Fostering synergy between transit and Autonomous Mobility-on-Demand systems: A dynamic modeling approach for the morning commute problem

Mélanie Cortina, Nicolas Chiabaut, Ludovic Leclercq

Fostering synergy between transit and Autonomous Mobility-on-Demand systems: a dynamic modeling approach for the morning commute problem

Mélanie Cortina[a,*], Nicolas Chiabaut[b], Ludovic Leclercq[a]

[a] *Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT, F-69518, Lyon, France*
[b] *Citec Ingénieurs Conseils SAS, F-69342, Lyon, France*

Abstract

Autonomous Mobility-On-Demand (AMoD) provides new options for the morning commute problem. The flexibility of AMoD could help boost Public Transportation (PT) attractiveness and accessibility. Intermodal AMoD systems could become a competitive alternative to personal cars. However, considering the convenience, comfort, and expected low fares of autonomous vehicles, the risk of competition between privately operated AMoD and PT exists. The joint design of PT and AMoD can foster their cooperation. This study investigates the joint PT-AMoD design problem in a many-to-one multimodal corridor where three transportation alternatives are available: full personal car on a congested freeway, walk and massive rapid transit (MRT), or autonomous vehicle (AV) and MRT. We introduce a simple dynamic model incorporating time-dependent mode and route choice subject to user equilibrium (UE) constraints. The presented model: (i) accounts for how UE settles and evolves, (ii) provides insight on PT-AMoD cooperation opportunities and competition risks depending on the design choices, (iii) is compatible with design optimization heuristics. We apply the model to a realistic scenario based in the city of Lyon (France). The number of MRT stations, their locations, the number of AVs fleets, and their coverage zone boundaries are optimized with a metaheuristic. The optimization is conducted under three policies regarding AMoD (protectionism, opportunism, liberalism) and three priority objectives (maximize MRT usage, minimize travel times, mitigate cars). By evaluating the potential benefits of each policy, we formulate recommendations for the transportation authority.

Keywords: morning commute, corridor, Autonomous Mobility-on-Demand (AMoD), user equilibrium (UE), intermodal trips, design

Highlights

- We compare a static and a dynamic model on their ability to capture PT-AMoD cooperation or competition.
- UE principles help to explain design impact on mobility patterns.
- Direct or long first-mile schemes emerge from individual usages depending on the corridor design.
- One AV fleet covering the whole corridor naturally competes with MRT.
- Segmenting AMoD coverage zone reduces competition with MRT.
- Protectionism, opportunism, and liberalism policies toward AMoD are compared.
- Opportunism performs the best, whatever the priority objective of the city.

*Corresponding author:
Address: 3 rue Maurice Audin, 69120, Vaulx-en-Velin, France
Tel: +33 (0) 4 72 04 77 11
Email address: melanie.cortina@entpe.fr

1. Introduction

Commuting journeys represent about 17% of daily trips (Ermans et al., 2018) and lead to significant delays. They heavily impact citizens' health (Titos et al., 2015) and contribute to global warming (Nicolas et al., 2012; André and Vieira da Rocha, 2020). Private companies are developing new mobility concepts and technologies to answer these considerations (e.g., Bagloee et al. (2019)). Among prospective solutions, Autonomous Mobility-on-Demand (AMoD) seems promising (Pavone, 2015).

AMoD presents advantageous features for the morning commute. As autonomous vehicles (AVs) are shared, the number of vehicles operating in the city could decline (Bischoff and Maciejewski, 2016; Wilson, 2015). More flexible than scheduled-based feeder transit services, AMoD could help extend the catchment area of mass transit stations (Mounce and Nelson, 2019; Basu et al., 2018). Thanks to lower operating costs compared to human-driven taxis (Becker et al., 2020), continuous rebalancing is conceivable and could improve the level of service. The compliance of AVs with the centralized operator eases the control of the fleet (Yao et al., 2020). Tientrakool et al. (2011); Litman (2021); Greenblatt and Saxena (2015); Kondor et al. (2019); Fagnant and Kockelman (2015) discuss many other potential benefits of AVs in terms of traffic, parking, safety, energy, pollution, and social inclusion.

Negative impacts have also been highlighted, such as the vehicles-miles traveled growth (Dang et al., 2021; Zwick et al., 2021) or the question of interactions with public transportation (PT), which is already an essential concern in the context of MoD (Hall et al., 2018; Rayle et al., 2016; Sadowsky and Nelson, 2017). Competition with PT will be even more relevant with AMoD, considering their large coverage zones, convenience, comfort, and lower fares Gurumurthy et al. (2020).

Regulation strategies are needed to prevent direct competition between AVs and PT (Zardini et al., 2022). For example, some studies consider pricing schemes. Salazar et al. (2018, 2020); Zgraggen et al. (2019) show that an intermodal AMoD can achieve significant benefits in terms of travel times, costs, and emissions compared to an isolated AMoD. While Zgraggen et al. (2019) suppose a centrally managed system for routing AVs and compliant travelers, Salazar et al. (2018, 2020) use an optimized pricing and tolling scheme to push the system closer to the social optimum. In Simoni et al. (2019), a congestion pricing scheme applying to both AVs and conventional cars improves social welfare. Gurumurthy et al. (2020) indicate that high fares for AVs mitigate AMoD-PT competition. Reck and Axhausen (2019) and Zhu et al. (2021) analyze the impact of subsidizing first and last-mile trips.

Other studies consider AMoD design and operation rules as levers to foster synergy between AMoD and PT. For example, in Zhou et al. (2019), AMoD approves only first-mile, last-mile, or door-to-door trips, depending on the day period. Instead, Militão and Tirachini (2021) limit the performance of the service by licensing a reduced number of AVs. Gurumurthy et al. (2021) evaluate the impact of disaggregating the service into several fleets operating within bounded areas.

The transportation authority (TA) holds a decisive control parameter by delivering licenses to AVs in a given geographical area. We focus on this regulation measure and go further by jointly optimizing AMoD and PT configurations. Indeed, intermodal trips require efficient PT, with short waiting times and few transfers (Gallotti and Barthelemy, 2015; Zhu et al., 2020), to become viable alternatives to personal cars. AMoD and PT designs must be connected to guarantee that both systems complement each other.

A crucial question is how to jointly design PT and AMoD to foster cooperation and reach environmentally-driven policy goals. Our framework, presented in Figure 1, tackles the many-to-one corridor case where AMoD serves the first-mile in connection with a massive rapid transit (MRT).

The PT design question in corridors has been mainly addressed using static approaches. Older studies do not include congestion (Vuchic and Newell, 1968) while many works employ simple volume delay functions (Wang et al., 2004; Liu et al., 2009). They miss the distribution of demand over time and departure rate evolution. These approaches either ignore or poorly reproduce congestion dynamics which is essential during peak hours. The chosen objective functions mainly focus on performance by minimizing the total travel time or cost (Vuchic and Newell, 1968; Wirasinghe et al., 1977; Shen and Zhang, 2009) but rarely consider current objectives of cities, such as PT usage maximization (Vuchic, 1969).

Several studies evaluate a PT-AMoD system under different design parameters in a sensitivity analysis manner, but they often focus on AMoD only. The AMoD system serves a predetermined (exogenous) demand emanating from or targeting a PT line. There is no representation of intermodal trips: the PT-AMoD connection is assumed. For example, Scheltes and de Almeida Correia (2017) use a travel demand survey to isolate the AMoD market. In contrast, Shen

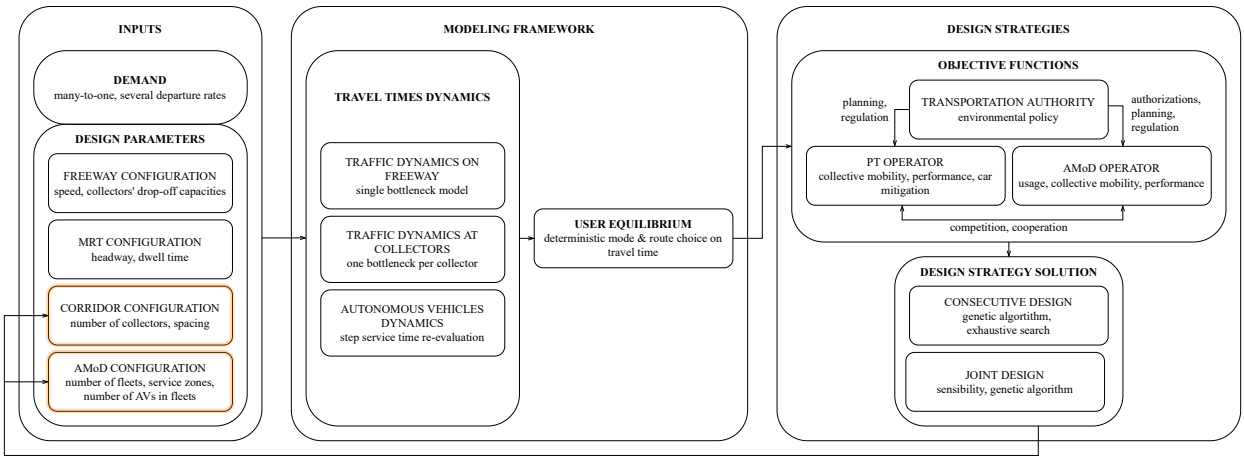


Figure 1: General framework of the paper. Our modeling contribution allows to answer the MRT-AMoD design problem with corridor and AMoD configurations as decision variables.

et al. (2018) and Rifki et al. (2021) respectively use a transit smart card data-driven analysis and a macro-economic model. In these works, the coupling between supply and demand is lacking.

In PT-AMoD design literature, optimization and simulation approaches exist. Most of the optimization models proposed are static. They do not account for congestion, except Wei et al. (2022). Liu and Ouyang (2021) optimize several design parameters with a constrained non-linear program. The study focuses on the integrated PT-AMoD system and considers no other transportation alternative or mode-route choice model. In contrast, Basciftci and Van Hentenryck (2021) consider the car alternative and induced demand. They formulate a bi-level optimization problem, including a simple mode choice model. However, they lack an (A)MoD model by neglecting the waiting time for being picked up by an on-demand vehicle. Shan et al. (2021) make the same assumption. Kumar and Khani (2022) use a queuing model to approximate this waiting time. They include assignment variables in a mixed integer non-linear program to capture travelers' behavior in the network, but the assignment problem formulated does not account for the selfishness of commuters. It is not a user equilibrium (UE).

Few studies tackle the PT-AMoD design with a simulation approach under the dynamic UE. While Pinto et al. (2020) work in a cooperation context, Mo et al. (2021) question a competition scheme where AVs are unregulated and profit-oriented. These works ignore the competition with cars. They use agent-based simulation, which relies on a computationally greedy convergence loop (e.g., exact swapping or heuristics such as the method of successive average, evolutionary algorithms (Ameli, 2019)) to approximate the stable state of the system. The PT-AMoD design optimization process, on top of dynamic traffic assignment, is costly, and its convergence is uncertain. The complexity of the resolution algorithm limits the number of decision variables investigated (transit routes removal from an initial set, transit headway, and number of AVs). For these reasons, the approach is hardly compatible with the joint optimization problem. Moreover, it suffers from the black box effect, providing no knowledge about how UE settles and evolves.

State of the art lacks methods compatible with design optimization processes to solve and deeply understand the dynamic UE. Our work partly fills this gap by proposing a simple but dynamic model for calculating UE in the morning commute corridor context. Its computational efficiency makes it easy to couple with a design optimization heuristic. We use this model to solve the MRT-AMoD design problem with three different objectives for the TA: minimizing travel times, maximizing MRT usage, and minimizing car usage. The MRT design parameters are the number of MRT stations and their locations. The AMoD design parameters are the number of fleets and the boundaries of their coverage zones. If the corridor abstraction is restrictive, it is sufficiently generic to apply to many urban areas with little data. It allows drawing general conclusions on the interactions between MRT and AMoD. It helps analyze the impact of the TA policy toward AMoD on the Western Lyon (France) corridor. We compare three policies where TA prohibits AVs (protectionism), authorizes AVs with (opportunism)) or without (liberalism) regulation.

The paper is organized as follows. Section 2 presents the modeling framework. It introduces the problem set-

ting and assumptions, describes the proposed dynamic model, including how UE settles and evolves, and justifies its computational efficiency. Section 3 shows the advantages of our model in the context of PT-AMoD design. Even if it is restrictive regarding intermodality and network, it provides general insight into PT-AMoD cooperation opportunities and competition risks depending on the design. It shows more cooperation-competition schemes compared to an equivalent static model. Understanding how UE settles and evolves allows a deep analysis of design influence on these schemes. Section 4 applies our dynamic model to the MRT-AMoD design problem in the Western Lyon case. It presents the numerical results obtained. Section 5 formulates recommendations for the TA and discusses the limitations of our approach.

2. Modeling framework

This section presents the modeling framework. It introduces the problem setting and assumptions, describes the proposed dynamic model, including how UE settles and evolves, and justifies its computational efficiency. The reader can refer to appendix A for notations.

2.1. Problem setting and assumptions

We focus on the morning commute in a corridor where the many-to-one mobility pattern is problematic because an unbalanced competition occurs between the freeway and an MRT. While cars provide a convenient door-to-door trip, the MRT suffers from accessibility issues. Consequently, car usage is substantial, and the freeway is congested daily.

The TA considers accrediting an AMoD private operator. The improvement of the MRT service (by extending the line, building new stations, or updating the train service pattern) and the regulation of AMoD are jointly studied to mitigate cars and improve MRT attractivity while satisfying a performance criterion.

Figure 2 presents the assumptions made to tackle this problem and the associated parameters. The following sub-sections describe each of them.

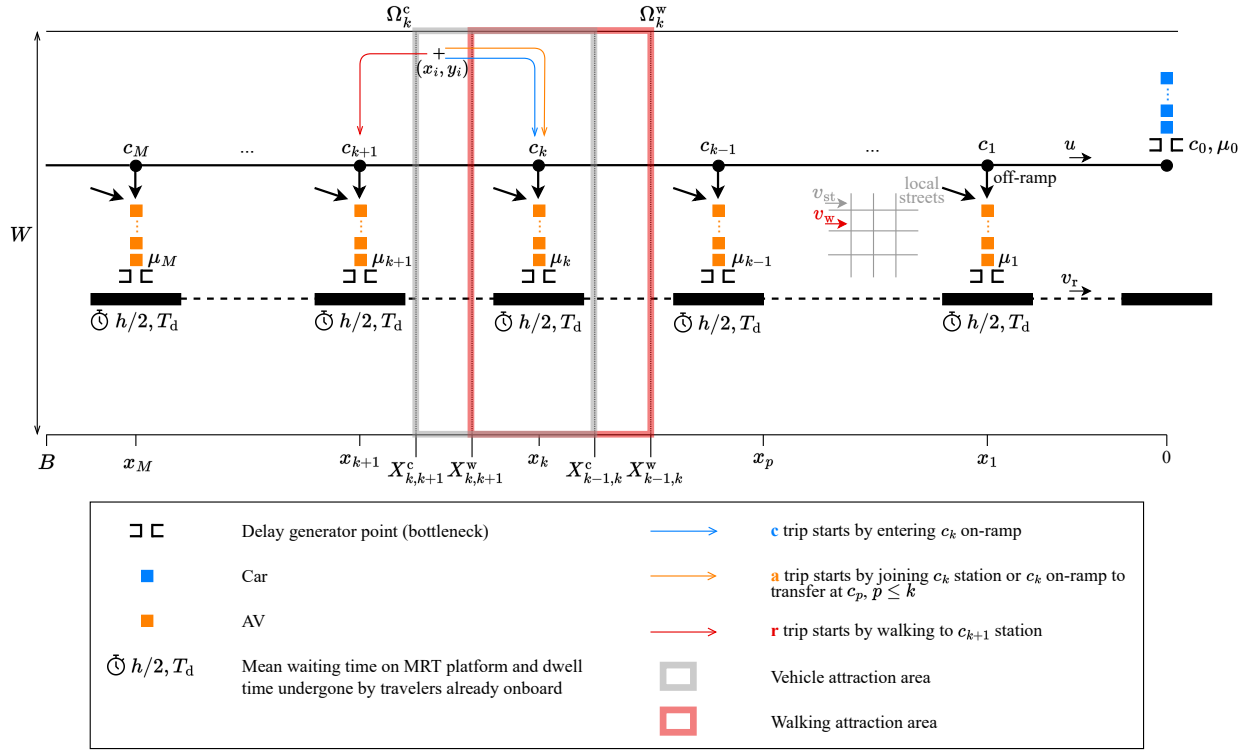


Figure 2: Modeling framework.

2.1.1. *Multimodal corridor infrastructure*

Let us consider the part of a monocentric city surrounding a linear transportation axis composed of a freeway and an MRT system (e.g., suburban rail, subway, or tram) as represented in Figure 3a. This corridor goes from the city boundary B to the center of the business district (CBD) located at $x = 0$ and has a given width W. Both freeway and MRT line are assumed to overlap at $y = 0$ and to be connected through $M + 1$ collectors c_k located along the corridor at x_k, $0 \leq k \leq M$. A collector is composed of a freeway on-ramp, an off-ramp, and an MRT station. Collector c_0 corresponds to the terminus station and the off-ramp leading to CBD.

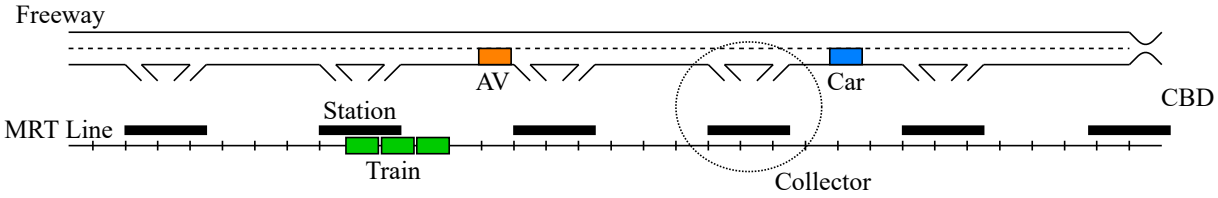
2.1.2. *Demand spatial profile*

Commuters depart from home and want to join the CBD. There are many origins for one destination. The origins of travelers distribute along the x and y axes. Traveler i's origin is located at (x_i, y_i). Collectors are connected to homes by an uncongested local streets grid that can be traveled at speed v_{st} by vehicle and at speed v_{w} by walk.

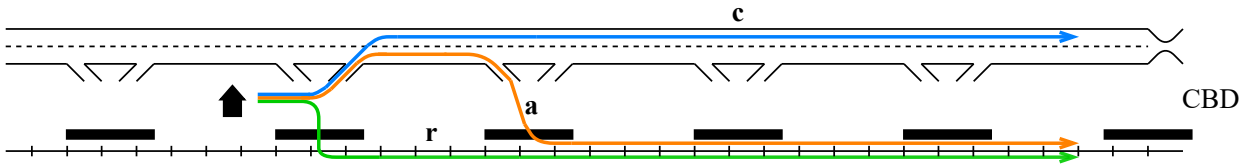
2.1.3. *Deterministic mode and route choice based on travel time*

Four modes of transportation are available: car, MRT, AV, and walk. Three options, represented in Figure 3b, are considered: car-only (**c**), combined walk and MRT (**r**), intermodal AV and MRT (**a**). Each traveler chooses the mode and route that provides minimal travel time. Backward movements on the freeway are prohibited. A traveler accessing infrastructures by c_k has to choose between $k+2$ itineraries:

- one corresponding to **c**: traveler drives until c_k, enters the freeway and keeps driving until c_0
- one corresponding to **r**: traveler walks to c_k station and takes the MRT until the terminus
- k corresponding to **a**: traveler rides an AV which joins c_k, then chooses to transfer to MRT at c_p, where $0 < p \leq k$. Indeed, AMoD operates within the corridor only, and AVs are not allowed to drop off commuters directly in the CBD.



(a) Freeway and MRT line are the two infrastructures in the corridor and are connected through collectors.



(b) Available options **c**, **r**, **a** respectively correspond to complete personal car, walk and MRT, AV and MRT trips.

Figure 3: Corridor sketch.

2.1.4. *Delay generator points*

Localized delay generator points exist in this network. A vehicle entering the freeway runs at speed u ($v_{\text{st}} < u$) in free-flow conditions. It may be caught in congestion due to a single capacity reduction point. The queuing delay experienced by a car exiting the freeway at c_0 is due to a demand rate higher than the destination off-ramp fixed capacity μ_0. Queues form at AV-to-MRT transfer points due to the fixed number of stopping spots and the fixed time T_{f} needed for the drop-off maneuver. These delays are the only ones AVs face. Continuum approximation of the drop-off times comes down to modeling AV-to-MRT transfer capacity at a station by a fixed value μ_k. In practice, adding drop-off spots will increase this capacity.

2.1.5. MRT operation

All stations identified by a collector c_k belong to the same MRT line: a train stops at each one. This line is assumed to have an unlimited passenger capacity, i.e., there is no left-behind on platforms because of saturated trains. Let us neglect the MRT rolling stock kinematics. Since the wasted time due to limited acceleration and braking capacities is ignored, no minimal spacing between two stations is required. The cruising speed of trains is v_r. We also neglect the relation between the number of travelers boarding at a station and the dwell time. This assumption is consistent with the unsaturated line hypothesis. Dwell time is therefore constant, equal for all stations, and noted T_d. The more collectors there are, the longer the total dwell time, and the lower the MRT line commercial speed. The design parameter M impacts the MRT line travel time to the CBD. Line headway h (in seconds) is considered stable during the morning commute. The time a traveler needs to wait at a station is approximated by $\frac{h}{2}$ (Fu et al., 2012).

2.1.6. Access to infrastructure

Each traveler has a fixed access collector by vehicle and a fixed access collector by walking. An access collector is chosen to minimize the travel time in free-flow conditions. The pool of commuters accessing the MRT line by walking to station k is called the walking attraction area of c_k and noted Ω_k^w. It forms a rectangle around c_k, gathering the origins of all commuters joining this collector rather than another by walking. Walking attraction areas boundaries can be computed by considering a traveler i departing from an origin located on the line defined as the intersection $\Omega_k^\mathrm{w} \cap \Omega_{k+1}^\mathrm{w}$ or as $x = X_{k,k+1}^\mathrm{w}$. On this line, travel times to c_k by accessing it directly or via c_{k+1} are equal.

Similarly, a rectangle vehicle attraction area Ω_k^c surrounds c_k. It gathers all commuters accessing the freeway through this collector by vehicle (car or AV) or directly accessing station k by AV. On the frontier $x = X_{k,k+1}^\mathrm{c}$ between Ω_k^c and Ω_{k+1}^c, travel times to c_k by entering freeway at on-ramp k or $k+1$ are equal for mode $\mathbf{c}$. For mode $\mathbf{a}$, with a given transfer collector c_p, travel times to c_p by accessing infrastructure through c_k or c_{k+1} are equal on the frontier $x = X_{k,k+1}^\mathrm{c}$. Figure 2 provides an example of vehicle and walking attraction areas boundaries betweenc_k and c_{k+1} where $v_\mathrm{w} < v_\mathrm{st} < u$.

2.1.7. AMoD operation

AMoD comprises one or several fleets of single-seat AVs. A fleet operates in a specific coverage zone, defined as a union of vehicle attraction areas. The fleet can provide service to every commuter departing from its coverage area and necessarily drops her off at a collector included in this zone. Booking is disabled: a commuter sends a request at the moment when she departs from home. The service time undergone by a traveler represents the time she has to wait before being picked up by an AV from when she issues a request. In section 2, we consider a unique fleet covering the whole corridor length and width for travel time expressions readability. This assumption is removed in section 4.

2.1.8. Equilibrium

We study the system at equilibrium. Travelers have adapted their mode and route choices after MRT design refinement and AMoD deployment. We assume the system respects the UE defined by Wardrop's first principle (Wardrop, 1952). No one can reduce their travel times by unilaterally choosing another route of the same origin-destination pair. Note that this is a pure route choice UE problem. Departure times are parameters.

Agent-based simulation approaches usually account for this process through a computationally greedy iterative loop. Our model implements a fast converging loop to refine AMoD service time (see section 2.2.6). At one iteration, the equilibrium can be directly derived (see section 2.2.5).

2.2. A computationally efficient dynamic model for the MRT-AMoD design problem

2.2.1. Dynamic demand

In the dynamic model, Demand is time-variant. Three trip generation rates reproduce congestion onset and offset. They correspond to three phases of the morning commute, as shown in Figure 4: low loading, high loading, and unloading phases.

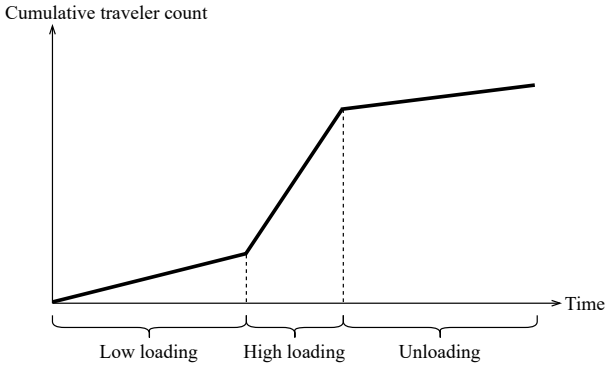


Figure 4: Three trip generation rates are considered over time on the corridor.

2.2.2. *Point-queue model for congestion*

A point-queue model provides traffic delays. Introduced by Vickrey (1969), this model first dealt with the departure time choice of commuters on a single bottleneck-constrained one-to-one corridor (Arnott et al., 1990). It has been extended to many-to-one (Fosgerau and de Palma, 2012), multiple bottlenecks (Akamatsu et al., 2015), and multimodal corridor (Sean Qian and Michael Zhang, 2011; Wu and Huang, 2014; Chiabaut et al., 2018). It is one of the simplest models accounting for congestion dynamics. It can deal with aggregated (flow) and disaggregated (traveler-specific) points of view, allowing analytical and numerical resolution of UE.

Vertical queues allow modification of corridor configuration while preventing perturbation of upstream off-ramps by spillback congestion. Cars and AVs flows do not interact in our model. This assumption is reasonable under two conditions. First, c_1 should be sufficiently far from c_0 compared to queue length, or AVs can run on a dedicated lane on the freeway. Second, cars and AVs use different roads or lanes on the local network nearby MRT stations.

Delays given by the point-queue model are equivalent to the LWR model with spreading congestion. It is a suitable model for design purposes. Considering that traffic obeys a first-in-first-out (FIFO) rule at each bottleneck μ_k, $0 \leq k \leq M$, waiting delay can be formulated by:

$$w_k(t) = \max\left(0, \frac{A_k(t) - A_k(t_k)}{\mu_k} - (t - t_k)\right) \tag{1}$$

where $A_k(t)$ denotes the cumulative number of travelers arrived at bottleneck μ_k by t, t_k represents the time at which congestion starts at bottleneck μ_k. For $1 \leq k \leq M$, $w_k(t)$ includes T_f, the time for an AV to park at a drop-off spot and for the passenger to leave the vehicle in security. The application of the point-queue model allows writing delays as simple functions of time.

2.2.3. *Travel times*

The free-flow travel time by **c** for a commuter i departing from (x_i, y_i) is:

$$T_i^\mathrm{c} = \frac{|x_i - x_k| + |y_i|}{v_\mathrm{st}} + \frac{x_k}{u} \tag{2}$$

where c_k located at x_k is the vehicle access collector of i.

The free-flow travel time by **a** when i transfers at c_p $(1 \leq p \leq k)$ is:

$$T_{i,p}^\mathrm{a} = \frac{|x_i - x_k| + |y_i|}{v_\mathrm{st}} + \frac{x_k - x_p}{u} + \frac{h}{2} + \frac{x_p}{v_\mathrm{r}} + (p-1)T_\mathrm{d} \tag{3}$$

Travel time by **r** is time-invariant:

$$T_i^\mathrm{r} = \frac{|x_i - x_{k'}| + |y_i|}{v_\mathrm{w}} + \frac{h}{2} + \frac{x_{k'}}{v_\mathrm{r}} + (k'-1)T_\mathrm{d} \tag{4}$$

where $c_{k'}$ is the walking access collector of i.

Outside of free-flow conditions, the travel time of i by **c** is:

$$\tau_i^{\mathrm{c}} = T_i^{\mathrm{c}} + w_0(t_i^0) \tag{5}$$

where t_i^0 is the theoretical arrival time of i at c_0 (when i travels by car in free-flow conditions).

Outside of free-flow conditions, the travel time by **a** when i transfers at c_p ($1 \leq p \leq k$) is:

$$\tau_{i,p}^{\mathrm{a}} = T_{\mathrm{s}}(t_i^{\mathrm{req}}) + T_{i,p}^{\mathrm{a}} + w_p(t_i^p) \tag{6}$$

where t_i^p is the arrival time of i at c_p, t_i^{req} is the time at which i sends a request to AMoD, $T_{\mathrm{s}}(t_i^{\mathrm{req}})$ is the AMoD service time experienced by traveler i (see section 2.2.6).

Additionally, free-flow extra travel times by **r** and **a** compared to **c** are $\Delta_i^{\mathrm{r}} = T_i^{\mathrm{r}} - T_i^{\mathrm{c}}$, and $\Delta_{i,p}^{\mathrm{a}} = T_{\mathrm{s}}(t_i^{\mathrm{req}}) + T_{i,p}^{\mathrm{a}} - T_i^{\mathrm{c}}$ respectively.

2.2.4. *Attraction areas boundaries*

Vehicle and walking attraction areas boundaries are:

$$X_{k,k+1}^{\mathrm{w}} = \frac{1}{2}\left[x_{k+1} + x_k + (x_{k+1} - x_k)\,\frac{v_{\mathrm{w}}}{v_{\mathrm{r}}} + T_{\mathrm{d}} v_{\mathrm{w}}\right] \tag{7}$$

$$X_{k,k+1}^{\mathrm{c}} = \frac{1}{2}\left[x_{k+1} + x_k + (x_{k+1} - x_k)\,\frac{v_{\mathrm{st}}}{u}\right] \tag{8}$$

Boundaries are well defined between two collectors when the spacing s_k between c_k and c_{k+1} verifies $s_k > T_{\mathrm{d}} v_{\mathrm{w}} v_{\mathrm{r}} / (v_{\mathrm{w}} - v_{\mathrm{r}})$.

2.2.5. *UE principles*

Let service time T_{s} be an exogenous constant function here. The next section tackles the extension to an endogenously computed time-variant service time.

FIFO property. As all travelers have the same AMoD service time and have a defined vehicle access collector, the order of arrivals at bottleneck μ_k is the same as at bottleneck μ_0. The route choice of a traveler depends only on the choices made by travelers that have arrived before her at μ_0 in free-flow conditions.

UE properties. The FIFO property ensures UE existence (Tampere et al., 2010). It allows computing the exact equilibrium with no need for a convergence loop by processing the itinerary choice of travelers in the order of free-flow arrival at the CBD off-ramp. Hence, the resolution process always finds a solution: it is robust.

Our network has, at most, one bottleneck per route. Consequently, route cost functions are monotonous with respect to their traffic flows, and the relaxed uniqueness of UE is ensured (Iryo, 2013). It has two additional characteristics that ensure the strict uniqueness of UE:

- the routes travel costs in free-flow conditions are strictly ordered
- when a route travel cost does not vary on traffic flows variation, it is necessarily compared with a route containing a bottleneck working over capacity (i.e., that has a strictly monotonous travel cost at that time). It occurs for **r** route, which has a constant travel cost, and for **c** and **a** routes when the bottleneck of the route works under capacity.

Resolution process. The resolution process is inspired by Laval (2009). The original method performs in a one-to-one corridor with cars only. We developed an extended method to deal with a many-to-one corridor with three modes.

Complete knowledge about how UE settles and evolves is available. UE principles (diversion patterns, bottlenecks synchronization, and out of sync) help describe the system's dynamics. They are analytically defined and illustrated in the two simplified examples of appendix B, where travelers' origins stick to the locations of collectors. Figure 5 shows the equilibrium for the same instance but with uniform distribution of travelers' origins along axes.

Stable states and transition conditions. Bottleneck μ_0 is switched on (used) from the beginning of peak hours since **c** is the best itinerary for all travelers in free-flow conditions. The theoretical arrival curve at c_0, noted $A^{\text{th}}(t)$, can be computed by assuming that all travelers choose mode **c**. It is the sum of the theoretical arrival curve at c_k transposed by the respective x_k/u. The theoretical arrival curve at c_k gathers all commuters within Ω_k^{c}.

A queue forms at c_0's bottleneck when demand exceeds μ_0. It is the transition condition to pass from system stable state 1 (**a** diversion pattern 1 or **a**-div1) to stable state 2 (**a**-div2). In **a**-div1, the arrivals rate is lower than the sum of switched-on bottlenecks capacities. In **a**-div2, used bottlenecks are saturated.

As $w_0(t)$ increases, **c** travel times will progressively equalize the travel times on other itineraries. Traveler's extra travel times Δ_i^{r} and $\Delta_{i,p}^{\text{a}}$ $(p \leq k)$, can be sorted in ascending order. As soon as $w_0(t)$ equals one of the extra travel times, the commuter has a new worthwhile diversion itinerary.

In practice, $\Delta_{i,p}^{\text{a}} = \Delta_p^{\text{a}}$ does not depend on i since free-flow in-vehicle access times by **c** and **a** are equal and T_{s} is a constant function here. When $w_0(t)$ equals one of the Δ_p^{a}, the theoretical arrival flow at c_0 splits: one part keeps using **c** while the other part diverts to another itinerary. It is the transition condition from **a**-div2 to **a**-div1. Opening a new diversion itinerary relieves the previously switched-on bottlenecks: they now work at capacity.

When $w_0(t)$ reaches Δ_i^{r}, mode **r** starts to be worthwhile for traveler i. She chooses r as long as travel times through downstream bottlenecks are strictly greater than Δ_i^{r}. This is **r**-div1. Diversion pattern **r**-div1 starts at c_k when $x_k/u + w_0(t) = h/2 + x_k/v_r + (k-1)T_{\text{d}}$. It is different from **a**-div1 because it is continuous. Progressively, more and more travelers within Ω_k^{w} divert to **r** while $w_0(t)$ increases. There is no **r**-div2 since mode **r** is uncongested.

The system simultaneously undergoes a continuous **r**-div1 and an alternation between **a**-div1 and **a**-div2.

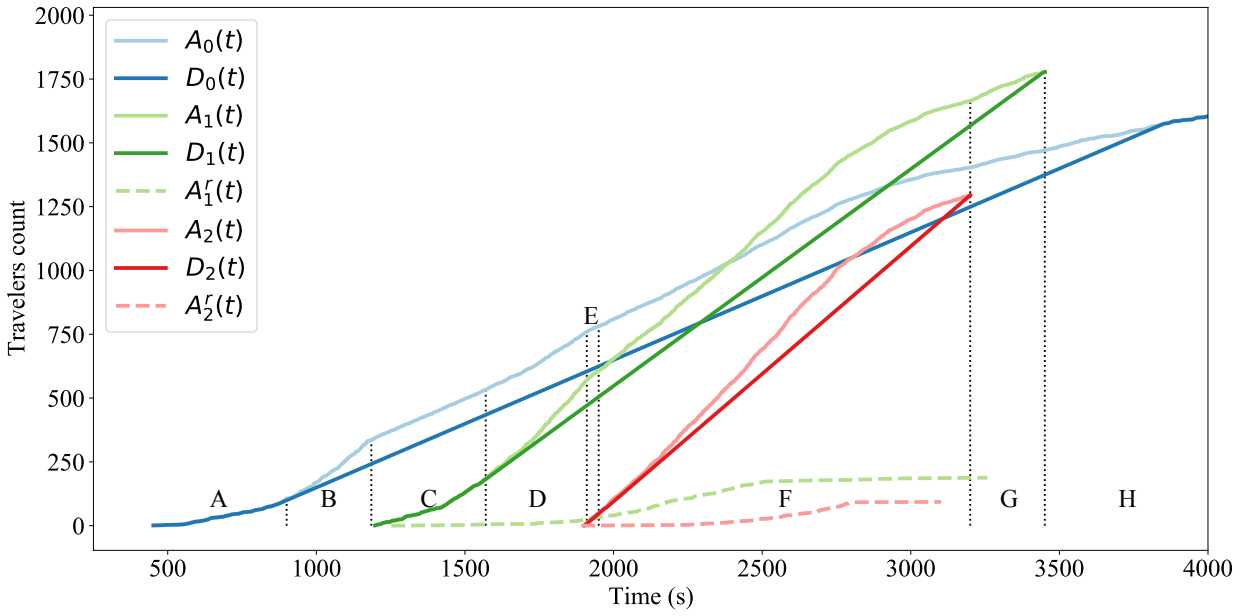


Figure 5: Equilibrium arrival ($A_k(t)$) and departure ($D_k(t)$) curves at bottlenecks in a corridor where $M = 2$. $A_k^{\text{r}}(t)$ designates the cumulative number of travelers within Ω_k^{w} diverting to **r**. Note that all curves have been aligned in time on c_0 to highlight the diversion patterns better. Vertical lines show the state shifts. Congestion on μ_0 starts at the beginning of period B. In period C, **a**-div1 occurs using c_1 as a transfer collector. Pattern **a**-div2 occurs in period D as μ_1 becomes congested. The states shifts continue in periods E and F till demand decrease leads to the desertion of c_2 in period G, of c_1 in period H. Diversion to **r** progressively concerns more travelers as waiting times at bottlenecks rise.

Synchronization and out of sync conditions. Bottleneck μ_0 is used for the whole period. It acts as a regulator and other bottlenecks will synchronize with it. When travel times by several bottlenecks are equal, these bottlenecks are said synchronized. With a one-to-one demand pattern, all bottlenecks are always synchronized. With a many-to-one demand pattern, bottlenecks can get out of sync.

Indeed, travelers have different diversion options. Those departing from downstream collectors have fewer possibilities than those from upstream collectors. When the commuters from downstream collectors have no more uncon-

gested bottlenecks to divert to, they may increase the travel times by the bottlenecks they are currently using and push them out of sync with upstream bottlenecks. Travel time equality does not hold anymore between the downstream and upstream groups. As a result, upstream travelers divert by routes belonging to the upstream group.

2.2.6. *AMoD service time*

The constant service time assumption is acceptable when fleet sizes are large enough. In general, AMoD service time depends on dispatching rules, the number of vehicles m in the fleet, and the amount of work to achieve, i.e., the distance to serve customers and relocate idle vehicles.

AMoD dispatching strategy. A traditionally used dispatching heuristic in MoD systems is the *nearest-idle-vehcile* (Maciejewski et al., 2016). It respects the FIFO rule by treating requests in the order of arrival. The first traveler that has ordered an AV is the first to be assigned a vehicle by the dispatcher. We consider such a dispatching strategy here. As no booking is allowed, $t_i^{\text{req}} = t_i^{\text{dep}}$ for each traveler i choosing $\mathbf{a}$, where t_i^{dep} is the time at which traveler i departs from home. The first traveler to depart is also the first to send a request. Service time is a function of request/departure time.

MSA process. To approximate the service time profile, we solve the fixed point problem $\widehat{T}_{\text{s}}(\text{UE}(T_{\text{s}})) = T_{\text{s}}$ by the method of successive averages (MSA) detailed in algorithm 1.

In step 1 of the algorithm, one should choose an initial service time profile. It can be defined arbitrarily or computed in a prior iteration. During this prior iteration, the service time is computed online by considering that AMoD receives requests in the order of theoretical arrival at c_0. We make the following strong assumption: $t_i^{\text{req}} = f(t_i^0)$ where f is a linear function. Each time a traveler chooses a route, the service time is re-computed following equation 9. Service time is updated only if the difference between the new and last values is above a certain threshold. It forms a step function. Once all travelers have chosen a route, they are sorted in the order of departure time, which is the effective order in which AMoD receives requests. Route choices being unchanged, the effective service time profile $\widehat{T}_{\text{s}}$ is computed considering this order and following equation 9. It serves as the reference initial service time profile (or predicted service time profile) for the subsequent iterations of MSA ($T_{\text{s}} \leftarrow \widehat{T}_{\text{s}}$ at the end of the prior iteration).

Step 3 of the algorithm launches the MSA loop which contains at least one iteration. The quality criteria used as stopping conditions for the loop are detailed in the ***UE properties*** paragraph below.

In step 5, the predicted service time profile is a parameter for UE resolution. Each traveler chooses an itinerary knowing the service time she will experience on $\mathbf{a}$ mode.

In step 6, effective service time profile $\widehat{T}_{\text{s}}$ is computed similarly as in the prior iteration, considering the distribution of travelers on itineraries resulting from step 5.

In step 7, the predicted service time for iteration $K+1$ of MSA is computed based on predicted and effective service time profiles of iteration K.

Service time computation. Equation 9 is used to compute the reference initial service time profile in the prior iteration and the effective service time profile in step 6.

$$T_{\text{s}}(t_j) = 2E_2 + E_3 - \left[t_i^{\text{req}} + \frac{\Delta t}{m} - E_1 \right] \tag{9}$$

It expresses the service time of traveler j as a function of E_1, E_2, E_3 and traveler i's parameters, i and j being processed consecutively. It is an approximation based on the last m requests $\{l\}$ received by the fleet during Δt. If i is the last traveler processed, $\Delta t = t_i^{\text{req}} - \min_l(t_l^{\text{req}})$. As AMoD treats requests in order of reception, the next traveler choosing this fleet will necessarily ride one of the AVs serving $\{l\}$, a mean predecessor request, representative of $\{l\}$, is built. With $E[.]$ being the expected value, the mean request is picked up at $E_1 = E[t_l^{\text{req}} + T_{\text{s}}(t_l^{\text{req}})]$, rides for $E_2 = E\left[\frac{|x_l - x_{k(l)}| + |y_l|}{v_{\text{st}}} + \frac{x_{k(l)} - x_{p(l)}}{u}\right]$ and waits for drop-off for $E_3 = E[w_{p(l)}]$ where $k(l)$ and $p(l)$ respectively label vehicle access collector and transfer collector for traveler l. AV relocates for E_2 to finally serve the next request, which is expected to be received by AMoD at $t_i^{\text{req}} + \frac{\Delta t}{m}$. AVs are initially located at the m first requests locations so that the service time of a fleet remains null until the mth request is received.

Algorithm 1: MSA process

1 Initialize T_s with an arbitrary step function or with a prior iteration (reference initial service time profile);
2 $K \leftarrow 1$;
3 **while** $K = 1$ *or UE(T_s) does not meet the quality criteria* **do**
4 $\quad K \leftarrow K + 1$;
5 $\quad$ Compute UE(T_s) by processing travelers one by one, taken in order of increasing t_i^0;
6 $\quad$ Compute $\widehat{T}_s(\text{UE}(T_s))$ based on equation 9 by processing travelers by increasing $t_i^{\text{req}} = t_i^{\text{dep}}$;
7 $\quad T_s \leftarrow \frac{1}{K}\widehat{T}_s + (1 - \frac{1}{K})T_s$;
8 Return UE(T_s);

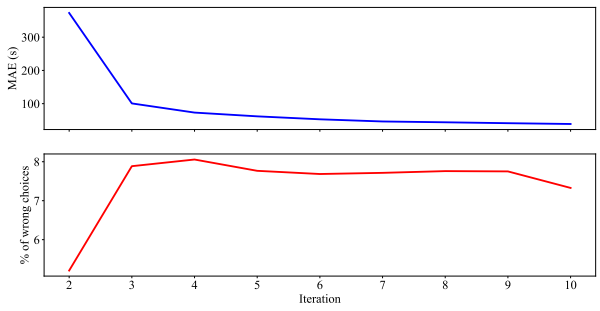
UE properties. With a dynamic endogenous service time, the formal demonstration of UE existence and uniqueness is challenging.

As service time is a continuous function of departure time, it is a discontinuous function of theoretical arrival time at c_0. Service time profile impacts the waiting times at drop-off bottlenecks w_k ($1 \leq k \leq M$). When service time increases between two travelers i and j taken in the order of theoretical arrival at c_0 ($t_i^0 < t_j^0$), this order is maintained at drop-off bottlenecks and the predicted **a** travel time equals the realized one. When service time decreases, travelers may be re-ordered: a customer requesting AV later may arrive earlier at the transfer collector. A few commuters choose a non-optimal route because their waiting times at the AV-to-MRT transfer are greater than expected due to order rearranging.

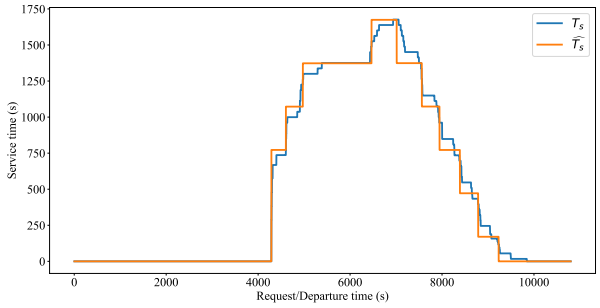
Discontinuities in the service time profile are local and bounded by the corridor parameters: the service time globally follows the loading of the network. It gives us confidence that jumps in service time have limited impact compared to the periods when T_s is constant and that existence and uniqueness of UE still hold.

Numerically, three criteria assess the convergence of the process and the quality of the solution found. The quality criteria (step 3) are the mean absolute error (MAE) between $\widehat{T}_s$ and T_s, quartiles of the difference $\widehat{T}_s - T_s$, and the percentage of travelers that have made a wrong route choice due to local order rearranging. The process found a solution for each experiment achieved with the following criteria: MAE below 40s, 1st and 3rd quartiles respectively greater than -5min and lower than 5min, less than 10% wrong choices. Moreover, replacing the reference initial conditions (computed during the prior iteration of MSA as described above) with different initial service time profiles led to the same equilibrium.

Figures 6 and 7 present one of the tests for a generic scenario. The scenario deals with a theoretical monocentric city where a high-frequency subway serves the whole corridor. The demand level is high, with 13.4k travelers in 3 hours, leading to intense congestion from the second hour. Discontinuity threshold on $\widehat{T}_s$ is 5min. Table 1 provides the scenario parameters.

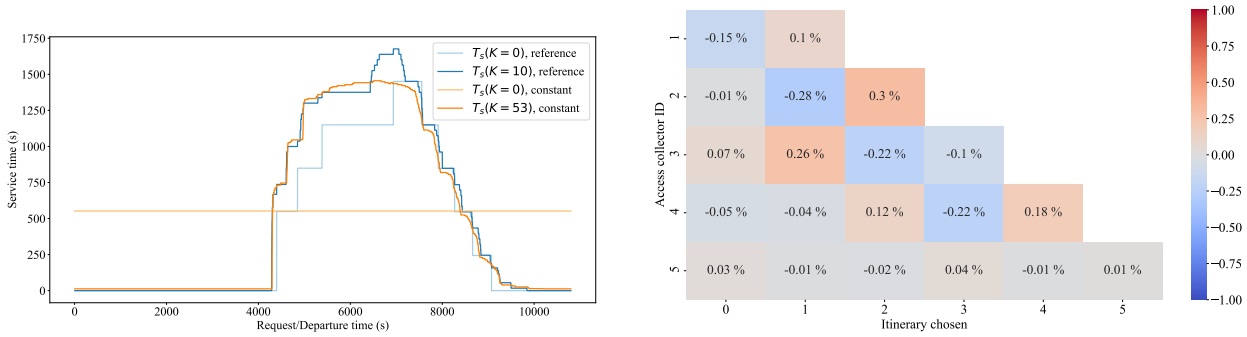


(a) Evolution of MAE($\widehat{T}_s - T_s$) and percentage of travelers that have made a wrong route choice.



(b) Difference between $\widehat{T}_s$ and T_s is acceptable at iteration 10 of the MSA process.

Figure 6: Convergence of the endogenous time-variant service time with the initial reference conditions.



(a) Initial and final service time profiles.

(b) Relative differences in travelers distribution are marginal.

Figure 7: Comparison of equilibrium obtained with reference and constant initial service time profiles.

Figure 7b compares the distribution of travelers obtained after convergence for the reference initial service time profile and a constant profile. Figure 7a shows the initial and final service time profiles. Differences are marginal. The MAE between final profiles is 57s. The maximum variation of travelers volume on a route represents only 0.28% of the total volume. Other initial service time profiles tested (translated reference, null, one-step function, three-steps function) led to similar results.

The number of iterations required to meet quality criteria differs depending on initial conditions. The prior iteration used to compute the reference initial conditions provides a good starting point. The process reaches the quality criteria in only ten iterations, as shown in Figure 6a. This fast convergence enhances compatibility with the design optimization framework.

Scenario name	**Monocentric city**	**Western Lyon**
Description	Wide monocentric compact city with its suburbs	Small suburban towns external to main urban area
Distribution of travelers	Uniform	BD TOPO adresses
MRT type	High freq. subway	Medium freq. regional trains
Corridor dimensions (km)	20*3	20*5
Morning peak duration (h)	3	5 2015 Lyon Area OD matrix
Number of travelers	15.5k	14k
v_w **(m/s)**	1.2	1.2
v_{st} **(m/s)**	9	14
u **(m/s)**	18	18
v_r **(m/s)**	14	25
h **(s)**	240	900
T_d **(s)**	45	45
$\mu_k, 0 < k \leq M$ **(veh/s)**	0.2	0.2
μ_0 **(veh/s)**	0.6	0.4
Nominal AVs ratio	10%	5%
Nominal number of collectors	5	7
Nominal spacing (km)	4	2,2,2,3,2,5,3
T_s **step size (s)**	300	150
α_c, β_c	222, 4	-
α_a, β_a	1, 2	-
Exogeneous constant T_s **(s)**	920	-
Pool dimensions (m^2)	300 * 300	-

Table 1: Scenarios parameters.

In this section, we have introduced our simple dynamic model for the morning commute in a many-to-one corridor with three combinations of modes. We have shown that exact UE is easily computable under an exogenous constant AMoD service time. The analysis of UE on an example has provided a deep understanding of how UE settles and

evolves. Under an endogenous time-variant service time, we have found that a few iterations of MSA lead to good quality UE, making our approach computationally efficient for design optimization.

3. Influence of design on MRT-AMoD cooperation or competition relations: the interest of dynamics and UE principles

In this section, we demonstrate that dynamics in our model are crucial to capture the influence of design on MRT-AMoD cooperation or competition. This influence is analyzed regarding the UE principles introduced in section 2.2.5.

3.1. Why are dynamics crucial?

3.1.1. A static model for benchmark

Transportation studies usually use static approaches with time-invariant supplies and demand flows to deal with design. We compare our model with a static model to show how crucial considering dynamics is.

Commuters aggregation. The corridor splits into N equally proportioned zones aggregating the origins of travelers in an abstract origin O_i. The access travel time from O_i to c_k by vehicle (resp. walk) is defined as the average vehicle (resp. walking) access time of aggregated commuters. O_i's vehicle (resp. walking) access collector corresponds to the one with the minimal vehicle (resp. walking) access time.

Multimodal network directed graph. The corridor is equivalent to a directed graph $G = (V, E)$ where V is the set of vertices and E the set of edges. Vertices include origins $\{O_i \mid 1 \leq i \leq N\}$, destination $\{D\}$ and three different nodes for each collector $\{C_k^{\mathrm{c}}, C_k^{\mathrm{r}}, C_k^{\mathrm{a}} \mid 1 \leq k \leq M\}$. Commuters in Ω_k^{c} (resp. Ω_k^{w}) choosing **c** (resp. **r**) use C_k^{c} (resp. C_k^{r}). Commuters choosing **a** and transferring at c_k use C_k^{a}. Edges of the graph link:

- origins to their vehicle and walk access collectors,
- origins to transfer nodes associated with collectors located downstream of their vehicle access collector,
- **a** nodes of collectors to their **r** nodes,
- **c** nodes of collectors to C_1^{c},
- **r** nodes of collectors to destination,
- C_1^{c} to destination.

Figure 8 represents an example of G.

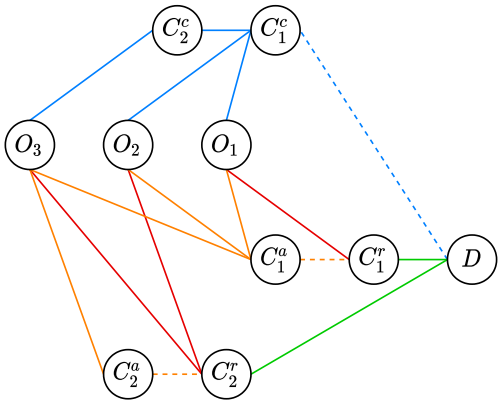


Figure 8: Corridor graph model G example where $N = 3$, $M = 2$, c_1 is the vehicle access collector for O_1, O_2, walk access collector for O_1 while c_2 is the vehicle access collector for O_3, walk access collector for O_2, O_3. Solid edges carry constant travel times, while volume-dependent travel times are associated with dashed edges. Colors represent the mode used to travel the edge (car in blue, AV in orange, MRT in green, and walk in red).

Travel times. Travel times are constant except for edges passing through a delay generator point, namely (C_1^{r}, D) and $(C_k^{\mathrm{a}}, C_k^{\mathrm{r}})$. The BPR volume delay function (Maerivoet and De Moor, 2005) gives the travel times on (C_1^{r}, D) and $(C_k^{\mathrm{a}}, C_k^{\mathrm{r}})$ links. Two parameters sets $(\alpha_{\mathrm{c}}, \beta_{\mathrm{c}})$ and $(\alpha_{\mathrm{a}}, \beta_{\mathrm{a}})$ determine the shape of functions for both delay generator point types. Volumes and travel times are respectively noted $v(\text{edge})$ and $\tau(\text{edge})$. AMoD service time T_{s} is assumed constant.

$$\tau(O_i, C_k^{\mathrm{c}}) = \sum_{j \in O_i} \frac{|x_j - x_k| + |y_j|}{\mathrm{card}(O_i) v_{\mathrm{st}}} \tag{10}$$

$$\tau(O_i, C_k^{\mathrm{r}}) = \sum_{j \in O_i} \frac{|x_j - x_k| + |y_j|}{\mathrm{card}(O_i) v_{\mathrm{w}}} \tag{11}$$

$$\tau(O_i, C_p^{\mathrm{a}}) = T_{\mathrm{s}} + \sum_{j \in O_i} \frac{|x_j - x_k| + |y_j|}{\mathrm{card}(O_i) v_{\mathrm{st}}} + \frac{x_k - x_p}{u} \tag{12}$$

$$\tau(C_k^{\mathrm{a}}, C_k^{\mathrm{r}}) = T_{\mathrm{f}} \left(1 + \alpha_{\mathrm{a}} \left(\frac{v(C_k^{\mathrm{a}}, C_k^{\mathrm{r}})}{\mu_k} \right)^{\beta_{\mathrm{a}}} \right) \tag{13}$$

$$\tau(C_k^{\mathrm{r}}, D) = \frac{h}{2} + \frac{x_k}{v_{\mathrm{r}}} + (k-1) T_{\mathrm{d}} \tag{14}$$

$$\tau(C_k^{\mathrm{c}}, C_1^{\mathrm{c}}) = \frac{x_k - x_1}{u} \tag{15}$$

$$\tau(C_1^{\mathrm{c}}, D) = \frac{x_1}{u} + \alpha_{\mathrm{c}} \left(\frac{v(C_1^{\mathrm{c}}, D)}{\mu_0} \right)^{\beta_{\mathrm{c}}} \tag{16}$$

Solving process. The equilibrium of this system is computed through MSA using the conventional descent gradient based on an all-or-nothing assignment and step size $\frac{1}{K}$ where K is the iteration number (Sheffi and Powell, 1982).

Static model calibration. We calibrate static model parameters on the dynamic model outputs to maintain a certain consistency and facilitate results comparison. N is chosen to keep pools of travelers small enough to account for access times diversity (necessary smaller than attraction areas) and wide enough to gather a sufficient number of commuters for the flow balance procedure. T_{s} value equals the mean service time computed by the dynamic model. The congestion term in $\tau(C_1^{\mathrm{c}}, D)$ (16) is intentionally independent of the free-flow travel time on edge (C_1^{c}, D), which is controlled by the design parameter x_1. It prevents unwanted modification of the congestion function when changing the corridor design. The free-flow travel time on (C_1^{c}, D) in the nominal monocentric city is used for α_{c}. The retained values for the other parameters (α_{a}, β_{a}, β_{c}) provide consistent total travel time and waiting times at bottlenecks regarding the dynamic model. Values for all parameters are provided in table 1. The MSA algorithm stops when the relative gap is below 10^{-4}. The tolerance for travel times equality is 30s.

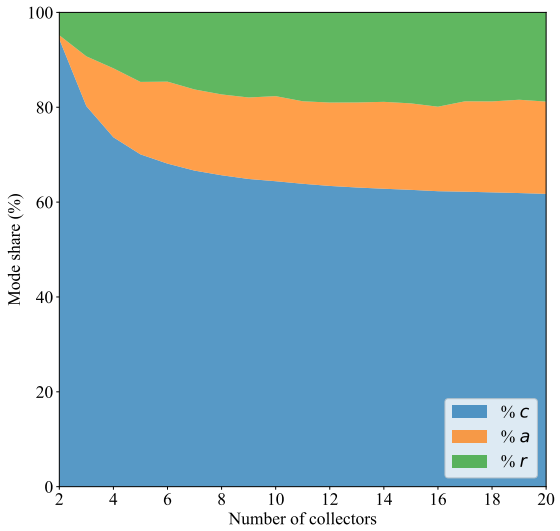
3.1.2. Sensitivity of design on MRT-AMoD cooperation or competition with static and dynamic models

Compared indicators. We evaluate competition and cooperation between modes based on system-level and link-level indicators. The system-level indicators observed are volume-based and distance-based modal shares. A distance-based mode share corresponds to the total distance traveled by this mode over the total distance traveled by all modes. The link level indicators are the flows on each mode-route alternative.

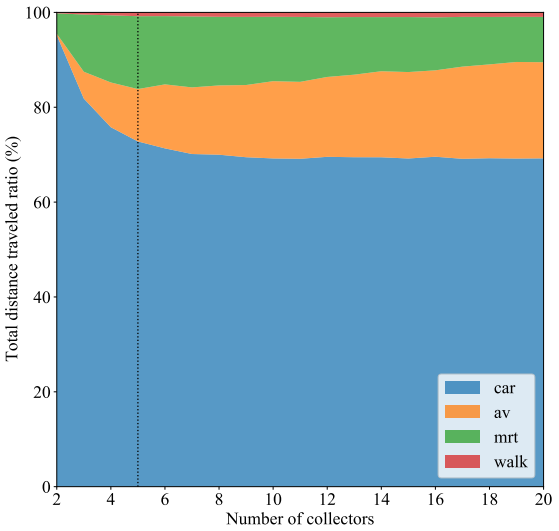
Design parameter. A crucial design parameter in our study case is the number of collectors and their locations along the corridor. An *a priori* fair design strategy regarding a uniform distribution of travelers, as in the monocentric city scenario, is to fix the spacing between two collectors. With constant spacing, we analyze the sensitivity of the number of collectors M on the chosen indicators. We consider that the TA does not regulate AMoD here. AMoD comprises a single fleet operating on the whole corridor.

Results analysis. Figure 9 shows the results for dynamic and static models.

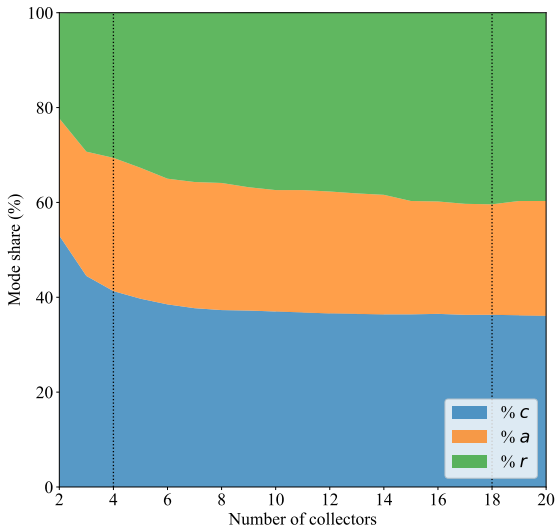
They both present a decreasing **c** volume-based mode share profile. The most significant modal shifts from car to other modes are achieved for a small number of collectors. Over eight collectors, accessibility improvement after adding a new collector has a limited impact on **c** volume-based mode share. The travelers who choose **c** whatever the number of collectors are "captives" to car mode. Their number directly relates to the CBD off-ramp capacity (μ_0) as **c** remains the most efficient mode of the system in free-flow conditions (we have $\Delta_i^{\mathrm{r}} > 0$ and $\Delta_{i,p}^{\mathrm{a}} > 0$). As expected, the static model underestimates congestion and overestimates **c** volume-based mode share compared to the dynamic model.



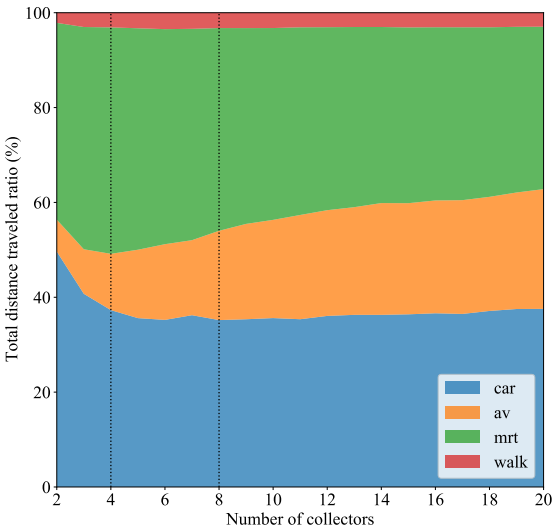
(a) Volume-based mode shares computed by the static model show only one cooperation behavior between **r** and **a** against **c**.



(b) Distance-based mode shares computed by the static model show cooperation then competition between **r** and **a**.



(c) Volume-based mode shares computed by the dynamic model show cooperation then competition between **r** and **a**.



(d) Distance-nased mode shares computed by the dynamic model show cooperation followed by competition between **r** and **a**.

Figure 9: Impact of the number of collectors with a constant spacing on volume-based and distance-based mode shares for static and dynamic models. The dashed vertical lines capture the boundaries between two competition-cooperation schemes.

Models show different MRT-AMoD interactions. In terms of volume-based mode shares, the dynamic model highlights three competition-cooperation schemes over M (Figure 9c).

- $M \leq 4$: a new collector leads to modal shift from **c** to **r** and **a**. MRT and AVs cooperate against cars.
- $4 < M \leq 18$: **r** benefits from a higher number of collectors and steals market share from **c** and **a**. MRT and AMoD compete to the advantage of MRT.
- $18 < M$: MRT and AVs still compete but **r** now looses market share to **a**.

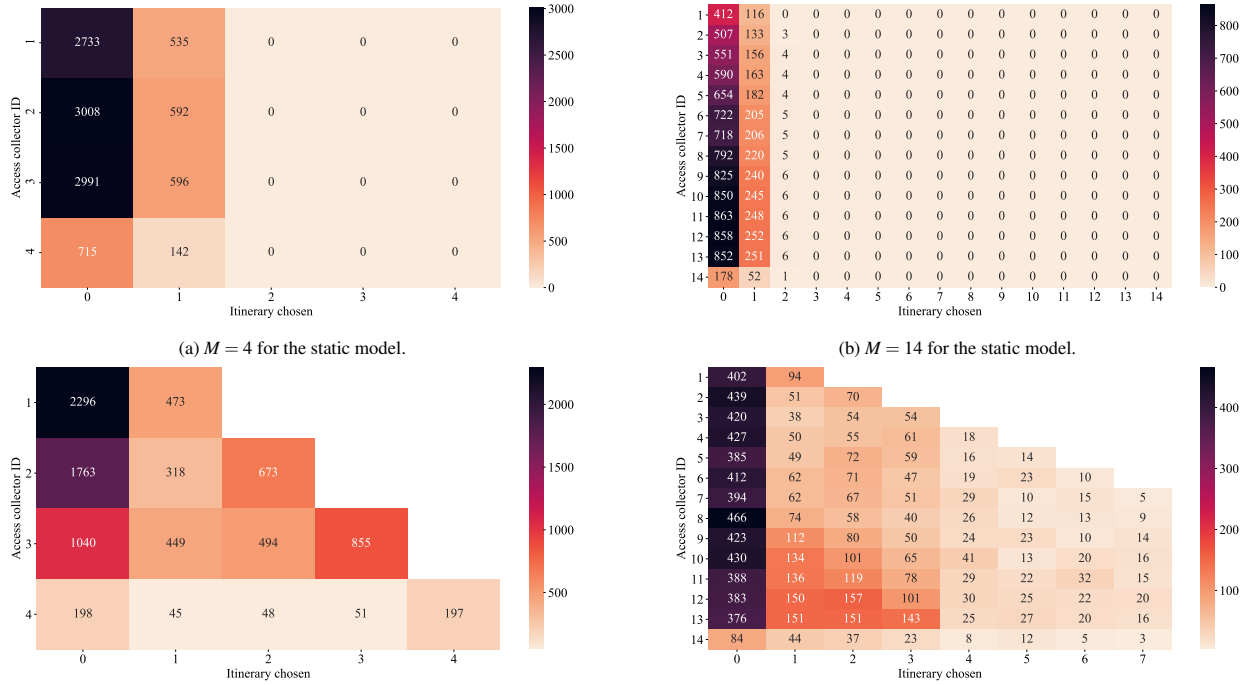
Opportunities to exploit **r**-**a** cooperation come for small numbers of collectors since competition starts from $M = 4$. The static model only shows the first of these schemes: MRT and AVs cooperate. This behavior is stable as M increases. The static model does not identify competition between **a** and **r** in terms of volume-based mode share.

Distance-based mode share measures the usage of a mode regardless of the trip in which it intervenes. In terms of distance-based mode share, the dynamic model reveals three schemes:

- $M \leq 4$: a new collector leads to lower car usage and higher MRT, AV, and walking usage. A positive synergy between MRT, AV, and walk occurs in this scheme.
- $4 < M \leq 8$: **a** benefits from a higher number of collectors. AVs steal kilometers from **r**, and to a lesser extent from **c**. AVs compete mainly with the MRT and secondarily with cars.
- $8 < M$: MRT usage declines against both AVs and cars. The usage of AVs and cars grows significantly.

The static model only captures the first two schemes, as shown in Figure 9b.

Figure 10 confirms that the static model misses the influence of design on travelers' distribution over itineraries. Travelers only divert to c_1, the closest collector to the CBD off-ramp, which is available for everyone. Collector c_2 is only used marginally for high values of M). For low M values, close and distant travelers similarly load c_1. As M increases, the provenance of AV riders and car drivers progressively imbalances. Distant travelers load more c_0 and c_1 than close travelers. (Figure 10b). Car mode share is greater in upstream vehicle attraction areas, while MRT mode share is greater in downstream walking attraction areas. Indeed, the MRT suffers from an increasing cumulative time lost due to train dwelling when M grows.



(a) $M = 4$ for the static model.

(b) $M = 14$ for the static model.

(c) $M = 4$ for the dynamic model.

(d) $M = 14$ for the dynamic model.

Figure 10: Impact of the number of collectors with a constant spacing on the distribution of commuters over itineraries for static and dynamic models. Only travelers that have chosen **c** or **a** appear here. The dynamic model shows an evolution in AMoD usage from a direct first-mile to a long first-mile pattern. The static model misses this behavior change.

The dynamic model leads to a more complex diversion pattern. Transfers take place at more than one collector. All collectors from c_1 to c_7 are loaded when $M = 14$. The distribution of commuters over transfer collectors is heterogeneous. As M increases, the distribution pattern evolves from the direct (Figure 10c) to the long first-mile (Figure 10d) type. In the direct first-mile pattern, most commuters in Ω_k^c request an AV to join c_k, their vehicle access collector. In the long first-mile pattern, more distant travelers use AVs to join the most downstream collectors. Cooperation-competition schemes and distribution patterns are consistent. The direct first-mile pattern is a cooperation scheme, while the long first-mile pattern is a competition scheme.

Considering dynamics allows for capturing the influence of design on cooperation-competition schemes and distribution patterns. Moreover, the dynamic model can better describe the highlighted schemes and patterns based on the UE principles introduced in section 2.2.5.

3.2. Analysis of the dynamic model schemes and patterns regarding UE principles

To discuss the influence of design on cooperation-competition schemes in the same conditions as in section 2.2.5, we relaunch the sensitivity analysis of M with a constant T_s function. T_s equals the mean service time obtained through the MSA process. Schemes and patterns of Figure 11 are similar to the ones of Figures 9c, 9d, 10c and 10d.

Boundaries between schemes are different. Notably, the volume-based first competition scheme shortens. It extends from $M = 4$ to $M = 18$ with an endogenous time-variant service time and from $M = 6$ to $M = 12$ with a constant service time. A more significant distance-based modal shift occurs to the advantage of AVs. The lack of a feedback loop on service time favors AMoD.

Regarding the distribution of travelers over itineraries, the direct and long first-mile patterns are respectively more and less intense than in the time-variant service time model. For low M values, a larger part of AV riders transfers at their vehicle access collector. For high M values, the distribution of distant AV riders on downstream collectors is more spread.

At $M = 2$, Ω_1^c is wide enough to saturate both μ_0 and μ_1: c_0 and c_1 are synchronized while c_2 is out of sync. Travelers in Ω_2^c divert through their vehicle access collector only: the direct first-mile pattern is almost perfect, with just 22 of them transferring at c_1.

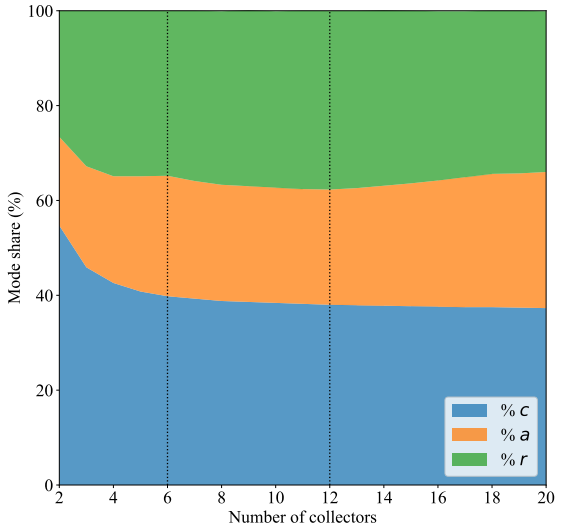
From $M = 3$ to $M = 5$, the first $M - 1$ bottlenecks are synchronized. A smaller vehicle attraction area surrounds c_M. This bottleneck is used but not overloaded: $w_M(t)$ is null. It is out of sync with downstream bottlenecks. The more distant commuters favor the direct first-mile pattern.

As M increases, attraction areas narrow, first diversions start earlier, and extra travel times values are closer to each other, i.e., successive bottlenecks switch on more quickly. Travelers better spread on the $M - 1$ first bottlenecks. The direct first-mile pattern here is less intense than in out of sync conditions.

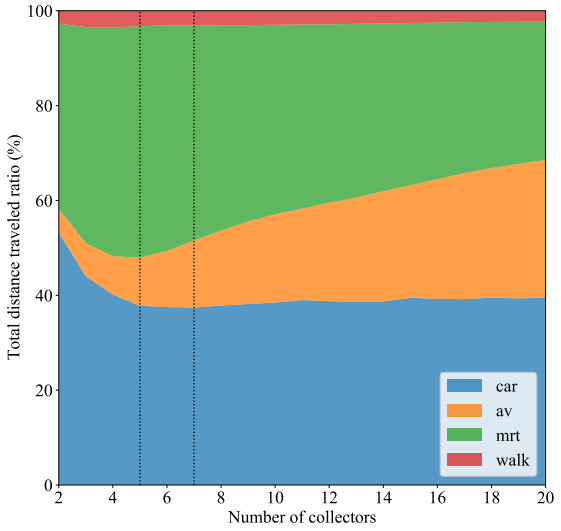
At $M = 5$, out of sync of c_5 still guarantees that the more distant travelers follow the direct first-mile pattern. However, the acceleration of successive bottleneck switching-on and the reduced size of vehicle attraction areas generates the long first-mile pattern.

From $M = 5$, an additional collector will not necessarily attract diversions: the system has reached the maximum potential for diversion. The long first-mile pattern intensifies since the first five collectors are getting closer to the destination. These collectors attract fewer surrounding travelers and more upstream travelers who depart near bottlenecks with excessive extra travel times. M increment is equivalent to moving attractive diversion alternatives away from travelers on average. Car distance-based mode share increases during this scheme because $\mathbf{r}$ suffers from a growing cumulative dwelling time. Travelers departing from upstream attraction areas are more affected by the cumulative dwelling time increase than travelers departing from downstream attraction areas. Consequently, more drivers come from more distant vehicle attraction areas.

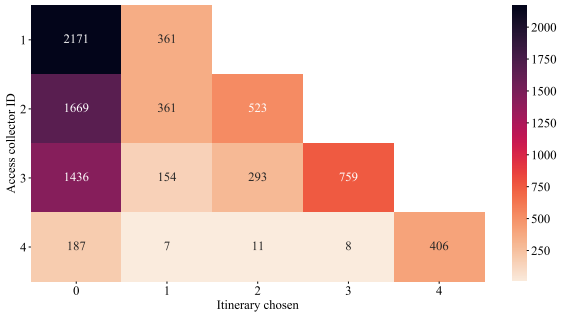
A good design should consider all the explained dynamics and behaviors regarding Ω_k^c size, the number of travelers diverting to $\mathbf{r}$, the maximum potential for diversion, and how extra travel times chain. The out of sync phenomenon forces the strict application of the direct first-mile pattern but might indicate that the overall diversion capacity is insufficient. A moderate long first-mile pattern is needed in practice to reach better car mitigation and MRT usage indicators.



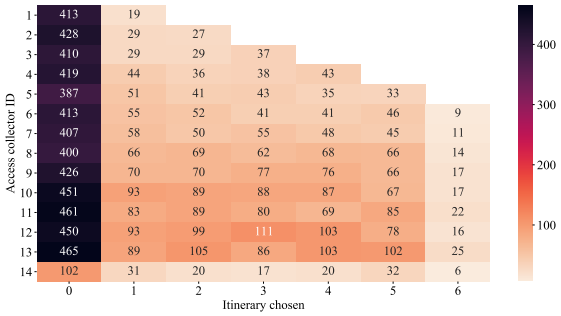
(a) Volume-based mode shares with constant T_s show a smaller M range where **r** and **a** compete to the advantage of **r**.



(b) Distance-based mode shares with constant T_s show a greater modal shift toward AV.



(c) Itineraries for $M = 4$ show a more intense direct first-mile.



(d) Itineraries for $M = 14$ show a less intense long first-mile.

Figure 11: Impact of the number of collectors with a constant spacing on volume-based, distance-based mode shares, and travelers distribution over itineraries for a "semi" dynamic model with a fixed service time.

4. Adressing the MRT-AMoD design problem

In this section, design can change. We apply an optimization metaheuristic in a realistic study case to find good designs under six different policies. We check how each design exploits the UE principles and make recommendations to the TA.

4.1. Scenarios

4.1.1. Western Lyon corridor

We apply our framework to a scenario derived from the West part of the city of Lyon, France (Figure 12).

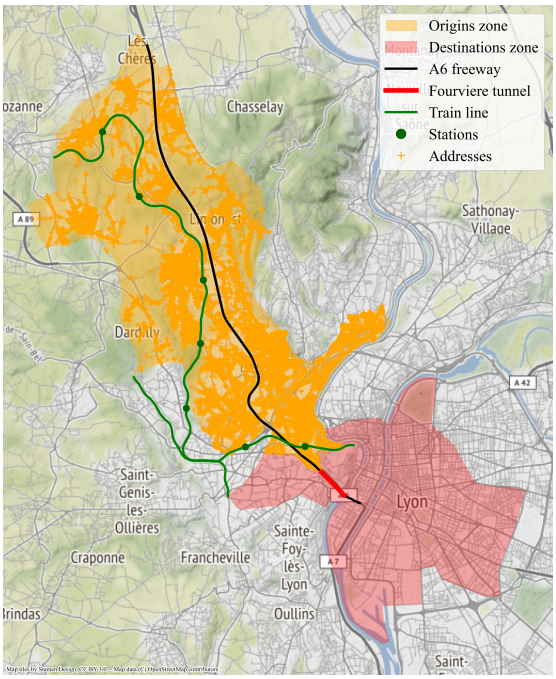


Figure 12: Lyon corridor is composed of 9 towns surrounding the A6 freeway and the regional train line from Marcilly-d'Azergues to Lyon-Gorge-de-Loup.

Freeway. The A6 freeway links several suburban towns with Lyon city center. The Fourvière tunnel, the final part of A6 road, is the entry point for the center. It acts as a bottleneck, causing significant congestion to arise daily.

MRT. A low-frequency regional train line operates along the corridor from Marcilly-d'Azergues to Lyon-Gorge-de-Loup. At Lyon-Gorge-de-Loup station, travelers can access the Lyon PT meshed network (subway, buses). In practice, the train line is underused because of its low frequency (waiting time can be up to 30 minutes in the morning) and lack of convenient access mode. Feeder buses operate in the downstream part of the corridor, but their routes are parallel to the train line. Their coverage is insufficient within this 5km-wide corridor. Moreover, its upstream part is out of the agglomeration bus network. The regional buses serving there are scarce.

Relevancy of the territory. The context is favorable to AMoD deployment. Today, car mode share from the corridor to the center is 64% against 31% for PT (Urba Lyon, 2018). At the local level, elected representatives request better first-mile solutions (Cerema, 2021). At the metropolitan area level, the political will is to improve the attractiveness of the train and change its role from a regional to an urban-suburban line. Public authorities have identified this corridor as an opportunity to develop new transportation systems. Several facilities favoring ridesharing have been deployed recently, including carpool areas and an HOV lane. The active entrepreneurial and research fabric on autonomous vehicles has already launched pilot projects in Lyon Confluence (Navly) and Meyzieu (Mia). AMoD may contribute to rebalancing mode shares in this corridor by providing on-demand home-to-station service to commuters.

Realism of parameters. Parameters for the Western Lyon case study are available in table 1.

The demand scenario has been built on the Lyon Area OD matrix (2015) and BD TOPO addresses (2021). The first database allowed the computation of hourly trip generation rates for 5 hours in the corridor. We distributed departures according to a Poisson process in time and uniformly on addresses. In total, 14k commuting trips occur. Such a disaggregation process is sufficient to compare policies. One could obtain a more precise design using more detailed data sources that render the relationship between origins locations and departure times.

Our model requires the separation of cars and AVs flows. This assumption is acceptable here as AVs could run on the HOV lane to escape the potential congestion spillback on A6. Two lanes are available for all vehicles in the

Fourviere tunnel. The considered demand represents only a part of the potential traffic passing through the tunnel. The capacity value chosen is half the maximal flow observed with inductive-loop detectors. The capacity value chosen for the other bottlenecks corresponds to 6 drop-off spots in front of a station and a drop-off time (T_f) of 30 seconds.

We assume a fixed headway that makes the MRT line attractive enough to justify the need for feeder service. A 15 minutes headway corresponds to the highest frequency currently operated.

Finally, to ensure the realism of the scenario, the number of AVs operating in the corridor is limited to 700.

4.1.2. Investigated policies

We benchmark three distinct policies regarding AMoD deployment and three priority objectives for the TA. Protectionism consists in refusing to accredit any AV and focusing on MRT design only (no AMoD). Opportunism relies on AMoD regulation to foster cooperation between MRT and AVs (regulated AMoD). Liberalism focuses on MRT design while allowing the AMoD operator to serve its interest, i.e., seek its profit increase (unregulated AMoD). Here, we consider a simple version of service pricing and cost schemes. The profit of AMoD is proportional to the amount of work achieved (in passengers * kilometers). AMoD operator's objective is reduced to maximize the usage of AVs. The distance-based mode share associated with a mode measures its usage.

The priority objective of the TA may be to: maximize MRT usage, minimize car usage, or minimize the average travel time per traveler during the morning commute.

Table 2 presents the set of scenarios.

Policy	Protectionism		Opportunism			Liberalism		
Regulation strategy	Forbid AVs		Liscense AVs by fleet and coverage zone			No regulation		
Design optimization	Unilateral		Joint			Consecutive		
Piority objective	Max MRT usage	Min average travel time	Max MRT usage	Min average travel time	Min car usage	Max MRT usage	Min average travel time	Min car usage

Table 2: Summary of the policy scenarios.

4.2. Optimization framework

4.2.1. Degrees of freedom

Figure 1 lists the design parameters. This section justifies the choice for four of them as degrees of freedom: the number of collectors, their locations, the number of AVs fleets, and their coverage zones.

Corridor configuration. As shown in the monocentric city scenario, the number of collectors is a decisive parameter in fostering MRT-AMoD cooperation. If spacing was constant in section 3.1.2, it is not relevant for the Western Lyon scenario, which has an inhomogeneous distribution of origins. The only remaining constraint on spacing is a lower bound. It prevents unrealistic configurations where two collectors are very close to each other.

AMoD configuration. Enforcing a boundary to one fleet coverage zone can mitigate the long first-mile pattern as shown by the sensitivity analysis presented in Figure 13. Two groups of consecutive bottlenecks supplied by two different fleets stay out of sync, so direct first-mile holds. One can obtain similar effects by decreasing (resp. increasing) downstream (resp. upstream) stations' drop-off capacities. However, this solution is less flexible than coverage zone geofencing, which could be demand-responsive and reviewed daily.

Geofencing AMoD can have other positive effects such as service time reduction and reliability (Shen and Quadrifoglio, 2013), total distance traveled by empty AVs drop (Fagnant et al., 2016), equity gains with more uniform service times over space (Gurumurthy et al., 2021). Gurumurthy et al. (2021) evaluate only three fleets with predetermined boundaries. Here, we optimize these parameters.

Fixed design parameters. Freeway speed, bottlenecks capacities, MRT headway, and AVs ratio are other levers to lower car performance but remain untouched in this study.

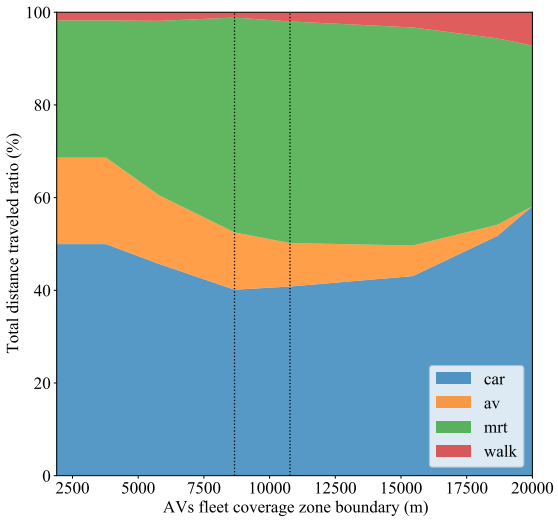


Figure 13: Sensitivity analysis of the downstream boundary of one fleet coverage zone on distance-based mode shares for the Western Lyon scenario. Compared to a situation without AV, a unique fleet mitigates car usage, whatever its coverage zone, and improves MRT usage when its coverage zone includes up to 6 upstream collectors. Three schemes emerge: (i) MRT-AV cooperation against cars for coverage zones extending from c_7 to at most c_4, (ii) AV competition with MRT and cars for coverage zone extending from c_7 to c_3, (iii) AV-cars competition with MRT for wider coverage zones.

4.2.2. Optimization loop

Under protectionism, the TA unilaterally optimizes the MRT design. Under liberalism, this optimization assumes that AMoD is composed of one fleet serving the whole corridor. Then, the unregulated AMoD adjusts its configuration to maximize its profit. Under opportunism, the TA jointly optimizes MRT and AMoD designs.

As the nature of the function that links our freedom degrees with each objective is unknown (not necessarily convex nor linear), we choose to use a genetic algorithm (GA) when exhaustive research is not possible. Such a metaheuristic does not guarantee global optimality. Stopping criteria, such as the stability of the fitness function over a large number of iterations and the small number of distinct individuals in the current population, are signs that the algorithm has reached a local optimum. Local optimality is sufficient for this study's goal to highlight the benefits of a joint MRT-AMoD design. Finding a good design instead of the optimal one still allows evaluation of the lower bound of the gains achieved under each policy. Since a fleet coverage zone is compact and extends on one or several vehicle attraction areas, an exhaustive search for the optimal AMoD configuration is possible for $M < 10$.

Applying the GA to the MRT design requires corridor discretization into a given number of location spots. A spot length is sufficiently large to have meaningful configurations (two freeway ramps/MRT platforms cannot be too close to each other) and sufficiently small to keep precise locations. Lower and upper bounds constrain the problem to account for the current corridor configuration and the finite investment budget. Applying the GA to the AMoD design requires a ternary encoding where chromosome length equals the number of collectors. For the joint MRT-AMoD design, the problem is quaternary-encoded, and chromosome length equals the number of location spots. Figure 14 provides an example of encoding.

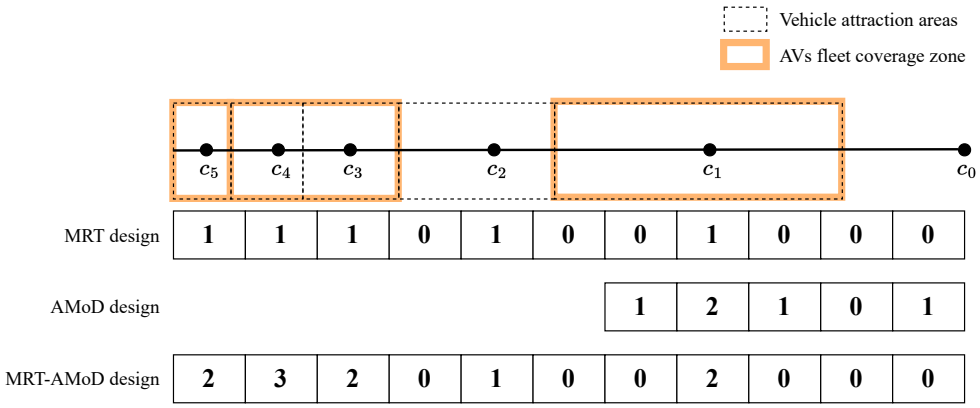


Figure 14: The MRT design is binary-encoded: 0 and 1, respectively, stand for an empty spot and a collector. The AMoD design is ternary-encoded: 0, 1, and 2, respectively, stand for an unsupplied attraction area, the downstream extremity of a new coverage zone, and the continuity of a coverage zone. The MRT-AMoD design is ternary-encoded: 0, 1, 2, and 3, respectively, stand for an empty spot, an unsupplied attraction area, the downstream extremity of a new coverage zone, and the continuity of a coverage zone.

4.3. Numerical results

Figure 15 presents the best designs found by the GA. Table 3 gathers the numerical results, and Figure 16 compares the scores obtained by each policy regarding five indicators: the average travel time per commuter (including those using streets only and located between c_1 and destination), MRT usage, car mitigation (sum of MRT, AVs and walk distance-based mode shares), the average number of commuters served per AV, the Gini coefficient of travelers waiting times (including service time and waiting time at bottleneck).

Scenario	**Western Lyon**							
Priority objective	Maximize MRT usage			Minimize average travel time			Minimize car usage	
Policy	Protectionism	Opportunism	Liberalism	Protectionism	Opportunism	Liberalism	Opportunism	Liberalism
M	15	9	6	17	13	12	5	10
Number of fleets	0	8	1	0	2	2	1	1
c volume-based mode share	51.1	48.7	48.6	51.2	49.0	49.6	48.9	48.5
r volume-based mode share	48.9	5.0	9.3	48.8	0.0	0.0	25.2	12.9
a volume-based mode share	0.0	46.3	42.1	0.0	51.0	50.4	25.9	38.5
Car distance-based mode share	39.4	33.6	47.1	54.3	38.1	41.8	32.8	35.5
MRT distance-based mode share	51.0	57.2	37.4	40.7	48.3	45.1	51.2	43.8
AV distance-based mode share	0.0	8.2	14.6	0.0	13.7	13.1	8.6	19.1
Walk distance-based mode share	9.5	1.0	0.9	4.9	0.0	0.0	7.4	1.6
Avg. TT per trav.	36:37	21:45	22:13	28:23	15:28	15:31	33:21	25:46
Gini coef.	0.68	0.68	0.51	0.68	0.63	0.63	0.57	0.52
AV occupancy	-	15.6	9.8	-	10.3	10.4	10.9	7.8

Table 3: Western Lyon results.

Maximize MRT usage. When the priority of the TA is to maximize MRT usage, opportunism performs better than other policies on all indicators except delays uniformity (Figure 16a). The MRT distance-based mode share reaches 57.5%, which is 6% more than under protectionism, and 20% more than under liberalism. The sum of AV and walk distance-based mode shares is similar in protectionism and opportunism scenarios (9.5%). It increases by 6% under liberalism, indicating the presence of long first-mile rides. These long AV trips compete with the MRT rather than the car: the car mitigation score of liberalism is lower than protectionism. Liberalism leads to counterproductive designs.

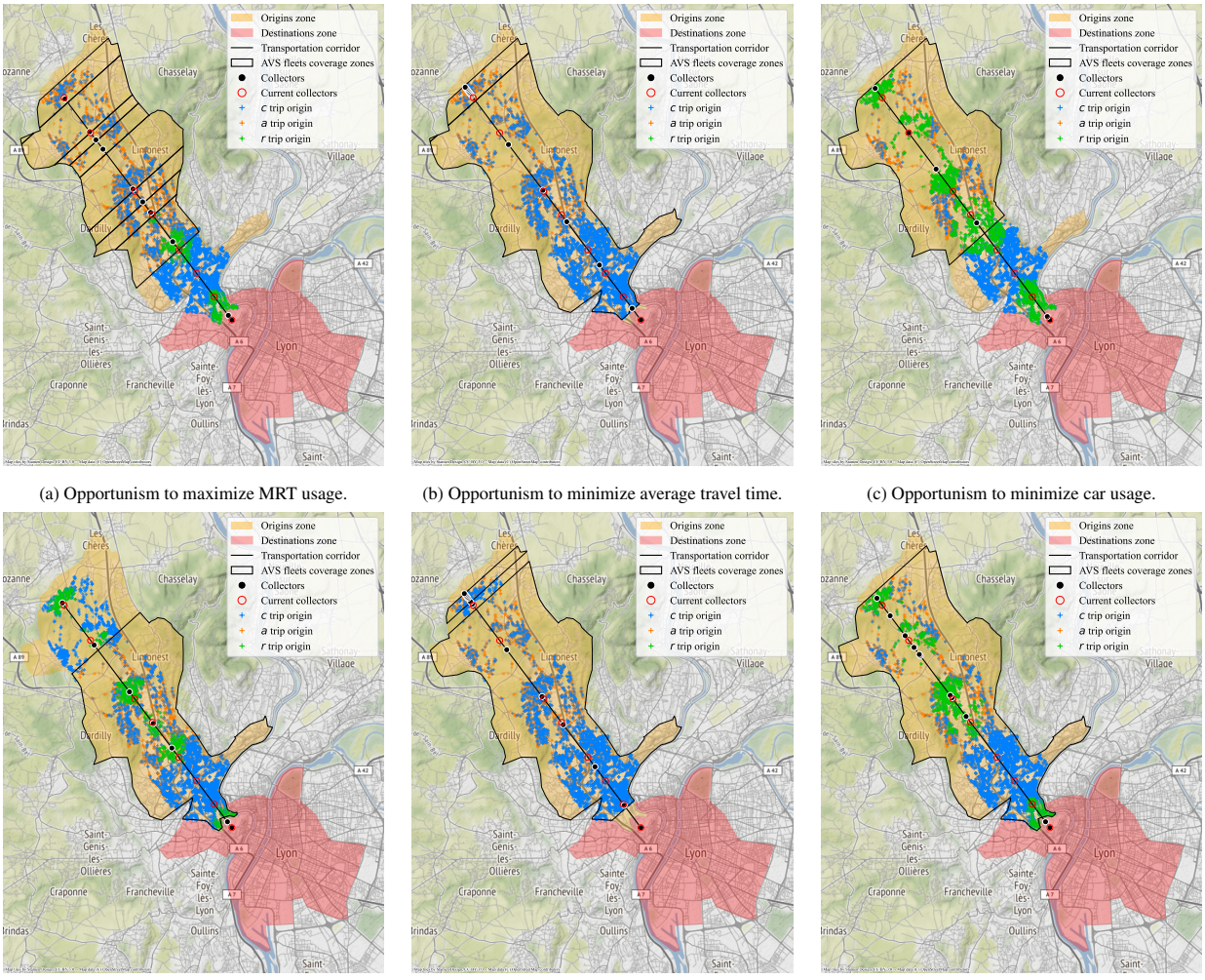
However, the MRT design found under liberalism is the closest to the current one (Figure 15d). It follows the urban geography, with one station per town. The deployment of AVs and the refinement of the MRT service pattern (skip stops c_2 and c_3) can bring benefits to commuters, notably reduced travel times and more uniformly distributed waiting times.

The liberalism to maximize the MRT usage scenario reaches the highest score regarding delays uniformity. Few travelers experience a null delay. Commuters departing from the downstream part of the corridor undergo waiting

time at the CBD off-ramp, just as those departing from the upstream part, which AMoD does not supply. Commuters departing from the middle experience non-null T_s and w_k.

On the contrary, 20% of travelers undergo 70% of the total waiting time under opportunism. Drivers mainly come from the downstream part of the corridor and undergo great w_0. As AMoD splits onto small coverage zones, T_s and w_k are null or small for those diverting to **a**.

Protectionism reaches the same score as opportunism regarding delays uniformity. The MRT riders do not undergo any delay, while drivers do.



(a) Opportunism to maximize MRT usage.

(b) Opportunism to minimize average travel time.

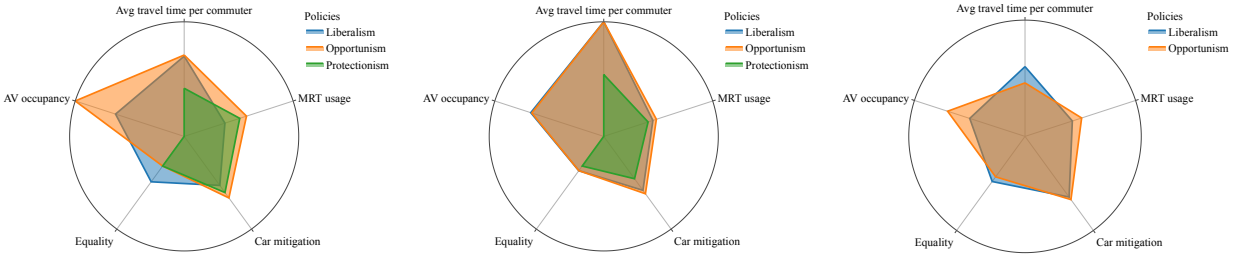
(c) Opportunism to minimize car usage.

(d) Liberalism to maximize MRT usage.

(e) Liberalism to minimize average travel time.

(f) Liberalism to minimize car usage.

Figure 15: Resulting designs for opportunism and liberalism scenarios.



(a) Maximize MRT usage objective.

(b) Minimize average travel time objective.

(c) Minimize car usage objective.

Figure 16: Scores depending on priority objective and policy for the Western Lyon corridor.

Under opportunism and liberalism, the optimization of design led to substantial s_2 (spacing between c_1 and c_2). This inter-station plays a considerable role in the distribution of travelers over itineraries since expanding Ω_1^c results in overloading bottleneck μ_0 with travelers close to the CBD. In this way, long-distance trips are subject to diversion.

A one fleet per collector strategy emerges from the joint optimization of MRT and AMoD designs, as shown by Figure 15a. This strategy avoids the long first-mile pattern since bottlenecks remain out of sync.

Moreover, it allows for keeping short service times. The service times of all fleets except those serving Ω_2^c and Ω_8^c remain below 150s. In Ω_2^c, travelers close to c_2 prefer to walk while travelers far from c_2 request an AV. Relocation and serving efforts being substantial, the service time in Ω_2^c is higher than 150s. Similarly, relocation and serving distances are high in Ω_8^c because travelers' origins are far from c_8. The attraction area extends on a sparsely populated territory between Lissieu-Dommartin and Civrieux d'Azergues towns.

Finally, it improves the AV utilization rate. An AV serves 15.6 customers on average under opportunism, which is higher than in the other scenarios. The indicator is homogeneous across fleets, except for the fleet surrounding c_8, which has a slightly smaller score (12.8).

The geometry of coverage zones under liberalism is very different. One fleet covers a huge downstream part of the corridor. AMoD operator has better deploy one fleet on a wide area to increase AVs usage. In this way, it takes advantage of synchronization between bottlenecks and can exploit the total capacity of each **a** diversion itinerary. The exclusion of attraction areas where AMoD's market share is limited (here Ω_5^c and Ω_6^c) allows for keeping its service time as small as possible in other areas.

If protectionism performs well on MRT usage maximization, it has the worst average travel time among all scenarios (36:37). AMoD deployment leads to a 15min reduction, AVs being regulated or not. When regulated, AVs utilization rate is higher, with, on average, five commuters transported per hour instead of 3.

Minimize average travel time. When the main objective is to minimize the average travel time, liberalism and opportunism have similar results and outperform protectionism on all indicators (Figure 16b). Deploying AVs reduces the average travel time by 46%, reaching the value of 15:30. Regulation is not necessary here: improving performance comes down to increasing AMoD usage. MRT has been designed considering one fleet serving the whole corridor, which is not so far from what emerges through the profit maximization process by AMoD (Figure 15e). The car (resp. MRT) distance-based mode share is slightly smaller (resp. higher) under opportunism. If liberalism provides satisfying results, a joint design helps to reach an even better equilibrium that benefits travelers, AMoD, and the TA.

The positioning of collectors is similar in both scenarios, except for c_1, which is closer to the destination under opportunism (Figure 15b). Thus, more downstream travelers are included in Ω_1^c and contribute to congestion at μ_0. It can explain that opportunism performs slightly better in car and MRT distance-based mode shares than liberalism. The obtained AMoD configurations both have one downstream fleet, which coverage zone ends just before Civrieux d'Azergues (the most upstream town).

Since $v_r > u$ here, $\Delta_{i,p}^a$ decreases as p increments. All extra travel times are positive given the substantial headway, but traveler $i \in \Omega_k^c$ has better diverting in priority as upstream as possible, i.e., at c_k. The order of extra travel times

favors the direct first-mile pattern. The long first-mile pattern has almost entirely disappeared. Travelers use AV to join one station downstream of their access collector at most.

Consequently, service time remains below the discontinuity threshold, and extra travel times by AVs are independent of i ($\Delta^{a}_{i,p} = \Delta^{a}_{p}$). The positioning of the collectors is so that all Δ^{a}_{p} are very close to each other. Then, **a**-div1 starts on all collectors supplied by AVs nearly simultaneously. It occurs before **r**-div1, so walking distance-based mode share is null. Bottlenecks are synchronized until the unloading phase. The design allows the system to enter as soon as possible and remain in a global **a**-div1 state. As a result, w_0 and $w_k (0 < k \leq 5)$ remain small and stable (5min for w_0 and 60s for w_k).

Upstream, around Civrieux d'Azergues, all designs share an accumulation of collectors. For scenarios with AVs, among these collectors, none is used as an access point by walk, and only the most downstream one is chosen for transfer from AV to MRT. Consequently, no traveler there undergoes a deterrent cumulative dwelling time. The multiplication of collectors improves the travel times of drivers and AV riders. Keeping only the most downstream of these collectors produces a marginal increase in the average travel time per commuter.

Minimize car usage. When car mitigation is the main objective, opportunism once again reaches the best score among all scenarios (Figure 16c). It relies more on walking than liberalism (+5.8%). It can explain the higher average travel time score (-8min).

Liberalism performs even worse than the "opportunism to maximize MRT usage" scenario. Though, the resulting AMoD configuration is precisely the one considered by the TA to design the MRT line (Figure 15f). It is a sign that MRT-AMoD cooperation is necessary to steal market share from **c**, especially concerning the drivers that cannot be attracted to **r** whatever the MRT design.

5. Conclusion

In this paper, we proposed a dynamic model and an optimization framework to tackle the MRT-AMoD design problem in a morning commute context. Under an exogenous constant AMoD service time, the FIFO rule characterizes the system. This property enables specifying the dynamics of the system equilibrium, which is outside the scope of DTA in more generic networks. We characterized the different stable states of UE and the conditions to pass from one to another (UE principles). Under an endogenous time-variant AMoD service time, the proposed MSA process approximates UE correctly in a few iterations, which makes our approach computationally efficient for design optimization.

Comparison with a static model showed that considering dynamics allows capturing richer cooperation-competition schemes between MRT and AMoD. Two extreme patterns emerged from individual choices depending on the MRT design. In the long first-mile pattern, travelers favor downstream stations to transfer, so AV legs of trips are long. In the direct first-mile pattern, travelers tend to use AVs on smaller distances to join their access collector and transfer there. These patterns have been explained in light of UE principles.

The application of our model to the Western Lyon corridor showed that the joint design of MRT and AMoD can foster their cooperation. Indeed, opportunism provides the best results for each priority objective of the TA. A single fleet per collector design is found in the Western Lyon corridor when the objective is to maximize MRT usage. However, such regulation offers travelers fewer choices for diversion, less flexibility, and uniformity in experienced delays. Liberalism is less reliable in achieving the given objectives but ensures greater uniformity of delays. Protectionism may perform adequately depending on the distribution of origins but never reaches the highest scores.

The one fleet per collector design strategy emerging from the opportunism policy is consistent with the paradigm of Autonomous Mobility District (AMD). AMD refers to using AV technology in a limited geographical area that generally includes an MRT station. AMD could answer several urban planning goals. Hou et al. (2018) list the proposed benefits of an AMD, including the reduction of car mode share, parking lots, and pedestrian-oriented land use. Many publications have studied intra-AMD mobility with simulation (Huang et al., 2021; Shen et al., 2018; Scheltes and de Almeida Correia, 2017), but few have studied inter-AMD mobility.

The presented model includes several strong assumptions and is quite restrictive regarding intermodality and network. However, simplifications make UE dynamics explicit and allow escaping the black-box effect that characterizes less restrictive frameworks. The network structure is generic enough to apply to several urban areas, and the approach

works with little data. When more precise data is available, refining the model parameters improves scenario realism. Our model could be extended to more intermodal options. Extensions should maintain the network properties (single bottleneck per route and combination of modes) and the ability of our UE resolution procedure to meet the quality criteria in a few iterations. They should either preserve the theoretical arrival order at all bottlenecks (as in our model under exogenous constant T_s) or generate local bounded order rearranging (as in our model under endogenous time-variant T_s). For example, one can add a car + MRT option with capacitated park-and-rides at collectors. A micro-mobility + MRT option would be similar to $\mathbf{r}$.

Another limitation of our model is the reality gap that stems from the deterministic route choice based on travel time only. We did not include the monetary aspect because it involves making additional assumptions on AMoD and MRT fare schemes. Such assumptions prevent us from accessing the primary cooperation-competition schemes between MRT and AMoD. We did not include a transfer penalty to account for the discomfort of changing mode because transfer between AV and MRT is already penalized with a wait time for drop-off. Adding a constant transfer penalty to $T^{\mathrm{a}}_{i,p}$ translates the extra travel times $\Delta^{\mathrm{a}}_{i,p}$ and changes the times at which the system changes state but does not modify the states themselves. In order to address the reality gap in the Western Lyon corridor, future research will evaluate the designs found in this study through simulation on agent-based/activity-based platforms, which are more accurate in reproducing travelers' choices and AMoD operation.

CRediT author statement

Cortina Mélanie: Funding acquisition, Visualization, Conceptualization, Methodology, Software, Writing- Original draft preparation. **Chiabaut Nicolas**: Funding acquisition, Conceptualization, Methodology, Supervision, Writing- Review & Editing. **Leclercq Ludovic**: Conceptualization, Methodology, Supervision, Writing- Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by "Lyon Urban School" (ANR-17-CONV-0004) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). Ludovic Leclercq has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement n°953783 (DIT4TRAM).

Appendices

A. Table of notations

The notations used in this paper are summarized below in the order of appearance.

Nomenclature

Assumptions

B	Corridor length.
W	Corridor width.
M	Number of collectors along corridor without counting the CBD off-ramp and terminus MRT station as a collector.
c_k	k^{th} collector of the corridor when $k > 0$, c_0 designates the CBD off-ramp and terminus MRT station.
x_k	Location of collector c_k.
(x_i, y_i)	Traveler i's origin coordinates.
$v_{\text{st}}, v_{\text{w}}$	Speed of cars in streets, speed of pedestrians.
$\text{c}, \text{r}, \text{a}$	Superscripts designating car only, walk and MRT, AV and MRT options. In the core of the text we note **c**, **r**, **a** in bold for readability.
u	Free-flow speed on freeway.
μ_k	Capacity of delay generator point associated with c_k.
T_{f}	Time for the drop-off maneuver (for an AV to park at a drop-off spot and for the passenger to leave the vehicle in security).
v_{r}	Speed of the MRT.
T_{d}	Dwell time at a MRT station.
h	Headway of MRT service during peak hours in seconds.
Ω_k^{w}	Set of travelers departing from the walking attraction area of c_k.
$X_{k,k+1}^{\text{w}}$	Frontier between two consecutive walking attraction areas ($\Omega_k^{\text{w}} \cap \Omega_{k+1}^{w}$).
Ω_k^{c}	Set of travelers departing from the vehicle attraction area of c_k.
$X_{k,k+1}^{\text{c}}$	Frontier between two consecutive vehicle attraction areas ($\Omega_k^{\text{c}} \cap \Omega_{k+1}^{\text{c}}$).

Dynamic model

$w_k(t)$	Waiting time at bottleneck associated with c_k for a traveler arriving at t.
$A_k(t)$	Cumulative travelers count arrival curve at bottleneck μ_k.
t_k	Time at which congestion starts at bottleneck μ_k.
T_i^{c}	Free-flow travel time of traveler i by mode **c**.
T_i^{r}	Travel time of traveler i by mode **r**.
$T_{i,p}^{\text{a}}$	Free-flow (nor waiting for AMoD nor waiting at bottleneck μ_p are included) travel time of traveler i by mode **a** with a transfer at c_p.
t_i^0	Theoretical arrival time of traveler i at c_0.
τ_i^{c}	Travel time of traveler i by mode **c** outside of free-flow conditions.
$t_i^{\text{req}} = t_i^{\text{dep}}$	Time at which traveler i sends a request for AV, it equals traveler i's departure time.
t_i^p	Time at which traveler i arrives at bottleneck μ_p when she chooses mode **a**.
T_{s}	Service time function, $T_{\text{s}}(t_i^{\text{req}})$ is the service time experienced by traveler i when it requests an AV.
$\tau_{i,p}^{\text{a}}$	Travel time of traveler i by mode **a**.
Δ_i^{r}	Extra travel time experienced by traveler i when she takes mode **r** instead of free-flow **c** mode.
$\Delta_{i,p}^{\text{a}}$	Free-flow extra travel time experienced by traveler i when she takes mode **a** with a transfer at c_p instead of free-flow **c** mode (noted Δ_p^{a} when it does not depend on i's origin which is the case when T_{s} is a constant function).
s_k	Spacing distance between c_{k-1} and c_k.
$A^{\text{th}}(t)$	Theoretical global arrival curve at bottleneck μ_0.

$A_k^{\mathrm{r}}(t)$	Cumulative travelers count arrival curve at c_k by walk.
m	Number of AVs in a fleet.
K	K^{th} iteration of the MSA process.
$\widehat{T}_{\mathrm{s}}$	Effective service time profile.
Δt	Time elapsed while a fleet receiving the m last request.
E_1	Expected pick up time for the mean request.
E_2	Expected running time for the mean request.
E_3	Expected waiting time at bottleneck for the mean request.
Static model	
N	Number of travelers pools.
O_i	Abstract origin aggregating several commuters.
$G(V,E)$	Graph representing the corridor.
D, C_k^{c}, C_k^{r}, C_k^{a}	Destination, on-ramp, station and transfer nodes in G.
α_{c}, β_{c}, α_{a}, β_{a}	BPR parameters respectively associated to freeway final off-ramp and drop-off delay generator points.
$v(\text{edge})$	Traffic volume on edge of G.
$T(\text{edge})$	Travel time on edge of G.

B. Simplified examples to illustrate UE principles and network characteristics

In order to better illustrate the notions associated with UE computation (diversion patterns, synchronization), we use the constant T_{s} assumption in this appendix and limit the travelers origins to collectors locations. Thus, we can treat travelers with flows. We note $A_k^{\mathrm{th}}(t)$ the demand from c_k, T_k^{c} and $T_{k,p}^{\mathrm{c}}$ the travel times by **c** and **a** for all travelers departing from c_k, Δ_k^{r} the extra travel time experienced by travelers departing from c_k when they divert to **r**. Example B.1 treats the case when only **c** and **a** modes are available, example B.2 deals with the three modes.

*B.1. Example with modes **c**, **a***

To demonstrate desynchronization phenomenon, **r** is unavailable here. Extra travel times on diversion routes are monotonically increasing from downstream to upstream collectors ($\Delta_1^{\mathrm{a}} < ... < \Delta_M^{\mathrm{a}}$).

***M* = 2**. Figure B.1 presents the resulting equilibrium.

1. Initially, every commuter chooses **c**. Congestion on freeway starts from the moment when $\dot{A}^{\mathrm{th}}(t) > \mu_0$ where $A^{\mathrm{th}}(t) = A_1^{\mathrm{th}}(t) + A_2^{\mathrm{th}}(t)$ denotes global demand as the sum of demands from both collectors. A queue forms and grows till $w_0 = \Delta_1^{\mathrm{a}}$. c_1 is the most downstream collector before destination, all travelers have a new alternative to c_0 itinerary.

2. If $\dot{A}^{\mathrm{th}} \leq \mu_0 + \mu_1$, **a** starts to be used by both travelers flows with a transfer at c_1 but μ_1 remains uncongested. According to the UE principle, travel times of all routes should be equal: $T_k^{\mathrm{c}}(t) = T_{k,1}^{\mathrm{a}}(t), \forall k \in \{1,2\}$, which can be derived into $\dot{w}_0(t) = 0$. Replacing w_0 by its expression given by 1 leads to: $\dot{A}_0(t) = \mu_0$ which characterizes diversion pattern 1 (div1). See period A in Figure B.1.

3. As soon as $\dot{A}^{\mathrm{th}} > \mu_0 + \mu_1$, μ_1 starts to be congested, and diversion pattern 2 (div2) is observed (period B). It is characterized by $\dot{w}_0(t) = \dot{w}_1(t)$: waiting times at μ_0 and μ_1 increase at the same pace till $w_0(t) = w_1(t) + \Delta_1^{\mathrm{a}} = \Delta_2^{\mathrm{a}}$.

4. At this time, we have $\mu_0 + \mu_1 < \dot{A}^{\mathrm{th}} \leq \mu_0 + \mu_1 + \mu_2$. Route c_2 starts to be used. This new alternative is only accessible to travelers departing from c_2. Two cases are possible:

   - If $\dot{A}_1^{\mathrm{th}} \leq \mu_0 + \mu_1$, travelers flow from c_2 still uses μ_0 and μ_1 bottlenecks. Waiting times $w_0(t)$ and $w_1(t) + \Delta_1^{\mathrm{a}}$ are maintained equal to Δ_2^{a}. This is a div1 where μ_0 and μ_1 work at capacity while μ_2 absorbs the surplus of demand coming from c_2.

$$\begin{cases} \dot{A}_0(t) = & \mu_0 \\ \dot{A}_1(t) = & \mu_1 \\ \dot{A}_2(t) = \dot{A}^{\mathrm{th}}(t) - \mu_0 - \mu_1 < \mu_2 \end{cases} \tag{B.1}$$

- If $\dot{A}_1^{\text{th}} > \mu_0 + \mu_1$, div2 is maintained for travelers flow departing from c_1 only, while c_2 route is chosen by all travelers departing from c_2. Waiting time on μ_2 gets out of sync with other bottlenecks: $w_0(t) = w_1(t) + \Delta_1^{\text{a}} > \Delta_2^{\text{a}}$ and $w_2(t) = 0$ (period C).

$$\begin{cases} \dot{A}_0(t) = \frac{\mu_0}{\mu_0+\mu_1}\dot{A}_1^{\text{th}}(t) \\ \dot{A}_1(t) = \frac{\mu_1}{\mu_0+\mu_1}\dot{A}_1^{\text{th}}(t) \\ \dot{A}_2(t) = \quad \dot{A}_2^{\text{th}}(t) \end{cases} \tag{B.2}$$

5. If global demand is still increasing, we get to a time when $\dot{A}^{\text{th}}(t) > \mu_0 + \mu_1 + \mu_2$. Again, several cases arise, depending on synchronization between bottlenecks and demand per origin collector.

   - If bottlenecks are all synchronized ($w_0(t) = w_1(t) + \Delta_1^{\text{a}} = \Delta_2^{\text{a}}$), $\dot{A}_1^{\text{th}}(t) \leq \mu_0 + \mu_1$ or $\frac{\dot{A}_1^{\text{th}}(t)}{\mu_0+\mu_1} < \frac{\dot{A}_2^{\text{th}}(t)}{\mu_2}$, a div2 is observed:

$$\begin{cases} \dot{A}_0(t) = \frac{\mu_0}{\mu_0+\mu_1+\mu_2}\dot{A}^{\text{th}}(t) \\ \dot{A}_1(t) = \frac{\mu_1}{\mu_0+\mu_1+\mu_2}\dot{A}^{\text{th}}(t) \\ \dot{A}_2(t) = \frac{\mu_2}{\mu_0+\mu_1+\mu_2}\dot{A}^{\text{th}}(t) \end{cases} \tag{B.3}$$

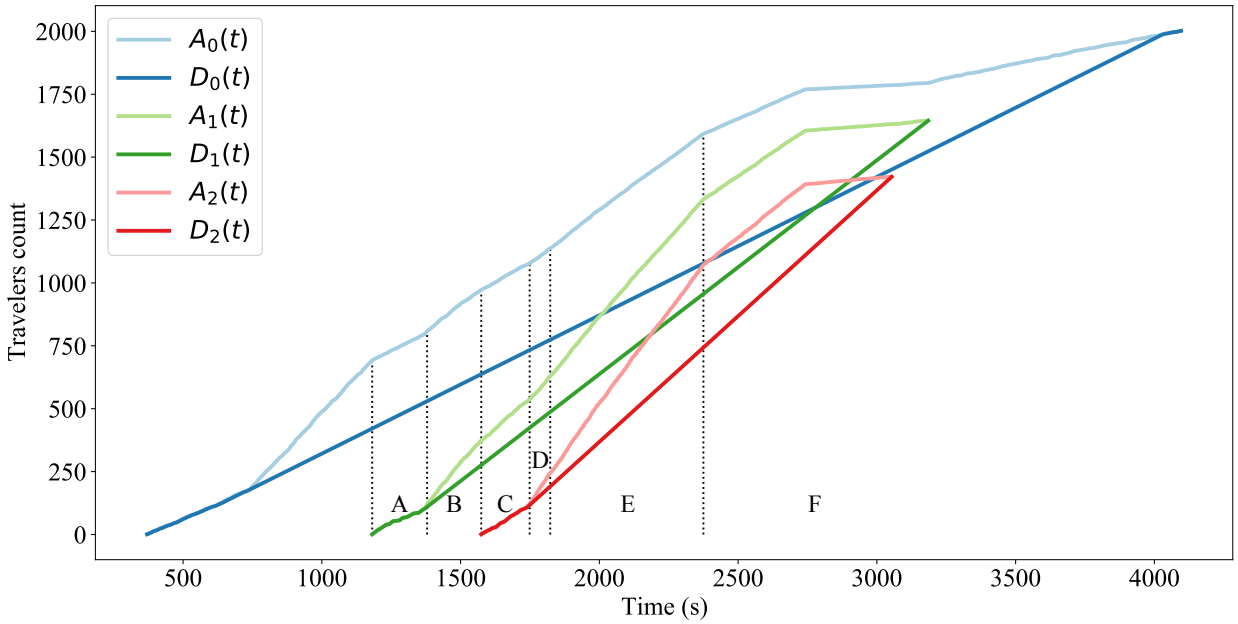
   - If μ_2 is not synchronized with others ($w_0(t) = w_1(t) + \Delta_1^{\text{a}} > \Delta_2^{\text{a}}$), and $\dot{A}_1^{\text{th}}(t) \leq \mu_0 + \mu_1$ or $\frac{\dot{A}_1^{\text{th}}(t)}{\mu_0+\mu_1} < \frac{\dot{A}_2^{\text{th}}(t)}{\mu_2}$, div2 continues for travelers flow departing from c_1 while c_2 route is chosen by others till re-synchronization. Re-synchronization takes place when $w_0(t) = w_1(t) + \Delta_1^{\text{a}} = w_2(t) + \Delta_2^{\text{a}}$, at the end of period D, from when a div2 follows (period E).
   - If $\frac{\dot{A}_1^{\text{th}}(t)}{\mu_1+\mu_2} > \frac{\dot{A}_2^{\text{th}}(t)}{\mu_2}$, whatever the synchronization state of bottlenecks, div2 continues for travelers flow departing from c_1, while c_2 route is chosen by others.

6. When global demand decreases, two possibilities emerge.

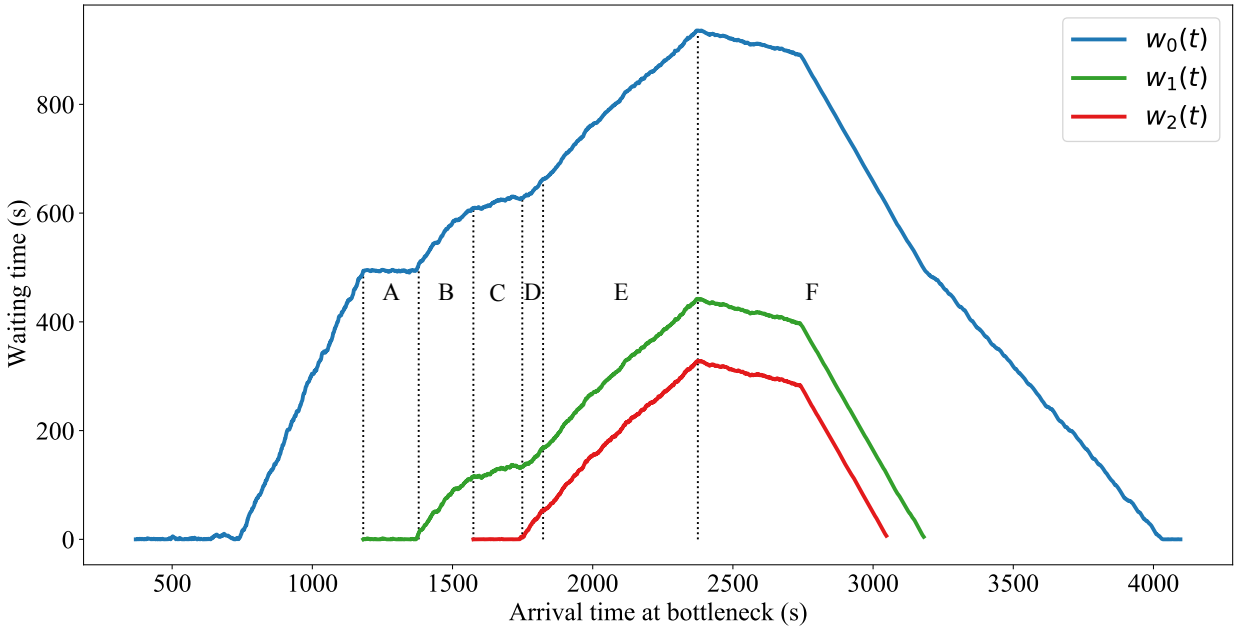
   - If bottlenecks are synchronized, div2 continues (period F). A bottleneck ceased to be used as soon as its queue vanishes.
   - If they are not, re-synchronization takes place during the demand decrease phase.

M > 2. We note p^* the index of the most downstream bottleneck synchronized with bottleneck μ_p.

- Bottlenecks μ_0 and μ_1 are always synchronized because they are accessible to every commuter.
- Bottleneck μ_{p+1} starts to be used when $w_p + \Delta_p^{\text{a}} = \Delta_{p+1}^{\text{a}}$.
- Bottleneck μ_{p+1} gets out of sync with bottlenecks μ_{p^*}, ..., μ_p when:
  - $$\begin{cases} \sum_{k=p^*}^{M} \dot{A}_k^{\text{th}}(t) \leq \sum_{k=p^*}^{p+1} \mu_k \\ \sum_{k=p^*}^{p} \dot{A}_k^{\text{th}}(t) > \sum_{k=p^*}^{p} \mu_k \end{cases} \tag{B.4}$$
  - or
  $$\begin{cases} \sum_{k=p^*}^{M} \dot{A}_k^{\text{th}}(t) > \sum_{k=p^*}^{p+1} \mu_k \\ \frac{\sum_{k=p^*}^{p} \dot{A}_k^{\text{th}}(t)}{\sum_{k=p^*}^{p} \mu_k} > \frac{\sum_{k=p+1}^{M} \dot{A}_k^{\text{th}}(t)}{\mu_{p+1}} \end{cases} \tag{B.5}$$
- Re-synchronization between bottlenecks μ_{p^*}, ..., μ_p and μ_{q^*}, ..., μ_q takes place when:
  - demand for μ_{p^*}, ..., μ_p decreases to reach condition $\sum_{k=p^*}^{p} \dot{A}_k^{\text{th}}(t) < \sum_{k=p^*}^{p} \mu_k$. Waiting times w_{p^*}, ..., w_p decline till $w_p + \Delta_p^{\text{a}} = ... = w_q + \Delta_q^{\text{a}}$.
  - demand for μ_{q^*}, ..., μ_q increases or demand for μ_{p^*}, ..., μ_p decreases so that $\frac{\sum_{k=p^*}^{p} \dot{A}_k^{\text{th}}(t)}{\sum_{k=p^*}^{p} \mu_k} < \frac{\sum_{k=q^*}^{q} \dot{A}_k^{\text{th}}(t)}{\sum_{k=q^*}^{q} \mu_k}$. The gap between waiting times on these two bottleneck groups is reduced till $w_p + \Delta_p^{\text{a}} = ... = w_q + \Delta_q^{\text{a}}$.



(a) Arrivals and departures at bottlenecks.



(b) Waiting times at bottlenecks.

Figure B.1: UE solving with M=2, modes **c**, **a** in a linear corridor. Div1 takes place during period A, div2 during period B. μ_2 gets out of sync with downstream collectors for period C and re-synchronizes for period D. Div2 continues on period E and F till the end on peak hours. Note that drop-off curves have been transposed in time to be aligned with CBD off-ramp bottleneck and a better highlight of patterns.

B.2. *Example with all modes*

M = 2. Mode **r** is now available. We have $\Delta_k^{\mathrm{a}} = \Delta_k^{\mathrm{r}} + T_{\mathrm{s}}$, $k \in \{1,2\}$. Parameters are so that $\Delta_1^{\mathrm{r}} < \Delta_1^{\mathrm{a}} < \Delta_2^{\mathrm{r}} < \Delta_2^{\mathrm{a}}$. Figure B.2 represents UE solution.

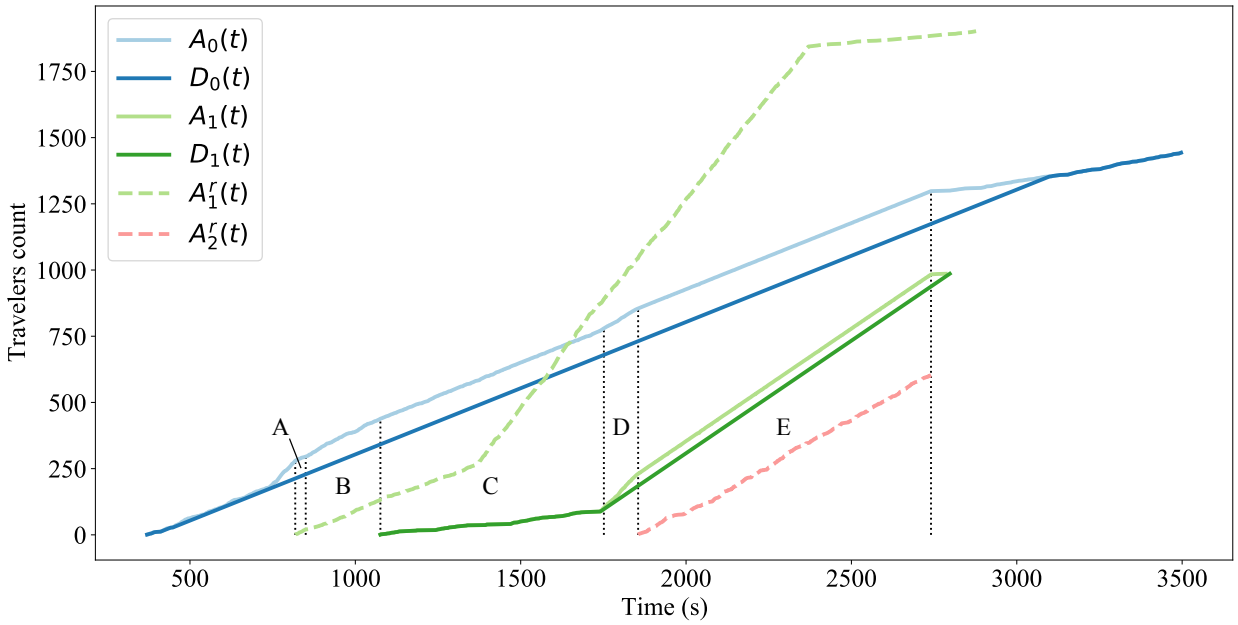


Figure B.2: Arrivals and departures at bottlenecks with $M = 2$, modes **c**, **a**, **r**. Diversion by **r** at c_1 starts from period A and it applies to all travelers departing from c_1 during following periods. Div1 happens on period C, E and div2 on period D.

1. Just as in the previous case, commuters choose **c** till $\dot{A}^{\mathrm{th}}(t) > \mu_0$ and $w_0 = \Delta_1^{\mathrm{r}}$.

2. Then travelers from c_1 start to use **r**. If $\dot{A}^{\mathrm{th}}(t) - \dot{A}_1^{\mathrm{th}}(t) \leq \mu_0$, div1 happens. A part of demand uses **c** while another part uses **r** (period A). We note $A_k^{\mathrm{r}}(t)$ the cumulative travelers arrival count at c_k by walking.

$$\begin{cases} \dot{A}_0(t) = & \mu_0 \\ \dot{A}_1^{\mathrm{r}}(t) = \dot{A}^{\mathrm{th}}(t) - \mu_0 \end{cases} \tag{B.6}$$

3. When $\dot{A}^{\mathrm{th}}(t) - \dot{A}_1^{\mathrm{th}}(t) > \mu_0$, all travelers from c_1 choose **r** and waiting time on freeway CBD off-ramp bottleneck increases till $w_0 = \Delta_1^{\mathrm{a}}$ (period B).

$$\begin{cases} \dot{A}_0(t) = \dot{A}^{\mathrm{th}}(t) - \dot{A}_1^{\mathrm{th}}(t) \\ \dot{A}_1^{\mathrm{r}}(t) = & \dot{A}_1^{\mathrm{th}}(t) \end{cases} \tag{B.7}$$

4. From this moment, it is worth for travelers from c_2 to use **a** with a transfer at c_1. If $\dot{A}^{\mathrm{th}}(t) - \dot{A}_1^{\mathrm{th}}(t) \leq \mu_0 + \mu_1$, we observe a div1 with the particularity that commuters from c_1 are out of the game (period C).

$$\begin{cases} \dot{A}_0(t) = & \mu_0 \\ \dot{A}_1(t) = \dot{A}^{\mathrm{th}}(t) - \dot{A}_1^{\mathrm{th}}(t) - \mu_0 \\ \dot{A}_1^{\mathrm{r}}(t) = & \dot{A}_1^{\mathrm{th}}(t) \end{cases} \tag{B.8}$$

5. As soon as $\dot{A}^{\text{th}}(t) - \dot{A}_1^{\text{th}}(t) > \mu_0 + \mu_1$, diversion pattern 2 applies similarly (period D) till $w_0 = w_1 + \Delta_1^{\text{a}} = \Delta_2^{\text{r}}$.

$$\begin{cases} \dot{A}_0(t) = \frac{\mu_0}{\mu_0+\mu_1}(\dot{A}^{\text{th}} - \dot{A}_1^{\text{th}}) \\ \dot{A}_1(t) = \frac{\mu_1}{\mu_0+\mu_1}(\dot{A}^{\text{th}} - \dot{A}_1^{\text{th}}) \\ \dot{A}_1^{\text{r}}(t) = \quad \dot{A}_1^{\text{th}}(t) \end{cases} \tag{B.9}$$

6. A last div1 concerning only travelers from c_2 happens (period E) and lasts till demand decrease and queues vanishing.

$$\begin{cases} \dot{A}_0(t) = \quad \mu_0 \\ \dot{A}_1(t) = \quad \mu_1 \\ \dot{A}_1^{\text{r}}(t) = \quad \dot{A}_1^{\text{th}}(t) \\ \dot{A}_2^{\text{r}}(t) = \dot{A}^{\text{th}}(t) - \dot{A}_1^{\text{th}}(t) - \mu_0 - \mu_1 \end{cases} \tag{B.10}$$

M > 2. Let us note p, q, the indexes of the most upstream collectors at which **r** and **a** are worth respectively. Index p is so that $\Delta_p^{\text{r}} \leq w_0 < \Delta_{p+1}^{\text{r}}$. Index q is so that $\Delta_q^{\text{a}} \leq w_0 < \Delta_{q+1}^{\text{a}}$. Because $\Delta_k^{\text{r}} < \Delta_k^{\text{a}}$ ($1 \leq k \leq M$), we do not find the same kind of desynchronization effects due to **a** as in the only-two-modes case. All $\mu_k, 0 \leq k \leq q$ are synchronized.

- Diversion at c_{p+1} by **r** starts as soon as $w_0 = ... = w_p + \Delta_p^{\text{a}} = \Delta_{p+1}^{\text{r}}$.
- MRT station at c_p can be involved in a div1 with downstream limited capacity bottlenecks if $\sum_{k=p+1}^{M} \dot{A}_k^{\text{th}} \leq \sum_{k=0}^{q} \mu_k$.
- Desynchronization of A_p^{r} takes place when waiting times on μ_0, ..., μ_q increase, i.e. when $\sum_{k=p+1}^{M} \dot{A}_k^{\text{th}} > \sum_{k=0}^{q} \mu_k$.

References

Akamatsu, T., Wada, K., Hayashi, S., 2015. The corridor problem with discrete multiple bottlenecks. Transportation Research Procedia 7, 474–498. doi:10.1016/j.trpro.2015.06.025.

Ameli, M., 2019. Heuristic methods for calculating dynamic traffic assignment simulation-based dynamic traffic assignment.

André, J., Vieira da Rocha, T., 2020. Émissions de gaz à effet de serre et de polluants atmosphériques en France. Technical Report. Citepa. https://www.citepa.org/fr/secten/.

Arnott, R., de Palma, A., Lindsey, R., 1990. Economics of a bottleneck. Journal of Urban Economics 27, 111–130. doi:10.1016/0094-1190(90)90028-L.

Bagloee, S.A., (Avi) Ceder, A., Sarvi, M., Asadi, M., 2019. Is it time to go for no-car zone policies? braess paradox detection. Transportation Research Part A: Policy and Practice 121, 251–264. doi:10.1016/j.tra.2019.01.021.

Basciftci, B., Van Hentenryck, P., 2021. Capturing travel mode adoption in designing on-demand multimodal transit systems URL: https://arxiv.org/abs/2101.01056, doi:10.48550/ARXIV.2101.01056. publisher: arXiv Version Number: 2.

Basu, R., Araldo, A., Akkinepally, A.P., Nahmias Biran, B.H., Basak, K., Seshadri, R., Deshmukh, N., Kumar, N., Azevedo, C.L., Ben-Akiva, M., 2018. Automated mobility-on-demand vs. mass transit: A multi-modal activity-driven agent-based simulation approach. Transportation Research Record: Journal of the Transportation Research Board 2672, 608–618. URL: http://journals.sagepub.com/doi/10.1177/0361198118758630, doi:10.1177/0361198118758630.

Becker, H., Becker, F., Abe, R., Bekhor, S., Belgiawan, P.F., Compostella, J., Frazzoli, E., Fulton, L.M., Guggisberg Bicudo, D., Murthy Gurumurthy, K., Hensher, D.A., Joubert, J.W., Kockelman, K.M., Kröger, L., Le Vine, S., Malik, J., Marczuk, K., Ashari Nasution, R., Rich, J., Papu Carrone, A., Shen, D., Shiftan, Y., Tirachini, A., Wong, Y.Z., Zhang, M., Bösch, P.M., Axhausen, K.W., 2020. Impact of vehicle automation and electric propulsion on production costs for mobility services worldwide. Transportation Research Part A: Policy and Practice 138, 105–126. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856420305772, doi:10.1016/j.tra.2020.04.021.

Bischoff, J., Maciejewski, M., 2016. Simulation of city-wide replacement of private cars with autonomous taxis in berlin. Procedia Computer Science 83, 237–244. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050916301442, doi:10.1016/j.procs.2016.04.121.

Cerema, 2021. Prise de la compétence d'organisation des mobilités par les communautés de communes : panorama au 15 juillet 2021.

Chiabaut, N., Comemale, A., Laval, J.A., 2018. Analysis of dynamic multimodal user equilibrium for the single destination problem. Unpublished.

Dang, L., von Arx, W., Frölicher, J., 2021. The impact of on-demand collective transport services on sustainability: A comparison of various service options in a rural and an urban area of switzerland. Sustainability 13, 3091. doi:10.3390/su13063091.

Ermans, T., Brandeleer, C., Hubert, M., Lebrun, K., Sieux, F., 2018. Travel between home and work: current situation and perspectives for action for companies: BSI synopsis. Brussels Studies doi:10.4000/brussels.1696.

Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice 77, 167–181. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856415000804, doi:10.1016/j.tra.2015.04.003.

Fagnant, D.J., Kockelman, K.M., Bansal, P., 2016. Operations of shared autonomous vehicle fleet for austin, texas, market. Transportation Research Record: Journal of the Transportation Research Board 2563, 98–106. doi:10.3141/2536-12.

Fosgerau, M., de Palma, A., 2012. Congestion in a city with a central bottleneck. Journal of Urban Economics 71, 269–277. doi:10.1016/j.jue.2011.12.002.

Fu, Q., Liu, R., Hess, S., 2012. A review on transit assignment modelling approaches to congested networks: A new perspective. Procedia - Social and Behavioral Sciences 54, 1145–1155. doi:10.1016/j.sbspro.2012.09.829.

Gallotti, R., Barthelemy, M., 2015. Anatomy and efficiency of urban multimodal mobility. Scientific Reports 4, 6911. doi:10.1038/srep06911.

Greenblatt, J.B., Saxena, S., 2015. Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. Nature Climate Change 5, 860–863. doi:10.1038/nclimate2685.

Gurumurthy, K.M., Auld, J., Kockelman, K., 2021. A system of shared autonomous vehicles for chicago: Understanding the effects of geofencing the service. Journal of Transport and Land Use 14, 933–948. doi:10.5198/jtlu.2021.1926.

Gurumurthy, K.M., Kockelman, K.M., Zuniga-Garcia, N., 2020. First-mile-last-mile collector-distributor system using shared autonomous mobility. Transportation Research Record: Journal of the Transportation Research Board , 036119812093626doi:10.1177/0361198120936267.

Hall, J.D., Palsson, C., Price, J., 2018. Is uber a substitute or complement for public transit? Journal of Urban Economics 108, 36–50. URL: https://linkinghub.elsevier.com/retrieve/pii/S0094119018300731, doi:10.1016/j.jue.2018.09.003.

Hou, Y., Young, S., Garikapati, V., Chen, Y., Zhu, L., 2018. Initial assessment and modeling framework development for automated mobility districts. URL: https://www.osti.gov/biblio/1420369.

Huang, Y., Kockelman, K.M., Garikapati, V., Zhu, L., Young, S., 2021. Use of shared automated vehicles for first-mile last-mile service: Micro-simulation of rail-transit connections in austin, texas. Transportation Research Record: Journal of the Transportation Research Board 2675, 135–149. URL: http://journals.sagepub.com/doi/10.1177/0361198120962491, doi:10.1177/0361198120962491.

Iryo, T., 2013. Properties of dynamic user equilibrium solution: existence, uniqueness, stability, and robust solution methodology. Transportmetrica B: Transport Dynamics 1, 52–67. URL: http://www.tandfonline.com/doi/abs/10.1080/21680566.2013.779793, doi:10.1080/21680566.2013.779793.

Kondor, D., Zhang, H., Tachet, R., Santi, P., Ratti, C., 2019. Estimating savings in parking demand using shared vehicles for home–work commuting. IEEE Transactions on Intelligent Transportation Systems 20, 2903–2912. doi:10.1109/TITS.2018.2869085.

Kumar, P., Khani, A., 2022. Planning of integrated mobility-on-demand and urban transit networks. Transportation Research Part A: Policy and Practice 166, 499–521. URL: https://linkinghub.elsevier.com/retrieve/pii/S0965856422002841, doi:10.1016/j.tra.2022.11.001.

Laval, J.A., 2009. Graphical solution and continuum approximation for the single destination dynamic user equilibrium problem. Transportation Research Part B: Methodological 43, 108–118. doi:10.1016/j.trb.2008.05.009.

Litman, T., 2021. Autonomous vehicle implementation predictions: Implications for Transport Planning. Technical Report. Victoria transport policy institute. https://www.vtpi.org/avip.pdf.
Liu, T., Huang, H., Yang, H., Zhang, X., 2009. Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice. Transportation Research Part B: Methodological 43, 692–707. doi:10.1016/j.trb.2009.01.001.
Liu, Y., Ouyang, Y., 2021. Mobility service design via joint optimization of transit networks and demand-responsive services. Transportation Research Part B: Methodological 151, 22–41. doi:10.1016/j.trb.2021.06.005.
Maciejewski, M., Bischoff, J., Nagel, K., 2016. An assignment-based approach to efficient real-time city-scale taxi dispatching. IEEE Intelligent Systems 31, 68–77. URL: http://ieeexplore.ieee.org/document/7389909/, doi:10.1109/MIS.2016.2.
Maerivoet, S., De Moor, B., 2005. Transportation Planning and Traffic Flow Models. Technical Report 05. SISTA. http://arxiv.org/abs/physics/0507127. arXiv:physics/0507127.
Militão, A.M., Tirachini, A., 2021. Optimal fleet size for a shared demand-responsive transport system with human-driven vs automated vehicles: A total cost minimization approach. Transportation Research Part A: Policy and Practice 151, 52–80. doi:10.1016/j.tra.2021.07.004.
Mo, B., Cao, Z., Zhang, H., Shen, Y., Zhao, J., 2021. Competition between shared autonomous vehicles and public transit: A case study in singapore. Transportation Research Part C: Emerging Technologies 127, 103058. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X21000863, doi:10.1016/j.trc.2021.103058.
Mounce, R., Nelson, J.D., 2019. On the potential for one-way electric vehicle car-sharing in future mobility systems. Transportation Research Part A: Policy and Practice 120, 17–30. doi:10.1016/j.tra.2018.12.003.
Nicolas, J., Verry, D., Longuar, Z., 2012. Évolutions récentes des émissions de CO2 liées à la mobilité des français : analyser les dynamiques à l'œuvre grâce aux enquêtes nationales transports de 1994 et 2008. Economie et statistique 457, 161–183. doi:10.3406/estat.2012.9970.
Pavone, M., 2015. Autonomous mobility-on-demand systems for future urban mobility, in: Maurer, M., Gerdes, J.C., Lenz, B., Winner, H. (Eds.), Autonomes Fahren. Springer Berlin Heidelberg, pp. 399–416. URL: http://link.springer.com/10.1007/978-3-662-45854-9_19, doi:10.1007/978-3-662-45854-9_19.
Pinto, H.K., Hyland, M.F., Mahmassani, H.S., Verbas, I., 2020. Joint design of multimodal transit networks and shared autonomous mobility fleets. Transportation Research Part C: Emerging Technologies 113, 2–20. doi:10.1016/j.trc.2019.06.010.
Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? a survey-based comparison of taxis, transit, and ridesourcing services in san francisco. Transport Policy 45, 168–178. URL: https://linkinghub.elsevier.com/retrieve/pii/S0967070X15300627, doi:10.1016/j.tranpol.2015.10.004.
Reck, D.J., Axhausen, K.W., 2019. Subsidized ridesourcing for the first/last mile: How valuable for whom? , 18 p.URL: http://hdl.handle.net/20.500.11850/354730, doi:10.3929/ETHZ-B-000354730. artwork Size: 18 p. Medium: application/pdf Publisher: ETH Zurich.
Rifki, O., Chiabaut, N., Nicolas, J., 2021. Design of an autonomous ride-sharing service through a graph embedding integrated 2 dial-a-ride problem: Application to the last-mile transit in lyon city, in: TRB Annual Meeting, https://annualmeeting.mytrb.org/OnlineProgram/Details/15713.
Sadowsky, N., Nelson, E., 2017. The impact of ride-hailing services on public transportation use: A discontinuity regression analysis. Economics Department Working Paper Series URL: https://digitalcommons.bowdoin.edu/econpapers/13.
Salazar, M., Lanzetti, N., Rossi, F., Schiffer, M., Pavone, M., 2020. Intermodal autonomous mobility-on-demand. IEEE Transactions on Intelligent Transportation Systems 21, 3946–3960. URL: https://ieeexplore.ieee.org/document/8894439/, doi:10.1109/TITS.2019.2950720.
Salazar, M., Rossi, F., Schiffer, M., Onder, C.H., Pavone, M., 2018. On the interaction between autonomous mobility-on-demand and public transportation systems, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, https://ieeexplore.ieee.org/document/8569381/. pp. 2262–2269. doi:10.1109/ITSC.2018.8569381.
Scheltes, A., de Almeida Correia, G.H., 2017. Exploring the use of automated vehicles as last mile connection of train trips through an agent-based simulation model: An application to delft, netherlands. International Journal of Transportation Science and Technology 6, 28–41. doi:10.1016/j.ijtst.2017.05.004.
Sean Qian, Z., Michael Zhang, H., 2011. Modeling multi-modal morning commute in a one-to-one corridor network. Transportation Research Part C: Emerging Technologies 19, 254–269. doi:10.1016/j.trc.2010.05.012.
Shan, A., Hoang, N.H., An, K., Vu, H.L., 2021. A framework for railway transit network design with first-mile shared autonomous vehicles. Transportation Research Part C: Emerging Technologies 130, 103223. doi:10.1016/j.trc.2021.103223.
Sheffi, Y., Powell, W.B., 1982. An algorithm for the equilibrium assignment problem with random link times. Networks 12, 191–207. doi:10.1002/net.3230120209.
Shen, C.W., Quadrifoglio, L., 2013. Evaluating centralized versus decentralized zoning strategies for metropolitan ADA paratransit services. Journal of Transportation Engineering 139, 524–532. URL: https://ascelibrary.org/doi/10.1061/%28ASCE%29TE.1943-5436.0000529, doi:10.1061/(ASCE)TE.1943-5436.0000529.
Shen, W., Zhang, H., 2009. On the morning commute problem in a corridor network with multiple bottlenecks: Its system-optimal traffic flow patterns and the realizing tolling scheme. Transportation Research Part B: Methodological 43, 267–284. doi:10.1016/j.trb.2008.07.004.
Shen, Y., Zhang, H., Zhao, J., 2018. Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in singapore. Transportation Research Part A: Policy and Practice 113, 125–136. doi:10.1016/j.tra.2018.04.004.
Simoni, M.D., Kockelman, K.M., Gurumurthy, K.M., Bischoff, J., 2019. Congestion pricing in a world of self-driving vehicles: An analysis of different strategies in alternative future scenarios. Transportation Research Part C: Emerging Technologies 98, 167–185. URL: https://linkinghub.elsevier.com/retrieve/pii/S0968090X1830370X, doi:10.1016/j.trc.2018.11.002.
Tampere, C., Viti, F., Immers, L., 2010. New Developments in Transport Planning. Edward Elgar Publishing. URL: http://www.elgaronline.com/view/9781848449633.xml, doi:10.4337/9781781000809.
Tientrakool, P., Ho, Y., Maxemchuk, N.F., 2011. Highway capacity benefits from using vehicle-to-vehicle communication and sensors for collision avoidance, in: 2011 IEEE Vehicular Technology Conference (VTC Fall), IEEE, http://ieeexplore.ieee.org/document/6093130/. pp. 1–5. doi:10.1109/VETECF.2011.6093130.
Titos, G., Lyamani, H., Drinovec, L., Olmo, F., Močnik, G., Alados-Arboledas, L., 2015. Evaluation of the impact of transportation changes on air quality. Atmospheric Environment 114, 19–31. doi:10.1016/j.atmosenv.2015.05.027.

Urba Lyon, 2018. Pratiques de déplacements sur les bassins de vie de Scot de l'agglomération lyonnaise: enquête déplacements 2015. Technical Report. SYTRAL. https://www.scot-agglolyon.fr/les-documents-du-scot/.

Vickrey, W.S., 1969. Congestion theory and transport investment. The American Economic Review 59, 251–260.

Vuchic, V.R., 1969. Rapid transit interstation spacings for maximum number of passengers. Transportation Science 3, 214–232. doi:10.1287/trsc.3.3.214.

Vuchic, V.R., Newell, G.F., 1968. Rapid transit interstation spacings for minimum travel time. Transportation Science 2, 303–339. doi:10.1287/trsc.2.4.303.

Wang, J.Y., Yang, H., Lindsey, R., 2004. Locating and pricing park-and-ride facilities in a linear monocentric city with deterministic mode choice. Transportation Research Part B: Methodological 38, 709–731. doi:10.1016/j.trb.2003.10.002.

Wardrop, J.G., 1952. Road paper. some theoretical aspects of road traffic research 1, 325–362. doi:10.1680/ipeds.1952.11259.

Wei, K., Vaze, V., Jacquillat, A., 2022. Transit planning optimization under ride-hailing competition and traffic congestion. Transportation Science 56, 725–749. URL: http://pubsonline.informs.org/doi/10.1287/trsc.2021.1068, doi:10.1287/trsc.2021.1068.

Wilson, S., 2015. Urban Mobility System Upgrade: How shared self-driving cars could change city traffic. Corporate Partnership Board Report. ITF. https://www.itf-oecd.org/urban-mobility-system-upgrade-1.

Wirasinghe, S.C., Hurdle, V.F., Newell, G.F., 1977. Optimal parameters for a coordinated rail and bus transit system. Transportation Science 11, 359–374. doi:10.1287/trsc.11.4.359.

Wu, W., Huang, H., 2014. Equilibrium and modal split in a competitive highway/transit system under different road-use pricing strategies. Journal of Transport Economics and Policy (JTEP) 48.

Yao, F., Zhu, J., Yu, J., Chen, C., Chen, X.M., 2020. Hybrid operations of human driving vehicles and automated vehicles with data-driven agent-based simulation. Transportation Research Part D: Transport and Environment 86, 102469. URL: https://linkinghub.elsevier.com/retrieve/pii/S1361920920306568, doi:10.1016/j.trd.2020.102469.

Zardini, G., Lanzetti, N., Pavone, M., Frazzoli, E., 2022. Analysis and control of autonomous mobility-on-demand systems. Annual Review of Control, Robotics, and Autonomous Systems 5, 633–658. URL: https://www.annualreviews.org/doi/10.1146/annurev-control-042920-012811, doi:10.1146/annurev-control-042920-012811.

Zgraggen, J., Tsao, M., Salazar, M., Schiffer, M., Pavone, M., 2019. A model predictive control scheme for intermodal autonomous mobility-on-demand, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE. pp. 1953–1960. URL: https://ieeexplore.ieee.org/document/8917521/, doi:10.1109/ITSC.2019.8917521.

Zhou, Y., Li, Y., Hao, M., Yamamoto, T., 2019. A system of shared autonomous vehicles combined with park-and-ride in residential areas. Sustainability 11, 3113. doi:10.3390/su11113113.

Zhu, Z., Qin, X., Ke, J., Zheng, Z., Yang, H., 2020. Analysis of multi-modal commute behavior with feeding and competing ridesplitting services. Transportation Research Part A: Policy and Practice 132, 713–727. doi:10.1016/j.tra.2019.12.018.

Zhu, Z., Xu, A., He, Q.C., Yang, H., 2021. Competition between the transportation network company and the government with subsidies to public transit riders. Transportation Research Part E: Logistics and Transportation Review 152, 102426. doi:10.1016/j.tre.2021.102426.

Zwick, F., Kuehnel, N., Moeckel, R., Axhausen, K.W., 2021. Agent-based simulation of city-wide autonomous ride-pooling and the impact on traffic noise. Transportation Research Part D: Transport and Environment 90, 102673. doi:10.1016/j.trd.2020.102673.