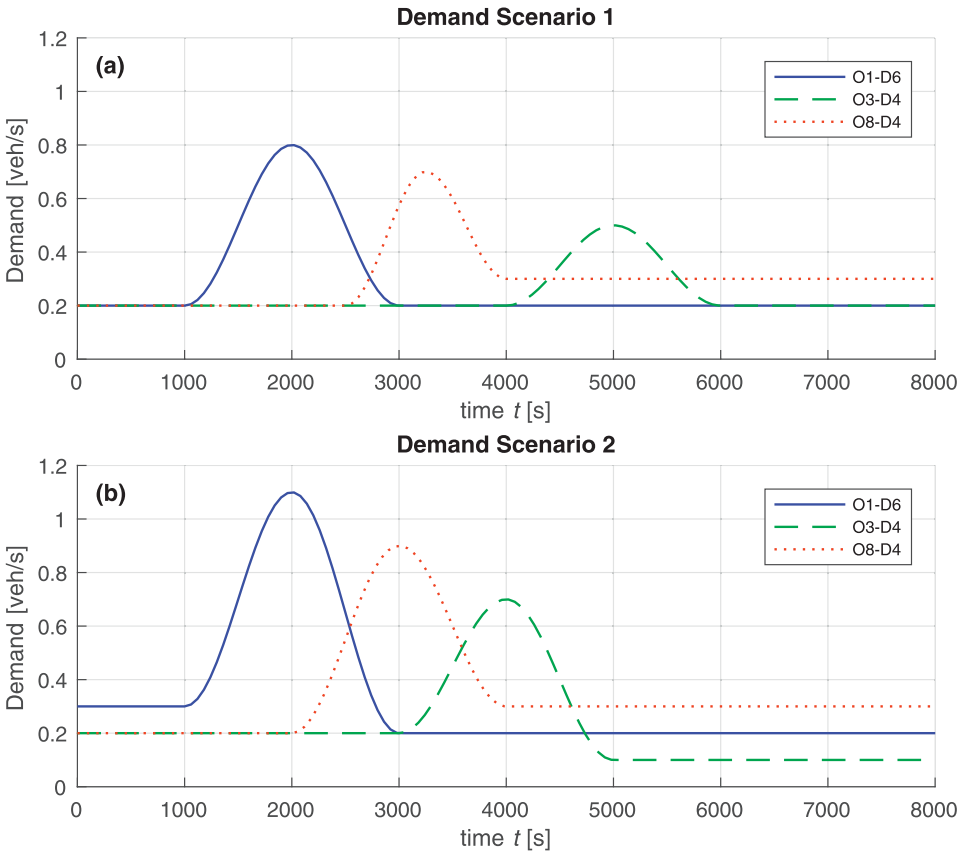

Fig. 11. Demand scenarios 1 and 2.

(Sbayti et al., 2007) and a maximum number of descent iterations N_{max}. The MSA converges when $Gap \leq tol$, $tol = 10^{-2}$ is a pre-defined tolerance. We also consider $N_{max} = 100$. The relative *Gap* represents the difference between the regional path travel time and the equilibrium travel time. Under DUE conditions, the relative *Gap* is 0, showing that all drivers choose the minimal travel time regional paths. The network equilibrium is estimated every 250 s for the whole simulation period of 8000 s. For the trip-based MFD model, we consider the trip lengths estimated in the previous Section 3, for the regional OD matrix *Z*. These trip lengths are also updated for the demand profiles illustrated in Fig. 11 (b) following the methodological framework proposed in Section 3.4. Under DUE conditions, only regional paths $p = \{1256\}$, $p = \{324\}$ and $p = \{8764\}$ are used. The average trip lengths for these three regional paths are listed in Table 7. The modeled traffic states are analyzed in Section 4.3.

4.2. Influence of trip length tuning when the regional path flow distributions are known

In this section, we investigate how the trip lengths tuning affect the modeled traffic states using a trip-based MFD model (Mariotte et al., 2017) and when the full trip patterns in the city network are available. This means that we know the regional path flow distributions and that the total travel distances for each region is similar for all four levels of information M_1 to M_4. Fig. 12 shows the traffic states modeled for Scenario 1 and for the eight regions of the 6th district of the Lyon network (see Fig. 4). The traffic states are represented by the accumulation $n(t)$ inside each region, for the four levels of information M_1 to M_4. Note that the average trip lengths to be traveled by the vehicles inside each region and for each of the four regional paths are listed in Table 7.

As shown, the tuning of the trip lengths yields different traffic states inside the eight regions. One important aspect is that the traffic dynamics predictions of M_3 are as close as those of M_4. We now analyze the predicted traffic states inside each region for the four levels of information in more detail. In region 1, we observe an accumulation peak between $\sim$2000 and 3700 s that is the demand peak of the OD pair 1–6 (see Fig. 11 (a)). We observe that the predicted accumulation is higher for M_1 than for the other three levels of information M_2 to M_4. In fact, the average trip length for M_1 is 205 m, longer than for the other three cases M_2 to M_4 and for both regional paths $p = \{1256\}$ and $p = \{12576\}$. Therefore, since the speed-MFD is the same for all vehicles traveling in region 1, they need more time to complete their trips for M_1 as the trip distance is longer, increasing the accumulation. In this region 1, we also observe another interesting accumulation peak between

Table 7
Average regional trip lengths (m) for the four levels of information M_1 to M_4 and for the regional paths used in both Scenario 1 and 2.

Scenario 1									
Method	Regional path	Region							
		1	2	3	4	5	6	7	8
1	1-2-5-6	205	537	~	~	318	336	~	~
2	1-2-5-6	175	465	~	~	308	209	~	~
3	1-2-5-6	175	465	~	~	308	216	~	~
4	1-2-5-6	172	467	~	~	308	216	~	~
1	1-2-5-7-6	205	538	~	~	319	337	284	~
2	1-2-5-7-6	174	465	~	~	366	209	271	~
3	1-2-5-7-6	175	465	~	~	366	119	104	~
4	1-2-5-7-6	197	440	~	~	366	119	104	~
1	3-2-4	~	538	376	317	~	~	~	~
2	3-2-4	~	626	349	304	~	~	~	~
3	3-2-4	~	653	349	335	~	~	~	~
4	3-2-4	~	653	349	358	~	~	~	~
1	8-7-6-4	~	~	~	316	~	337	284	266
2	8-7-6-4	~	~	~	304	~	411	271	252
3	8-7-6-4	~	~	~	280	~	411	287	252
4	8-7-6-4	~	~	~	280	~	411	287	264

Scenario 2									
Method	Regional path	Region							
		1	2	3	4	5	6	7	8
1	1-2-5-6	312	365	~	~	284	259	~	~
2	1-2-5-6	277	373	~	~	276	246	~	~
3	1-2-5-6	253	441	~	~	280	258	~	~
4	1-2-5-6	191	455	~	~	285	236	~	~
1	3-2-4	~	365	346	311	~	~	~	~
2	3-2-4	~	387	353	268	~	~	~	~
3	3-2-4	~	695	363	243	~	~	~	~
4	3-2-4	~	695	370	179	~	~	~	~
1	8-7-6-4	~	~	~	311	~	259	304	299
2	8-7-6-4	~	~	~	268	~	304	180	290
3	8-7-6-4	~	~	~	266	~	411	276	269
4	8-7-6-4	~	~	~	295	~	411	273	267

$\sim$5000 and 6000 s. Between these time instants, we have vehicles traveling on regional paths $p = \{1256\}$ and $p = \{324\}$ and that have to cross region 2. The question now is which vehicles cross first this region. The trip-based MFD model is solved following an event-based scheme (Mariotte et al., 2017; Mariotte and Leclercq, 2018). The latter relies on a vehicles ranking that is set depending on the remaining travel distances in the region. In free-flow conditions, this vehicles ranking allows to calculate the next event, i.e. when the vehicle enters or exits the region. In congestion conditions, there are entry capacicity constraints that are applied. Therefore, how do we handle two vehicles traveling on different regional paths and that want to enter the same region that is already congested. We adopt classical merging and diverging schemes from Traffic Flow Theory (Mariotte et al., 2017; Mariotte and Leclercq, 2018) to calculate which is the next vehicle that is allowed to enter the region. In the case of region 2, the first vehicles to enter are traveling on regional path $p = \{324\}$. On the other hand, vehicles traveling on regional path $p = \{1256\}$ have to travel at low speed in region 1 before entering region 2. These vehicles queuing in region 1 increase the accumulation between $\sim$5000 and 6000 s. In region 3, we observe an accumulation peak between $\sim$4000 and 6000 s originating from the demand peak of OD pair 3-4 (see Fig. 11 (a)). The predicted traffic states for all four levels of information are similar as the average trip lengths for regional path $p = \{324\}$ inside region 3 are also close. In region 2, we observe two accumulation peaks between $\sim$1500 and 3700 s and between $\sim$4500 and 6000 s. The first accumulation peak corresponds to the vehicles traveling on OD pair 1-6. The second accumulation peak corresponds to the vehicles traveling on regional path $p = \{324\}$. We can observe that the accumulation is lower for M_1 than for the other three levels of information M_2 to M_4. Indeed, the travel distance given by M_1 for region 2 is 538 m, shorter than the 626 m for M_2 and 653 m for both M_3 and M_4. Since the travel distance is shorter, vehicles complete their travels faster inside region 2 for the case of M_1, leading to less accumulation. Region 8 is an origin region for the OD pair 8-4. We observe the presence of an accumulation peak between $\sim$3000 and 4000 s that represents the demand of vehicles traveling on regional path $p = \{8764\}$.

We now analyze regions 5, 6 and 7 in conjunction. In region 7, we observe an accumulation peak between $\sim$3000 and 4500 s due to the vehicles circulating on regional paths $p = \{12576\}$ and $p = \{8764\}$. The peaks are much higher for M_1 and M_2 than for M_3 and M_4. Moreover, the predicted traffic states are different for M_1 and M_2. We will now dissect the sources

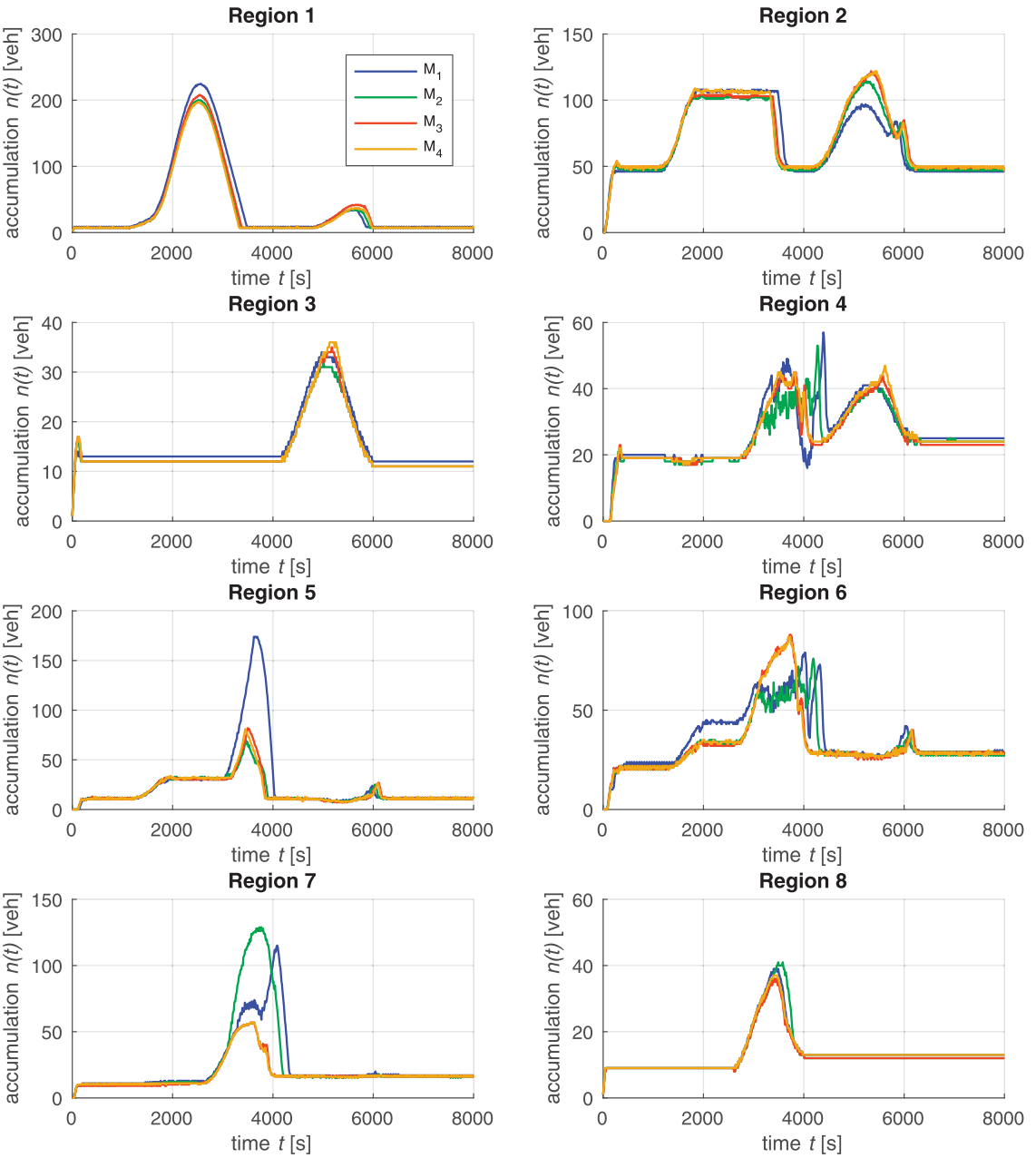

Fig. 12. Evolution of $n(t)$ during the modeling period for Scenario 1, considering the four methods for calculating the regional trip lengths.

of these different traffic predictions. We first focus on regional path $p = \{8764\}$. The average travel distance of this regional path inside region 7 is $\sim$280 m for all four levels of information M_1 and M_4. Therefore, all vehicles will have similar travel times in region 7. The predicted traffic states for all four levels of information are similar and do not explain the observed differences in the accumulation peaks for region 7. We now focus on regional path $p = \{12576\}$. The average travel distances inside region 7 are $\sim$280 m for M_1 and M_2 and $\sim$100 m for M_3 and M_4. Vehicles need less time to complete their crossing of region 7 for M_3 and M_4 compared to M_1 and M_2. This explains why the accumulation peak is lower for M_3 and M_4 than the other two cases. Nevertheless, this does not explain the differences between the traffic state predictions for M_1 and M_2 in region 7. In fact, for the case of M_1, there is competition between both regional paths $p = \{12576\}$ and $p = \{8764\}$ for which vehicles are allowed to enter first region 7 (Mariotte et al., 2017; Mariotte and Leclercq, 2018). Vehicles traveling on regional path $p = \{8764\}$ are the first ones to travel on region 7. Thus, the vehicles traveling on the regional path $p = \{12576\}$ queue inside region 5, waiting to be allowed to enter region 7. This also explains the much larger accumulation of vehicles in region 5 and for M_1, compared to the other three levels of information. In region 6, we observe an increase in the accumulation of

vehicles between 2000 and 3000 s. This is due to the arrival of vehicles at their destination, that are traveling on OD pair 1–6. Table 7 shows that for both regional paths $p = \{1256\}$ and $p = \{12576\}$, the average travel distance inside region 6 is higher for M_1 compared to the other three levels of information. As the average travel distance is longer for M_1, vehicles need more time to complete their trips. Consequently, the accumulation increases as observed in Fig. 12, compared to the other three levels of information. Between 3000 and 4000 s, there is an interaction between the vehicles traveling on OD pair 1-6 that are arriving at their destination and those traveling on regional path $p = \{8764\}$. The accumulation of vehicles in region 6 and between 3000 and 4000 s, increases due to two reasons. First, the demand peak of OD pair 8-4 crosses region 6 between these instants. Second, the average travel distance for the regional path $p = \{8764\}$ inside region 6 is much longer than for the other two regional paths $p = \{1256\}$ and $p = \{12576\}$ (see Table 7). Therefore, vehicles need more time to complete their trips and the accumulation increases. We also observe a similar evolution trend of the predicted traffic states for M_1 and M_2, between 3000 and 4000 s. Despite the fact that the average travel distance for M_2 is longer than for M_1 for the regional path $p = \{8764\}$, this is compensated by the opposite trend of both regional paths $p = \{1256\}$ and $p = \{12576\}$ (see Table 7). This effect leads to very close predictions of the traffic states for M_1 and M_2. Region 4 is a destination region for both OD pairs 3-4 and 8-4. We observe the presence of two accumulation peaks between $\sim$3000 and 4000 s and $\sim$4500 and 6000 s, that represent the arrivals of vehicles from regions 8 and 3, respectively.

We observe that the tuning of the trip lengths has a non-negligible influence on the predicted traffic states by a trip-based MFD model. In fact, we have access to the full information on the city network trip patterns. Therefore we know the regional path flows distributions and the total travel distance for each region is identical for all four levels of information M_1 to M_4. Nevertheless, the traffic dynamics inside the regions is different, especially during more congested states or when capacity restrictions at the regions entries arise.

4.3. Influence of trip length tuning under the regional network DUE

In this section, we investigate the effect of the trip length setting on the predicted traffic states under the regional network Deterministic User Equilibrium. Fig. 13 depicts the predicted traffic states for Scenario 2 for the eight regions of the 6th district Lyon network (see Fig. 4). As in the previous section, the traffic states are represented by the temporal evolution of vehicle accumulation $n(t)$ inside each region for the four levels of information M_1 to M_4. The average trip lengths for the three regional paths used by drivers, are listed in Table 7.

It can be seen that the traffic dynamics is clearly influenced by the tuning of the trip lengths inside the regions. Moreover, M_3 gives traffic state predictions similar to M_4, which are very different from those for M_1 and M_2. We now dissect the traffic dynamics inside each region in more detail. We first focus our attention on the predicted traffic states for origin regions 1 and 3 of regional paths $p = \{1256\}$ and $p = \{324\}$, respectively. In region 1, we observe an accumulation peak between $\sim$2000 and 4000 s, whereas in region 3, we observe an accumulation peak between $\sim$3500 and 6000 s. These peaks are caused by the increase of demand on OD pairs 1-6 and 3-4 (see Fig. 11 (b)) between these instants. Nevertheless, we observe that these peaks are much higher for M_3 and M_4 than for M_1 and M_2 for both regions 1 and 3. This might be due to longer trip lengths assigned by M_3 and M_4, leading to the existence of a potential bottleneck (i.e., a reduction of the inflow capacity for this region and specific regional path) of the regional paths. However, this is not the case, as observed by the average trip lengths for regional paths $p = \{1256\}$ and $p = \{324\}$ inside regions 1 and 3, respectively, listed in Table 7 for M_3 and M_4. To investigate the source of the higher accumulation peaks for M_3 and M_4 in regions 1 and 3, we now focus on the predicted traffic states for region 2. We observe the presence of two accumulation peaks. The first accumulation peaks appear between $\sim$1000 and 3000 s for M_1 and M_2 and between $\sim$1000 and 4000 s for M_3 and M_4. This first accumulation peak is generated by the vehicles traveling on regional path $p = \{1256\}$. The average trip lengths for M_3 and M_4 are longer than for M_1 and M_2, for both regional paths $p = \{1256\}$ and $p = \{324\}$ inside region 2 (see Table 7). Therefore, vehicles need more time to cross region 2 and the accumulation peak lasts longer for M_3 and M_4. For these two cases, region 2 becomes a bottleneck of regional path $p = \{1256\}$. Vehicles will be blocked at the entry of region 2 and queue in region 1, increasing the accumulation as observed. On the other hand, the second accumulation peak is due to the inflow of vehicles traveling on regional path $p = \{324\}$. This second accumulation peak appears between $\sim$3500 and 4000 s for M_1 and M_2, whereas for M_3 and M_4, this second accumulation peak appears between $\sim$4500 and 6000 s. This is because, vehicles traveling on regional path $p = \{1256\}$ are the first to cross region 2. Vehicles traveling on regional path $p = \{324\}$ queue in region 3, while waiting to be allowed to enter region 2. This increases the accumulation in region 3 for M_3 and M_4. The traffic dynamics in region 2 strongly influences the traffic states on the other regions. One example is region 5. For this region, all four levels of information yield similar average trip lengths (see Table 7) for regional path $p = \{1256\}$. However, the traffic dynamics is very different. The first accumulation peak observed in region 2 for M_1 and M_2 is then observed in region 5 between $\sim$1000 and 3000 s. The inflow of vehicles in region 5 for M_3 and M_4 will be lower thanks to the bottleneck in region 2. Consequently, the accumulation peak is lower and lasts longer (between $\sim$1500 and 4000 s).

We now focus our attention on the traffic dynamics in region 6. This region is crossed by regional path $p = \{8764\}$ and is the destination region for regional path $p = \{1256\}$. The average trip lengths for all four levels of information are similar for regional path $p = \{1256\}$ in region 6 (see Table 7). However, region 6 is a potential bottleneck for regional path $p = \{8764\}$ and for M_3 and M_4. These trip lengths are longer for the latter two cases than for M_1 and M_2. A longer trip length means that vehicles need more time to complete their trips and therefore the accumulation peak lasts longer for M_3 and M_4 than for M_1 and M_2. Vehicles traveling on regional path $p = \{1256\}$ are the first to cross region 6. Therefore, several vehicles

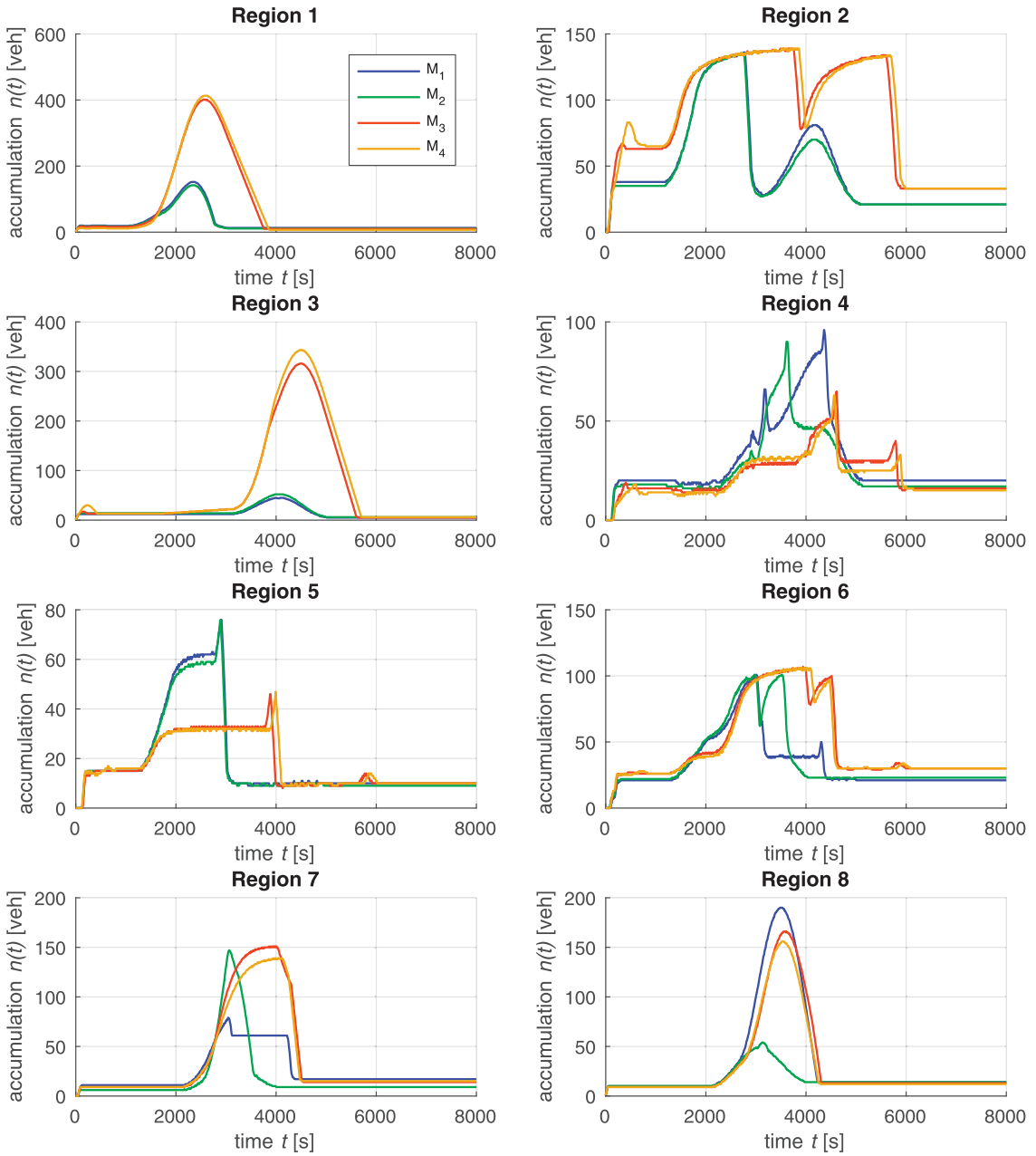

Fig. 13. Same as Fig. 12, but for Scenario 2.

traveling on regional path $p = \{8764\}$ are waiting in region 7 to be allowed to enter region 6, before $\sim$4000 s. After this instant, most of the vehicles traveling on regional path $p = \{1256\}$ have already completed their trips and the accumulation drops. This allows the vehicles queuing in regions 7 and 8 to pursue their trips in region 6 and the accumulation increases again until $\sim$4500 s. In the case of M_1, the largest fraction of vehicles finishes its trips earlier, around $\sim$3000 s. This allows vehicles traveling on regional path $p = \{8764\}$ to cross this region, without queuing in region 7. In the case of M_2, the regional path $p = \{8764\}$ has a longer trip length in region 6 than in 7 (see Table 7). Therefore, region 6 is a potential bottleneck and vehicles queue inside region 7. This is evidenced by the accumulation peak verified in region 7 between $\sim$2500 and 3500 s. Moreover vehicles traveling on regional path $p = \{1256\}$ are allowed to complete their trips first inside region 6. After $\sim$3000 s most of the vehicles traveling on the latter regional path have completed their trips and the accumulation peak drops. Vehicles traveling on regional path $p = \{8764\}$ and queuing in region 7 are then allowed to enter region 6, again increasing the accumulation peak until $\sim$3500 s.

The traffic dynamics in both regions 2 and 6 clearly influence the traffic states in region 4, which is a destination region for both regional paths $p = \{324\}$ and $p = \{8764\}$. The accumulation peak for M_2, between $\sim$3500 and 4000 s, represents the arrivals of vehicles traveling on regional path $p = \{8764\}$ to their destination region. The accumulation peak for M_1, between $\sim$3500 and 4500 s represents the arrival of vehicles traveling on regional path $p = \{324\}$. The increase of the accumulation between $\sim$3000 and 6000 s for M_3 and M_4 represents the arrival of vehicles traveling on regional path $p = \{324\}$, whereas the sudden increase of the accumulation verified for M_3 and M_4, between $\sim$4000 and 4500 s, is due to the arrival of vehicles traveling on regional path $p = \{8764\}$.

We observed that the definition of the trip lengths at the regional level clearly modifies the traffic dynamics. The four levels of information proved to be of crucial importance for the tuning of trip lengths for multi-regional MFD-based applications. In particular, M_3 and M_4 give similar predictions of the traffic states inside the regions.

5. Conclusions

To our knowledge, in this paper we proposed for the first time in the literature, a methodological framework to explicitly tune trip lengths for multi-regional MFD-based model applications and to investigate the effect on MFD dynamics. This framework is based on a set of virtual trips on the city network as well as its partitioning. It differentiates the definition of the trip lengths into four levels of information of the regional network. We analyzed the proposed framework on the 6^{th} district of the Lyon network, which is divided into eight regions.

Initially, we investigated how to properly define the set of virtual trips. We showed that the criterion $N_{cov}^{nodes}(N_{od})$ gives good prior guidance for the calibration of the virtual trips set Γ. We performed an in-depth analysis of the influence of the N_{od} setting on the calculation of the trip lengths through the four levels of information. M_3 and M_4 were the most sensitive to the N_{od} value, for which we identified several outliers in the ρ distributions. These outliers came from distributions that were not statistically significant. One way to circumvent this is to filter the non-significant regional paths from the set Ψ.

Secondly, we highlighted the importance of considering more detailed levels of information such as M_3 and M_4. We showed that the average trip length estimated by M_1 is not representative of all the possible average trip lengths estimated by M_4 for one region. We also discussed another important limitation of M_1. We showed that two regional paths defined by the same regions, but in a reverse order, are assigned similar average trip lengths by M_1. Nevertheless, this was not the case of M_4 which captured important information from the directed links of the city network.

Thirdly, we proposed a methodology to quickly update the trip lengths for M_1 to M_3, for updated regional OD matrices and without the need to reestimate the set of virtual trips. The methodology provided good estimation results at a low computational cost.

Lastly, we analyzed the importance of properly tuning the trip lengths for multi-regional trip-based MFD model applications. We considered two different model scenarios. In one scenario, we knew the full information on the trip patterns in the city network. Therefore, the regional path flow distributions are known and the total travel distances in each region is the same independent of the level of information M_1 to M_4. In the other scenario, we did not know the real trip patterns in the city network and reached a solution for the regional network Deterministic User Equilibrium. We showed that the tuning of the trip lengths changes the traffic dynamics in the region for both scenarios. For the first scenario, the differences in the traffic dynamics particularly appear during more congested states or when capacity restrictions at the regions entries arise. We also showed that the predicted traffic states for M_3 and M_4 were similar and in several regions different from those predicted for M_1 and M_2.

We conclude that M_3 and M_4 should be the preferred levels of information for calculating trip length distributions for multi-regional MFD-models applications.

In future research, we plan to investigate the time-dependence of trip lengths and propose robust and computationally light methodologies for updating them according to the traffic dynamics in the regional network.

Acknowledgments

We thank the anonymous reviewers for their comments and suggestions and have improved this manuscript. This project is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no 646592 - MAGnUM project; and grant agreement no 338205 - METAFERW project). S. F. A. Batista also acknowledges funding support by the region Auvergne-Rhône-Alpes (ARC7 Research Program) and from the University of Lyon supported by the PALSE mobility program.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.trb.2019.02.009.

References

Aboudolas, K., Geroliminis, N., 2013. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. Transp. Res. Part B Methodol. 55, 265–281. doi:10.1016/j.trb.2013.07.003.

Ambühl, L., Menendez, M., 2016. Data fusion algorithm for macroscopic fundamental diagram estimation. Transp. Res. Part C Emerg. Technol. 71, 184–197. doi:10.1016/j.trc.2016.07.013.
Arnott, R., 2013. A bathtub model of downtown traffic congestion. J. Urban Econ. 76, 110–121. doi:10.1016/j.jue.2013.01.001.
Azevedo, J., Santos Costa, M., Silvestre Madeira, J., Vieira Martins, E., 1993. An algorithm for the ranking of shortest paths. Eur. J. Oper. Res. 69, 97–106. doi:10.1016/0377-2217(93)90095-5.
Batista, S.F.A., Leclercq, L., 2018. Introduction of multi-regional MFD-based models with route choices: the definition of regional paths. In: Proceedings of the PLURIS Eight LUSO-BRAZILIAN CONGRESS for Urban, Regional, Integrated and Sustainable Planning. Coimbra, Portugal.
Casadei, G., Bertrand, V., Gouin, B., Canudas-de-Wit, C., 2018. Aggregation and travel time calculation over large scale traffic networks: an empiric study on the Grenoble city. Transp. Res. Part C Emerg. Technol. 95, 713–730. doi:10.1016/j.trc.2018.07.033.
Daganzo, C., 2007. Urban gridlock: macroscopic modeling and mitigation approaches. Transp. Res. Part B Methodol. 41, 49–62. doi:10.1016/j.trb.2006.03.001.
de la Barra, T., Perez, B., Anez, J., 1993. Multidimensional path search and assignment. In: Proceedings of the Twenty-First PTRC Summer Annual Meeting. Manchester, England.
Ekbatani, M., Papageorgiou, M., Papamichail, I., 2013. Urban congestion gating control based on reduced operational network fundamental diagrams. Transp. Res. Part C Emerg. Technol. 33, 74–87. doi:10.1016/j.trc.2013.04.010.
Flötteröd, G., Bierlaire, 2013. Metropolis-hastings sampling of paths. Transp. Res. Part B 48, 53–66. doi:10.1016/j.trb.2012.11.002.
Fosgerau, M., 2015. Congestion in the bathtub. Econ. Transp. 4, 241–255. doi:10.1016/j.ecotra.2015.08.001.
Geroliminis, N., 2015. Cruising-for-parking in congested cities with an MFD representation. Econ. Transp. 4 (3), 156–165. doi:10.1016/j.ecotra.2015.04.001.
Geroliminis, N., Daganzo, C., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transp. Res. Part B Methodol. 42, 759–770. doi:10.1016/j.trb.2008.02.002.
Geroliminis, N., Haddad, J., Ramezan, M., 2013. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: a model predictive approach. IEEE Trans. Intell. Transp. Syst. 14, 348–359. doi:10.1109/tits.2012.2216877.
Geroliminis, N., Sun, J., 2011. Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks. Proc. Soc. Behav. Sci. 17, 213–228. doi:10.1016/j.sbspro.2011.04.515.
Godfrey, J.W., 1969. The mechanism of a road network. Traffic Eng. Control 11, 323–327.
Haddad, J., 2017. Optimal perimeter control synthesis for two urban regions with aggregate boundary queue dynamics. Transp. Res. Part B Methodol. 96, 1–25. doi:10.1016/j.trb.2016.10.016.
Herman, R., Prigogine, I., 1979. A two-fluid approach to town traffic. Science 204, 148–151. doi:10.1126/science.204.4389.148.
Ji, Y., Geroliminis, N., 2012. On the spatial partitioning of urban transportation networks. Transp. Res. Part B Methodol. 46, 1639–1656. doi:10.1016/j.trb.2012.08.005.
Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. Transp. Res. Part B Methodol. 46 (10), 1393–1403. doi:10.1016/j.trb.2012.06.008.
Kouvelas, A., Saeedmanesh, M., Geroliminis, N., 2017. Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. Transp. Res. Part B Methodol. 96, 26–45. doi:10.1016/j.trb.2016.10.011.
Lamotte, R., Geroliminis, N., 2016. The morning commute in urban areas: Insights from theory and simulation. In: Proceedings of the Transportation Research Board Ninety-Fifth Annual Meeting, pp. 16–2003.
Lamotte, R., Murashkin, M., Kouvelas, A., Geroliminis, N., 2018. Dynamic modeling of trip completion rate in urban areas with MFD representations. In: Proceedings of the Ninety-Seventh Annual Meeting Transportation Research Board. Washington DC, USA.
Leclercq, L., Parzani, C., Knoop, V.L., Amourette, J., Hoogendoorn, S., 2015. Macroscopic traffic dynamics with heterogeneous route patterns. Transp. Res. Part C 59, 292–307. doi:10.1016/j.trb.2015.05.006.
Leclercq, L., Sénécat, A., Mariotte, G., 2017. Dynamic macroscopic simulation of on-street parking search: a trip-based approach. Transp. Res. Part B Methodol. 101, 268–282. doi:10.1016/j.trb.2017.04.004.
Loder, A., Ambühl, L., Menendez, M., Axhausen, K.W., 2017. Empirics of multi-modal traffic networks using the 3D macroscopic fundamental diagram. Transp. Res. Part C Emerg. Technol. 82, 88–101. doi:10.1016/j.trc.2017.06.009.
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, 1–11. doi:10.1038/s41598-017-14237-8.
Mahmassani, H., Williams, J.C., Herman, R., 1984. Investigation of network-level traffic flow relationships: some simulation results. Transp. Res. Rec. J. Transp. Res. Board 971, 121–130. doi:10.3141/2315-16.
Mariotte, G., Leclercq, L., 2018. MFD-based simulation: spillbacks in multi-reservoir networks. In: Proceedings of the Ninety-Seventh Annual Meeting Transportation Research Board. Washington DC, USA.
Mariotte, G., Leclercq, L., Laval, J.A., 2017. Macroscopic urban dynamics: analytical and numerical comparisons of existing models. Transp. Res. Part B 101, 245–267. doi:10.1016/j.trb.2017.04.002.
Nielsen, O.A., 2000. A stochastic transit assignment model considering differences in passengers utility functions. Transp. Res. Part B 34 (5), 377–402. doi:10.1016/S0191-2615(99)00029-6.
Prato, C.G., Bekhor, S., 2006. Applying branch and bound techniques to route choice set generation. Transp. Res. Rec. 19–28. doi:10.3141/1985-03.
Ramezani, M., Haddad, J., Geroliminis, N., 2015. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. Transp. Res. Part B 74, 1–19. doi:10.1016/j.trb.2014.12.010.
Saeedmanesh, M., Geroliminis, N., 2016. Clustering of heterogeneous networks with directional flows based on “snake” similarities. Transp. Res. Part B Methodol. 91, 250–269. doi:10.1016/j.trb.2016.05.008.
Saeedmanesh, M., Geroliminis, N., 2017. Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. Transp. Res. Proc. 23, 962–979. doi:10.1016/j.trb.2017.08.021.
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. Transp. Res. Rec. J. Transp. Res. Board 2029, 22–30. doi:10.3141/2029-03.
Sheffi, Y., 1985. Urban Transportation Networks: Equilibrium Analysis With Mathematical Programming Methods. Prentice Hall Inc., United States of America. Chapter 10 and 11
Sirmatel, I.I., Geroliminis, N., 2018. Economic model predictive control of large-scale urban road networks via perimeter control and regional route guidance. IEEE Trans. Intell. Transp. Syst. 19, 1112–1121. doi:10.1109/TITS.2017.2716541.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. Inst. Civ. Eng. 1, 325–362. doi:10.1680/ipeds.1952.11259.
Yang, K., Zheng, N., Menendez, M., 2018. Multi-scale perimeter control approach in a connected-vehicle environment. Transp. Res. Part C Emerg. Technol. 94, 32–49. doi:10.1016/j.trc.2017.08.014.
Yildirimoglu, M., Geroliminis, N., 2014. Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams. Transp. Res. Part B Methodol. 70, 186–200. doi:10.1016/j.trb.2014.09.002.
Yildirimoglu, M., Ramezani, M., Geroliminis, N., 2015. Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. Transp. Res. Part C Emerg. Technol. 59, 404–420. Special Issue on International Symposium on Transportation and Traffic Theory. doi: 10.1016/j.trc.2015.05.009.
Yildirimoglu, M., Sirmatel, I.I., Geroliminis, N., 2018. Hierarchical control of heterogeneous large-scale urban road networks via path assignment and regional route guidance. Transp. Res. Part B Methodol. 118, 106–123. doi:10.1016/j.trb.2018.10.007.
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. Transp. Res. Part B Methodol. doi:10.1016/j.trb.2017.09.008.