UNIVERSITÀ DEGLI STUDI DI CATANIA
Dipartimento di Ingegneria Civile e Architettura

Giovanni Calabrò

ADAPTIVE TRANSIT DESIGN AND OPERATION VIA AGENT-BASED SIMULATIONS AND ANALYTICAL MODELS

Doctoral dissertation in

Evaluation and Mitigation of Urban and Territorial Risks

XXXIV cycle

Tutor:	Prof. Ing. Giuseppe Inturri
Co-tutors:	Prof. Ing. Matteo Ignaccolo
	Prof. Ing. Michela Le Pira
	Prof. Alessandro Pluchino
PhD coordinator:	Prof. Ing. Massimo Cuomo

Abstract

Urban mobility is experiencing an exceptional season of change, under the impulse of the disruptive innovations brought by information and communications technologies. Innovative *demand-responsive transport* services, allowing the intelligent matching between demand and supply and enabling travellers to request shared rides in real-time via mobile applications, are spreading in most urban areas. They are used as an alternative to public transport or as an access/egress leg to mass transit stations, i.e., acting as a feeder service. This thesis focuses on planning, design and operation, in a multimodal transport network, of innovative forms of *flexible transit*, a term denoting a group of shared mobility solutions able to combine a high level of shareability, typical of conventional public transport, with an adequate flexibility of route and schedule. A specific attention will be paid to the role of flexible transit in low-demand areas, where it is difficult to provide an effective and cost-efficient public transport, thus resulting in an extensive use of private vehicles. We propose different methodological approaches and tools in order to plan, design and simulate flexible transit services, exploring operating strategies, optimization algorithms and performance indicators. We are also interested in understanding which spatial and temporal demand characteristics are more favourable to flexible operations rather than fixed-route public transport in providing feeder services toward mass transit. Eventually, we will devise a layout of urban transit networks able to be *adaptive*, i.e., to vary its operating features and optimally deploying conventional and demand-responsive strategies according to the spatial and temporal variations of the demand. The research questions raised in this thesis are addressed focusing on the operational, tactical and strategic decision levels of transit problems and using different methodological approaches accordingly. In particular, agent-based modelling and simulation and analytical models based on continuous approximation techniques are employed. Different applications of the proposed models, based on real case studies or synthetic networks, are presented to analyse and support the design of flexible transit services and their integration with mass transit. The main findings of our work highlight that optimally employing fixed-route and flexible feeder operations can prevent the passenger accessibility to the transit system from degrading in low-demand areas and off-peak periods, thus providing a proper quality of service without burdening on operating costs.

Sommario

La mobilità urbana sta attraversando un'eccezionale stagione di cambiamento, sotto l'impulso delle innovazioni introdotte dalle tecnologie dell'informazione e della comunicazione. In molte aree urbane si stanno diffondendo servizi innovativi di mobilità *on-demand* di tipo flessibile, che consentono l'incontro intelligente tra domanda e offerta e permettono ai viaggiatori la prenotazione di viaggi condivisi tramite applicazioni mobili in tempo reale. Tali servizi sono utilizzati come alternativa al trasporto pubblico o come soluzione per coprire il primo e ultimo miglio di un viaggio, fungendo cioè da servizio di alimentazione (*feeder*) delle stazioni del trasporto collettivo. Il presente lavoro di tesi si concentra sulla pianificazione, progettazione e gestione, all'interno di una rete di trasporto multimodale, di forme innovative di trasporto collettivo flessibile (*flexible transit*), ovvero di un insieme di soluzioni di mobilità condivisa in grado di coniugare un elevato livello di condivisibilità, tipico del trasporto pubblico convenzionale, con un'adeguata flessibilità di percorso e orari. Un'attenzione specifica sarà dedicata al ruolo del *flexible transit* nelle aree a domanda debole, dove è difficile fornire un trasporto pubblico efficace ed economico, implicando così un ricorso estensivo alla mobilità privata. Diversi approcci metodologici e strumenti per la pianificazione, progettazione e simulazione di servizi di trasporto flessibile verranno proposti, esplorando diverse strategie operative, algoritmi di ottimizzazione e indicatori di performance. Il nostro interesse è inoltre volto a comprendere quali caratteristiche spaziali e temporali della domanda siano più favorevoli al trasporto flessibile piuttosto che al trasporto pubblico a percorso fisso nel fornire servizi *feeder* verso il trasporto di massa e, infine, a come progettare una rete di trasporto collettivo *adattiva*, in grado cioè di variare la sua caratteristiche operative e impiegare in modo ottimizzato servizi di trasporto fisso e flessibile in base alle variazioni della domanda nello spazio e nel tempo. Le domande alla base della ricerca verranno affrontate concentrandosi sui livelli decisionali operativo, tattico e strategico dei problemi progettuali riguardanti il trasporto condiviso e impiegando di conseguenza diversi approcci metodologici, nello specifico, modelli di simulazione ad agenti e modelli analitici basati su tecniche di *continuous approximation*. Diverse applicazioni dei modelli proposti, basate su casi di studio reali o reti ideali, verranno presentate con l'obiettivo di analizzare e supportare la progettazione di servizi di trasporto collettivo flessibile e la loro integrazione con il trasporto pubblico di massa.

I principali risultati del nostro lavoro evidenziano che l'utilizzo ottimale di trasporto fisso e flessibile è in grado di prevenire il degradarsi dell’accessibilità al sistema di trasporto pubblico, nelle aree a domanda debole e nei periodi non di punta, fornendo così un’adeguata qualità del servizio senza gravare eccessivamente sui costi operativi.

Table of contents

List of figures

List of tables

Glossary of terms

Here follows a glossary with the definitions of some recurring terms:

Adaptive transit	A multi-modal transit network which consists of a mass transit system and a feeder service provided by bus. Depending on the sub-region of the urban area and the period of the day, the system changes the feeder operation, between fixed-route and demand-responsive, in order to adapt to the spatial and temporal variation of the demand density (p. 142).
Agent-based modelling and simulation	A computer-based modelling technique where a system is modelled as a collection of autonomous decision-making entities called agents. Each agent individually assesses its situation and makes decisions on the basis of a set of rules, which determines its behaviours. Emergent phenomena result from the interactions of individual agents and complex dynamics (out of the reach of pure mathematical methods) can be explored (Bonabeau, 2002) (pp. 16, 35).
Continuous Approximation	It is an analytical modelling technique able to provide high-level guidelines for network planning, consisting in treating the transport demand and supply across a study area as a parametric environment, where input data and decision variables are density functions over time and space. The results from such models often bear closed-form analytical structures and, compared with discrete models, incur less computational burden, require less accurate input data, and can conveniently reveal managerial insights (Ansari et al., 2018) (p. 16).
Demand responsive transport	A transport service which is available to the general public, provided by low-capacity road vehicles such as small buses, vans or taxis, responding to changes in demand by either altering its route and/or its timetable and charging fares on a per-passenger and not a per-vehicle basis (Davison et al., 2014) (p. 6).
Feeder bus service	The use of an access mode to a rail rapid public-transport (Kuah and Perl, 1989), often by means of small public transport vehicles providing short-distance connections to feeder rail transit stations (Zhu et a., 2017) (p. 43).

Flexible transit	A group of shared mobility solutions combining a high level of shareability, typical of traditional public transport, with an adequate flexibility of route and schedule, since passengers are often asked to walk a short distance to access/egress the service. Flexible transit can be provided by either a private company or a public agency, or be intended for a specific group of people, e.g., disabled and elderly (p. 6).
Mobility-as-a-Service	A user-oriented integrator of transport services enabling searching, booking, and payment through a single digital platform (e.g., a single smartphone application), offering a tailored mobility package for door-to-door customized trip solutions (Jittrapirom et al., 2017; Le Pira et al., 2021) (pp. 7, 108, 191).
Shared Mobility	An innovative transportation strategy that enables users to have short-term access to a transportation mode (e.g., vehicle, bicycle, or other travel mode) on an as-needed basis. Shared mobility includes various services and transport modes that meet the diverse needs of travellers (Shaheen et al., 2017) (p. 4).

List of Abbreviations

ABM	Agent Based Model/Modelling
ACO	Ant Colony Optimization
AVs	Autonomous Vehicles
CA	Continuous Approximation
DRF	Demand Responsive Feeder service
DRT	Demand Responsive Transport
FBNDP	Feeder Bus Network Design Problem
FCFS	First-Come-First-Served
FMLM	First-Mile and Last-Mile
FRF	Fixed Route Feeder
GIS	Geographic Information System
ICT	Information and Communications Technology
MaaS	Mobility as a Service
MRT	Mass Rapid Transit
PT	Public Transport / Collective Transport / Transit
QoS	Quality of Service
TNDP	Transit Network Design Problem
TRNDP	Transit Route Network Design Problem
VoT	Value of Time
VRP	Vehicle Routing Problem

CHAPTER 1

Introduction

Sustainable mobility is one of the main challenges of our century. It implies the joint planning of land use and transport system in order to guarantee a better accessibility to a wide range of opportunities, e.g., workplace, education, healthcare, leisure, etc., and reduce the carbon footprint of human activities. Accessibility-oriented transport planning pays particular attention on active mobility, public transport and promotes multi-modal integration, i.e., combining of various transport modes throughout a trip from origin to destination. While remaining the most sustainable way to provide high transport capacity in dense urban settlements, conventional transit shows its inefficiency in low-demand areas, where mobility needs are spatially and temporally diverse. The recent technological innovations paved the way for the emerging of innovative forms of flexible, on-demand transport services, which are currently operated by private companies, often in competition with transit operators.

This thesis focuses on the first/last mile problem of passenger trips, identifying the provision of feeder bus routes for mass rapid transit as main solution. We intend to provide a methodological framework for the design, planning and operation of feeder services with various levels of flexibility, from fixed-route to on-demand door-to-door transit, with the ultimate objective of devising a demand-adaptive transit system, able to optimally integrate conventional and innovative shared transport modes in different urban contexts and time of the day. In this introductory chapter, we explain the motivation that underlie the thesis. Then, we establish the research objectives and questions and briefly outline the theoretical and methodological background. We further introduce our research approach, highlighting the main research contributions and some policy implications. Finally, we present the thesis outline.

1.1 Motivation of the research

Mobility has a significant impact on people's quality of life. The ever-increasing growth of urban areas, which would not have been possible without the impressive technological development of transport systems, has definitely widened the action radius of citizens' daily activities. If on one side this process increased the number of territorial opportunities, like the choice of residence or workplace locations, on the other side it exacerbated the need of fast, flexible and reliable transport modes.

Except from those forms of mobility meant as leisure and sport activities, transport would not exist alone, for the sake of movement. Instead, transport represents a derived demand, i.e., related to other human activities. People employ a significant part of their daily routine travelling between the places where they live, work, study, buy, or have social relations.

Transport is also a crucial element of the "sustainable development", as first recognized at the 1992 United Nation's Earth Summit. Recently, in 2015, the United Nations general Assembly set up 17 interlinked global goals, named Sustainable Development Goals (SDGs), which are included in "the 2030 agenda for sustainable development". Sustainable transport is mainstreamed across several goals and targets, especially those related to food security, health, energy, economic growth, infrastructure, cities and human settlements. Specifically, sustainable transport is the focus of the 11th SDG, target 2, which states: *"By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons".*

Although sustainable mobility is almost a 30-years-old concept, not enough has been done to reduce the carbon footprint of transport activities and the desirable modal shift towards active and shared mobility is still far from being achieved in most cities, although the recent diffusion of the Information and Communication Technologies (ICT) in transport applications offers a chance to boost the emergence of innovative, accessible and flexible forms of shared mobility.

1.1.1 The risks of the car-oriented development

In the last decades, most cities have witnessed the uncontrolled expansion of low-density, single-use suburban development (La Greca et al., 2011), which is, almost always, characterized by a low level of accessibility. With reference to the passenger transport, Geurs and Van Wee (2004) defined accessibility as "the extent to which land use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)". Biazzo et al. (2019) described accessibility as "the capacity of cities to allow people to move efficiently by guaranteeing equity and equal access to personal and professional opportunities".

The car ownership is a factor that radically changed the shape and size of modern cities: thanks to its characteristics of speed and flexibility, the car has favoured the increase of the average distance of the daily trips and the consequent decline of the residential density. So, the ability to make long journeys has been increasingly becoming a necessary condition for accessing the territorial opportunities. Since low-density areas are located far away from the city core and often poorly served by public transport (PT), they became over the years strongly dependent on motorized private transport, increasing the congestion of the city centre as well as the time and monetary cost of transport activities, and thus contributing to the social, economic and environmental unsustainability of urban mobility (Newman and Kenworthy, 2006; Banister, 2008, Ignaccolo et al., 2016).

The modal imbalance towards the use of private vehicles in urban areas involves a series of costs and externalities for the community, including:

- The reduction in safety conditions of soft mobility: as reported by the World Health Organization (2018), road traffic injuries are one of the ten leading causes of death and at least 30% of road crashes occurs in urban areas. Pedestrians, cyclists and motorcyclists should be regarded as "vulnerable" road users, and their number is rising also thanks to the recent worldwide diffusion of the electric micro-mobility.
- The increased levels of atmospheric and noise pollution: transport is the major contributor to local and global pollution of air, soil and water, accounting for almost a quarter of energy-related global greenhouse gas emissions (IEA, 2018), undermining the efforts to meet the global challenge of climate change.
- The uneven access to mobility services and territorial opportunities: a combination of land use and transport planning firstly directed at the motorized private mobility causes an unequal accessibility to public and private services and instances of social exclusion (van Wee and Geurs, 2004).
- The congestion of road networks: road congestion is detrimental for travellers to the extent that it increases the time and money cost of accessing destinations (Levine and Garb, 2002).

- The inefficiency of PT: urban sprawl and car-oriented development gave rise to low-density areas, where the need of private vehicles becomes endemic, and PT is not able to ensure at the same time an extensive coverage and an adequate quality of service (QoS).

Various risks are closely related to these costs for the society, first of all: (i) the road crashes and injuries in urban areas; (ii) the premature deaths attributable to air pollution; (iii) the negative contribution to the climate change and (iv) the rise of new forms of social exclusion. This last is defined by Van Wee and Geurs (2011) as the fact that some people or population groups are excluded from a certain minimum level of participation in location-based activities, in which they wish to participate.

Under this respect, PT should aim at reducing the mobility gap experienced by several people to have a good accessibility to all the opportunities and services throughout a metropolitan area, e.g.: workplaces, education and healthcare services, leisure activities. From this perspective, by offering a mobility potential to all categories of users, PT stands as a determining factor for social inclusion policies (Lucas, 2006; Giuffrida et al., 2017).

1.1.2 Shared mobility and sharing economy

Urban mobility is facing a remarkable season of change. The sharing economy paradigm applied to transport services allows a shift from a travel behaviour based on private vehicle ownership to a new one based on sharing services and assets on as-needed basis (Ambrosino et al., 2016). Such alteration is sustained by the disruptive innovations brought by ICT, which enable new flexible, on demand transport services spreading as complementary to conventional public transport or in substitution to it (Cohen-Blankshtain and Rotem-Mindali, 2016).

Examples like carsharing, ridesourcing, bikesharing, electric micromobility, microtransit show how these concepts are emerging in the ecosystem of urban mobility and rapidly sprawling worldwide (Shaheen and Chan, 2016; Machado et al., 2018). Not only passenger transport but also logistics processes and related freight transport flows are rapidly changing, due to progress in information technology and unparalleled growth of consumer involvement in supply chains (Tavasszy, 2019; Marcucci et al., 2017a).

Figure 1.1 outlines the innovative transport services for urban mobility, according to their level of flexibility, sustainability and shareability. Referring to shared mobility, a first classification can be made on the object of sharing. Therefore, we distinguish those services enabling the sharing of a vehicle and those allowing for the sharing of passenger rides.

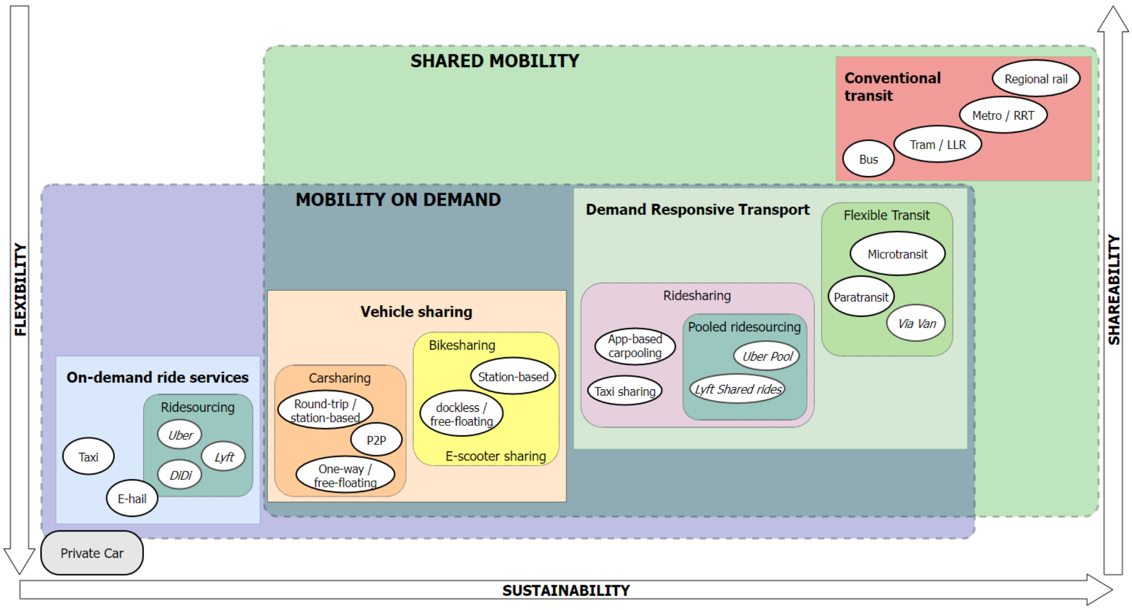


Figure 1.1. Classification of mobility on demand and shared mobility services (adapted from Inturri et al., 2019)

Carsharing, bikesharing and, more recently, e-scooter sharing, with their different operation policies, belong to the first category. These services offer the flexibility of the individual trip planning and vehicle usage without the burden of the ownership (e.g., fuel, maintenance, insurance), since users pay a usage- and/or membership-based fee (Shaheen and Chan, 2016). E-scooter and bike sharing have proven to be effective in allowing a modal shift towards PT (Ma et al., 2020), by providing a first/last mil access to PT lines.

The sharing of a passenger ride is a more recent concept in the shared mobility framework. In fact, shared transport services accommodating users with similar trip origin and destination exist since the late 1960s and include carpooling and vanpooling (Chan and Shaheen, 2012). However, the new on-demand ride services are made possible by the spreading of ICT, which dynamically matches supply and demand and allow travellers to request rides in real-time via mobile applications, nowadays accessible to the majority of individuals. Ridesourcing, i.e., transport services connecting community drivers with passengers via mobile applications (Shaheen and Chan, 2016), has been growing rapidly: Lyft (North America), Didi (China), Uber (worldwide) are the most famous examples of how smartphone apps are able to change the urban transport ecosystem. Ridesourcing consists in real-time coupling potential passengers with drivers via ad hoc dispatching algorithm, under a dynamic pricing model and has often outperformed the most traditional form of on-demand individual mobility, like taxis (which also have adopted "e-hail" apps to book a taxi) in an uncertain regulatory and policy climate (Shaheen and Chan, 2016). According to Sadowsky and Nelson (2017), ridesourcing services initially tended to complement PT coverage, offering themselves as last mile solution in low density urban areas and during night hours, where and when PT operations are reduced. However, the increasing competition between ridesourcing companies for the market share, the growth in vehicle supply and thus the fares' reduction caused, in the long term, a detrimental effect for PT.

The subsequent step towards shared mobility is represented by the pooled ridesourcing services (Uber pool, Lyft shared ride, etc.) which offer a cheaper option thanks to accommodating different trip requests with similar origin and destination.

Finally, with the term "flexible transit" we denote a group of shared mobility solutions combining a high level of shareability, typical of traditional PT, with an adequate flexibility of route and schedule, since passengers are often asked to walk a short distance to access/egress the service. Flexible transit can be provided by either a private company (e.g., Via Van) or a public agency, or be intended for a specific group of people (e.g., paratransit services for disabled and elderly).

For the scope of this thesis, we group together ridesharing and flexible transit under the definition of Demand Responsive Transport (DRT) given by Davison et al. (2014), i.e., a transport service which is available to the general public, provided by low capacity road vehicles such as small buses, vans or taxis, responding to changes in demand by either altering its route and/or its timetable and charging the fare on a per-passenger and not a per-vehicle basis.

It is not difficult to imagine a near future in which traditional transit modes will be replaced by DRT services also in large cities, also considering the disruptive potential of employing fully automated vehicles, but it has been shown (Basu et al., 2018) that only the mass transit can effectively provide high transport capacity by accommodating large numbers of commuters, while reducing the congestion on the urban street network. Therefore, policymakers and transit agencies should benefit from the sharing economy paradigm and devise the optimal operation range of DRT services, i.e., where and when they should be deployed to increase the modal share of PT, prevent the dependence on private car and promote sustainable mobility.

1.1.3 Research context

Shared mobility in general and DRT in particular can find diverse potential applications according to the spatial context and the specific constraints of the built environment (Shaheen et al., 2017). Although literature and real applications typically devote most of the attention on ridesharing services in dense urban contexts, we believe that DRT services have a great potential also (and especially) in low-density regions, whether they are suburban areas connected with a mass transit network, or small cities during off-peak hours, or even rural areas (which we do not consider in this work) where PT services would be unfeasible, to offer an alternative to the exclusive use of private vehicles.

In Figure 1.2 we propose a simplified three-dimensional scheme (the "DRT operational cube") representing the DRT operating context, according to:

- The type of urban context. We distinguish the city centre, defined by the highest concentration of jobs (central business districts) and services, from the periphery, characterized by a low-density residential development and a commuter demand mostly directed to/from the inner urban area.
- The time of the day. DRT operations should be managed in a different way when dealing with peak demand rather than serving off-peak demand.
- The size of the city. Large urban areas usually have a MRT network acting as backbone of the transit system, while in small cities a bus network including few strategic lines is often provided.

To pursue a decisive modal shift towards shared mobility, future cities cannot rely only on traditional forms of PT. Instead, the technological innovations are favouring the transition from a culture where consumers own assets, i.e., vehicles, toward the Mobility-as-a-Service (MaaS) culture (Jittrapirom et al., 2017; Le Pira et al., 2021), where consumers share access to assets.

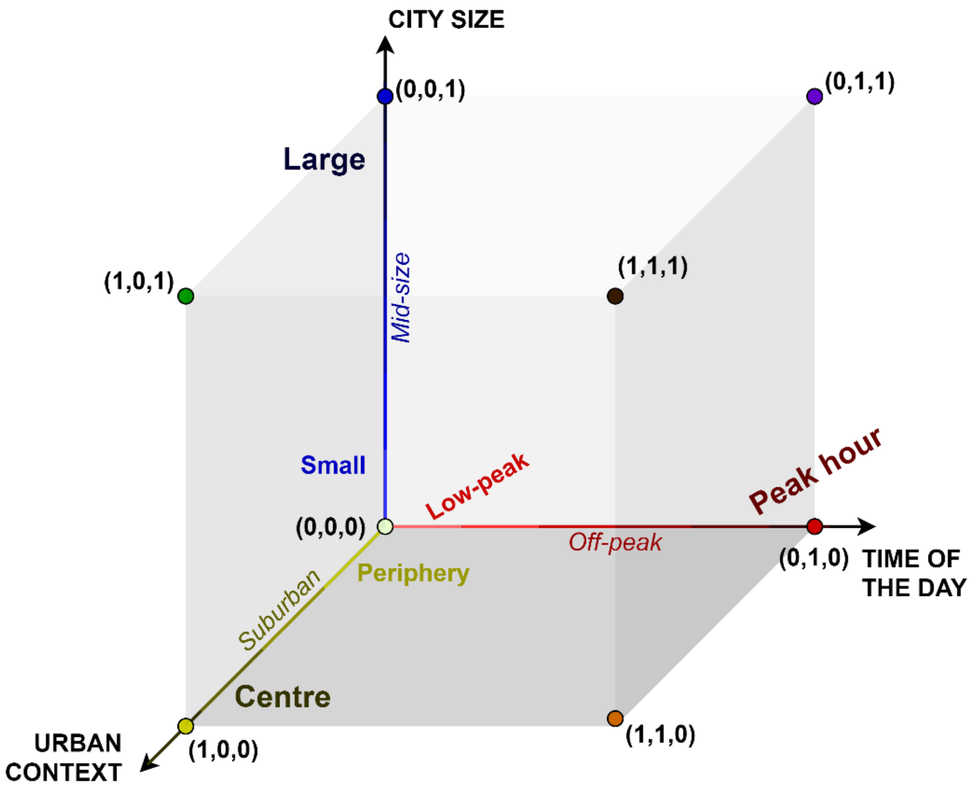
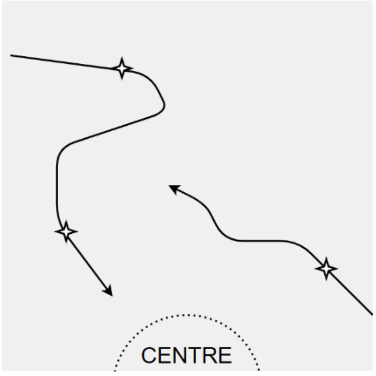
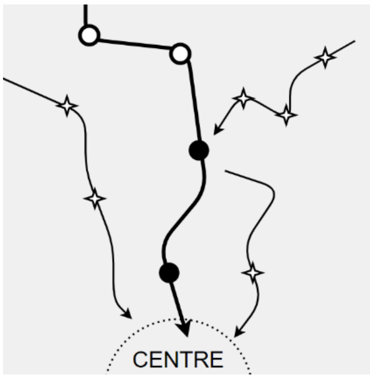
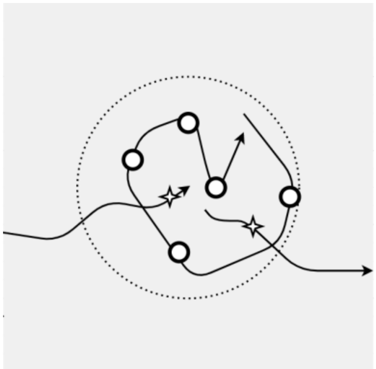
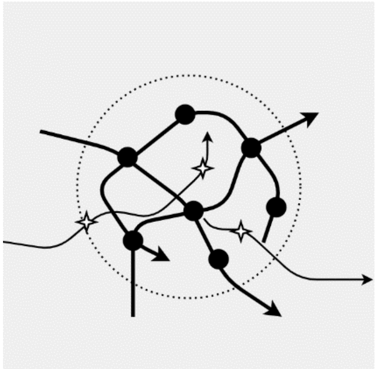
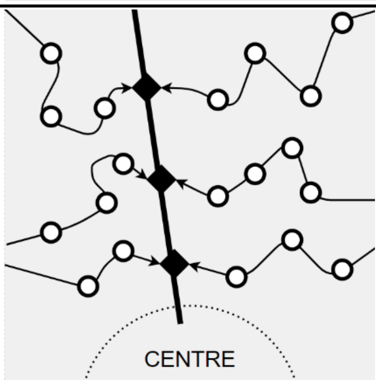
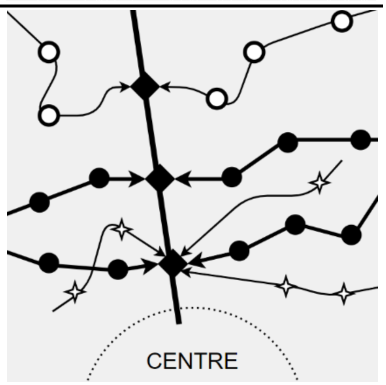
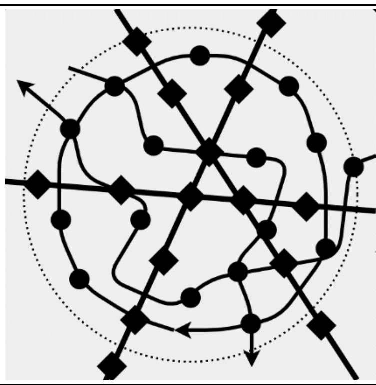
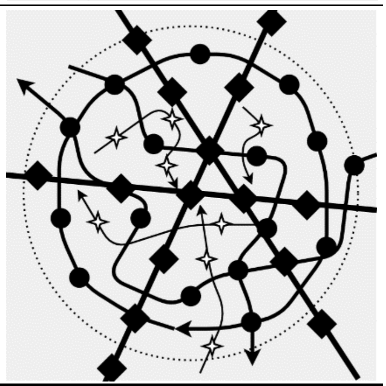


Figure 1.2. The DRT operational cube, considering the urban context, the time of the day and the size of the city.

Table 1.1 identifies some possible solutions of how DRT services can properly complement or substitute PT according to the eight different (simplified) combinations depicted Figure 1.2. We highlight that these schemes are purely qualitative and do not account for a wide set of intermediate situations. However, they are intuitive insights into the integration between conventional and innovative forms of shared mobility.

First let us consider the case of a small city. During peak hours, some bus routes with a relatively high ridership can be provided. They can ensure a good coverage of the city centre and the denser peripheral areas, where the uncovered zones can be served by ridesharing services by means of low-capacity vehicles. Also, the bus operations could adopt a flexible routing strategy in certain peripheral areas. During off-peak hours, instead, providing conventional PT would be too expensive, due to low ridership, and it could be substituted by RS. Besides, in the city centre, a flexible transit (via minibuses or vans) could be provided. To summarize, in small cities DRT can act as (i) an “independent” service with respect to PT, (ii) a substitute of conventional (fixed) transit especially in off-peak and low-peak hours, and (iii) a complement of the fixed-route transit network, to connect those areas uncovered by PT with transfer stations.

Table 1.1. Possible integration between Mass Rapid Transit (MRT), Conventional bus, Flexible Transit (FT) and Ridesharing (RS), according to different spatial and temporal contexts.

LEGEND ◆ MRT ○ FT ● BUS ✧ RS		Off-peak	Peak hour
Small city	Periphery		
	Centre		
Large city	Periphery		
	Centre		

Large urban areas often count on a MRT system and an extensive bus network. Currently, ridesharing providers are almost always competitors of PT. However, with the future improvement of MaaS platforms the two mobility options would be integrated in multi-modal trip chains and DRT could play a different role according to the spatial and temporal context. In peak hours, fixed transit modes operate with high frequencies and wide coverage. Suburban areas are connected to MRT via feeder bus routes, but in the distant periphery the amount of demand could justify the use of flexible transit policies. Ridesharing services mostly complement MRT, serving shorter trips and allowing for a good shareability. During off-peak hours, in the city centre the PT structure does not change, but a less frequent service is provided, while suburban and peripheral areas generate demand levels low enough to favour flexible transit operations feeding the MRT system.

The arguments discussed so far allow us to state that a different approach in planning transit systems, considering the integration between on-demand and conventional transport services, would make it possible for transit agencies to save capital cost and gain in ridership. This concept is qualitatively outlined in Figure 1.3. Transport demand for PT services usually exhibit two daily peaks following commuting patterns. The fleet for transit operations in urban areas, i.e., the number of trains, buses, etc. and the vehicle size, is determined by the characteristics of demand during the peak hour, while in off-peak periods, where the demand is quite lower, frequencies are reduced consequently (Jara-Díaz et al., 2017). However, when DRT services are optimally integrated with the transit system, the PT agency could avoid oversizing the fleet of conventional buses and reduce the number of bus lines in the periphery (see Section 7.5.3), since a share of trip requests is served by DRT both for first/last mile connections (complement) and shared trips between underserved areas (substitute). Besides, the introduction of DRT services would be beneficial also in terms of the new latent demand induced by the increased attractiveness of the transit system, mostly consisting in trips which were made by private car due to the poor QoS of conventional transit. In off-peak hours, DRT has mainly a complementary role, acting as feeder of MRT. During the evening and night hours, DRT services are the most efficient way to serve the low and sparse demand which otherwise cannot be satisfied by conventional transit.

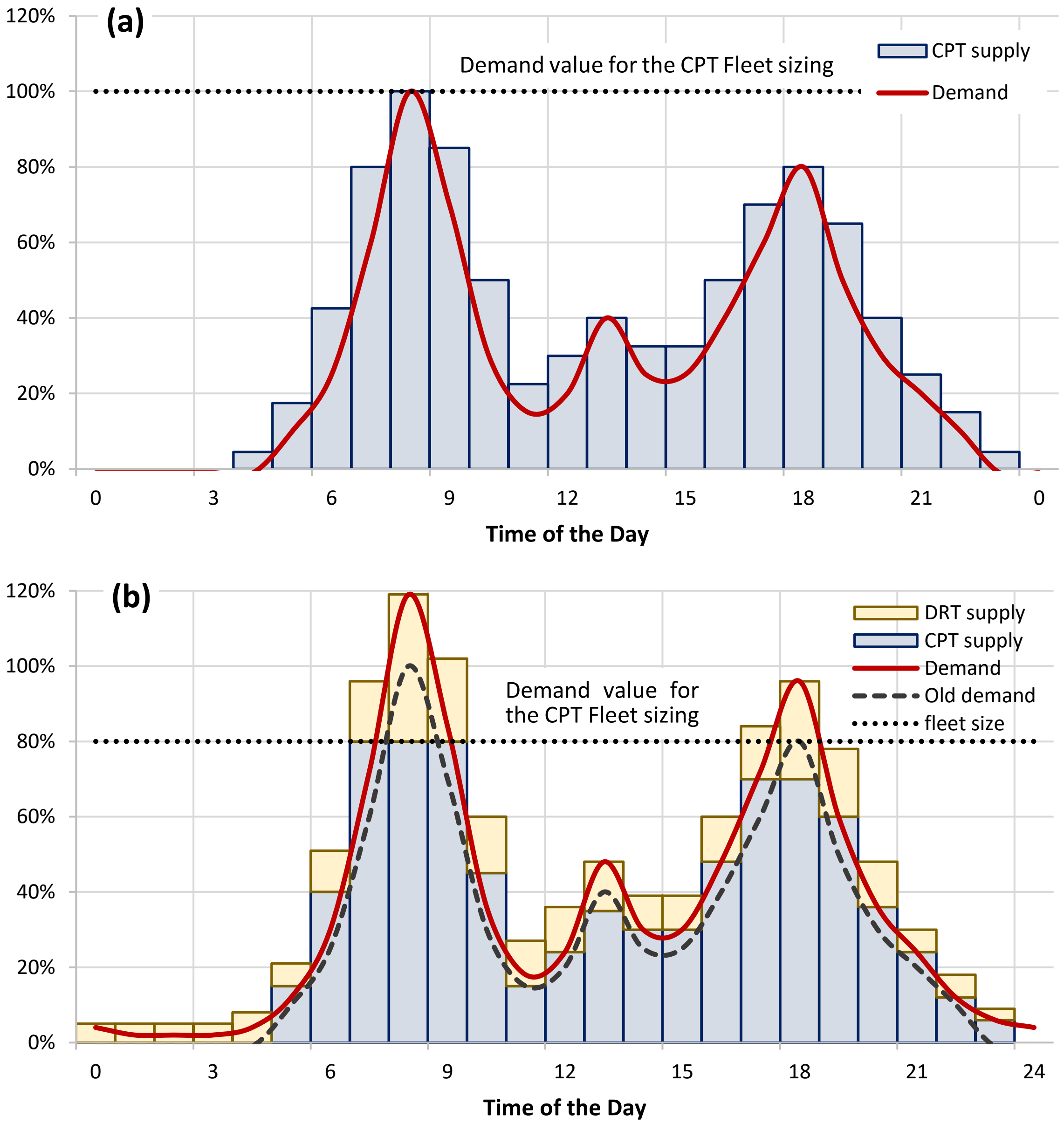


Figure 1.3. Transit demand satisfaction when (a) only conventional public transport (CPT) is provided and (b) demand-responsive transport (DRT) services complement CPT (own setup).

1.2 Research scope

This thesis focuses on planning, design and operation of innovative forms of flexible transit in multimodal transport networks. A specific attention will be devoted to the role of flexible transit in low-demand areas (also called “weak demand areas”), i.e., those urban and suburban settlements characterized by a significant dispersion of the mobility demand. The attributes “low” and “weak” do not refer only to the number of trips generated or attracted by the zone itself, but also to the variability of this number over the time of the day (or day of the week, etc.) and the heterogeneity of the demand density across the space.

We believe that low-demand areas would benefit from a hierarchical transit network model having the mass rapid transit (MRT) as “backbone” of the travellers’ trip chain and bus feeder services to cover the first and last mile leg of transport (“feeder-trunk” scheme, depicted in Figure 1.4). The latter can consist of fixed-route microtransit, performed by buses with a schedule known in advance, or demand-responsive flexible services, which often deploy smaller vehicles (minibuses, vans, etc.) and offer a tailored alternative to private mobility. Multimodal integration along a trunk-feeder transit corridor is crucial to expand the service coverage of PT without exceeding in operational cost (e.g., vehicle-kilometres travelled) and still guaranteeing an accessibility level competitive with that of the private automobile.

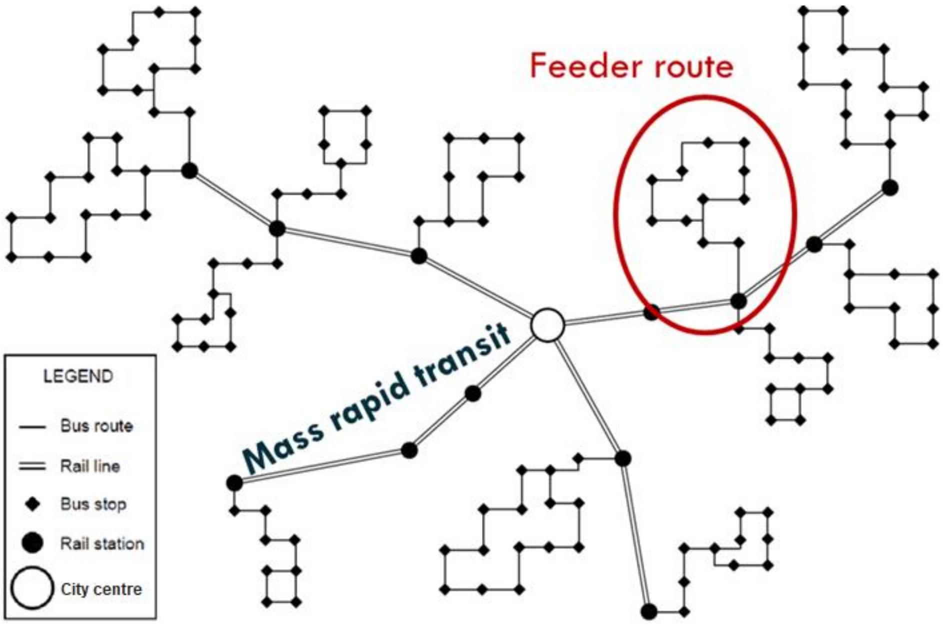


Figure 1.4. Representation of a feeder-trunk transit scheme

1.2.1 Research gaps

The present thesis aims at addressing some research gaps (*G*) related to the design of flexible transit services. Low-demand areas usually lack of an effective and cost-efficient PT service, implying an exclusive use of private vehicles, though restricted to who can afford it. Conventional PT is well suited to serve high-density and high-demand mobility corridors; nevertheless, fixed-route and scheduled transit services can result in low quality of travel experience and scarce load factors where the demand is sparse and highly variable, since it is particularly challenging to ensure both good coverage and high ridership (Walker, 2012). In order to keep a good QoS, it would be necessary for PT agencies to operate a huge number of lines without penalizing too much the service frequency, which would result in an excessive agency cost, with respect to the few passengers served per-vehicle.

Besides, the fast development and diffusion of ICT do not make difficult to transform PT by making it “adaptive”, i.e., able to modify its operational features (fleet size and composition, routing strategies, service areas, etc.) to the spatial and temporal variations of transport demand, and thus bridging a performance gap between expensive (but flexible) on-demand ridesharing and cheap (but fixed) conventional bus transport. Flexible transit services should take advantage of ICT allowing passengers to book and pay for a ride and get real-time travel information and transit operators to manage the fleet, update the route and communicate with the users or the other vehicles.

In medium and large cities, where a MRT system operates, an effective planning and design of feeder routes (with variable level of flexibility) can be identified as a solution to the first and last mile transport problem related to suburban low-density zones. Since Kuah and Perl (1989) defined the feeder bus network design problem, many authors have proposed analytical (Chang and Schonfeld, 1991) and heuristic (Wang, 2019) models aimed at supporting the design of feeder services, either in the network perspective (Mohaymany and Gholami, 2010; Kuan et al., 2006; Almasi et al., 2018) or in the optimal routing (Xiong et al., 2013) and scheduling (Lu et al., 2015) of such services. However, none of the previous studies adopts accessibility measures obtained, e.g., from socio-demographic and mobility datasets (*G1*).

Moreover, a gap exists between modelling sophistication, which often consists in abstract, complex approaches difficult to understand by non-experts (Park et al., 2019), and the practicality of urban transport planning, which should benefit from handy and flexible tools allowing for a detailed spatial representation of transport network, demand data and geographical constraints (*G2*).

Let us focus on the opportunity to provide flexible transit services for the first/last mile coverage. There is a broad literature on how to design flexible transit systems (Koffman, 2004, Cordeau and Laporte, 2007), from the almost outdated form of dial-a-ride, where the reservation for the service should take place in advance via call centres, to the new challenges brought by real time internet-based ridesharing services (Ho et al., 2018). Nonetheless, a small percentage of this research field is devoted to flexible feeder services. Among them, many authors have proposed (a) efficient heuristics, to address the related dynamic routing and scheduling problems (Pan et al., 2015; Li et al., 2018), and (b) analytical models (Quadrifoglio and Li, 2009; Chandra et al., 2013; Huang et al., 2020) to compare the performance of fixed-route (FR) vs. demand responsive (DR) feeder options. However, the former do not allow for a comparison between FR and DR policies under different demand levels or fleet configurations, while the latter do not explore intermediate flexible strategies between a purely fixed and a purely door-to-door service, also due to the simplifications related to analytical modelling (*G3*).

Besides, most of the work on flexible transit focuses on vehicle dispatching and routing strategies to meet users' requests exactly where they originate, i.e., without considering a walking time for passengers, thus implying longer detours to satisfy the demand. Only a few papers face the design of ridesharing services in which users are requested online to walk towards access/egress their stop locations (Li et al., 2018; Fielbaum et al., 2021). However, the aforementioned works consider a "stand-alone" system involving small vehicles (i.e., low shareability) without any modal integration with other transit modes (*G4*).

Eventually, much research has been devoted to transit network design problem (TNDP) (Guihaire and Hao, 2008; Farahani et al., 2013; Ibarra-Rojas et al., 2015). This problem can be modelled as a discrete optimization problem (Baaj and Mahmassani, 1995; Zhang et al., 2014) to determine the set of transit lines given an OD matrix and then predicted travel times on the network and general performance indicators of the system. An alternative to this approach is to represent the TNDP through analytical models where passenger demand (Ouyang et al., 2014) and transit supply (Chen et al., 2015) are represented via a continuous approximation (CA) approach, i.e., as function over a geographical space. These CA models deal with fixed PT (Daganzo, 2010; Badia et al., 2014), flexible transit (Nourbakhsh and Ouyang, 2012; Shi and Gao, 2020), or the integration of both (Chen and Nie, 2017; Luo and Nie, 2019). However, none of the previous works has considered how to vary (spatially and temporally) the layout of the whole transit network (composed of FR and DR transit modes) across the entire urban area, thus making the network capable of optimally accommodating flexible transit modes in a given time of day and in certain zones of the urban area (*G5*).

1.2.2 Research questions

The research questions (*Q*) underlying the thesis work arise from the gaps we identified in the literature on flexible transit, as explained in pre previous sub-section. The research questions, which are listed below, will be addressed in different chapters of the thesis, as follows:

Q1: How can sprawled urban and suburban areas be effectively served by public transport, thus contrasting the massive use of private vehicles? (Addressing *G1* and *G2*; Chapters 3 and 5).

Q2: What can be a useful, practical and flexible tool to implement, reproduce and analyse the performance of on-demand transit services with different levels of flexibility? (Addressing *G1, G2* and *G3*; Chapters 4, 5 and 6).

Q3: How flexible and demand responsive feeder transport services can effectively match the demand with the supply in real-time, trying to maximize shareability, minimize operator cost and contain passenger travel time? (Addressing *G3* and *G4*; Chapter 5 and 6).

Q4: When is it more convenient to adopt a fixed-route policy rather than a flexible one in providing feeder services toward mass rapid transit? (Addressing *G3* and *G4*; Chapters 6 and 7).

Q5: Where, when and how the layout of the urban transit network should adapt to the spatial and temporal demand variations? (Addressing *G5*; Chapter 7).

The research questions we mentioned above are the driving force behind the objectives and goals we set in present thesis, as discussed in the next sub-section.

1.2.3 Research goals

The scope of our research comes within the overarching objective of promoting the use of public transport and the multimodal integration, with particular reference to low-demand areas. The specific goals this thesis aims to pursue are the following:

1. To develop a practical, customizable but rigorous tool able to guide the design of feeder bus lines in low-density and suburban contexts.

2. To show the suitability of *ad-hoc* agent-based simulation environments to reproduce and compare the performance of different operating strategies for flexible transit, involving a range of service flexibility based on the demand variations.

3. To optimize the operation of flexible transit services integrated with mass rapid transit, in a multimodal feeder-trunk scheme. Such services are aimed

at dynamically matching the transport supply (vehicles and facilities) to the demand (users' trip requests).

4. To propose a methodology for the design of "adaptive" transit networks for large urban areas, i.e., a transit system deploying demand responsive modes to complement fixed transit according to spatial characteristics and the time of day, therefore with the aim of providing high transport capacity, while at the same time ensuring high levels of service even in peri-urban and low demand areas. Two objectives that would otherwise be in conflict when recurring only to conventional transit.

1.2.4 Methodological approach

In order to pursue the aforementioned research goals, two different methodological approaches have been chosen:

<u>Agent-Based Modelling</u> (ABM). Several modelling techniques have been used to design and reproduce the main characteristics of transport systems. ABM is a computer-based computational method enabling researchers to simulate the behaviour of a complex system (such as transport systems), involving numerous autonomous agents governed by given decision rules and interacting with the environment. ABM adopts a "bottom-up" approach: the agents (bottom level) are able to interact with the other agents, having a partial, localized knowledge of the surrounding environment and possibly learn and adapt their behaviour over time. By simulating such agents, it is possible to understand the comprehensive patterns emerging from the system (upper level) and evaluate the resulting outcomes. In transit-related models, the major advantage of ABM over mathematical and also microsimulation approaches is the ability to provide insight into the operation of a system (Ronald et al., 2015), e.g., taking passenger preferences or transit operator strategies into account.

<u>Continuous Approximation</u> (CA). In Section 1.2.1 we mentioned the distinction between discrete and continuous approaches to address the TNDP. By the use of CA we model the transport demand and supply across a study area as a parametric environment, where input data and decision variables are density functions over time and space. The key idea of this modelling technique is to provide high-level guidelines for network planning, especially for large-scale practical problems (Daganzo, 2010). The results from such models often bear closed-form analytical

structures and, compared with their discrete counterparts, incur less computational burden, require less accurate input data, and can conveniently reveal managerial insights (Ansari et al., 2018).

1.3 Thesis outline

This section presents an overview of the thesis structure. Chapter 2 examines the scientific background and the relevant literature on flexible transit and transit network design problems, with a focus on the methodological approaches. Chapters 3-7 form the core of this thesis: each of them addresses a problem related to a specific decision level in transit planning and design (see Farahani et al., 2013), i.e.:

1. Strategic. This level involves long-term decisions, mostly related to the transit infrastructure.
2. Tactical. This level includes those mid-term decisions regarding the use of infrastructures and facilities of the existing transit network, including bus routes and reserved lane allocation, transit service frequency, rolling stock type, etc.
3. Operational. This level encompasses only short-term decisions, which are mostly related to transit fleet control, demand management or scheduling problems.

According to the decision level addressed in each chapter and in accordance with the methodology used, the core of the thesis can be organised in two main parts, as shown in Figure 1.5:

I. The first part (Chapters 3-6) addresses the design (tactical level) and simulation (operational level) of flexible feeder transit services. The methodology is mainly based on the agent-based modelling (ABM) technique. Also, optimization algorithms are incorporated into the ABMs.

II. The second part (Chapter 7) mainly addresses the strategic decision level, by proposing a solution model of the adaptive transit network design problem. The methodological approach is analytical and based on the continuous approximation technique.

The five chapters constituting the thesis core are based on own research papers (already published or currently under review).

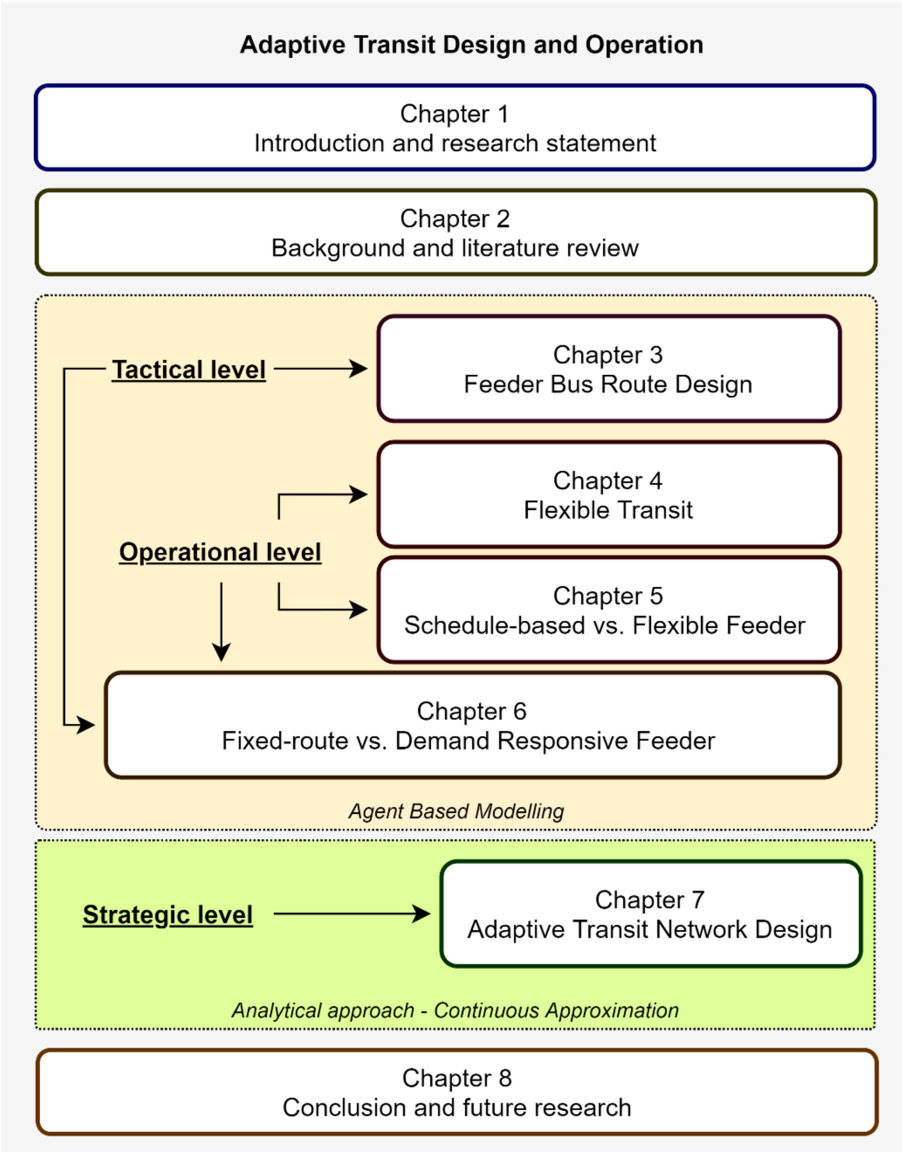


Figure 1.5. Overview of the thesis structure

Chapter 3 focus on the design of feeder bus routes conceived for covering the commuting demand of suburban areas and enhancing the ridership of mass transit system. We implemented an Ant Colony Optimization (ACO) algorithm into an ABM programming environment in order to find the optimal set of routes connecting the service area to the transfer stations and constrained by a maximum cycle time. The potential demand at a feeder bus stop is estimated according to accessibility indicators, using GIS-based demographic data. The model was applied to case studies in the city of Catania (Italy). This chapter is based on the following publication:

Calabrò, G., Inturri, G., Le Pira, M., Pluchino, A., and Ignaccolo, M. "Bridging the gap between weak-demand areas and public transport using an ant-colony simulation-based optimization". Transportation Research Procedia, vol. 45, pp. 234-241, 2020.

Chapter 4 presents an ABM for the simulation of a flexible transit service called MVMANT, experimented in the city of Dubai (United Arab Emirates) in 2019. The study evaluates the best vehicle dispatching and route choice strategies, also varying the capacity if the operating vehicles, from the points of view of both the operator and the community. This chapter is based on the following publication:

Giuffrida, N., Le Pira, M., Inturri, G., Ignaccolo, M., Calabrò, G., Cuius, B., ... and Pluchino, A. "On-demand flexible transit in fast-growing cities: The case of Dubai". Sustainability, vol. 12, no. 11, p. 4455, 2020.

In Chapter 5 the comparison between two alternative feeder transit services is conducted, i.e., fixed-route-and-schedule versus on-demand flexible bus. The objective is to identify the optimal demand ranges that justify the adoption of flexible rather than the fixed operating strategy. For this purpose, a new ABM integrated with GIS data is developed and tested on the real case study of a feeder service for a metro station located in Catania. The two operational strategies are simulated under various demand and supply conditions, with the aim of establishing a trade-off between passengers, operator and community needs. This chapter is based on the following publication:

Calabrò, G., Correia, G., Giuffrida, N., Ignaccolo, M., Inturri, G., and Le Pira, M. "Comparing the performance of demand responsive and schedule-based feeder services of mass rapid transit: an agent-based simulation approach". In 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS), pp. 280-285, IEEE, 2020.

Chapter 6 extends the scope of the previous one, by proposing a novel theoretical ABM which reproduces and allows to compare different operating policies of feeder services in a synthetic simulation environment. The conditions (demand patterns, shape of the service area, number and capacity of vehicles) making a demand-responsive feeder more attractive than a fixed-route feeder are explored, in order to find the critical demand density allowing the switch between the two operating strategies. This chapter is based on the following paper:

Calabrò, G., Le Pira, M., Giuffrida, N., Inturri, G., Ignaccolo, M., and Correia, G. "Fixed route vs. demand-responsive transport feeder services: an exploratory study using an agent-based model". Submitted to Journal of Advanced Transportation (under review).

In Chapter 7 an analytical method for the adaptive design of transit systems is presented. The objective is to propose a new planning tool to approach the design of a transit network which gets the best from fixed-route and demand-responsive transit, combining both, depending on the demand observed in each sub-region of the urban conurbation and time-of-day. The goal is to provide high transportation capacity while guaranteeing high QoS, two objectives that are instead conflicting if only classic fixed-schedule transportation is deployed, as in today's cities. This chapter is based on the following paper:

Calabrò, G., Araldo, A., Oh, S., Seshadri, R., Inturri, G., Ben-Akiva, M. "Adaptive transit design: optimizing fixed and demand responsive multi-modal transport for metropolitan areas" (to be submitted). Presented at the 100th Annual Meeting of the Transportation Research Board (TRB 2021).

Finally, we draw the overall conclusions of the thesis in Chapter 8, where we also discuss their practical implications. At the end of this last chapter, we also formulate recommendations for future research.

CHAPTER 2

Background

In this chapter, a conceptual, theoretical and methodological foundation of the thesis work is provided.

Section 2.1 examines the evolution of demand-responsive transport and flexible transit concepts, including their specific characteristics, operational features, and performance indicators. The development of new technologies has been changing the planning and modelling approach adopted by researchers, transit agencies and transport providers towards the design and operation of on-demand flexible services. From the first operating experiences of dial-a-ride services, where the request for a trip was supposed to be notified well in advance, to the current possibility of booking a shared ride in real time using a smartphone application, several routing and scheduling strategies as well as the vehicle dispatching algorithms to optimize DRT performance have been proposed. Moreover, in view of a future where shared mobility will be driven by the diffusion of autonomous vehicles (AVs), some relevant studies on the impact of autonomous mobility on-demand on urban transport systems are analysed in this chapter.

Section 2.2 reports a literature review of transit networks design problems, analysing the objectives, input parameters, decision variables, and solution methods characterizing them. Then, a specific attention to the design of feeder bus routes for first/last mile connection to mass transit is paid, focusing on different heuristic procedures employed to propose solutions to the problem.

This chapter also provides an overview on agent-based modelling and continuous approximation techniques to address the simulation of transport systems on various scales and the design of transit networks, respectively.

2.1 The evolution of Flexible transit

The popularity of Demand Responsive Transport (DRT) services has been growing since the end of the last century. Nowadays, the range of different transport services which can be grouped in the DRT family is considerably wider than twenty years ago.

An early definition of DRT is reported by Brake et al. (2004): "an intermediate form of public transport, somewhere between a regular service route that uses small low floor buses and variably routed, highly personalised transport services offered by taxis. Services are routed according to the needs of the customers, generally only stopping where passengers request collection or dropping off". One can note how DRT was originally devised as a tool to integrate or substitute conventional public transport (PT) under certain conditions, but still considered as a form of PT.

Moreover, probably the most significant difference between the first DRT services, also known under the name of dial-a-Ride, and those of today lies in the dynamic way in which trip booking, vehicle assignment and vehicle routing occur. In fact, in dial-a-ride services, all trip requests were known beforehand (from few days to few hours earlier) and vehicle routes were progressively adjusted to meet demand. Cordeau and Laporte (2007) provided a review of modelling and solution approaches to the Dial-A-Ride Problem (DARP), which consists of designing vehicle routes and schedules to serve a given number of users who specify origin and destination of their trips. The DARP has many similarities, from a modelling point of view, with two common variants of the Vehicle Routing Problem (VRP), namely:

- The Pickup and Delivery Vehicle Routing Problem (PDVRP), since serving each new DRT user involves two additional locations to visit, in accordance with vehicle capacity constraints.
- The Vehicle Routing Problem with Time Windows (VRPTW), since passengers may have constraints about the earliest or latest time to be served as well as the overall travel time.

As emphasised by Cordeau and Laporte (2007), what makes the DARP different from VRPs is the human perspective. Not only operating cost should be taken into account, but also (especially) reducing user inconvenience. In fact, two main approaches to the DARP can be adopted when balancing operator and user perspectives:

1. Minimize system costs subject to full demand satisfaction (no-rejection policy) and side constraints. In this case (a) the fleet size is variable, according to the demand, and/or (b) user time windows are sufficiently large or “flexible”.
2. Maximize satisfied demand subject to vehicle availability and side constraints. In this case (a) the fleet size is fixed and/or (b) time windows are “rigid”.

In its initial dial-a-ride form, DRT was a social service rather than a proper public service. As a matter of fact, it was first deployed to serve specific user categories (elderly, disabled, etc.), often at high social costs. Users who were interested in using the service would communicate their trip requests via telephone some days in advance to the operator. The objective of Dial-a-ride was to serve widely dispersed trip-patterns and to provide a service in low-density suburban areas for mainly non-work journeys but have often been criticised because of their relatively high cost of provision, their lack of flexibility in route planning and their inability to manage high demand (Mageean and Nelson, 2003).

Currie and Fournier (2020) identified three distinct phases of DRT development: (1) the early DRT dial-a-bus type development (1970–1984), (2) the paratransit/community transport DRT era (1985–2009), and (3) the ICT tech Micro-Transit DRTs (2010–2019). The authors showed that, statistically, high service costs are a major failure factor for DRT systems and that simpler operations (e.g., many-to-few/one, where patronage is concentrated at one of the two trip ends or close to it) had lower failure rates compared to more complex, difficult to manage, many-to-many service types.

Yet, an evolution of dial-a-ride which combine features of conventional (fixed-route) PT and purely demand-responsive (door-to-door) services is represented by “flexible transit services”. Koffman (2004) surveyed 28 flexible transit systems throughout North America, finding some common features: (1) flexible services operated in limited areas that are considered hard to serve for reasons of demographics, street layout, or community preferences; (2) they provide service in low-demand time periods; (3) they provide the entire transit service for a small city, low-density suburban area, or rural area. Moreover, flexible services had different functions based on the spatial and temporal context: in cities with an extensive transit network, flexible transit was likely to substitute for fixed-route operation in limited areas, while in other cities with more limited fixed-route network, flexible transit replaced the entire transit network at certain hours of the day.

Koffman (2004) first classified flexible transit operations from the less to the more flexible type. This classification is schematically depicted in Figure 2.1, where a distinction between two different cases is made, based on the demand pattern: in case origins and destinations of user trips are quite uniformly sprawled over the service area (many-to-many demand pattern), the flexible service travels from a departure terminal to an arrival terminal; instead, when a high percentage of ridership is directed to or come from a transfer station (many-to-one demand pattern) as in peak-hours commuting trips, the flexible service follows a cyclic route which “feeds” the transfer station. In the last case, flexible transit complements and expands the coverage of the fixed-route transit network.

We report the description of the main typologies of flexible transit, based on the survey of Koffman (2004), from those slightly less rigid than conventional bus services to those approaching to door-to-door solutions. Among the “many to many” operation strategies, we can cite:

- *Request stops*. Vehicles can serve a on-demand stops near their route.
- *Flexible-route segments*. Vehicles travel along a fixed-route but switch to demand-responsive operation for a limited portion of the route, without pre-defined stops.
- *Route deviation*. Vehicles operate on a regular schedule along a well-defined path, with or without mandatory bus stops, and deviate to serve demand-responsive requests within a zone around the path, which extent can be flexible.
- *Point deviation*. Vehicles serve demand-responsive requests within a service area and also serve a limited number of stops within the zone without any regular path between the stops.
- *Zone route*. Vehicles have established departure and arrival times at one or more locations but operate in demand-responsive mode along a corridor.

Regarding “many-to-one” (feeder) transit services, also known as demand-responsive connectors, the vehicle routes can also be based on demand-responsive stops (without a pre-defined route) or door-to-door operations, (the maximum level of flexibility). In both cases, the connection with one (or more) transfer station of the fixed transit network is ensured.

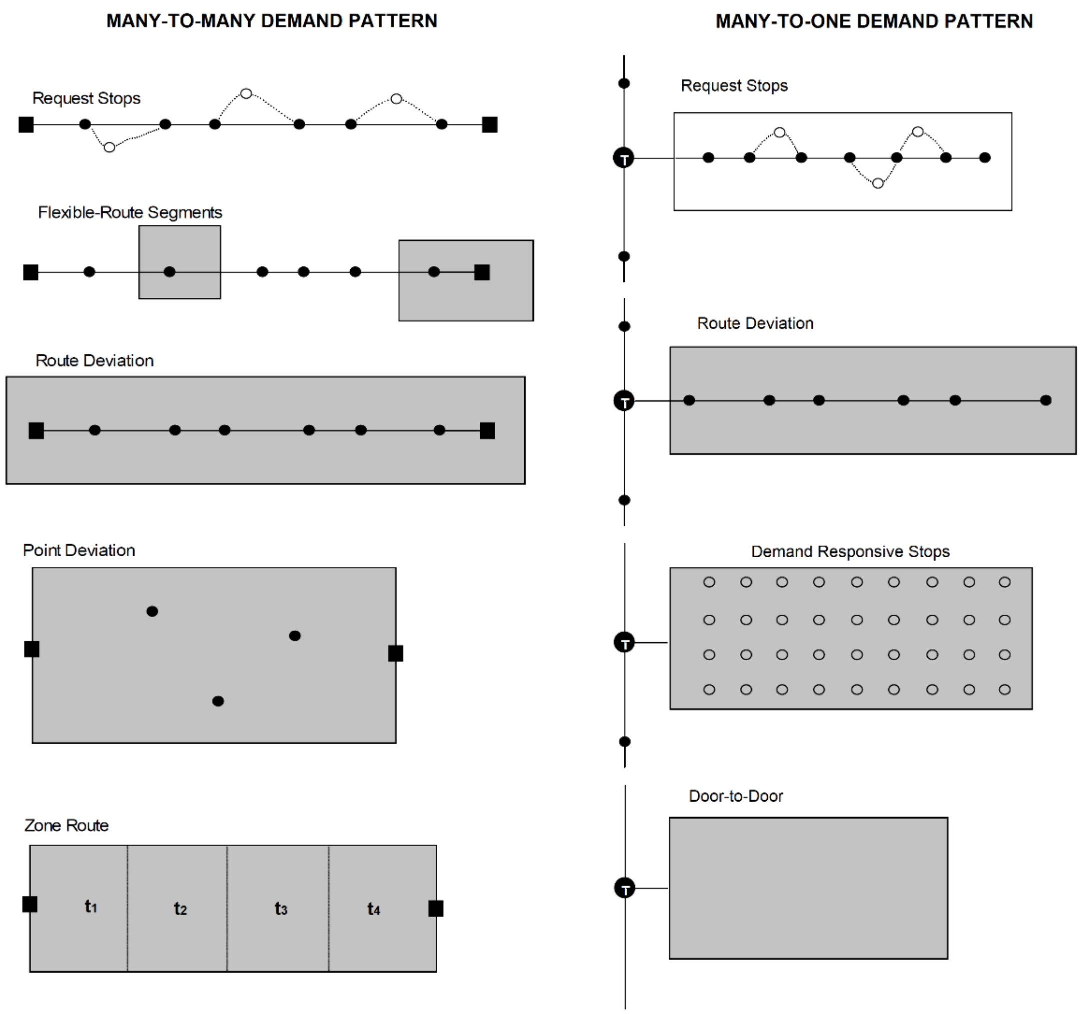


Figure 2.1. Flexible transit typologies (adapted from Koffman (2004))

As flexible services are being introduced, a critical issue is how to ensure the profitability of transport operators while providing a high QoS (Atasoy et al., 2015). The development of ICT in the transport applications has been providing new instruments to face this challenge and broaden the original restricted market of DRT services. Nowadays, ICT allows "on-demand" transport services booking a shared ride on a vehicle in real time. A transport operator can control a vehicle fleet via GPS, track the position of users through smartphones, monitor the service into a dynamic GIS, predict travel times and optimize the matching of drivers and riders with similar itineraries and schedule. Users can book, cancel, or easily change their reservation, pay using Internet tools, acquire information on transport modes, routes and expected travel times before and during the trip and expected arrival times. By the same token, new ICT facilitate coordination between operators, who can collect aggregate booking requests based on the place, time of departure and destination; select vehicle carriers, based on the number of passengers, tracing flexible routes and estimate travel times with high load factor and low driven distances; collect and storage of service's performance data. DRT services can assume different levels of sharing and flexibility, from a shared taxi service to a flexible transit service (Inturri et al., 2019), and can have a beneficial role in terms of social, environmental, technological and economic impacts (Liyanage et al., 2019).

Alonso-González et al. (2018) propose an assessment framework with the aim of evaluating the performance of DRT with focus on user accessibility (Figure 2.2). The authors describe five distinctive characteristics of DRT systems, which are explained below:

- Coverage and routing flexibility: the operation can range from a pure door-to-door service (maximum flexibility) to a fixed-route service which allows for small detours, and the extent of the operating area can either be fixed or change over time.
- Operating period: it is the span of the DRT service.
- Fleet composition: the capacity of DRT vehicles, which often consist in vans or minibuses, determines the flexibility, as well as the level of shareability of the operation. This in turn affects the service fare, the expected passenger travel time and thus the user cost.
- Booking system. Traditionally, dial-a-ride schemes relied on telephone reservations, but large-scale real-time DRT systems require the online processing of bookings and vehicle dispatching in real-time.
- Rejection policy. Even though a DRT could operate under a no-rejection policy, enabling the satisfaction of all traveller requests, such strategy is not always feasible. Therefore, vehicle availability or user time windows are common request acceptance criteria.

The authors then identify the main features of a DRT system from which specific indicators measuring the change in accessibility due to DRT derive. Regarding the travel distance, measuring the percentage of trips shorter than the walking or cycling thresholds (to be evaluated according to the trip purpose, the geographical area, the type of user, etc.) shows how much the DRT is indeed a competitor of active modes. Analysis of spatial and temporal usage reveals where and when DRT is more efficient, as a PT complement or substitute. In fact, long-distance trips can require connections with the mass rapid transit, where available, with DRT acting as a feeder. Measuring the share of such trips using DRT as first/last mile leg is proposed as another specific accessibility indicator.

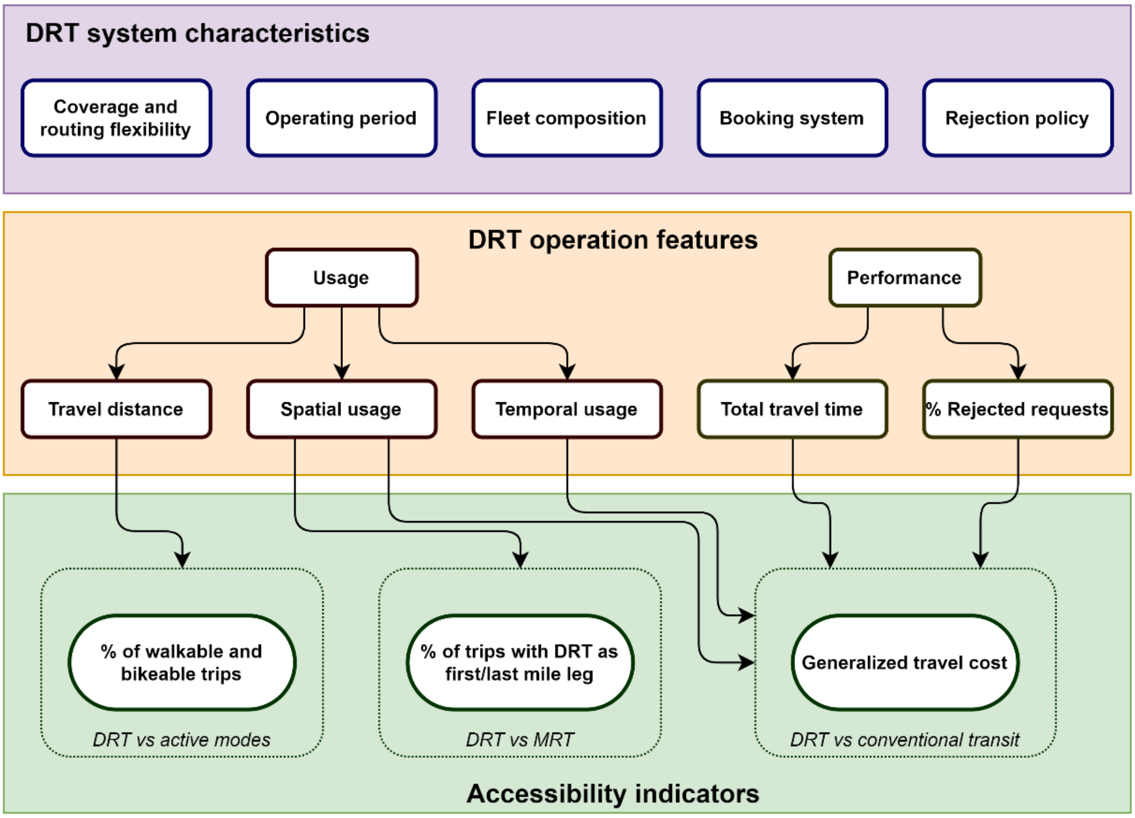


Figure 2.2. DRT assessment framework (adapted from Alonso-González et al., 2018).

The spatial coverage, the operating periods, as well as the performance of the system from the user perspective, including the travel time components and the percentage of rejected requests, determine the generalized travel cost experienced by users, which is an indicator used to compare the performance of DRT as a substitute of conventional transit.

2.1.1 Routing, scheduling and operation of flexible transit

In parallel with the conceptual development and paradigm change of DRT, literature has investigated the diverse problems that must be addressed when designing and operating ridesharing and flexible transit services. Among others, routing and scheduling problems have been widely addressed and various solution techniques have been proposed.

Mourad et al. (2019) provide an extensive survey of models and algorithms for optimizing DRT. Most of the times, optimization problems regarding DRT (often called ridesharing problems) are modelled as a variant of the VRP formulations, according to a set of additional constraints defining the problem, which are listed below:

- Routing constraints.
- Time constraints.
- Capacity constraints.
- Cost constraints.
- Synchronization constraints.

Optimization problems basically involve the minimization or the maximization of one (or more) objective function(s). In ridesharing problems, the objectives of the optimization procedure can be (a) operational objectives and (b) quality-related objectives. The former usually aim at ensuring a cost-effective service (limiting vehicle kilometres, minimizing the fleet size, maximize the number of served requests, etc.) and mainly reflect the operator's point of view. The latter seek to ensure a good QoS provided (minimizing walking, waiting and in-vehicle time, maximizing direct trips, etc.) and reflect the users' point of view. Operational and quality-related objectives regard both individual perspectives, which usually do not match with the collective benefit. A transport operator and, in a wider perspective, the transit agency, should adopt "system-wide", multi-objective approach.

Analytical methods were used to evaluate the performance of both pure door-to-door services in terms of passengers' waiting and riding time (Daganzo, 1978) and dial-a-ride systems with checkpoints able to cluster much of the demand (Daganzo, 1984; Quadrifoglio et al., 2006). While these initial studies assumed a many-to-many demand pattern, many further works focused on demand-responsive transit as a feeder for the First Mile/Last Mile in a feeder-trunk scheme.

On this account, Quadrifoglio and Li (2009) proposed an analytical model to estimate the demand density threshold for a feeder transit service, for which a demand-responsive, door-to-door policy starts to be more attractive (from a user perspective) than a fixed-route operation. In their model, a feeder services operates in a residential area of rectangular shape and homogeneous density, on the side of the major fixed transit line. Since such services are thought for commuters, no intrazonal trips were supposed (Many-to-One demand pattern). The best policy is chosen to maximize the service quality, defined as a weighed sum of customer walking time, waiting time and ride time.

As the similar work proposed by Li and Quadrifoglio (2010), these analytical models aim at assisting planners in the decision-making process when having to choose between a demand-responsive and a fixed-route operating policy and whether and when to switch from one to the other during the day. The authors found that critical demand densities range from 10 to 50 customers/mile^{2}h and that the DRT service performs increasingly better when the portion of drop-off customers increases, such as in afternoon peak hours.

Qiu et al. (2015) identified in the so-called "demi-flexible operating policies" a new category of transit operations able to fill the gap between door-to-door and conventional fixed-route systems in low-demand areas. The author proposed an analytical model to evaluate the transit system performance under the three transit policies, namely fixed-route, flag-stops and flex-route, when dealing with expected and unexpected demand levels. The passenger cost function, built on the three components of travel time (walking, waiting and riding), was treated as the performance measure of transit systems. In addition, the dynamic-station policy was introduced to assist the flex-route service to better deal with the uncertainties of travel demand in low-demand areas, allowing passengers whose deviation request were rejected to make use of accepted door-to-door stops, at the expense of a walking distance to cover. This strategy was found to be useful to reduce user costs at unexpectedly high demand.

Recently, Papanikolaou and Basbas (2021) addressed a similar problem, identifying the critical demand density making conventional bus more convenient than DRT in serving the transport demand from a low-density suburban area towards the city centre. The analytical model involves the minimization of a cost function including both operator and user cost.

Regarding the heuristics techniques applied to routing and scheduling problems of DRT, which constitute the most used approach in literature, Diana and Dessouky (2004) addressed the dial-a-ride problem with user time windows and showed the effectiveness of the proposed regret insertion heuristic in processing large number of requests without excessive computational burden. Their solution techniques outperformed classical heuristics when a high-quality service, i.e., narrow time windows, is provided to the passengers. Diana et al. (2006) proposed a solution method for the fleet size problem related to a DRT service. Their model determines the number of vehicles needed to ensure a predetermined quality for the user in terms of time windows, waiting time at the stops and maximum allowed detour, according to a probabilistic distribution of the demand over the service area.

Quadrifoglio et al. (2008) studied the impact on productivity of specific operating practices currently used by DRT providers. They investigated the effect of using a zoning vs. a no-zoning strategy and time-window settings on performance measures such as total trip miles, deadhead miles and fleet size. Moreover, they conducted this study through a simulation model of the operations of DRT providers on a network based on data for a real DRT service operating in Los Angeles County, although authors stated that the methodology is quite general and applicable to any other service area. Their results suggest the existence of a quasi-linear relationships between the performance measures and two service parameters, i.e., the time window size (increasing time windows allows to reduce the fleet size and the miles driven, despite a lower user satisfaction) and the zoning policy (a centralized control with smaller service areas carries operating cost savings).

Martinez et al. (2015) focused on urban shared-taxi services, proposing a new efficient organizational design and pricing scheme which aim at optimizing the usage of the vehicle capacity. It considers that the client is only willing to accept a maximum deviation from his or her direct route and establishes an objective function for selecting the best candidate taxi. The objective function considers the minimum travel time combination deriving from possible taxi-user matching, while establishing a policy of bonuses to competing taxis with certain number of occupants.

Pan et al. (2015) proposed a mathematical model and a heuristic approach to design the routing plans for a flexible transit system acting as feeder of a main transfer station and serving irregularly shaped communities. The authors developed a mixed integer linear programming model based on a gravity-based solution heuristic to optimize the service area covering a set of demand collection points and simultaneously plan the routes for all service trips. Moreover, the feeder buses are scheduled and dispatched in coordination with the timetable of urban rail transit. The authors assumed that pick-up requests are collected and processed before the beginning of each trip, the travel times between pick-up points and the transfer station are given and the fleet size is known. Finally, the model featured a two-level structure with the aim of (i) maximizing the number of served passengers (upper level) and (ii) minimizing the operational cost (lower level).

Given the user requests and the location of fixed stops, Lu et al. (2015) describe the three critical issues to be solved in constructing a DRT-type feeder network: (1) how to divide the target network into several sub-areas served by different feeder buses; (2) how to effectively design the optimal routing plan for each bus to serve both pre-determined stops and reserved passengers; (3) how to assign the reserved passengers to different feeder buses so as to balance the increasing of travel time along each bus route.

Mendes et al. (2016) proposed a multi-objective approach to VRP applied to DRT. Using the "aggregation tree" iterative methodology, the authors construct a bi-objective version for the problem starting from five different objective functions.

Shen et al. (2017) concentrates on the vehicle routing operation of a demand-responsive feeder system with on-demand stops. Authors propose a two-stage routing model to serve variable passenger demand minimizing the system (operator plus travellers) cost. The first-stage model aims at serving predetermined requests and the second stage model manages extra requests received after the reservation deadline, which could be accepted or rejected according to vehicle capacity and maximum travel time constraints.

Alonso-Mora et al. (2017) faced the large-scale ridesharing problem, where numerous travellers are matched in real time to a large fleet of shared vehicles (e.g., thousands of vehicles), showing how ride-pooling services can bring substantial improvement in urban transport systems. The authors introduced a highly scalable anytime optimal algorithm which processes requests collected during a time window and assigns them in batch to the different vehicles. The algorithm was then validated with 3 million rides extracted from the New York City taxicab public dataset, considering a ridesharing service with a capacity of up to 10 simultaneous passengers per vehicle. The authors found that the 98% of the demand can be served by the 15% of the taxi fleet (with vehicles of capacity 10) or the 23% of the taxi fleet (with vehicles of capacity 4), within a mean waiting time of about 3 min and trip delays between 2 and 4 min. Moreover, assuming that travellers accept a maximum delay of 5 min or more, the use of higher-capacity vehicles increase significantly the percentage of serviced requests, reducing the waiting time and the vehicle-distance travelled, so it is beneficial both for users and ridesharing operators.

The model of Alonso-Mora et al. (2017) was recently extended by Fielbaum et al. (2021), shows that a significant improvement (reductions in the vehicle-hours travelled and in the number of rejections) can be achieved by asking travellers to walk a short distance to reach the pick-up/drop-off points (PUDOs) assigned by the optimization algorithm. They proposed an insertion heuristic algorithm able to face large instances while saving computational time, with the objective of minimize the number of rejected requests and the expected travel times for passengers (including walking and waiting times), limiting the vehicles-hours-travelled (which have a relevant impact congestion). The simulation was performed both on a toy grid network and the real case study of Manhattan, where one hour of the operation was simulated using the publicly available dataset of taxi travels. The authors also found that passengers whose origins or destinations lie in the most demanded zones tend to have a worse QoS and are more likely to be requested to walk.

Li et al. (2018) proposed a mixed integer linear programming model for the route design of a ridesharing system with meet points, including practical constraints, e.g., the travellers' desired time windows and drivers' maximum acceptable travel time. The authors developed a tabu-search meta-heuristic algorithm to solve the problem and found that the introduction of meet points can save the total travel time by ~3% for small-scale ridesharing systems.

Araldo, Di Maria et al. (2019) also proved that limiting pick-up and drop-off points to a finite set of meeting points allows to increase the number of shared trips and the demand served. However, a trade-off between favouring high capacity or QoS should be evaluated on case-by-case basis, e.g., according to the fleet size and the amount of demand.

Stiglic et al. (2018) investigated and quantified the potential benefits of integrating ridesharing with PT as well as the ride-matching technology required to support it. The authors considered a transit service provider that receives a sequence of trip requests from commuters, so the transit service provider is able to three types of matches: (i) a ride-share match between a rider and a driver, (ii) a transit match in which the driver transports the rider to a transit station, and (iii) a park-and-ride match in which the driver parks his car and then uses public transport to reach his final destination. Authors showed that such multi-modal integration enhances the transport system performance by reducing total distance travelled by vehicles and thus the negative externalities associated with car travel.

Sayarshad and Gao (2020) proposed a dynamic optimization approach for smart transit service that switches between demand responsive vehicles and fixed-route minibus system, including a dynamic pre-positioning of idle vehicles in anticipation of new customer arrivals, and relocation of vehicles to rebalance the use of vehicles in the system. The proposed method was tested using a New York City transit dataset, showing that the dynamic switching and relocation approaches can improve social welfare by up to 32% and reduce vehicle miles travelled up to 53% over that of the current transit service, hence the positive impact on energy and environmental conservation.

An interesting change in perspective in improving the efficiency of the dynamic matching between demand and supply in DRT systems is given by the work of Wang et al. (2021), who integrated monetary incentives for passengers to motivate them to consolidate pick-up/drop-off locations and thus reduce the ineffective detours of vehicles. Although users' sensitivity towards incentives depends on various factors (e.g., trip purpose, socio-demographic characteristics, etc.), first results suggest that the proposed approach has the potential of reducing the number of deployed vehicles, shortening the onboard time, decreasing the total distance travelled and increasing the profits for the service provider.

2.1.2 Autonomous mobility on-demand

Another field of research that is developing fast and which could give further incentives to the implementation of innovative DRT services is that of Autonomous Vehicles (AVs). In the next future, fully autonomous vehicles, i.e., those which do not need a driver, are expected to bring a disruptive change in urban mobility, reducing travelling costs and providing a safer, more comfortable, and sustainable transport mode (Meyer et al., 2017). Recently, many authors have been interested in studying the impact of autonomous mobility on-demand (AMoD) on urban transport systems (Azevedo et al., 2016), as well as rural areas (Schlüter et al., 2021) and whether it can lead to a substitution or a complementary effect with respect to PT (Tak et al., 2021). In fact, while AMoD is often considered as PT competitor, many researchers envision a scenario in which AVs are integrated in the PT offer, thus allowing transit to dynamically adapt to travellers' demand, instead of following pre-determining fixed routes.

One of the first investigations about the operation of urban shared AVs and its benefits versus vehicle ownership was carried out by Fagnant and Kockelman (2014), who developed an ABM simulating the daily displacements of travellers and shared AVs within an ideal "gridded" city. The model estimates the impact of different demand patterns and vehicle relocation strategies on the QoS and the operating cost. A remarkable result is that a system of shared AVs is able to accommodate the demand served by a number of private vehicles ten times higher, while increasing of about 11% the user travel time.

AVs have also a potential as first mile/last mile connection to MRT modes. In this regard, Liang et al. (2016) proposed a mathematical model thought as planning tool for automated taxis operation used for the last mile of train trips, according to a profit maximization function. Scheltes and Correia (2017) developed an ABM to evaluate the performance of an automated electric fleet of small vehicles employing a dedicated dispatching algorithm able to distribute trip requests, having origin or destination at a train station, amongst the available vehicles. Requests are processed by the algorithm according to a first-come-first-served sequence a set of specified control conditions (e.g., travel time to reach a passenger). Several simulation scenarios were set up by considering different vehicle relocation strategies, pre-booking possibilities, vehicle cruising speeds and battery charging strategies, finding that such a system should be able to compete with the walking mode and, if higher cruising speed would be allowed, also with the cycling mode. Finally, the authors conclude that operational and economic advantages can be brought by using higher capacity vehicles and increasing the system shareability.

Shen et al. (2018) focused on the relationships between AV operators and conventional PT agencies and operators (Figure 2.3), identifying synergy strategies between AVs and the PT system according to the organizational structure and demand characteristics of Singapore. The authors simulated several integrated AV and PT scenarios for first-mile trips during morning peak hours, finding that the integrated system would improve the QoS, save resources, and be financially sustainable, with respect to the current bus operations.

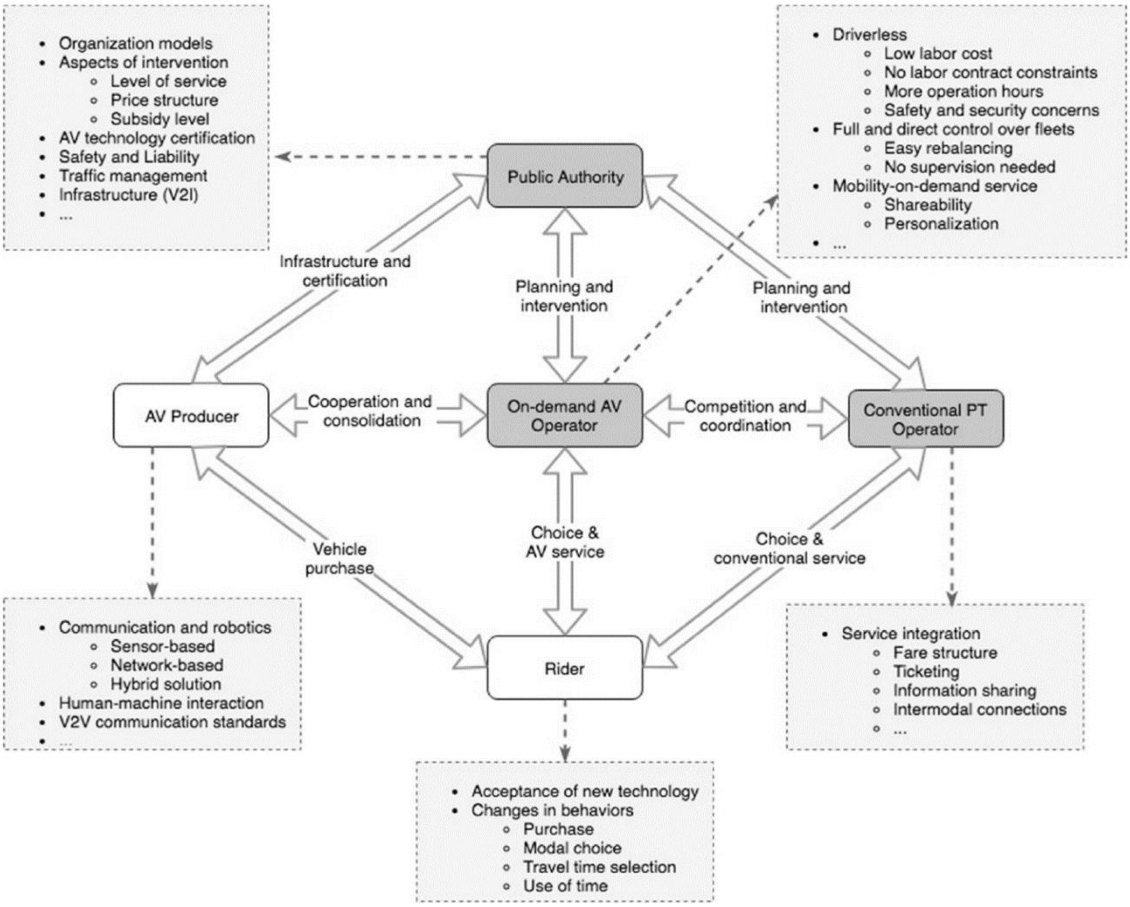


Figure 2.3. Stakeholders around AV-PT operation and AV characteristics from the perspectives of operation (AV and PT operators), governance (public authority), technology (AV producer), and consumption (AV and PT riders) (Shen et al., 2018).

A step towards the shared mobility concept was done by Fagnant and Kockelman (2018), who proposed a simulation model on dynamic ridesharing service provided by AVs. The system pools multiple travellers with similar travel patterns in the same vehicle, optimizing fleet sizing. Not surprisingly, results of this work suggested that automated ridesharing can play a crucial role in limiting vehicle-miles travelled and traffic congestion. Moreover, Basu et al. (2018) analysed the competitive and complementary relationship between MRT and AMoD via a multi-modal agent-based simulation platform. One significant finding regarded the inability of replacing mass transit with AMoD without compromising the user level of service.

Zhang et al. (2019) addressed the topic of automation in bus transit systems by proposing a cost model for bus operations under increasing levels of automation technology, from conventional bus to semi-autonomous and fully autonomous bus services, operating on a trunk-and-branches network. One of the main concerns raised by the authors regards the commercial speed of autonomous buses, which needs to be higher than current speed levels (related to safety issues) to lead to successful implementation.

Despite the attractive perspective of a near future where shared mobility is enabled and driven by vehicle automation, various barriers still remain for implementing AMoD in most areas, including technological and safety barriers, regulatory barriers and cost.

2.1.3 Agent-based simulation of demand responsive transport

Modern cities and their transport systems are regarded as complex systems (Bettencourt, 2015; Ettema, 2015). The state of the system, as well as its development over the time, are driven by the aggregation of individual behaviours, which interaction leads to highly non-linear emergent phenomena. The complexity of transport systems is difficult to address using analytical models, which aims at providing exact solutions based on a mathematical formulation of the problem under exam. Although analytical models are very suited for "high-level" design problems, where the whole system is entirely described and ruled by the parameters and the decision variables of the model, they require many simplifying assumptions to limit the number of parameters needed and reduce the computational burden, especially in large-scale instances.

The need to capture the stochastic dynamics of the interactions between the individual components (agents) of the system led to development of powerful simulation tools able to reproduce the behaviour of the agents and the emergence of macro-level outcomes (bottom-up approach) from micro-level interactions (M'hamdi and Nemiche, 2018). In this context, one computer-based simulation technique which has been gaining an ever-increasing popularity is the agent-based modelling (ABM). A system modelled via ABM is composed of autonomous decision-making entities called agents, each one behaving according to a set of rules which govern the interactions with other agents or the environment where they act. As stated by Bonabeau (2002) even a simple ABM can exhibit complex behavioural patterns out of the reach of pure mathematical methods and provide valuable information about the dynamics of the real-world system that it emulates. More sophisticated ABMs can incorporate learning techniques (e.g., neural networks, evolutionary algorithms, etc.) to allow realistic learning and adaptation. Contrary to traditional simulation modelling (where a model is implemented according to a centralized theory), in ABM, agents are provided with few rules of behaviour and emerging patterns through the simulation are investigated. Another difference stands in the level of aggregation: Even if traditional behavioural models can be disaggregated and detailed, they always presume limited options, rational behaviour and maximizing goals (Le Pira, Marcucci et al., 2017). Moreover, agent-based programming differs from object-oriented languages (e.g., Java, C++, Python, etc.) in the degree of autonomy and the behavioural flexibility towards the environment. In fact, while an object typically exhibits autonomy over its state, an agent has control also over its autonomous behaviour, which is also reactive (it can perceive and respond to changes in the environment), proactive (it can take initiatives to achieve given goals), and social (it can interact with other agents to satisfy given objectives), and has its own thread of control (Wooldridge, 2009).

ABMs have been widely proposed as a valid tool to study complex urban environments and they are suitable to understand the complexity of transport systems and the emergent phenomena deriving from the interaction between individual agents with different objectives and behaviours (Jennings and Wooldridge, 1998). Besides, demand models, such as discrete choice models, can be integrated in the design of ABM, allowing a realistic representation of agents' behaviour (Marcucci et al., 2017b). The use of ABM also provides a suitable environment where to test transport systems serving the same demand and evaluate their performance under different configurations, to understand the potential effectiveness of shared services and their applicability range.

Martinez et al. (2015) proposed an ABM to simulate a shared-taxi system and show its advantages over the current scenario without shared taxis. Clients and taxis are modelled as agents who take decisions according to their specific interests, while the optimization of space and time matching between the demand and the supply side is performed by an entity programmed to act in the interest of both the agents. Results showed that taxi passengers may benefit from significant fare and travel time savings at the expense of penalizing a little the productivity of the taxis in terms of revenue per travelled kilometre. The extension of the previous ABM considering the introduction of self-driving vehicles on large-scale was proposed by Martinez and Viegas (2017). The authors envisioned and simulated a scenario where all private mobility plus conventional bus services were replaced by shared autonomous taxis and autonomous minibuses, showing substantial reductions of travelled vehicle-kilometres and CO2 emissions.

Several platforms or software have developed and made available ABMs to transport modelling, even including public transport and shared mobility components. Among these platforms, we can cite MATSim and SimMobility.

MATSim[1] is an activity-based, extendable, multi-agent simulation framework implemented in Java (Axhausen et al., 2016). In MATSim, every agent repeatedly optimizes in a given number of iterations its daily activity schedule (plan) on the transportation infrastructure, while in competition for space-time slots with all other agents. Following an activity-based approach to demand generation, agents select a plan (composed of a daily activity chain) from their memory, according to the plan scores (econometric utilities), which are computed after each mobility simulation run. A certain share of the agents is then allowed to modify their plan, through the replanning module. The iterative process is repeated until the average population score stabilizes (Figure 2.5).

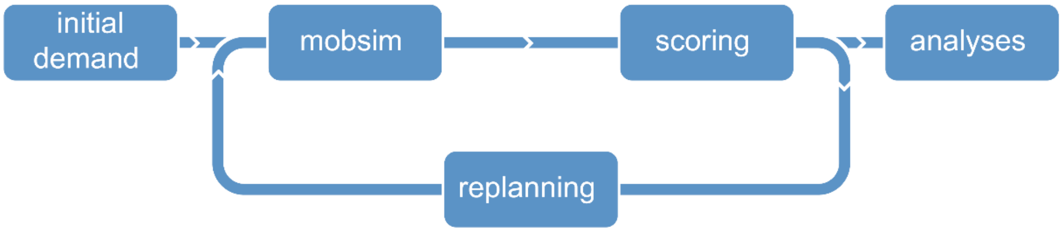


Figure 2.4. The MATSim loop (Axhausen et al., 2016).

Franco et al. (2020) employed the MATSim platform to model and predict the demand for DRT services using anonymised and aggregated mobile phone network dataset as well as additional land-use data sources for the city of Bristol (UK). The model aimed at helping plan flexible transit services integrated with mass transit in substitution of bus services where and when is needed. Narayan et al. (2020) propose a multimodal route choice and assignment model implemented in MATSim and applied for the network based on the city of Amsterdam, allowing travellers combining fixed and flexible transit services for their trips. The simulation study highlighted the benefits of multi-modal integration when flexible transit is used as access/egress mode to the fixed PT, rather than as exclusive mode. Recently,

[1] https://matsim.org/

MATSim has also been used to assess the advantages of MaaS and the potential change in modal share among urban mobility options (Cisterna et al., 2021).

SimMobility[2] is an agent-based modelling framework which integrates various mobility-sensitive behavioural models within a multi-scale simulation platform that considers land-use, transport and communication interactions, following the activity-based paradigm (Adnan et al., 2016). The framework incorporates three sub-models, each one dealing with different time horizons (Figure 2.5a) The Short-Term module works at the operational model and can simulate in detail movement, interactions and decisions of agents (pedestrians, private vehicles and PT). The Mid-Term simulator handles the transport demand and supply on a daily level, including route choice, mode choice, activity pattern and its iterative scheduling of departure time choice. This module is based on activity-based modelling in a multi-modal transport environment and with explicit pre-day and within-day behaviour. The Long-Term simulator captures the strategic decision process on a year-to-year basis, including house, job and economic activity location choices and land use development. The main advantage of this approach is that these three modules can be used individually and, at the same time, enable a seamless flow of information between them.

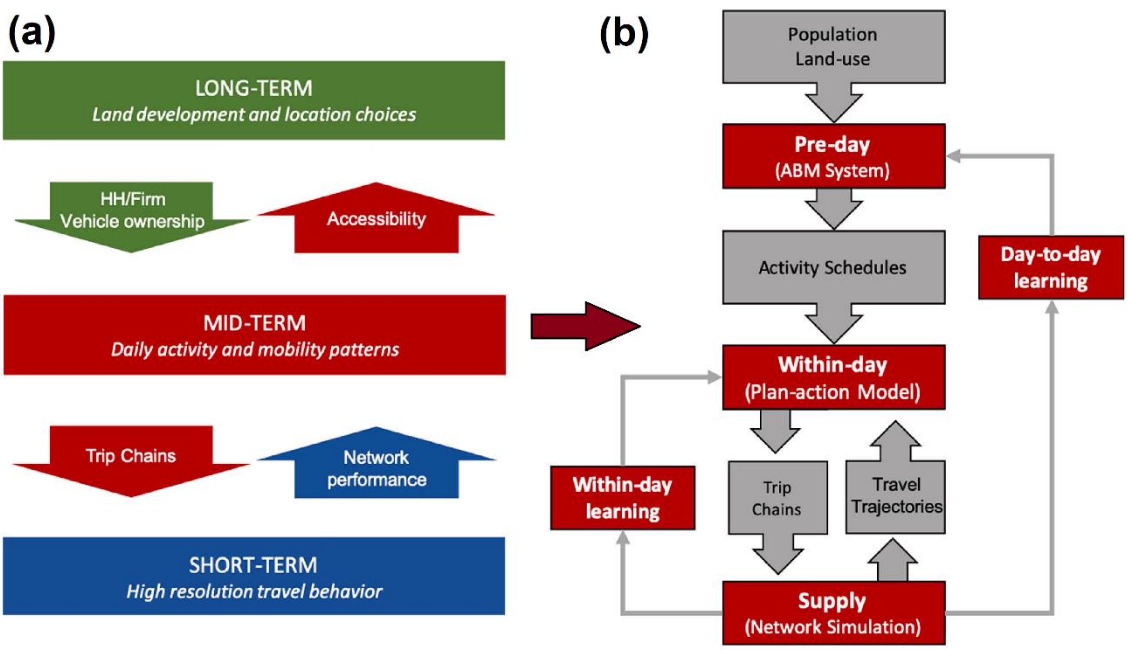


Figure 2.5. Simulation environment of SimMobility, including (a) the three temporal component of the simulator and (b) the three modules of the Mid-term simulator and their relationship (adapted from Oh et al. (2020)).

[2] https://github.com/smart-fm/simmobility-prod/wiki

Several extensions to SimMobility have been implemented, including autonomous mobility on-demand (AmoD) systems (Azevedo et al., 2016), urban logistics (Sakai et al., 2020) and energy estimation models (Fournier et al., 2018). As regards the scope of this thesis, it is worthwhile to cite the work of Basu et al. (2018), which used SimMobility to simulate different operations of AMoD modes and their potential impact on mass transit. The results over a mid-size virtual city suggest that AMoD could not replace mass transit without incurring in heavy congestion problems, while encouraging the introduction of AMoD for the first/last mile connectivity, with beneficial outcomes in terms of mass transit ridership. Moreover, Oh et al. (2021) evaluated the environmental impacts of AMoD in terms of energy and emissions, referring to the mobility network of Singapore reproduced in SimMobility. Results suggest that the introduction of AMOD may increase traffic congestion, travel delays and energy consumption, while reducing vehicle emissions, and that a large portion of AMOD demand would derive from PT with walk access.

The ABM platform used in most of research presented in this thesis, despite not being specifically thought to transport modelling, is NetLogo.[3] Introduced by Wilensky (1999), NetLogo is a multi-agent programming language and modelling environment for simulating natural and social phenomena, developed at the *Center for Connected Learning and Computer-Based Modeling* of Northwestern University. The program offers a fully programmable and customizable environment for studying complex systems that evolve over time (Tisue and Wilensky, 2004), allowing the dynamic visualization of their significant parameters. NetLogo provides to the modeler a meta-language that is simple and intuitive, but highly powerful at the same time, based on the original Logo syntax, and makes it possible, in real time (a) to interact with the simulated environment by modifying the control parameters through *buttons* and *sliders*, (b) to visualize variables, graphs and histograms related to the simulation in progress, (c) to import and save images and data from/to external files, and (d) to perform sets of experiments when the initial conditions or control parameters change. The simulation environment of NetLogo consists of a graphical interface (Figure 2.6) where simulations are managed, and a code section.

Another significant advantage in using NetLogo is its compatibility with the GIS environment, including the possibility of supporting vector and raster data within the model. This is useful to build simulations based on the actual population (residents and employees) of a certain area at census zone level, thus allowing the

[3] https://ccl.northwestern.edu/netlogo/

integration of theoretical models with real data sets (Inturri et al., 2019) and in this way to give the model a more immediate transferability to other contexts.

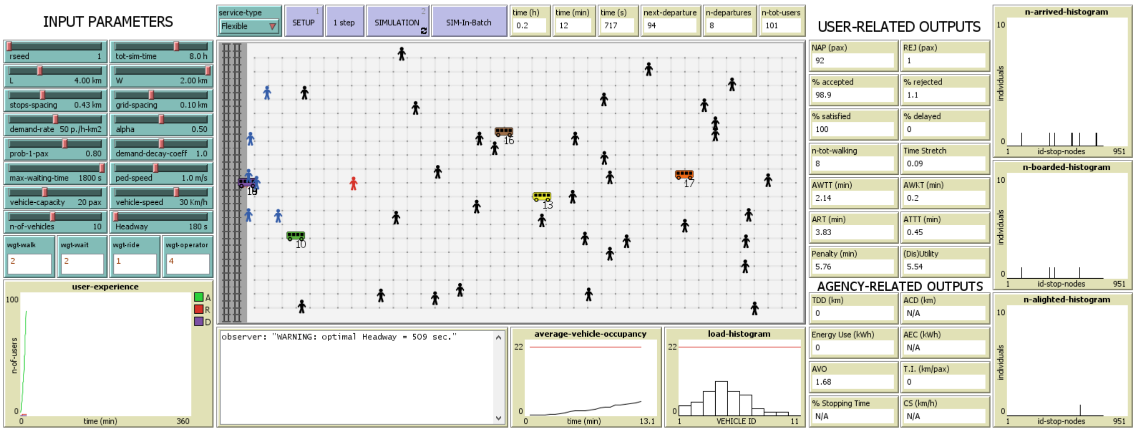


Figure 2.6. NetLogo graphical interface (snapshot from one of our models).

Few examples of ABMs on transport applications using NetLogo can be found in literature. They regard the simulation of dynamic carpooling operations (Armendáriz et al., 2011), the implementation of scheduling strategies for paratransit vehicles (Torkjazi and Huynh, 2019), the optimization of the matching between passenger demand and service supply to improve the capacity usage and the QoS of urban rail transit networks (Zhang (2021), or the comparison between flexible transit and taxi services (Inturri et al., 2021). Finally, an agent-based simulation approach was adopted by Chen and Nie (2018) in order to validate results from an analytical model of a demand-adaptive transit network integrating flexible and fixed transit operations.

2.2 Transit Network Design Problems

The optimal design of transit networks is a widely studied issue. During the last five decades much research has been carried out in the field of transit network design problem (TNDP) (Guihaire and Hao, 2008; Farahani et al., 2013, Ibarra-Rojas et al., 2015). TNDP consists in a hierarchy of decisions, based on three levels, i.e., strategic, tactical and operational.

TNDP is an optimization problem which has been addressed via different solution methodologies, from exact, analytical approaches, often applied to small-size or simplified instances, to heuristics or meta-heuristics, to tackle more complex large-scale problems. The objectives taken most frequently into account include user and/or operator cost minimization, total welfare maximization, transit capacity improvements, protection of the environment. Parameters, constraints, and decision variables regard the network structure, demand patterns, fleet characteristics, headway, route and stop spacing, etc.

According to Ceder and Wilson (1986), the global transit planning process can be divided into a sequence of five steps, i.e., (1) the design of transit routes; (2) the definition of frequencies; (3) the timetabling; (4) the vehicle scheduling; (5) the crew scheduling and rostering. The first step lies at the strategic planning level and is also known under the name of Transit Route Network Design Problem (TRNDP). The second and third steps belong to the tactical level, while the last two deal with operational decisions.

An extensive review on TRNDPs is provided by Kepaptsoglou and Karlaftis (2009). The authors disaggregated the TRNDP into three layers, namely (a) objectives; (b) parameters and (c) methodology (as shown in Figure 2.7).

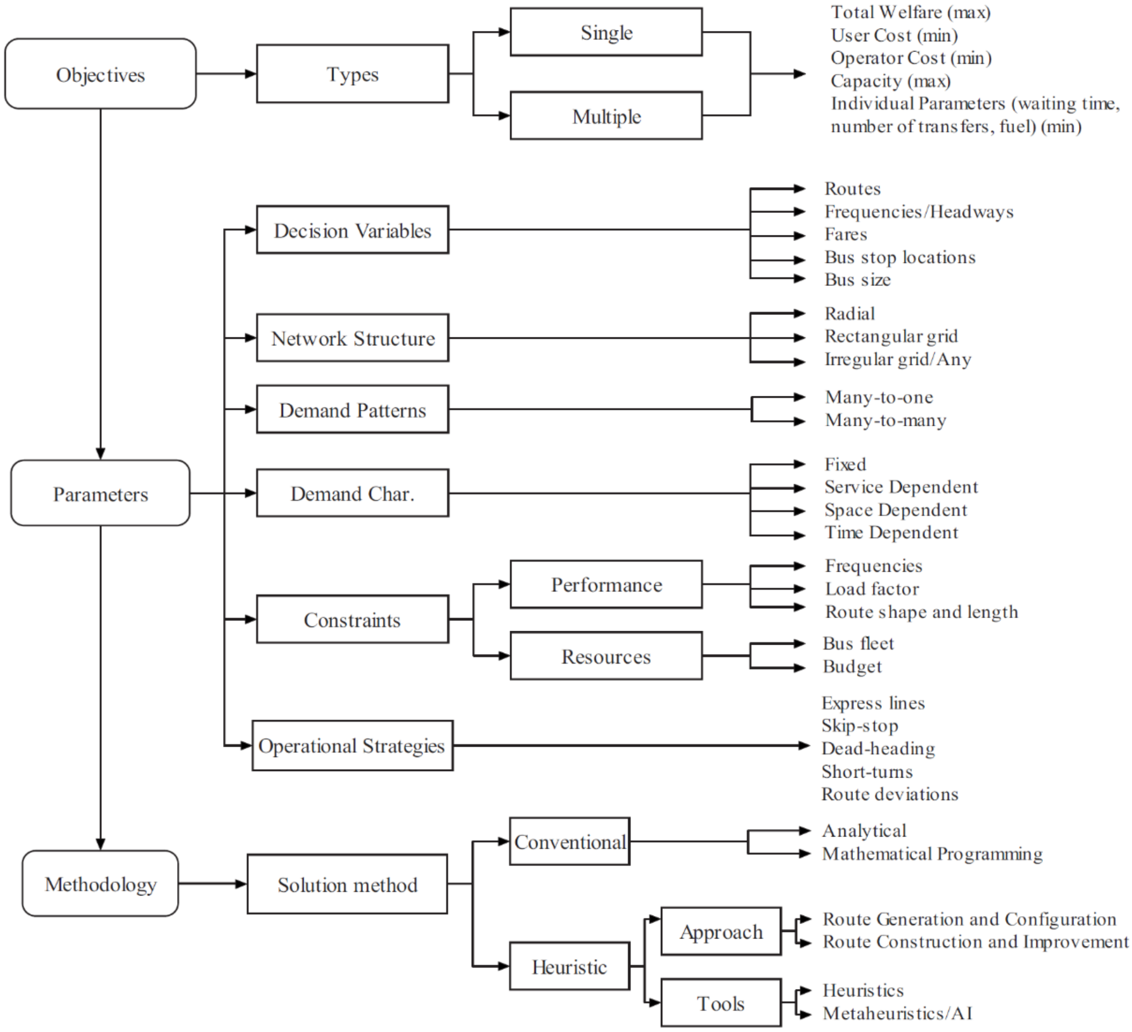


Figure 2.7. Key characteristics of TRNDP models (Kepaptsoglou and Karlaftis, 2009)

The definition of objectives largely depends on the decision-maker's policy for transit network, which might aim to minimize the costs of the system or maximize the QoS provided. Most times, it considers more than one objective, so the resultant problem is usually addressed using a weighted sum objective function. Ceder (2001) suggested four objectives for the TNDP, namely (i) minimize the total waiting time of passengers (user perspective), (ii) minimize the total unused seat capacity to improve the cost-effectiveness of the PT service (operator perspective), (iii) minimize the extra time (and related cost) if all the PT passengers are switched to the shortest path, e.g., using the private vehicle (both community and user perspective), and (iv) minimize the fleet size, i.e., the number of vehicles to carry on the desired frequencies (operator perspective). In general, as shown in Figure 2.7, a classification of design objectives can be done distinguishing between the point of view of the different stakeholders involved in transit operations, i.e., (1) user benefit, (2) operator cost, (3) total welfare, (4) system capacity; (5) environmental protection, and (6) other individual parameters.

The parameters of TRNDPs, according to Kepaptsoglou and Karlaftis (2009), are the decision variables (e.g., headways, stop and route spacings, fares, stop location, etc.) of the problem, the (simplified) structure of the network (e.g., grid-like, rectangular, radial, irregular, etc.), the demand patterns (many-to-one or many-to-many) and characteristics (elasticity, time- and space-dependency, etc.), the constraints and the operational strategies.

Eventually, the methodologies adopted for the solution of TNDPs and TRNDPs can be basically divided into two categories, namely conventional and heuristic approaches. The former includes analytical models and mathematical programming formulations, both providing exact solutions to the problem. However, given the significant number of variables involved in modelling the relationships between components of the public transportation network, these models are usually applied to small size networks with a simple or idealized structure (Ceder and Wilson, 1986). The latter includes heuristics and, more recently, metaheuristics. Heuristic algorithms encompass a broad family of innovative solution methods devised to obtain near-optimal solutions in reasonable computing time. Metaheuristics (Genetic Algorithms, Tabu Search, Simulated Annealing, Ant Colony Optimization, Particle Swarm Optimization, to mention a few) are algorithmic procedures, often inspired by natural processes, which explore the space of solutions to find the closest possible to the optimal one. Computation time, number of solutions explored, and accuracy of the best solution found, i.e., the closeness to optimality, are indicators of the goodness of heuristics.

2.2.1 Design of Feeder Bus Services

In large urban areas, where a mass rapid transit (MRT) system operates, i.e., one or more transit lines with high transport capacity and commercial speed, providing a hierarchical structure of the transit network, with the design of feeder bus lines complementing the MRT in medium-low demand areas, would increase the coverage and the ridership of the PT network. The definition for a feeder service in transit operations is not unitary in literature. Kuah and Perl (1989) defined it as the use of an access mode to a rail rapid public-transport, to improve multi-modal integration. According to Zhu et al. (2017) the term feeder bus refers to small public transport vehicles providing short-distance connections to feeder rail transit stations. It is also called community/feeder shuttle (Xiong et al., 2013; Ceder, 2013) or, when it carries out on-demand operations, it is sometimes referred as demand-responsive connector (Li and Quadrifoglio, 2010; Shen et al., 2017).

The problem of properly designing feeder bus routes for MRT usually belongs to the tactical planning level. A feeder line can be based on fixed routes and schedules, or it can provide a flexible service, able to vary its characteristics as transport demand changes. Most of the times, the problem is to examine which strategy is optimal for the context, according to the demand pattern or taking into consideration the real-time traffic situation.

Ceder (2013) examined different operational strategies and routing scenarios for an implementation of a feeder service (Figure 2.8). The system could be founded on a set of fixed routes, previously identified as optimal, with demand-driven schedules. These routes may be travelled by vehicles maintaining the same direction of travel and the same sequence of stops, or vehicles could be allowed to choose the direction according to real-time demand information (bi-directional routing strategy). Also, a flexible schedule makes possible the eventuality of short turns, i.e., when a vehicle can turn back to the terminal before reaching the end of the line, if there are no passengers waiting in the remaining part of the fixed route, to better accommodate the demand at transfer MRT station, and short cuts, i.e., when a vehicle does not continue its fixed route and uses the shortest path to arrive at the terminal station, e.g., due to loading constraints or synchronization criteria. The objective of the maximum coverage of potential user demand towards the transfer station is also constrained by travel time thresholds, which may not be exceeded to ensure adequate level of service.

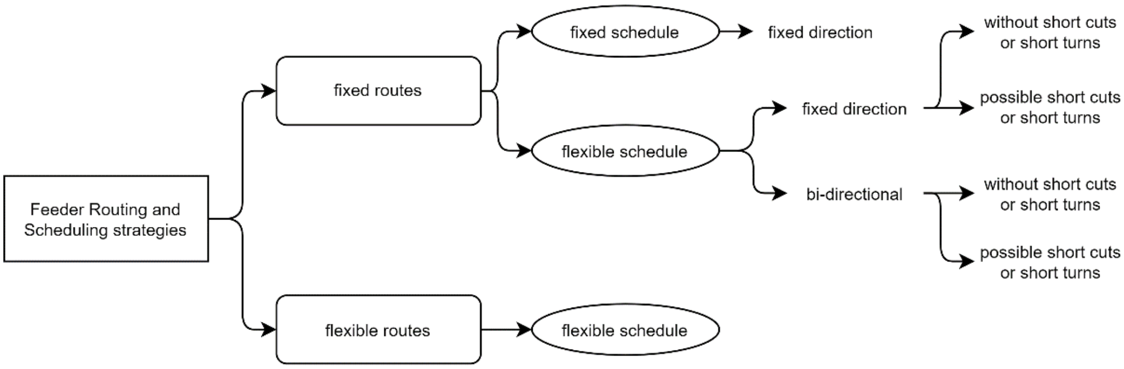


Figure 2.8. Routing and scheduling strategies of feeder service according to Ceder (2013).

The Feeder Bus Network Design Problem (FBNDP) was defined by the authors Kuah and Perl (1989). It consists in determining the optimal bus-route spacing, bus operating headway, and bus-stop spacing of a feeder bus system designed to access an existing rail transit network, in order to minimize the sum of user and operator costs. The mathematical model of FBNDP basically constitutes a multiple vehicle-routing type model, implying the satisfaction of a non-linear objective function, under a set of constraints. Given the rail public-transport network, the location of bus stops and the fleet of bus vehicles, also given the OD demand at each bus stop, the travel cost in the rail system between each pair of railway stations, the distance between each pair of bus stops and between every bus stop and every railway station, the problem is to design a set of feeder-bus routes and to determine the service frequency on each route so as to minimize the sum of operator and user costs.

The coordination of feeder services with mass transit alternatives has been widely studied in the public transport literature, most of times using vehicles with reduced capacity, such as minibus, picking-up users at pre-established stops and drop them at a rail transit station. Input data often encompass the area's topology (defined by the road network, bus stops and transfer stations), bus and train operating costs, fleet size and performances of various modes (commercial speed, capacity, etc.). To limit the complexity of the model, passenger demand is usually assumed fixed, or inelastic, as if users are not sensitive to service variables (comfort, prices, etc.), and may be viewed as hourly averages for a given time period.

As stated above, this class of problems have a multi-objective nature, so fitness to the criteria to be met is usually evaluated considering the trade-off between operator and passenger costs. Without considering fares, the operator looks at the number of vehicles involved, operating and maintenance costs, while users' costs are referred to the amount of time spent in the system (waiting, riding, and transfer time). Also, users would like to have a bus network with more coverage area (percentage of the estimated demand which can be served), more direct trips (if a trip requires more than two transfers, it is assumed that the user will switch to another mode of transport) and high accessibility in the service area. On the other hand, the operator's costs are reduced by keeping the total route length within a certain bound (Almasi et al., 2018).

It is well-known that FBNDP is a NP-hard, non-convex optimization problem, therefore it often requires a complex mathematical modelling and innovative solution methods, including exact (analytical) methods, heuristics and metaheuristics. Many of these heuristic optimization techniques (typical of operational research) have acquired an ever-increasing importance in the various phases of planning, management and monitoring of transport systems.

The design of feeder bus routes may be considered as a Vehicle Routing Problem (VRP) applied to passenger transport, which requires to satisfy the passenger demand using a fleet of vehicles departing from one or more terminal stations. The original version of the VRP was proposed by Dantzig and Ramser (1959) under the definition of Truck Dispatching Problem, which dealt with the calculation of a series of optimal routes for a fleet of trucks for petrol deliveries. This issue, in turn, may be considered as a generalization of the Travelling-Salesman Problem (TSP), which consists in finding the shortest route (or, in general terms, the lowest cost path) connecting all vertex of a graph, starting and finishing at a specified vertex after having visited each other vertex exactly once. Assuming that each pair of nodes in the graph is joined by a link, the total number of different routes through n points is ½n!, and even for small values of n the total number of routes is extremely large.

Several variants of the basic problem have been put forward and VRPs have been studied extensively in last decades (Toth and Vigo, 2002; Cordeau et al., 2007), thanks to their numerous practical implications, especially in logistics but also in passenger transport. One of the most studied members of the VRP family is the Capacitated Vehicle Routing Problem (CVRP), in which a fleet of identical vehicles has to be optimally routed from a central depot to supply a set of geographically dispersed customers with known demands (Baldacci et al., 2012). In Vehicle Routing Problems with Simultaneous Pick-up and Delivery (VRPSPD) the delivery demand and the pick-up demand are served at the same time for each customer (Zhang et al., 2008). In Vehicle Routing Problems with Time Windows (VRPTW) clients may need to be served according to time constraints, within desired time windows (Gambardella et al., 1999). In Dynamic Vehicle Routing Problems (DVRP), new orders and service requests are also handled during pick-up and delivery operations, although some of the orders are known in advance, so the schedule (previously defined) must be updated during run time and vehicle routes undergo continuous adjustments. In such problems, vehicles which are assigned new orders when they are already travelling, do not have to go back to the depot in order to process them. This formulation covers some families of DVRP, as stated by Montemanni et al. (2005), including feeder systems which typically are local dial-a-ride systems aimed at feeding mass transit at a particular transfer location. The authors proposed a solving technique based on the Ant Colony Optimization (ACO) paradigm, explaining how a DVPR can be tackled as a sequence of static VRP instances, by implementing a mechanism to transfer information about good solutions from a static VRP to the following one. There is an event manager collecting new orders and keeping trace of the already served orders and of the current position of each vehicle. This kind of information is used by the event manager to construct a sequence of static instances, which are solved heuristically by the ACO algorithm, to which the pheromone conservation procedure is strictly connected. To limit the time dedicated to each static problem, the concept of time slot has been introduced. However, the drawback of such an approach is that the time dedicated to each static problem would not be known in advance, and then optimization might be interrupted before a good local minimum is reached, producing unsatisfactory results.

VRP could be a useful optimization tool for the urban PT design (Borowska-Stefanska and Wisniewski, 2017). In this regard, a similarity between goods and passengers can be observed: the main problem related to the transport of goods concerns their distribution from the origins (e.g. production places) to destinations (e.g. distribution and / or sales centres), which should be both effective (capable to satisfy demand) and efficient (able to minimize transport costs). Passenger transport, likewise, deals with this problem, but also includes socio-psychological factors and other issues related with the subjective perception of the service offered, which enhance the complexity of the problem.

Mohaymany and Gholami (2010) developed an approach to solve the multi-modal feeder network design problem (with buses and vans offering connections to rail stations) with the objective of minimizing the total operator, user, and society costs. ACO was used to construct feeder routes, identify the best transport mode to deploy and determine service headways. Results showed how multimodal networks have a great potential to reduce passenger costs, hence, they are more likely to attract private vehicle users to use transit and thus improve the profit of transit operators.

ACO algorithms are, by their nature, suitable for constructing optimal paths (Dorigo and Stützle, 2004). Moreover, this family of algorithms can be used in hybrid approaches, together with other heuristics. Kuan et al. (2006) compared and combined different metaheuristic procedures, such as genetic algorithms and ACO to improve the initial solutions for FBND. One potential solution (i.e., the sub-optimal set of feeder-bus routes and related frequencies) of the problem is shown in Figure 2.9. The benchmark problem is taken from Kuah and Perl (1989), which includes 55 stops and 4 stations covering a service area of 5 mile^2, with a bus-stop density of 11 stops/ mile^2 and a hourly demand density of about 2200 pax/ mile^2. The objective function to be minimized represents the sum of passenger (total travel time) and operator (total length travelled by the buses) costs.

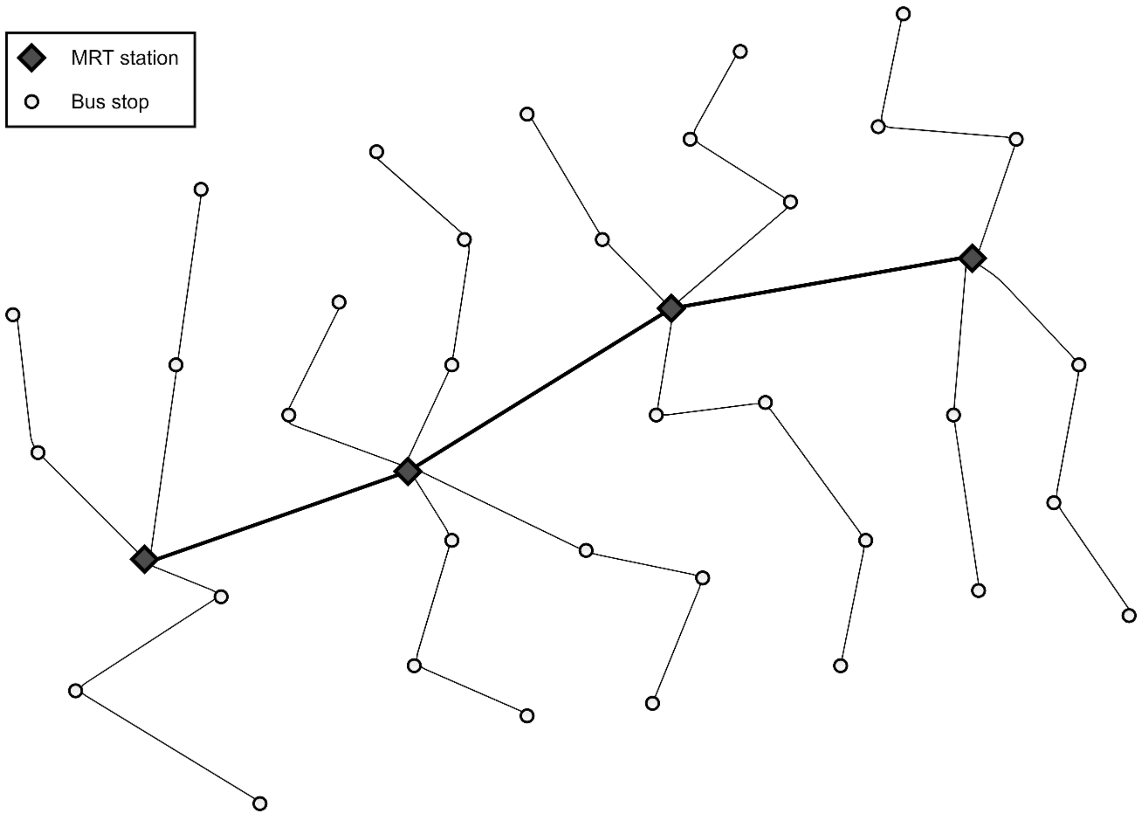


Figure 2.9. Scheme of a potential solution of FBNDP.

Ciaffi et al. (2012) suggested a methodology based on a heuristic route generation algorithm, in order to design a basin of different and complementary lines. Then, the authors found the sub-optimal network of feeder services and the relative frequencies by using a genetic algorithm which properly combines the candidate lines, with the objective of maximizing the service coverage while minimizing the total travel time.

Zhu et al. (2017) found that a circular route model could maximize the flexible advantage of a feeder bus, under the condition of shorter length for feeder bus lines, with the major objective of maximizing the potential demand. Optimal routes starting and ending at a rail station were generated and examined by a genetic algorithm. The potential demand of a link is assumed proportional to the traffic demand of the corresponding aggregated area and the distance between the link and a rail transit station, and inversely proportional to the average distance of passengers to be attracted to that link. The authors also introduced the potential demand reduction coefficient to reflect the relationship between the operating conventional bus lines and feeder bus lines.

When dealing with large and complex transport networks, the problem of determining optimal routes cannot be easily solved by exact algorithms, since computational times grow exponentially with problem size. Heuristics and metaheuristics can thus represent the only feasible options for complete design when the problem complexity is relevant. Instead, exact methods and analytical models are more suitable for suggesting transport policies, where less detailed information are required and general insights about the system performance are needed.

2.2.2 Analytical models and Continuous Approximation

Analytical models are widely used for studying and planning transport systems, allowing the comparison of different type of transit services (Nourbakhsh and Ouyang, 2012; Luo and Nie, 2019) or network structures (Badia et al., 2014; Chen et al., 2015). On one hand, addressing transit network design problems with discrete models can allow to obtain good sub-optimal solutions when dealing with large-scale instances. On the other hand, discrete models are often not robust to stochasticity and uncertainty of input data (Daganzo, 1987), thus not suitable for the strategic level of transit planning. Continuous Approximation (CA) techniques have been developed to overcome this limitation, modelling parameters and variables of the system under exam as continuous density functions over time and space. As reported by Ansari et al. (2018), who provide a comprehensive review of CA models for logistics and transport systems, the key idea is to approximate the objective into a functional (e.g., integration) of localized functions that can be optimized by relatively simple analytical operations. CA is thus used as a complement to discrete models in various contexts. The CA approach was first proposed by Newell (1973), who stated that "the approximate nonunique solution of an idealized problem may be more useful than the (possibly unique) exact solution", e.g., when evaluating strategic decision problems. The systems modelled via CA are defined by parametric schemes and the optimization of the decision variables can be solved with high computational speed. CA has been applied to various logistics problems including facility location, inventory management, vehicle routing, transit studies, integrated supply chain and logistics studies.

CA models of urban public transport networks with many-to-many (Chen et al., 2015; Badia, 2020) and many-to-one (Chandra and Quadrifoglio, 2013; Huang et al., 2020) demand patterns have been recently acquired a renewed interest. Aldaihani et al. (2004) used CA to model the integration between fixed-route transit services travelling along the lines of a grid network and a taxi service operating within the resulting sub-regions. The comparison of gridded-based fixed-route and demand-responsive transport networks was extended by Edwards and Watkins (2013) to a wide variety of street and transit layouts, finding that DRT services could handle trip requests towards transit stations with relatively low demand at off-peak hours and in a cost-effective way.

A milestone paper in analytical modelling of transit network design problems is that of Daganzo (2010), who described the network shapes and operating characteristics that allow a transit system to deliver an accessibility level competitive with that provided by the automobile. The transit network scheme proposed by Daganzo involves a "hybrid" structure, combining the advantage of both the grid and the hub-and-spoke layouts (Figure 2.10a). In the city centre, a double routes coverage is provided determining a transit grid, while radial routes branch toward the periphery. Such structure allows transit users to reach every destination by means of two trips with an intermediate transfer. The optimization problem only involves three decision variables, namely the stop spacing, vehicle headway and ratio between the side of the peripheral region and the side of the boundary of the double-coverage area, allowing to derive handy closed-form solutions.

Estrada et al. (2011) generalized the hybrid model of Daganzo (2010) for rectangular networks, distinguishing transfer stops from non-transfer stops and applied the model to the master plan for a high-performance bus network in Barcelona. Taking inspiration from the "polar network" of Vaughan (1986), Badia et al. (2014) applied the hybrid concept to a ring-radial route layout (Figure 2.10b), where ring lines are provided in the city centre while radial lines bifurcate towards the periphery.

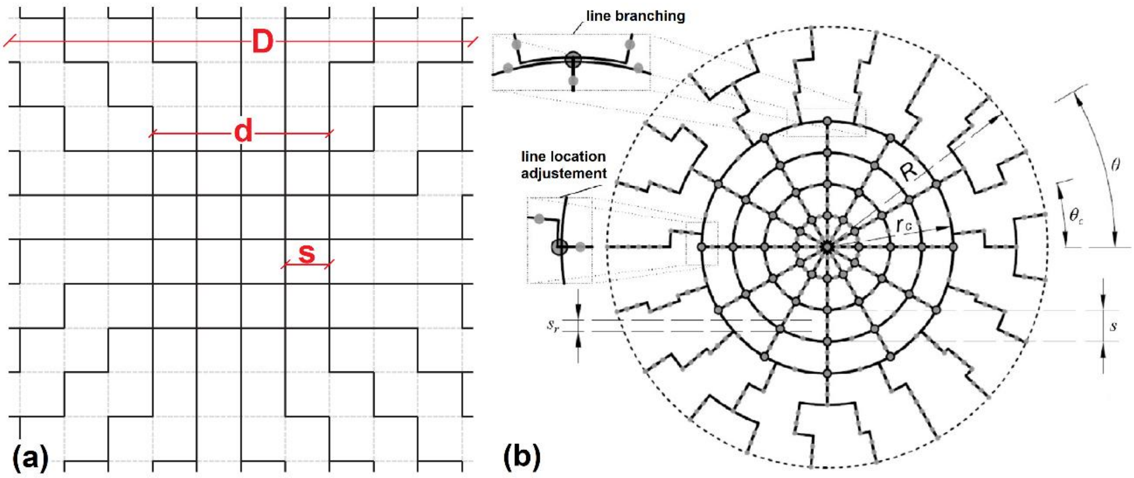


Figure 2.10. Hybrid networks with (a) a grid in the centre and branches in the periphery. (Adapted from Daganzo (2010)) and (b) a radial structure with ring routes in the centre and radial routes toward the periphery (Badia et al., 2014).

Chen et al. (2015) also proposed a ring-radial transit structure and compared it with the grid network design. The novelty is the decision variables of the model, as well as the objective function, including agency- and user-related costs, are based on the integration of localized functions of the distance from the city centre.

Recently, Badia (2020) faced the optimal design of bus networks taking into account the mobility patterns associated with different degrees of urban dispersion. The author investigated the area of applicability where different network structures, e.g., radial, direct-trip-based, hybrid (transfer-based) achieve the minimum total system cost. He found that the demand distribution is the most determinant factor to discern the optimal structure: high demand levels favour direct trips operations, while large cities would achieve a more competitive transit system under a transfer-based hybrid network where transfers between lines are optimally coordinated.

The spatial heterogeneity of demand was also considered by Luo and Nie (2020), who developed a design of paired-line hybrid transit systems, consisting in a combination between fixed-route and demand-adaptive transit services in a ring-radial network structure, to find a balance between accessibility and efficiency.

CA models have been also proposed for DRT systems, to evaluate the performance of both door-to-door services (Daganzo, 1978; Fu, 2002) and dial-a-ride systems with checkpoints clustering the demand (Daganzo, 1984; Quadrifoglio et al., 2006). More recent works focused on flexible transit as the first/last mile solution of a feeder-trunk scheme. Quadrifoglio and Li (2009) used CA to estimate the critical demand density which justifies the switching between demand-responsive (door-to-door) and conventional (fixed-route) operations for feeder transit service. Their analyses were extended (Li and Quadrifoglio, 2010) showing that DRT feeder services perform better with lower demand rates and when larger values are assigned to the weight for traveller walking time. Moreover, the identification of optimal cycle length range for demand-responsive connectors has been investigated by Chandra and Quadrifoglio (2013), highlighting the trade-off between service coverage and QoS.

The hybrid transit network proposed by Nourbakhsh and Ouyang (2012) involves an alternative flexible-route transit system, in which buses are allowed to travel across elongated areas (tube) while collectively form a hybrid (hub-and-spoke) structure resembling a grid network.

Multi-modal transit has also been studied using CA models, with the integration of fixed-route and demand-responsive strategies. Chen and Nie (2017) studied a grid and a radial network with fixed-route transit lines paired with demand-responsive lines connecting passengers to the stops of the former, showing that the paired lines design outperforms the two systems (only fixed-route or only demand-responsive) adopted individually, under a wide range of scenario configurations. The model was also extended to radial network structures (Chen and Nie, 2018), able to perform better than the grid route network. A dynamic system, switching from fixed to flexible transit operations for last mile services was proposed by Guo et al (2018). Luo and Nie (2019) compared six different transit schemes (most of them studied in the aforementioned works) using via CA, showing advantages and disadvantages of each of them.

Daganzo and Ouyang (2019) developed a general analytic framework to model demand-responsive door-to-door services, from non-shared taxis to dial-a-ride systems with increasing levels of shareability, providing approximated but closed-form formulas. Finally, Badia and Jenelius (2020) compared fixed-routes and door-to-door transport operations, carried out by means of autonomous vehicles, to provide first/last-mile solutions in suburban areas, showing how the reduction of operating cost brought by vehicle automation can broaden the range applicability of demand-responsive services.

CHAPTER 3

Feeder Bus Route Design via Ant Colony Optimization

In this chapter we present the first results of an agent-based model (ABM) aimed at designing feeder bus routes able to cover the gap between public transport coverage and ridership in weak demand areas. We approach the optimized design of feeder bus routes as a Vehicle Routing Problem applied to passenger transport, using Ant Colony Optimization (ACO) to find the minimum cost paths within a road network. The methodology proposed has been applied to the case of Catania (Italy), where a metro line is being extended to the city centre to peripheral areas. We used a GIS approach to build the road network, select all potential bus stops, and weight them via accessibility indicators, as a proxy of the potential transport demand. Then, we developed and implemented the ACO algorithm in *NetLogo*, a multi-agent programming and modelling environment for simulating complex systems, in order to find an optimal set of feeder bus routes, where the terminal is a given metro station. These routes are chosen to maximize the potential demand of passengers while complying with the constraint of a desired travel time. Different scenarios have been analysed by comparing a set of key performance indicators based on service coverage and ridership. First results highlight the validity of the method to find suitable routes to cover the gap between conventional public transport and weak demand urban areas and provide useful suggestions for the operation and design of a feeder service.

3.1 Introduction

Public transport (PT) has a key role in providing an extensive coverage of an urban area, preventing congestion phenomena arising from private motorised transport and offering an affordable mobility solution to all citizens.

However, conventional PT is unable to ensure both coverage and ridership in low demand areas, i.e., those areas characterized by low and heterogeneous residential density and high motorization rate (this is due to the poor QoS provided by the PT). Introducing feeder bus lines connecting low demand areas with mass rapid transit nodes could therefore help to shift passenger's mode of transport from individual to collective and shared mobility, thus enhancing the accessibility to urban facilities and services, and reducing the environmental impact. However, the effective design and operation of feeder services should take into account two conflicting objectives: on one side, trying to maximize the demand to serve; on the other side, guaranteeing a reasonable access time to the terminal stations they "feed".

The Feeder Bus Network Design Problem (FBNDP) was addressed by Kuah and Perl (1989) determining a set of feeder-bus routes and the related service frequency, in order to achieve the optimal balance between operator and user cost. The mathematical model of FBNDP basically constitutes a multiple vehicle-routing type model, in which a non-linear objective function must be satisfied, given a set of constraints. Due to the NP-hard nature of such optimization problems, conventional mathematical programming methods can be suitable only for small, simplified networks, while for real large networks, heuristic and, more recently, metaheuristic methods have been used and combined to find "good" suboptimal solutions to the FBNDP.

Under this respect, Kuan et al. (2006) proposed to improve initial solutions by using Genetic Algorithms (GA) and Ant Colony Optimization (ACO) algorithms. Mohaymany and Gholami (2010) developed an approach to solve a multimodal feeder network design problem based on minimizing costs of users, operator, and society. They used an ACO algorithm for the development of bus and van routes in a given service area, calculating frequency of all modes on each route. Martínez and Eiró (2012) considered the timetable of the mass rapid transit as a constraint to which a minibus feeder service has to adjust, taking into account commuters' time windows. Huo et al. (2014) proposed an optimization of the school bus routing operation based on ACO. The circular route model proposed by Zhu et al. (2017) included GA to find a route starting and ending at urban rail transit stations, with the major objective of maximizing the potential demand.

All these approaches proved valid to support an optimized design of feeder bus routes. However, there is still a gap between modelling sophistication and the practicality of urban transport planning. Under this respect, they usually act as “black boxes”, difficult to understand by non-experts, and do not include a detailed spatial representation of spatial constraints and transport demand.

Based on this premise, this paper presents a new modelling framework applied to a real case study for the optimized design of feeder bus routes, by integrating GIS data, accessibility indicators with real transport networks in a multi-agent programming environment. The main contribution of this work with respect to the existing literature is the use of ACO to determine the best set of feeder-bus stops to be served in order to strike a balance between area coverage and ridership.

The remainder of the chapter is organized as follows: Section 3.2 outlines the methodology with the details of the ABM we developed and of the ACO algorithm we implemented in it. Section 3.3 describes a first application of the model to the case study of the *San Nullo* metro station in Catania (Italy), whose accessibility should be improved by designing a feeder bus service, discussing these first results. Then, in Section 3.4 we solve the feeder routing problem involving multiple stations to be served and apply the model to two real case studies. Finally, Section 3.5 resumes the work, providing some considerations for further research.

3.2 Methodology

In this work, the optimized design of feeder bus routes has been approached as a Vehicle Routing Problem (VRP) applied to passenger transport. Simulations were carried out by using *NetLogo* (Wilensky, 1999), a multi-agent programming and modelling environment for simulating natural and social phenomena, which allows users to model and simulate complex systems, reproducing their main characteristics and allowing the visualization of their significant parameters in real time.

Although scientific literature devotes much space to the resolution of feeder bus routing problems, applications to real transport networks are rarely used and often replaced by small ideal graphs. In this regard, a significant advantage in using *NetLogo* lies in its integration with the GIS environment, thanks to the possibility of supporting vector and raster files within the model. It is therefore possible to integrate a GIS-based demand providing the model an easy transferability to other contexts (Inturri, et al., 2019). The interface of the model proposed in this chapter is shown in Figure 3.1.

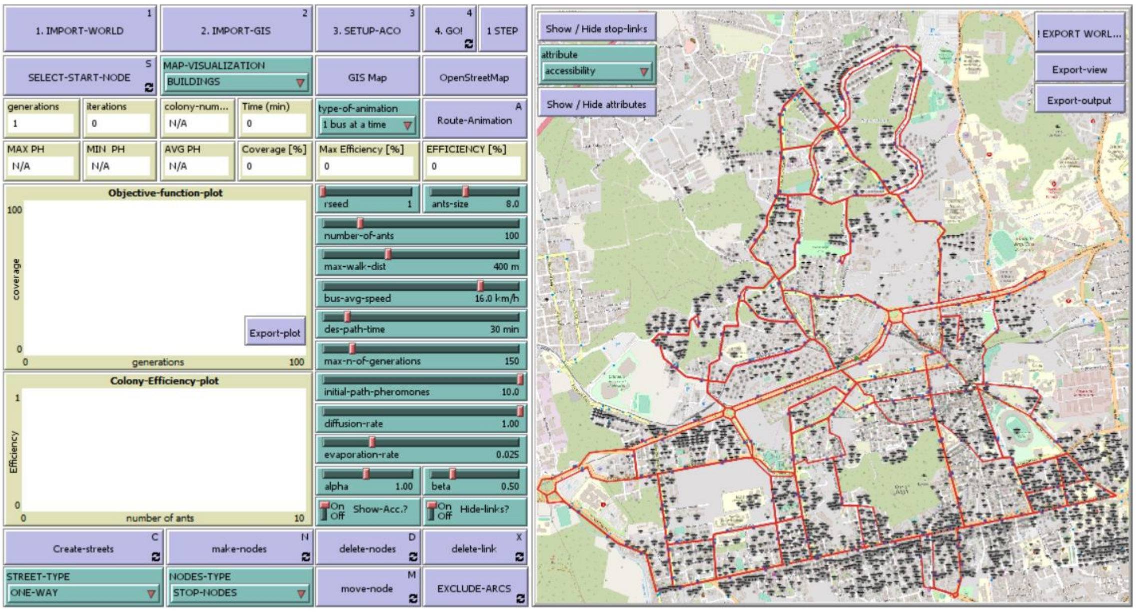


Figure 3.1. NetLogo interface (own setup).

The main features and the steps involved in the model creation are: (1) the construction of the road graph and the identification of the potential bus stops; (2) the estimation of the transport demand associated to bus stops; (3) the implementation of the vehicle routing algorithm. These steps are explained in detail in the following subsections.

3.2.1 Characteristics of the transport network

The construction of the transport network starts from the real street network made up of by arcs (street lanes), nodes (intersections) and stop-nodes (potentially served by the feeder route). Then, the transport network consists in a directed graph where the stop-nodes are connected via through non-congested links.

Stop-nodes are equipped with the value of the potential transport demand gravitating around it. Each link *ij* is characterized by a generalized cost (C_{ij}) that only includes the average travel time along the link itself, given by the ratio between the length (l_{ij}) and the assumed commercial speed of the vehicle (V_{bus}), be it a conventional bus, minibus, van, etc. Location and distance between stops should be studied appropriately considering different factors, including the travel time incurred by users walking from or to the stop, speed performances of vehicles (such as the commercial speed, which is also influenced by the time spent at each stop) and the potential overlap of stops' catchment area.

3.2.2 Transport demand and accessibility functions

We used a Many-to-One demand pattern (Kuah and Perl, 1989), with the following assumptions:

a. Passengers have the transfer station as common origin/destination.
b. Each stop-node can be served by one feeder-bus route.
c. Each bus route is linked to exactly one rail station.
d. Vehicles have standard commercial speed and no capacity constraints.

In order to estimate the potential demand, we also assume that:

i. The transport demand is concentrated at stop-nodes, around which a "passenger catchment area" of the feeder bus service gravitates.
ii. The potential demand of a stop-node is directly proportional to the transport demand of the areas around, thus depending on socio-demographic factors.
iii. The more a potential user is far from the nearest stop, the less she will be attracted to the feeder service.
iv. The more a stop is far from the terminal station, the more it will be attractive to passengers.
v. The presence of other PT lines serving the same stop reduces the potential demand for the feeder service, depending on their frequency and service coverage area.

In view of this, travel demand can be expressed in terms of accessibility of a stop-node. Most of accessibility measures depend on the amount of opportunities in a given zone and the generalized transport cost to reach it (Geurs and Van Wee, 2004). Among the different existing formulations, gravity-type accessibility indexes (Hansen, 1959) provide a continuous measure where the opportunities are weighted by a spatial impedance function, usually depending on travel time or travel distance d_{ij} between places i and j. Moreover, Cascetta (2009) makes a distinction based on land use between the active (origin) accessibility and the passive (destination) accessibility.

Regarding the choice of the impedance function, Kwan (1998) indicates the inverse power function $d_{ij}^{-\alpha}$, the negative exponential function $exp(-\beta d_{ij})$ and the modified Gaussian function $exp\,(-\,d_{ij}^2\,/\,\gamma)$. According to Ingram (1971), the latter has the advantage of having a slow rate of decline close to the origin, in comparison with the first two functions. Therefore, two impedance functions are used to estimate the number of potential users of a stop-node: the first (Equation 3.1) takes in account the effect of decreasing attractivity due to the increase of walking distance, while the second (Equation 3.2) considers the effect of increasing attractivity due to the greater distance from the metro station, based on:

$$f_b(d_{i,k}) = e^{\frac{{d_{i,h}}^2}{2\gamma_b^2}}, \quad if\ d_{i,h} < d_{wk,max} \tag{3.1}$$

$$f_t(d_{t,h}) = 1 - e^{\frac{(d_{t,h}-d_{min})^2}{2\gamma_t^2}}, \quad if\ d_{t,h} < d_{min} \tag{3.2}$$

where $d_{i,h}$ is the walking distance from the centroid of the traffic zone k to the nearest stop-node i, $d_{wk.max}$ is the radius of an area around stop-nodes where demand for feeder service can arise, $d_{t,h}$ is the distance from the terminal node (e.g. a MRT station), d_{min} is the radius of an area around the station with no demand for feeder service since, thanks to the short distance, travellers prefer to walk, γ_s and γ_t are constants to be calibrated. In addition, as suggested by Zhu et al. (2017), the model considers the potential demand reduction for the feeder service due to the other bus lines that are operating via the stop-node through a coefficient M_i.

Based on the above analysis, the potential demand for a feeder stop-node i can be calculated in terms of active and passive accessibility, as shown below:

$$A_{act,i} = M_i \cdot \sum_{h \in Z} f_b(d_{i,h}) \cdot f_t(d_{t,h}) \cdot\ R_h \tag{3.3}$$

where Z is the set of traffic zones gravitating around the stop-node i and R_h is the number of residents of zone k. By the same token, passive accessibility $A_{pas,i}$ is calculated by replacing R_h with W_h, or the number of employees of zone h, as follows:

$$A_{pas,i} = M_i \cdot \sum_{h \in Z} f_b(d_{i,h}) \cdot f_t(d_{t,h}) \cdot\ W_h \tag{3.4}$$

To assign to each stop-node a single accessibility value, we chose to express this value as a linear combination of active and passive accessibility, weighted respectively by w_{act} and w_{pas} coefficients, ranging from 0 to 1 (Equation 3.5). Furthermore, for each stop-node accessibility is normalised by dividing by the highest accessibility value A_{max} in the study area.

$$A_i = (w_{act}\ A_{act,i} + w_{pas}\ A_{pas,i})\ /A_{max} \tag{3.5}$$

3.2.3 The vehicle routing algorithm

The route design model we propose deals with an NP-hard problem in the field of operations research. For this type of instances, several heuristic procedures have been developed to find good suboptimal solutions with acceptable computational efforts. Such procedures are often inspired by natural mechanisms and known under the name of metaheuristics. Ant Colony Optimization (ACO) algorithms (Dorigo and Stützle, 2004) take inspiration from the social behaviour of certain ant species and their ability to find shortest paths between their nest and a food source, simply by exploiting communication based on pheromone trails, a volatile substance that ants may deposit on the ground and smell (Teodorović, 2008). In this family of metaheuristics, by extension, a certain number of simple artificial agents cooperate to build good solutions to hard combinatorial optimization problems via low-level based communications (Dorigo and Gambardella, 1997). Iteration after iteration, more pheromone is deposited on the more frequented trails and this brings out a learning mechanism: when constructing a solution of the problem, the probability of selecting a certain move is higher if this move has previously led to a better solution in previous iterations. This principle can be applied to find minimum cost routes compatible with a set of constraints, by employing a population of ants to jointly solve the optimization problem.

As stated by Dorigo and Di Caro (1999), the ACO metaheuristic can be applied to discrete optimization problems characterized by the following items.

- A finite set of components $C = \{c_1, c_2, \dots, c_{N_c}\}$.
- A finite set of possible connections $L = \left\{l_{c_i c_j} \mid \left(c_i, c_j\right) \in \tilde{C}\right\}, |L| \leq N_c^2$ among the elements of $\tilde{C}$, where $\tilde{C}$ is a subset of the cartesian product $C \times C$. It is preferable to simplify notation by referring to l_{ij}.
- A connection cost function $J_{c_i c_j} \equiv J\,(l_{ij}, t)$ associated to each $l_{ij} \in L$, possibly parameterized by some time instant t.
- A finite set of constraints $\Omega \equiv \Omega\,(C, L, t)$ assigned over the elements of C and L.
- A sequence over the elements of C (or, equivalently, of L) called a state of the problem $s = \langle c_i, c_j, \dots, c_k, \dots \rangle$. If S is the set of all possible sequences, the set $\tilde{S}$ of all the (sub)sequences that are feasible with respect to the constraints $\Omega\,(C, L, t)$ is a subset of S. The elements in $\tilde{S}$ define the problem's feasible states. The length of a sequence s (or the number of components in the sequence) is expressed by $|s|$.

Using the graph $G = (C; L)$, ACO algorithms can be applied to find minimum cost sequences (paths) feasible in respect of the constraints Ω, employing a population of ants to jointly solve the optimization problem. Given two states s_1 and s_2, the state s_2 is said to be a neighbour of s_1 if both states belong to S, and s_2 can be reached from s_1 in one logical step. This way we can define a neighbourhood structure of a state s, denoted by N_s. Each element of $\tilde{S}$ that satisfies all the problem's requirements is a solution ψ, to which a *cost* $J_\psi(L, t)$ is associated. This is a function of all the costs J_{ij} of all the single connections belonging to the solution ψ. To each connection (arc) l_{ij} two types of information are associated:

- The pheromone trail τ_{ij}, encoding a long-term memory about the whole ant search process.
- The heuristic value η_{ij}, which generally represents a priori information about the problem instance definition provided by an external source (different from the ants).

From the earliest applications, several variants of ACO have been proposed by different authors. Dorigo and Socha (2006) described the main features of three popular ACO algorithms: *Ant System*, *Ant Colony System* and *MAX-MIN Ant System (MMAS)*. The latter was developed by Stützle and Hoos (2000) in order to improve performance of the original Ant System algorithm, introducing the following changes:

1. To better exploit the best solutions found, after each iteration only the best ant is allowed to reinforce the pheromone trail.
2. To avoid stagnation of the search the range of possible pheromone values is limited to an interval $[\tau_{min}; \tau_{max}]$.
3. To achieve a higher exploration of solutions at the start of the algorithm, the pheromone trails are initialized to τ_{max}.

The algorithm implemented in our model derives from Ant Colony System (Dorigo and Gambardella, 1997), and particularly from MAX-MIN Ant System (Stützle and Hoos, 2000), which are the two main improvements of the first Ant System, originally applied to the Travelling Salesman Problem. Three subsequent steps are taken for one simulation, as schematically shown in Figure 3.2, i.e., (i) the initial setup of transport network and GIS dataset ($t = 0$); (ii) the setup of the ACO parameters ($t = 0$); (iii) the simulation run ($t > 0$).

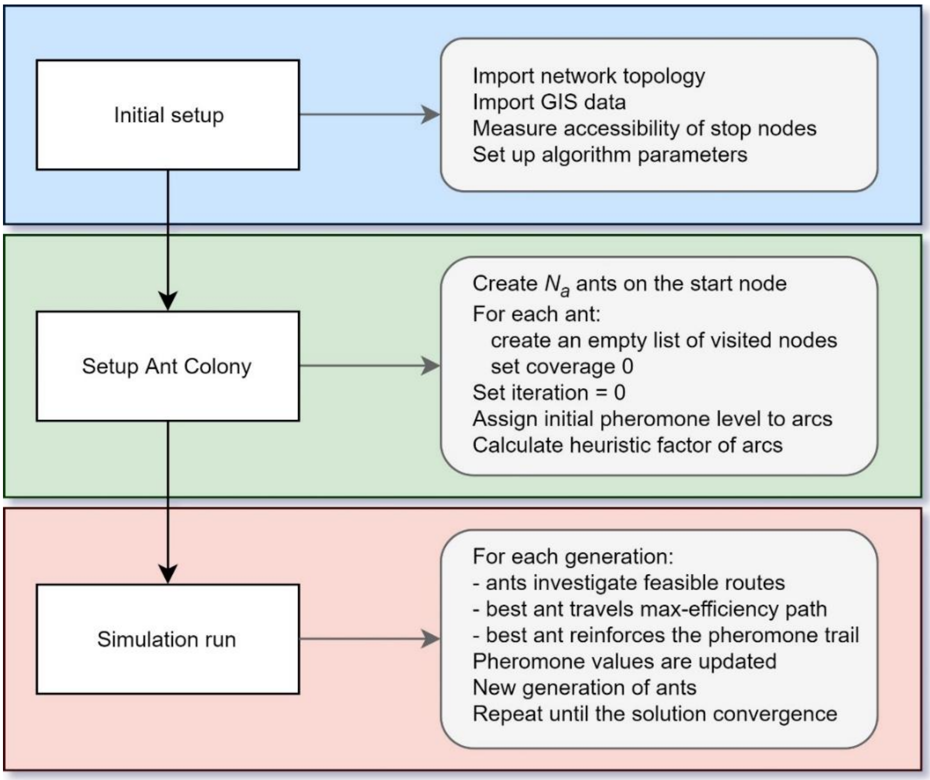


Figure 3.2. The three main steps of the simulation process.

In Figure 3.3 the flow chart of the routing algorithm implemented in our model is outlined. At the start of a simulation, a colony of m artificial ants is generated on the start-node (i.e., the metro station). At every iteration, ants explore the network recording the visited stop-nodes and the crossed links and updating the route travel time. After visiting a stop-node, each ant updates its own Coverage attribute CV_k, given by the sum of *Accessibility* of the visited stop-nodes. When all m ants have completed their tour, the best ant (i.e., the one with the highest objective function (Equation 3.6)) is selected and the pheromone updating rule is applied.

The constraints of the model can be summarized as follows:

1. Each stop-node can be included in at most one feeder route.
2. Each route must start and end at the same metro station.
3. If the travel time ($T_{path,k}$) exceeds a specified threshold (T_{des}) representing the desired route travel time (input parameter), the ant comes back to the start-node through the shortest path completing its tour.

The objective function is calculated for each ant as a route efficiency indicator E to be maximized, as follows:

$$Maximize \quad E_k(t) = CV_k - \frac{2\,CV_k}{T_{route,k}}\,\Delta T_{excess} = CV_k \cdot \left(1 - 2\,\frac{\Delta T_{excess}}{T_{route}}\right) \tag{3.6}$$

where $\Delta T_{excess} = (T_{path,k} - T_{des})$ if $T_{path,k} > T_{des}$, otherwise it is equal to zero. In this way, a penalty is given to the Coverage when the travel time of ant k exceeds the desired route travel time.

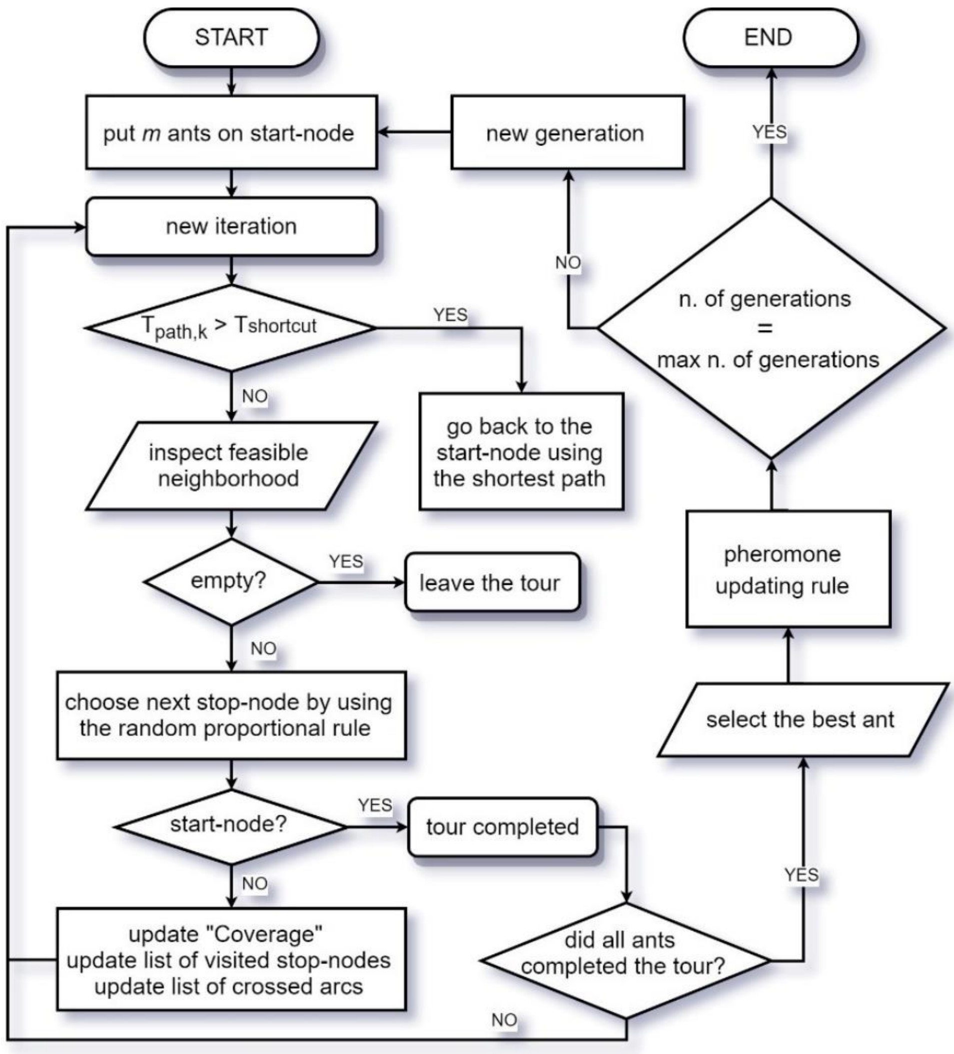


Figure 3.3. Flow chart of the ACO algorithm for the feeder bus routing problem.

As previously stated, each ant builds a route by applying a random proportional rule to decide the next stop-node to go. Therefore, the probability with which ant k, currently at stop-node i, chooses to go to stop-node j is given by:

$$p_{ij}^k = \frac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{l \in \mathcal{N}_i^k} [\tau_{il}(t)]^{\alpha} \cdot [\eta_{il}]^{\beta}} \qquad \text{if } j \in \mathcal{N}_i^k \tag{3.7}$$

where α and β are parameters that control the relative importance of the pheromone trail τ_{ij} versus the heuristic information η_{ij}, $\mathcal{N}_i^k$ is the feasible neighbourhood of ant k when being at stop-node i, i.e., the set of stop-nodes directly linked to i and not visited yet.

While the pheromone trail is updated after every generation *t,* the heuristic information is available a priori and is given by the ratio between the potential demand (accessibility) of stop-node i and the distance between i and j, or the length of link ij, measured on the road network: $\eta_{ij} = A_i/d_{ij}$.

If $\alpha = 0$, the process acts as a greedy algorithm and the closest stop-nodes are more likely to be chosen; on the other hand, if $\beta = 0$, only pheromone amplification is enabled, and this would lead to the rapid emergence of solutions stagnation and hence to strongly sub-optimal solutions.

Each generation of ants concurrently builds circular routes starting and ending at the metro station. Once all the m ants have completed the tour, only the "best" ant (i.e., the one that founds the solution that maximizes or minimizes the desired objective function) is allowed to reinforce the pheromone trail, to better exploit the best solutions found by every generation, by means of the following updating rule:

$$\tau_{ij}(t+1) = (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}^{best}(t) \quad \text{with } \Delta\tau_{ij}^{best}(t) = Q \cdot \frac{E^{best}(t)}{E^{global-best}} \tag{3.8}$$

where ρ is the evaporation rate, ranging from 0 to 1, $\Delta\tau_{ij}^{best}(t)$ is the amount of pheromone deposited on link ij used by the best ant at iteration t, E^{best} indicates the highest value of the objective function among all the m ants at generation t, $E^{global-best}$ is the best value of E found from the start of the simulation and Q is the diffusion rate. Moreover, the pheromone trail is lower bounded by τ_{min} to avoid stagnation of the search. Finally, if the number of generations N_{gen} is still lower than a given value, another generation of m ants is launched on the network searching for the best route, otherwise the algorithm stops and outputs the results.

3.3 Case study 1: single-station problem

Improving urban accessibility by PT service is a subject of significant interest for Catania, a city of about 300,000 inhabitants located in the eastern part of Sicily (Italy), which has been affected by a process of decentralization of housing settlements and commercial services to peripheral areas, modifying the citizens' travel patterns, increasing the amount of commuter traffic and producing air pollution. A metro line currently connects the city centre with the northwest zones of the city and it is subjected to further development.

The case study focuses on improving the accessibility of the *San Nullo* metro station (SN), thanks to the optimized design of feeder bus routes. The station is located in a suburban area, where the only access routes to the station are not equipped with sidewalks and present high slopes, constituting a real obstacle for pedestrians and therefore reducing the attractiveness of the station itself. Once the study area has been identified, the potential stops of the feeder service have been placed on the map, partially based on existing stops of public transport. The study area is divided up into a grid of patches (i.e., cells in *NetLogo*, each one having its own pairs of spatial coordinates), to which a number of residents aged 15 to 74 and of employees is assigned, based on demographic data provided by the most recent ISTAT database (dating back 2011) at census zone level. Demand (R_k and W_k) is generated only within a specified radius around stop-nodes, which accessibility A_i is calculated using Equation 3.5.

For a first test of the model, the design of two feeder routes has been carried out, seeking to obtain the widest coverage possible in accordance with travel time constraints. Various scenarios have been analysed and compared, by setting different desired travel times. Input parameters are shown in Table 3.1. They have been chosen after several tests which resulted in better computational times and simulation outcomes.

Table 3.1. Input parameters set for case study 1.

w_{act}	w_{pas}	d_{min} (m)	$d_{wk,max}$ (m)	γ_s	V_{bus} (km/h)	# ants	τ_{start}	Q	ρ	α	β
1	0.35	250	400	0.20	16.0	100	10.0	1.00	0.025	1.00	0.50

Figure 3.4 and Table 3.2 report the results of 7 scenario simulations, performed by increasing T_{des} from 10 to 40 minutes. Figure 3.4 shows the convergence process of the objective function E after each ants' generation, for given different T_{des}.

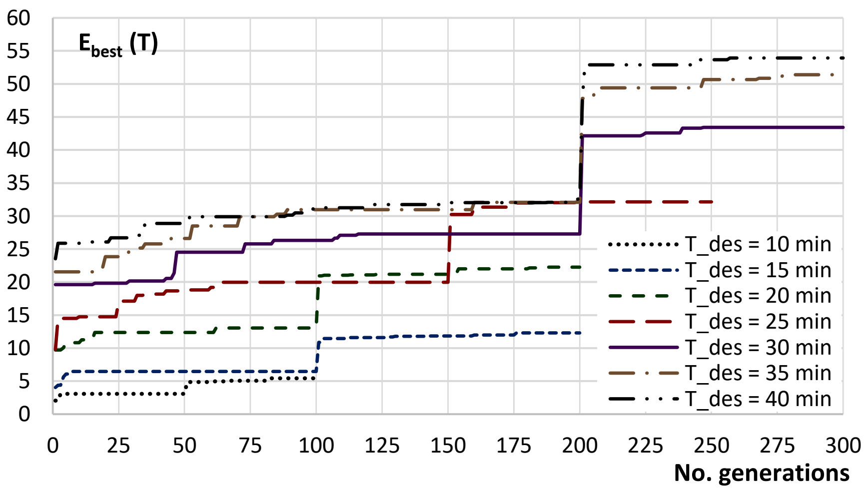


Figure 3.4. Convergence process of the objective function (best Efficiency).

As can be observed in Figure 3.5, as well as in the last column of Table 3.2, Efficiency shows higher growth rates when desired travel times are set between 20 and 30 minutes.

Table 3.2. Comparison of simulations results.

Scen. No.	Route No.	N_{gen}	T_{des} (min)	T_{route} (min)	CV (%)	CV_{tot} (%)	E (%)	E_{tot} (%)	ΔE / ΔT_{des} (min^{-1})	ΔE_{tot} / ΔT_{des} (min^{-1})
1	1	50	10	10.4	3.35	9.12	3.09	5.44	0.309	0.272
	2	50	10	14.2	5.77		2.35		0.235	
2	1	100	15	18.5	10.37	19.00	6.49	12.31	0.680	0.687
	2	100	15	17.9	8.63		5.82		0.694	
3	1	100	20	20.6	13.87	24.03	13.03	22.26	1.308	0.995
	2	100	20	21.0	10.16		9.23		0.682	
4	1	150	25	25.4	20.44	35.33	19.86	32.04	1.366	0.978
	2	100	25	27.5	14.89		12.18		0.590	
5	1	200	30	30.4	28.10	44.27	27.29	43.42	1.486	1.138
	2	100	30	30.0	16.17		16.13		0.790	
6	1	200	35	34.6	32.07	51.39	32.07	51.39	0.956	0.797
	2	100	35	34.9	19.32		19.32		0.638	
7	1	200	40	41.1	34.42	55.95	32.55	53.92	0.096	0.253
	2	100	40	40.2	21.35		21.37		0.410	

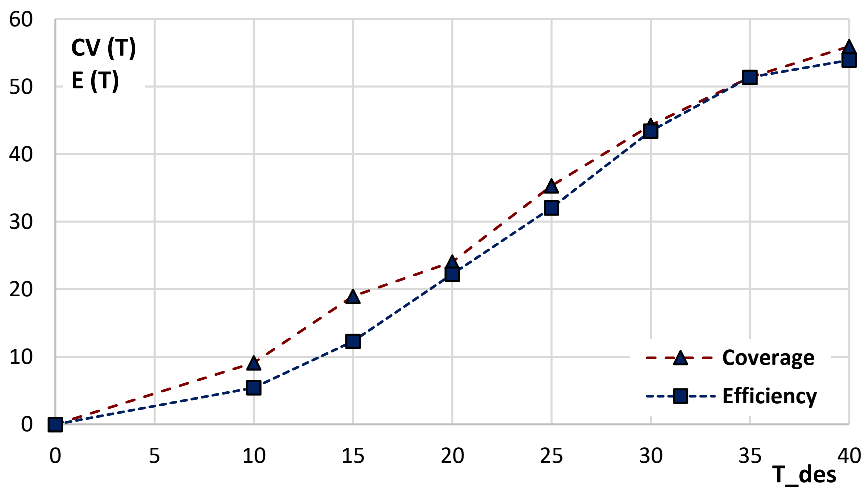


Figure 3.5. Total coverage and efficiency in each scenario.

Based on this result, the best compromise between demand coverage and operating costs is represented by the 5th scenario, involving two routes with a travel time of 30 minutes.

From the graphical output of the three simulations shown in Figure 3.6, it emerges that routes extend towards areas which are more distant from the terminal station, when increasing the desired travel time. It should be noted that if a simulation produces a wide circular route, this result could not be optimal for a feeder line, as it would force some passengers who get on the shuttle to bear unjustifiably high travel times before reaching the station. In this case, two circular routes that run in opposite directions could be provided.

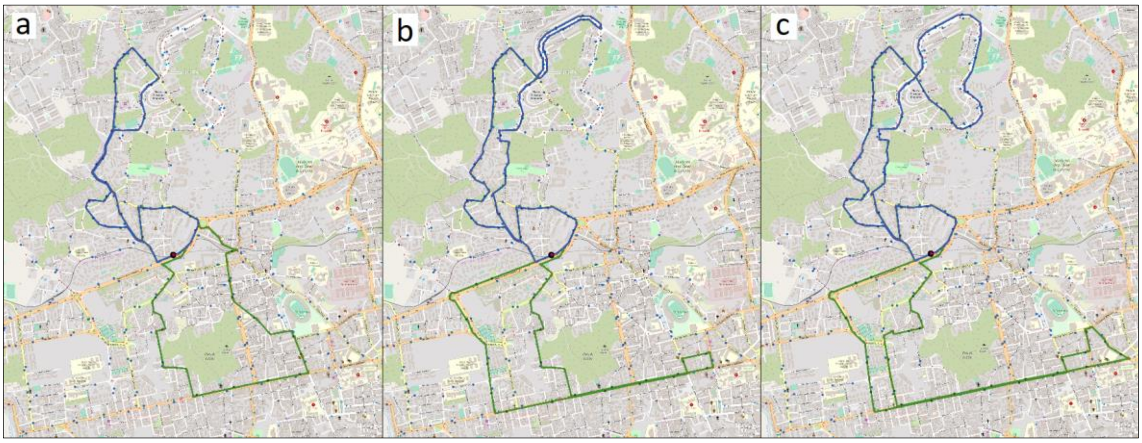


Figure 3.6. Graphical output of three simulations: couple of feeder routes setting T_{des} equal to (a) 20 min; (b) 25 min; (c) 30 min.

3.4 Case study 2: feeder bus routing problem with multiple stations

This section presents the first results of the described ABM improved by considering the multiple feeder lines connected to different MRT stations, assuming that each station is served by one feeder line.

With respect to the methodology explained in Section 3.2 the model was improved by considering the expected passenger travel time, using the feeder service, from the bus stop i to the terminal station ($t_{access,i}$) and vice-versa ($t_{egress,i}$), in relation with the corresponding travel times if using the shortest path ($t_{sp,access,i}$ and $t_{sp,egress,i}$). Specifically, every time the ants complete their routes, the accessibility A_i (Equation 3.5) of the stop-nodes served by one feeder route is recomputed considering a penalty factor, as shown in the following equation:

$$A_i^* = A_i \cdot \sqrt{\frac{t_{sp,access,i}}{t_{access,i}} \cdot \frac{t_{sp,egress,i}}{t_{egress,i}}} \tag{3.9}$$

The computation of the objective function (Equation 3.6) is affected by this penalty factor. Therefore, the optimization procedure is less likely to find routes with a "wide" shape.

To test the updated version of the model, two case studies were chosen: the first one focuses on the metro stations "Nesima", "San Nullo" and "Cibali", (opened between 2017 and 2021) while the second one relates to the urban rail stations "Ognina" and "Picanello" (opened between 2017 and 2018).

Both the case studies regard the MRT network of Catania, which currently does not have a feeder bus system complementing its stations. An overview of the geographic location of the stations is shown in Figure 3.7.

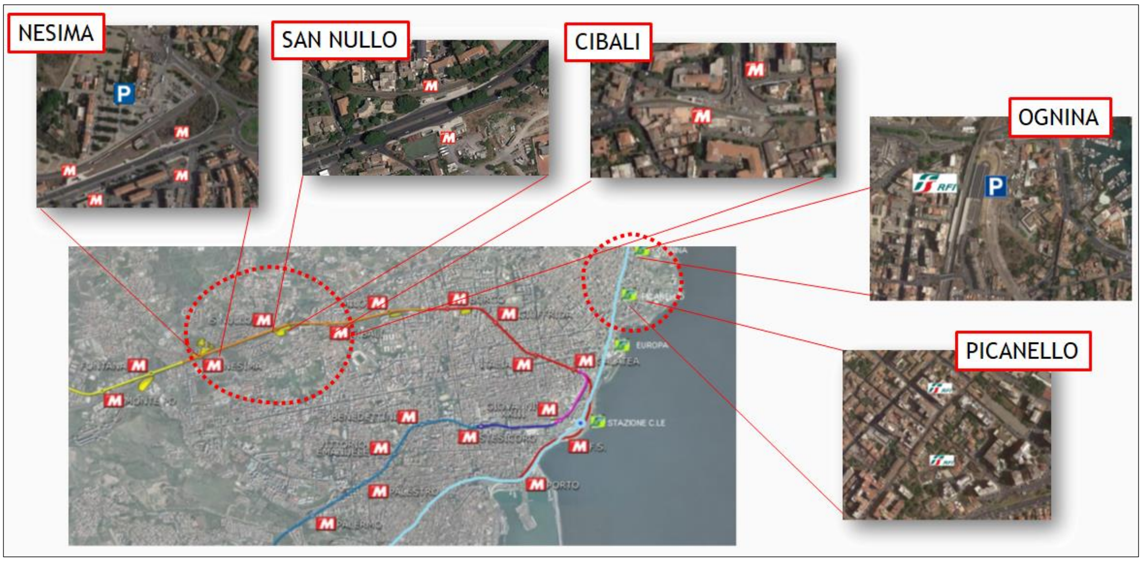


Figure 3.7. Location of the stations to serve.

The input parameters chosen for case study 2 are shown in Table 3.3. Input parameters set for case study 2.

Table 3.3. Input parameters set for case study 2.

w_{act}	w_{pas}	d_{min} (m)	$d_{wk,max}$ (m)	γ_s	V_{bus} (km/h)	# ants	τ_{start}	Q	ρ	α	β
1	1	200	400	0.128	16.0	100	10.0	1.00	0.025	1.00	0.50

We point out that the parameters of the impedance functions (Equations 3.1-3.2) are chosen considering the suburban context of the stations under exam, thus a rapid decline of the walking accessibility with the distance from the stop. From Figure 3.8

one can note that the impedance f_b is close to zero at distance d = 400 m, which is the maximum walkable distance users are willing to cover.

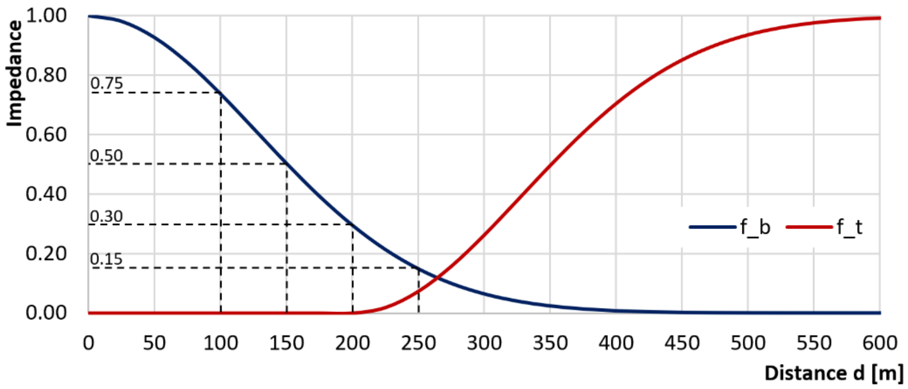


Figure 3.8. Values of the impedance functions f_b and f_t along the distance d from the terminal station, with d_{min} = 0.2 km and $\gamma_t = \gamma_b$ = 0.128.

3.4.1 Metro stations "Nesima – San Nullo – Cibali"

The service area related to the three stations under exam measures 7 km^2 and includes 210 stop-nodes. We recall that, in our model, not every stop-nodes have to be served by feeder bus routes, since we do not aim at covering the 100% of the area (it would be possible, but not cost-effective). About 41200 residents live and 11680 employees work within the study area. Figure 3.9 shows the graphical output of the three optimal routes found with T_{des} equal to 30 min. Figure 3.10 shows the convergence process of the objective function.

In detail, the first feeder route linked to Nesima has an expected travel time (cycle time) of 29.5 min, the second route linked to San Nullo has a shorter cycle time 23.2 min, while the third route connected with Cibali station has the higher cycle time, 31.6 min. The explanation of this difference in the cycle time lies in the accessibility value of the stop-nodes. In fact, increasing too much the route coverage is detrimental for the travel time experienced by passengers, and this impact is stronger when the stop-nodes of the route have a relatively low accessibility value (e.g., few users are in their catchment area). In other words, it is better to provide a more "direct" feeder service, avoiding circuitous routes serving a sparse demand, even though this means renouncing a wider coverage.

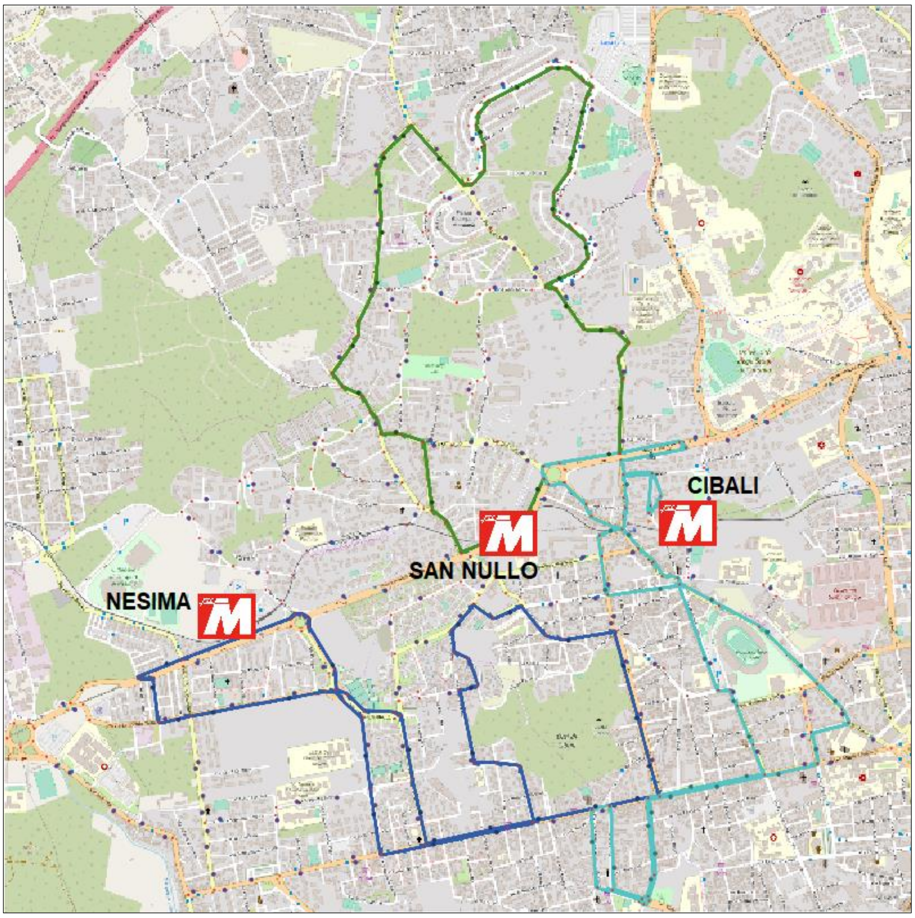


Figure 3.9. Feeder routes linked to the 3 metro stations (T_{des} = 30 min)

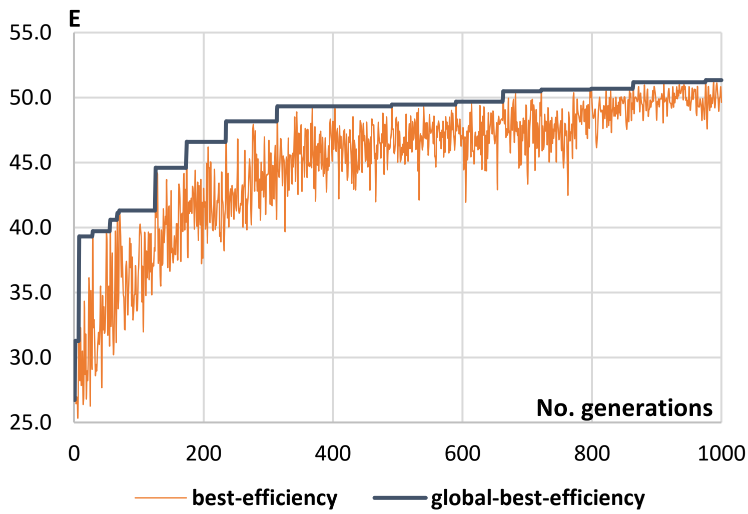


Figure 3.10. Convergence process of the objective function

3.4.2 Rail stations "Ognina – Picanello"

The service area of the two rail stations is smaller than the previous case: it measures 4 km² and includes 130 stop-nodes. Overall, 25800 residents live and 13430 employees work within the study area. The graphical output of the two optimal routes (with T_{des} = 30 min) is shown in Figure 3.11, while Figure 3.12 shows the convergence process of the objective function.

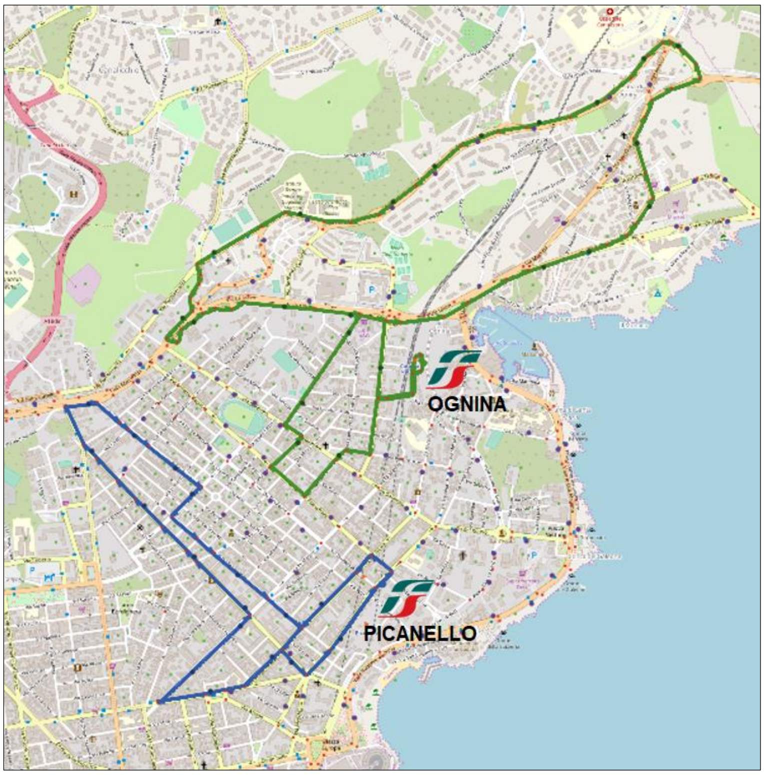


Figure 3.11.

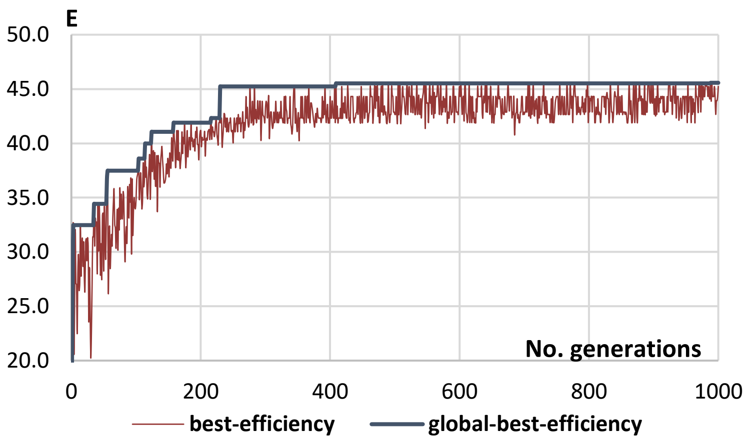


Figure 3.12. Convergence process of the objective function

In detail, the first feeder route linked to Ognina station has a cycle time of 28.0 min, while the second route linked to Picanello station has a shorter cycle time 17.8 min. Also in this case, a longer route would imply more detours and higher travel times to serve a low share of additional demand. Moreover, a short route has the advantage of reducing the service headway while keeping constant the number of vehicles assigned to the feeder route.

3.5 Discussions

This chapter has focused on the development of a solution model for the route optimization problem of feeder bus services, which mainly serve for the last mile leg of travellers, helping to build or reinforce a trunk-feeder scheme. We integrated an ACO algorithm into an ABM environment to design optimal transit routes in low-demand areas, where the transport demand is biased towards MRT stations, and thus, enhancing the effectiveness of the whole transit system.

Thanks to the possibility of dynamically interacting with the simulation environment and visualizing routes on the map, with an easy transferability to other contexts, our model can serve as a practical and flexible tool for public transport planners and companies.

First simulations were carried out regarding the case study of the San Nullo station of the Catania metro line. By varying travel times, several routes have been originated allowing some conclusions to be drawn, regarding both the service coverage and operating costs. In this regard, for the specific case study, a couple of routes with 30 minutes travel time seems a good compromise between the effectiveness and efficiency of the service. Then, the feeder route design problem involving multiple stations was addressed using an updated version of the model, enabling us to find the optimal combination of routes serving different MRT stations located in Catania.

A limitation of the present model is that it is able to suggest the optimal fixed-route configuration when most of the ridership is directed to or come from a transfer station, even though this is rarely valid in mixed-use urban zones. Also, suburban areas can significantly benefit from a flexible, demand-responsive feeder service, able to vary the flexibility of route and schedule according to the period of the day. In reality, most of flexible transit services have some fixed operating schedule (Koffman, 2004), typically limited to departure and arrival times at checkpoints, and the uncertain travel demand in low-demand areas makes it difficult to design reliable deviation services to meet all door-to-door requests, which explains why nowadays flexible operating policies are mostly limited to extreme low-demand areas. In this respect, the model presented in this chapter can be used to understand which locations (e.g., potential fixed bus stops) are the most relevant to be served and which, instead, could be served only via a demand-responsive operation.

CHAPTER 4

Simulating On-Demand Flexible Transit via Agent-Based Modelling

Increase in city population and size leads to growing transport demand and heterogeneous mobility habits. In turn, this may result in economic and social inequalities within the context of rapid economic growth. Provision of flexible transit in fast-growing cities is a promising strategy to overcome the limits of conventional public transport and avoid the use of private cars, towards better accessibility and social inclusion.

This chapter presents the case of Dubai (United Arab Emirates - UAE), where a demand-responsive transit service called MVMANT (a company based in Italy) has been tested in some low demand districts. The contribution of this work relies on the use of an agent-based model calibrated with Geographic Information System (GIS) real data to reproduce the service and find optimal configurations from both the perspective of the transport operator and the community. Different scenarios were simulated, by changing the vehicle assignment strategy and capacity, and comparing MVMANT with a ride-sharing service with smaller vehicles.

Results suggest that route choice strategy is important to find a balance between operator and user costs, and that these types of flexible transit can satisfy transport demand with limited total costs compared to other shared mobility services. They can also be effective in satisfying fluctuating demand by adopting heterogeneous fleets of vehicles. Finally, appropriate planning and evaluation of these services are needed to fully explore their potential in covering the gap between low-quality fixed public transport and unsustainable private transport.

4.1 Motivation and aim of the study

Modern cities are evolving into complex and fragmented systems where the proximity to activities, job places and other opportunities provides a social advantage and an increase of possibility of socialization. One of the main challenges of transport planning is trying to promote sustainable and shared transport modes, and the improvement of social equity, health, resilience of cities, urban-rural connection and productivity of rural and dispersed areas (Gudmundsson et al., 2016).

In cities characterized by fast economic growth, the growth city population and demand for mobility require tailored strategies to satisfy the ever-increasing and diverse transport demand. In these contexts, innovative on-demand mobility, i.e., the one provided by Demand Responsive Transport (DRT) services, can help to bridge the gap between shared low-quality public transport and unsustainable individual private transport. DRT can promote the socio-economic and territorial integration of the residents, favouring the connection with the reference area centres and even providing a more extended and frequent public transport, flexible mobility schemes and feeder services (Ambrosino et al., 2016).

There are some clear trade-offs that should be tackled when planning DRT services. From the operator's point of view, it is important to correctly dimension the fleet, and select the optimal strategy to assign vehicles to passengers' requests, so to perform high load factor (to increase revenues) and low driven distances (to reduce operation costs). This should be done while minimizing the additional time and distances travellers have to experience to assure the expected level of service (user point of view). Such an optimal design of a generic DRT can guarantee its effective management and operation. Therefore, it becomes fundamental to explore the variables of the system that can make the service successful. In this respect, simulation models can provide useful decision support.

Among others, agent-based models (ABM) have proven their suitability to reproduce complex systems characterized by individual agents acting and interacting with the other agents and with the environment (Tisue and Wilensky, 2004). Starting from the micro-interaction among the agents, it is possible to monitor the state of the system and find a balance between conflicting objectives.

The following sections present the results of agent-based simulations using GIS real data to support the planning and operation of DRT services in real-world contexts. In particular, the ABM proposed by Inturri et al. (2019) is applied to the case of a particular type of DRT, i.e., an on-demand flexible route transit service for one district of the city of Dubai.

The aim of the model is to find optimal configurations both from the perspective of the transport operator and of the community. Our contribution is twofold:

1. To prove the suitability of our ad-hoc agent-based simulation environment to reproduce different DRT services and adapt to different contexts.
2. To test an already existing service so to provide useful suggestions for DRT correct planning and management.

In this respect, we believe that this tool is suitable for both for ex-ante and for ex-post evaluation of flexible transport services. This is in line with sustainable urban mobility planning processes, implying continuous monitoring and evaluation of solutions from the beginning.

Since our attention is focused on transport services presenting both a demand-responsive fashion and a high shareability, we deal with levels of flexibility that are intermediate between a pure door-to-door individual transport and the conventional fixed-route public transport. The degree of flexibility affects both the operator and the users of the service. The former makes decisions about the fleet size and composition or the ICT facilities to be equipped with. The latter experience costs related to the different travel time components (not taking the fares into account), which can be strongly variable with the operational strategy. The importance of the user perspective in determining the success of flexible and demand-responsive transit services is underlined by Alonso-González et al. (2020), which estimate the time-reliability-cost trade-offs of users in terms of value of time and value of reliability, the latter being related to the time variability and uncertainty.

In this study, a particular type of DRT, i.e., an on-demand flexible route transit service for one district of the city of Dubai, is tested through an ABM, by simulating the interaction between vehicles travelling along with the road network and users willing to get the transport service to their destinations. Such service has been launched as a pilot by *The Roads and Transport Authority* and was provided by MVMANT (https://www.mvmant.com). It can be configured as a dynamic and flexible transit. It is dynamic since real-time arriving demand requests affect both the way routes are assigned to vehicles and the service of optional stops, resulting in a dynamic routing elaborated just in time. Moreover, the company designed it as flexible because it is composed of fixed and optional routes, which are currently travelled by the DRT only if demand is present, and which have been designed in advance together with the fixed-route, based on-demand patterns and infrastructure constraints. In fact, due to the short distances between the users' desired stops, reduced size of the circuit and constraints of roads' geometry, routes were chosen with a more static approach, making the service more efficient than a pure door-to-door service.

The presented ABM is based on the implementation of:

i. An ad-hoc GIS-based demand model integrated into the simulation environment.
ii. Different dispatching strategies allowing a dynamic matching between vehicles and passengers.
iii. A flexible simulation interface that easily changes variables and monitor the state of the system in real-time.

The model has been presented in Inturri et al. (2019) and applied to the case study of the small-medium city of Ragusa (Italy), where a similar service was tested and where an optimal operation range was found. In this respect, the model can be easily adapted so to reproduce different DRT services and applied to other contexts, allowing to find optimal operation configurations, by monitoring ad-hoc performance indicators. The ABM is able to capture the performance of the DRT, which depends on how vehicles are allocated to passengers, routes and schedules, on the particular topology of the road network and on the spatio-temporal pattern of passenger demand. In the following, the main characteristics of the ABM for the test of MVMANT service in Dubai are presented.

4.2 Methods and materials

The context reference for the case study is the city of Dubai, the largest and most populous in the UAE, laying on the southeast coast of the Persian Gulf and capital of the Emirate of Dubai (Figure 4.1a). In September 2016, the Dubai Future Accelerators selected MVMANT for a 12-week innovation and acceleration program aimed at transforming the city of Dubai into a global testing ground for cutting-edge ideas and technologies. MVMANT is an urban mobility platform which enables the deployment of a dense fleet of vehicles circulating on a fixed-route and the forecast of the mobility flow, coupled with the requests in real-time generated by customers. MVMANT worked closely with the local government transport authority to the implementation of its smart mobility solution in the city of Emirate. The service was tested in two districts located in the vibrant heart of the city - i.e., Al Barsha 1, one of the neighbourhood of the Al Barsha district, where life gravitates around the "terminus" Mall of Emirates (including both a commercial centre and a metro station), characterized by users moving mainly for shopping and leisure (Figure 4.1b); Dubai Internet City, which gravitates around the "terminus" Nakheel MS Seaside (a metro stop), refers to a business centre area, and is characterized by home-work journeys. Our work aims at reproducing MVMANT service via the ABM, evaluating the performance of different vehicle dispatching strategies and service configurations. Simulations have been performed for the district of Al Barsha 1, where the service covered an area of about 3 km² and has been particularly successful in connecting the different parts of the district with each other and with the main point of interest, i.e., Mall of Emirates.

The demographics available for the area, supplied by the local authorities to the MVMANT company, refer to 2014; the resident population of the area is approximately 15.000 inhabitants (Figure 4.1c), while the number of employees is higher (approx. 18.000, Figure 4.1d) mainly due to the presence of the Mall.

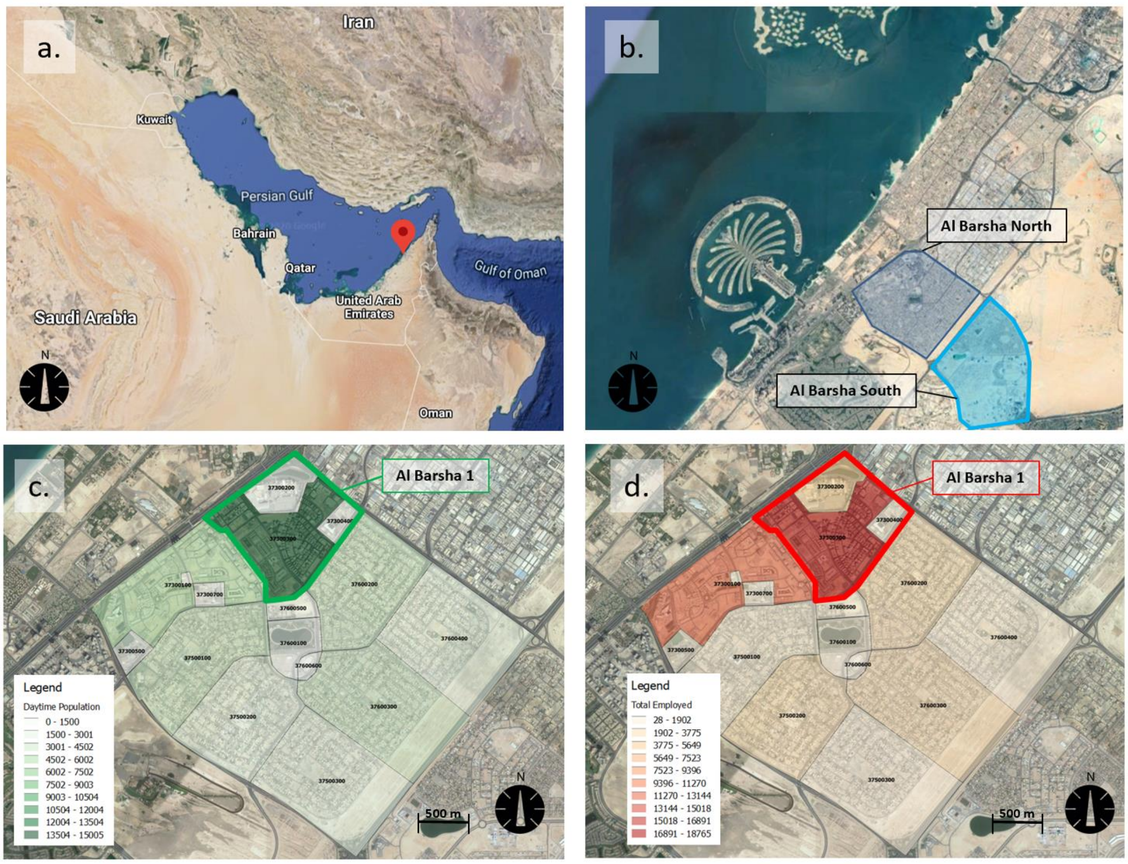


Figure 4.1. (a) Location of Dubai on the Arabian Peninsula; (b) Al Barsha district and its subdivision; population (c) and employees (d) in 2014 in Al Barsha 1.

The ABM used in this study has been built within the NetLogo agent-based programming and integrated modelling environment (Wilensky, 1999). The main features of the model are the (a) transport network, (b) demand model, (c) agent (user and vehicle) dynamics, (d) assignment of route choice strategies and (e) indicators to evaluate the model's performance.

4.2.1 Transport network and demand model

The model's network, based on the real road network, is built with a fixed-route and different optional routes, composed of links, stop nodes and diversion nodes (to skip from fixed to optional routes). Through the GIS extension of NetLogo, a GIS dataset is used both to build the actual road network and implement a georeferenced real dataset of origin-destination (OD) requests, collected during the MVMANT service's first weeks of tests.

The demand model is based on previous works developed by Inturri et al. (2019). However, in this work, the demand model has been improved by using real-world observed data for trips origin and destinations. A trip request is randomly generated in the network following the Poisson distribution with an average trip rate ATR, suited for simulating booking processes, particularly when demand rate is low, and requests can be considered independent (Coffman et al., 1999). The trip rate $TRij$ from an origin i to a destination j is calculated via a gravitationally distributed model as in Equation 4.1:

$$TR_{ij} = TR_i \times p_{ij} = \frac{TOi}{\sum_i TOi} \cdot ATR \cdot \frac{TD_i \cdot d_{ij}^{\alpha} \cdot e^{-\beta d_{ij}}}{\sum_k TD_k \cdot d_{ik}^{\alpha} \cdot e^{-\beta d_{ik}}} \tag{4.1}$$

where TRi is the generation trip rate of zone i, and p_{ij} is the probability that a trip with origin i has destination in j. TR_i is proportional to the density of trip origins (TO_i) and to an average trip rate (ATR) set at the beginning of the simulation (see Equation 4.2)

$$TR_i = \frac{TOi}{\sum_i TOi} \cdot ATR \tag{4.2}$$

while p_{ij} is calculated via Equation 4.3:

$$p_{ij} = \frac{TD_i \cdot d_{ij}^{\alpha} \cdot e^{-\beta d_{ij}}}{\sum_k TD_k \cdot d_{ik}^{\alpha} \cdot e^{-\beta d_{ik}}} \tag{4.3}$$

where TD_j is the number of trip destinations located at zone j, d_{ij} is the distance from i to j, α and β are the parameters of the decay function $f(d)$ in Equation 4.4:

$$f(d) = d^{\alpha} e^{-\beta d} \tag{4.4}$$

4.2.2 Agents (passengers and vehicles) dynamics

A trip request of a passenger group (with a maximum prefixed size) is stochastically generated, according to the demand model. The user group's trip request can assume different status:

- "Rejected", if the distances from the origin destination (OD) stops overcome a prefixed threshold.
- "Waiting", if the OD pair is within the distance range, and the users' group moves to the nearest stop.
- "Satisfied", when a vehicle with an appropriate number of available seats reaches the stop, each user boards and alights at the nearest stop to its required destination; this is an indicator of the satisfied demand of service within the accepted requests.
- "Unsatisfied", if no vehicle reaches the passenger group within a maximum waiting time.

At the beginning of the simulation, the number of vehicles, their seat capacity and their speed are set, and vehicles are generated at random stops. As far as the vehicle dynamics is concerned, each vehicle starts travelling along the fixed-route until it stops where waiting users are loaded following the first-come-first-served queue rule, updating available vehicle's seats.

4.2.3 Route Choice Strategies

In our model, vehicles can be assigned to different Route Choice Strategy (RCS), i.e.:

- FX - "fixed-route", each vehicle drives only on the fixed-route.
- FR - "fully random", each vehicle at a diversion node randomly chooses to go on the flexible route; approximately, half of the vehicles will drive on the flexible route, and the other half will keep on driving on the fixed-route, until the next diversion node is reached.

- AVAR - "all vehicles drive on all flexible routes", each vehicle is allowed to drive on a flexible route, but it chooses to do so only when passengers have to alight at a stop located along the flexible route, or she is waiting at such a stop, or according to a given random component (probability to choose the flexible route even though none of the previous two situations happens).
- EVAR - "each vehicle is assigned to a flexible route" with a prefixed percentage of vehicles assigned at random.

FR, AVAR and EVAR strategies can have a randomness component. This component refers to the possibility that the vehicle would follow a different route from the one indicated by the RCS. For example, in the case of AVAR strategy, where all vehicles always drive on the flexible route, it could happen that a percentage of vehicles' trips would skip driving the flexible route (even if demand is present). This is to test the role of randomness, since it has been demonstrated that it can increase the efficiency of social and economic complex systems (see Pluchino et al. (2010)).

4.3 Results

Dubai's network used for MVMANT service was reproduced in the ABM (Table 4.1) with fixed (dark blue) and flexible (pink) routes (Figure 4.2a).

Table 4.1. Values of the input variables in the different scenarios

Routing strategy	Length (m)	Time (s)
ER - extended route (fixed + optional)	3944	700
FX - fixed-route (fixed)	3053	580
Difference	891	120

From the analysis of OD data collected during the first weeks of MVMANT service (while it still was in operation), it was possible to derive a peak hour demand rate in the range between 50-60 requests per hour. For scenario simulations, a demand rate equal to 100 requests per hour was chosen to simulate a higher demand that, in principle, could be satisfied by the service. The GIS zoning implemented in the simulation is based on a squared grid layer (100m x100m) where each zone is assigned with the information related to the number of OD requests (Figure 4.2b); the demand model then takes into consideration the requests generated at a maximum of around 300 m from the centroid of the zones (Figure 4.2c,d).

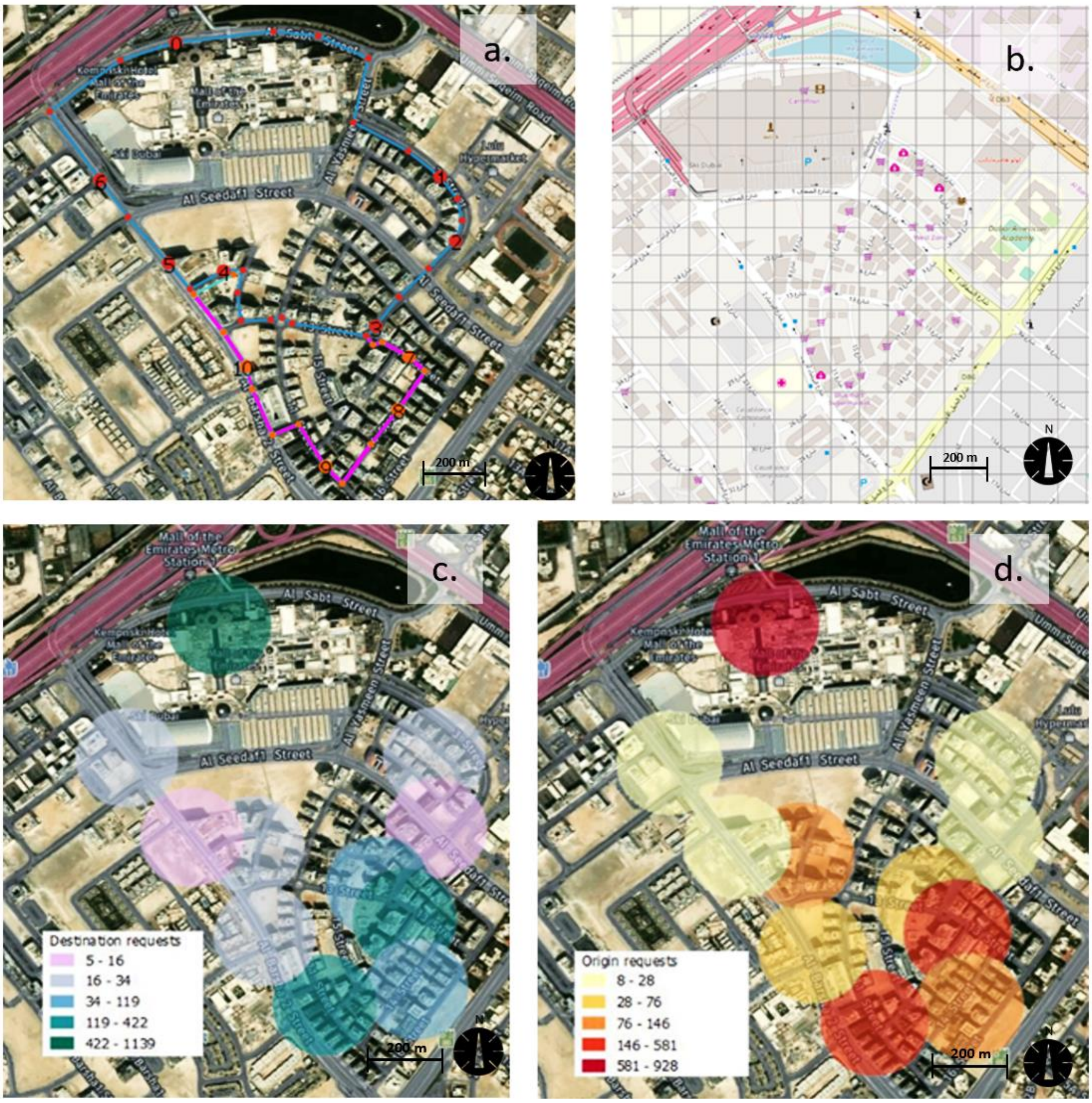


Figure 4.2. MVMANT service with fixed (dark blue) and flexible (pink) routes; (b) Geographic Information System (GIS) zoning; requests for the catchment area of (c) destination and (d) origin.

4.3.1 Scenario 1: test of system efficiency based on RCS

The first set of simulations considered all the different RCS with increasing levels of randomness, to test routes and the overall system performance. Three types of input variable were considered: service and demand variables, and RCS. They are summarized in Table 4.2.

A set of performance indicators (Table 4.3) has been analysed during the simulation to test the impact of different vehicle dispatching strategies on service efficiency and effectiveness.

Table 4.2. Input data used for Scenario 1.

Type of variable	Input variables	Input data
service variables	total simulation time (ST)	6 h
	number of vehicles (nv)	2
	vehicle maximum capacity (cap)	19 seats
	vehicle average speed	30 km/h
	vehicle operation cost (VOC)	0.9 €/h
	driver cost (DC)	20 €/h
demand variables	demand rate (request/hour)	100
	maximum no. of passengers per demand	1
	maximum waiting time (min)	10
	value of time (VOT)	10 €/h
RCS	Type of strategy (FX, FR, EVAR, AVAR) Level of Randomness (0-30%)	FX 0% ER 0% = EVAR 0% FR 100% AVAR 0–30%

Table 4.3. Performance indicators

Acronym	Indicator	Unit
%–r	percentage of rejected requests	%
%–s	percentage of satisfied users	%
NP	total number of passengers transported	adim.
TDD	total driven distance	km
APTD	average passenger travel distance	km
ALF	average vehicle load factor	adim.
AWT	average passenger waiting time	min
AoBT	average passenger on-board time	min
APTT	average total travel time	min
TPTT	total user travel time (with a penalty of 60 min for each unsatisfied user)	h
OC	Operation cost (depending on simulation time, TDD, driver cost, VOC)	€
TUC	total unit cost	€/pax

Indicators evaluate the quality of transport service, its sustainability and the overall benefit brought to society. They are chosen to capture the different objectives and points of views of the system actors - i.e., (1) a user is interested in reducing the trip cost (length, travel time, fare); (2) a private company providing the service is interested in maximizing the profit, by increasing the number of passengers within a prefixed travelled distance or, conversely, reducing the amount of travelled distance to serve a prefixed demand; (3) the community is interested in reducing transport externalities, i.e., pollution and congestion.

The total unit cost (*TUC*) indicator is evaluated according to Equation 4.5, taking into account the total passenger travel time total user travel time (*TPTT*) (h), the value of time *VOT* (€/h) for passengers, and the vehicle operation cost (*OC*) (€):

$$TUC = \frac{TPTT \cdot VOT + OC}{NP} \tag{4.5}$$

It considers the cost of users (with the *VOT* and the *TPTT*, which users would like to have as low as possible) and the cost of the operator per each transported passenger, as follows:

$$OC = n_V \cdot DC \cdot ST \; + \; \mathrm{TDD} \cdot \mathrm{VOC} \cdot (\mathrm{cap}/15) \tag{4.6}$$

Therefore, the *TUC* can be seen as a unit cost for the transport system (demand and supply) as a whole.

From the comparison among the different RCS, it is possible to evaluate the efficiency of each route in relation to the performance indicators.

<u>Operator profit</u> (total driven distance (*TDD*) and average vehicle load factor (*ALF*)). From the point of view of the operator, it is important to achieve high load factors (to increase revenues) and low driven distances (to reduce operation costs); those two factors can be measured through *TDD* and *ALF* indicators, which have their best values for AVAR strategy (Figure 4.3a). This means that, for the case study under consideration, and from the point of view of the operator, assigning all vehicles to the flexible route, but only when demand is present, would be the best way to satisfy more demand while minimizing the travelled distance. This will also have a direct impact in terms of carbon and local pollutants emissions. Note that the higher *TDD* occurs for the *FX* strategy, despite the shorter distance of the fixed route. This can be explained considering the smaller number of served passengers, resulting in a low *ALF* (Figure 4.3b) and in less dwell time at stops.

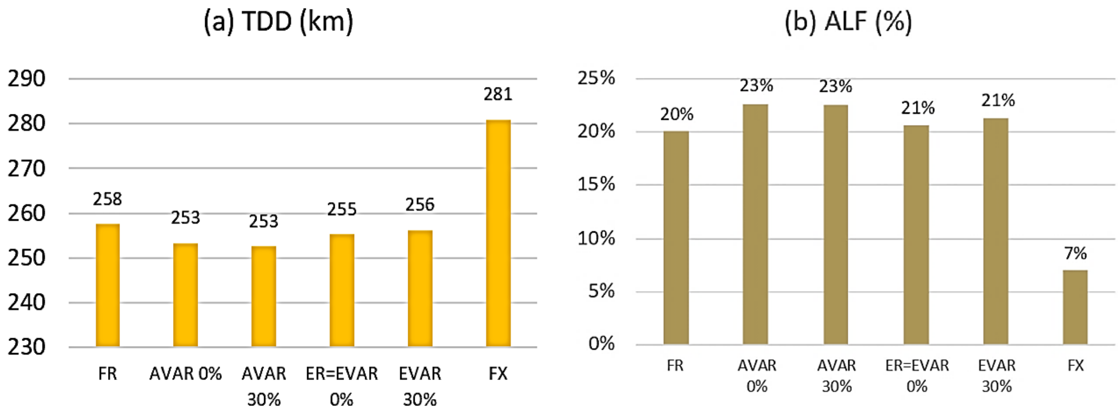


Figure 4.3. (a) Total driven distance and (b) Average Load Factor, according to the different strategies.

<u>Service coverage</u> *(%-r)*. With regard to the percentage of rejected requests (Figure 4.4a), it is possible to see how, in comparison to any other strategy, FX reports a potential loss of 33% of requests in the catchment area. This shows the importance of the introduction of the optional route in order to expand the accessibility of users in the study area.

<u>Demand satisfaction</u> *(%-s)*. Demand satisfaction keeps above the 95% for all strategies (Figure 4.4a), indicating the possibility of extending the service to a greater catchment area, without changing operation costs.

<u>Total unit cost</u>. If one considers both the point of view of the operator and of passengers, it results that the worst strategy in terms of TUC is FR, while the other performs similarly (Figure 4.4b). The low value assumed for the FX strategy clearly depends on the fact that routes are shorter, implying a lower value of total passenger travel time. While, if one considers the environmental impact, then AVAR should be preferred, since it implies less travelled distance and higher load factor. In this respect, some carbon credit mechanisms could be defined from policymakers to take into considerations environmental concerns (see Anand et al., 2019). However, the FX strategy would clearly provide more than 50% of rejected requests, as shown in Figure 4.4a, and this would be socially unsustainable. While, for an equally good value of coverage, the best strategy in terms of TUC is the EVAR with 0% of randomness, corresponding to ER.

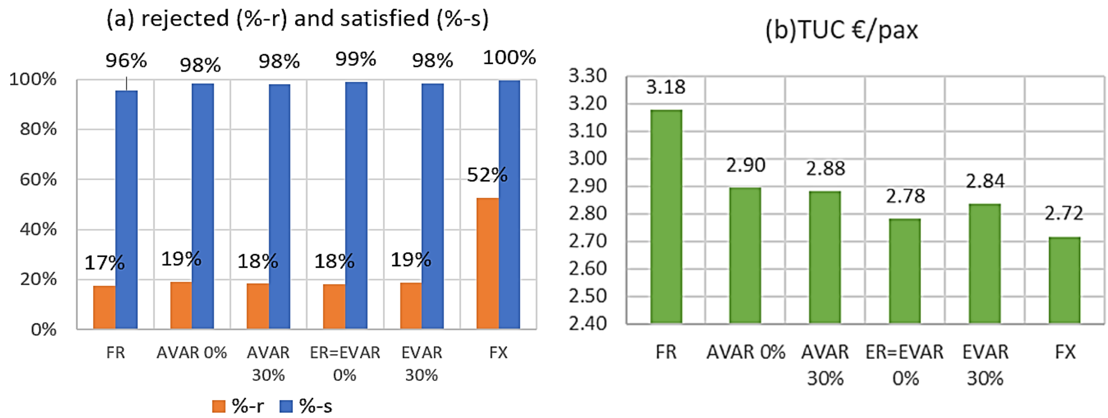


Figure 4.4. (a) Percentage of rejected requests and satisfied users and (b) Total Unit Cost, according to the different strategies.

4.3.2 Scenario 2: comparison to other DRT services

This second set of simulations tests different DRT services in terms of vehicle capacity and evaluates the economic efficiency of each system in relation to TUC. Scenario 2 takes into account two services with the following characteristics:

- A flexible transit with minibus (comparable to the actual MVMANT service).
- A ridesharing with small vehicles, capable of collecting up to two requests at a time (comparable to typical ride sourcing operators providing services like a shared taxi).

The main input data are summarized in Table 4.4; unlike the other tested scenarios, in Scenario 2, the number of vehicles will be one of the main outputs.

Table 4.4. Input data used for Scenario 2

Type of variable	**Input variables**	**Input data**	
		Ride-sharing service	**flexible transit service**
service variables	total simulation time (h)	6	6
	vehicle average speed (km/h)	30	30
demand variable	demand rate (request/hour)	100	100
	maximum waiting time (min)	6	10
	number of passengers per demand	1	1
performance constraint	percentage of satisfied users - %-s	> 85%	> 85%
route choice strategy	RCS	EVAR with 0% randomness	

Number of vehicles. The first output of simulations is that, in order for the ride-sharing service to achieve at least the 85% of satisfied users, six small vehicles are needed; two vehicles with a capacity equal to 19 seats are needed for the minibus system.

Additional time and Distance. In order to assure the expected level of service, operators should minimize the additional time and distances travellers have to experience. One can consider the additional travel distance (*ATD*), depending on average passenger on-board time (*AoBT*) and average vehicle speed (as a function of *TDD*, simulation time and number of vehicles), and the additional travel time (*ATT*), related to average passenger waiting time (*AWT*) and *AoBT*. In the comparison between the two services, passengers of the minibus travel on average the same distance than those travelling by the ride-sharing service, with an additional time of three minutes (Table 4.5). This implies that the two services are comparable in terms of the additional distance experimented by a passenger when sharing a trip with others, but the flexible transit achieves less *TDD* because of the reduced numbers of vehicles used. Of course, this can be easily converted into pollutant emissions.

Table 4.5. Additional travel distance and additional travel time and TUC according to the different services.

Service	N. of vehicles	ATD (km)	ATT (min)	TUC (€/pax)
Ridesharing	6	2.19	7.44	6.01
Flexible transit	2	2.26	10.43	2.90

Total Unit Cost. The *TUC* of the ride-sharing service assumes a value equal to more than the double of the same service performed via a minibus (Table 4.5). This implies that the supposed transport demand can be easily satisfied by a flexible transit service rather than a shared ride-sharing service with more vehicles of lower capacity.

4.3.3 Scenario 3: optimum vehicle capacity based on demand fluctuation

This third set of simulation analyses the efficiency of the system as the number of available seats varies, and according to the fluctuation of demand during the week. Based on the analysis of the real-word data coming from the pilot, a demand rate of

50% of the weekdays has been adopted for holidays. Vehicles with 19 seats (corresponding to the MVMANT service), eight seats and four seats have been used to perform the simulations. The main input data are summarized in Table 6.

Table 4.6. Input data used for Scenario 3.

Type of variable	Input variables	Input data
service variables	total simulation time (h)	6
	number of vehicles	2
	vehicle maximum capacity (seats)	4 – 8 – 19
	vehicle average speed (km/h)	30 km/h
demand variable	demand rate weekdays (request/hour)	100
	demand rate holidays (request/hour)	50
	maximum waiting time (min)	10
	number of passengers per request	1

Transport intensity (*TI*). This indicator represents kilometres travelled for each passenger (km/pax) and companies providing the service are interested in having low values of *TI*. The main result of this analysis is that a 19-seats vehicle is good for the weekday demand rate, while the 8-seats vehicle is the best solution during the holidays (Figure 4.5).

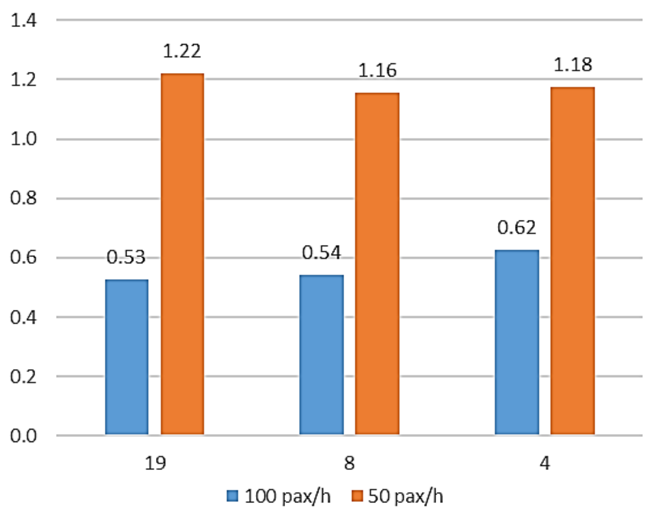


Figure 4.5. Travelled kilometres per passenger during weekdays and holidays.

This suggests that, in order to be efficient, the service could be tailored according to the fluctuating demand, by using a heterogeneous fleet of vehicles. This also has interesting implications in terms of environmental impact, with low-capacity vehicles emitting less pollution than high-capacity ones. Once again, incentive mechanisms or restrictive policies (i.e., a limited traffic zone) could be useful to include environmental concerns in the planning and operation of such services.

4.4 Discussion and conclusions

This chapter presented the results of agent-based simulations of the on-demand flexible transit service called MVMANT, experimented in 2019 in the city of Dubai. The aim was to test its performance by reproducing the service in the district of Al Barsha 1, changing important service characteristics, like route assignment and vehicle capacity, and comparing it with alternative shared transport services.

An already implemented ABM was used to this purpose, calibrating it via georeferenced real dataset including OD couple requests collected during the first weeks of the service. Results demonstrates the flexibility of the model to reproduce different services and contexts and provide useful suggestions for its correct planning and management. Results of the simulations of this flexible transit service highlighted:

1. The importance of introducing an optional route in order to expand the accessibility of users in the study area.
2. The possibility of extending the service to a greater catchment area, without changing operation costs.
3. The importance of dynamically scheduling vehicles to passengers and routes and its reflection on economic performance parameters.

Besides, the comparison of the minibus flexible transit service provided by MVMANT with a ride-sharing service with small vehicles showed that the former is more convenient than the latter in terms of a total unit cost accounting both for operator and user points of view, in the hypothesis of guaranteeing a minimum demand satisfaction level of 85%. Finally, fluctuation of demand during the week was considered in the simulations, showing the efficiency of a 19-seats minibus during the weekdays, while suggesting the adoption of a smaller vehicle (with eight seats) during holidays.

From an environmental point of view, some useful suggestions can be derived from the results. In this respect, policymakers should adopt push and pull measures so to include environmental concerns in the planning and operation of such flexible services. As an example, operators can decide to use a heterogeneous fleet based both on service efficiency and on policy restrictions.

Other useful policy implications could be derived by changing other service variables. As an example, it would be interesting to understand how this flexible service would perform if considering a "delay" time at the terminus "Mall of Emirates", as a standard bus. In this respect, from a simple evaluation based on real data from the pilot, it resulted that the introduction of a 3-minute delay time of the vehicles at the terminus would lead to a reduction of about the 15% of necessary rides (and thus, of travelled kilometres). The increase in service efficiency with the introduction of a delay time at the terminal could be due to the greater number of passengers that the service will be able to serve, but also because, being a low-demand area, the vehicle would travel less kilometres while empty. It would be interesting to verify this scenario via the ABM, evaluating its feasibility, taking into account the increase in passenger travel time.

Future research should also aim at improving demand estimation, by looking at the preferences of users for different service configurations (see Atasoy et al., 2015) and include more direct environmental evaluations, both related to global emissions (i.e., greenhouse gases) and to local pollutants (e.g., PM10, NOx). A combination with other evaluation approaches, like stated preference surveys or multicriteria decision-making techniques, would be suitable both to investigate agent preferences, and to evaluate different scenarios according to multiple criteria of judgment (environmental, social and economic). In this respect, an analysis of different options from the sustainability lens would be interesting, especially in comparison with private mobility, starting from a rough estimation of the travelled distances (per passenger) of private vehicles satisfying the same demand. Finally, it would be important to understand how this type of service should be integrated with other transport services, in order to guarantee a seamless transport system following the approach of mobility as a service (MaaS).

All these evaluations are useful to support decision-makers in planning, management, and optimization of innovative transit services, which have a great potential to reduce car dependency and increase accessibility - especially in cities characterized by fast economic growth, resulting in heterogeneous mobility needs and economic and social inequalities.

CHAPTER 5

Comparing Fixed-route and Flexible Transit Feeder Services: an Agent-Based Model

This chapter presents a new agent-based model able to simulate innovative flexible demand-responsive transport services, specifically thought to solve the last-mile problem of mass rapid transit (MRT). This is particularly needed in areas characterized by insufficient transit supply and lower sprawled demand, where conventional public transport is not always able to meet the population's need for accessibility to jobs, education, health, and other opportunities in terms of routes and frequencies.

However, new technologies have the potential to dynamically couple demand with supply, providing forms of flexible transit which are well suited to act as a feeder service to high-capacity transit lines.

The model compares the performances of two feeder services, one with fixed routes and stops, and the other with flexible routes and stops activated by the requests of users, satisfying the same demand.

The case study city is Catania (Italy), where such services could increase the ridership and coverage of a 9 km long metro line that connects the city centre to peripheral areas. Different scenarios have been analysed by comparing a set of key performance indicators based on service coverage and ridership.

The first results allows us to identify optimal operation ranges of flexible on-demand services and pave the way for further investigation needed to understand their acceptability and economic viability.

5.1 Introduction and related works

Transport systems are experiencing times of unprecedented changes. This can be attributed to the innovations brought by new information and communication technologies, which enable flexible services, spreading e.g., as complementary to conventional public transport or in substitution to it (Cohen-Blankshtain and Rotem-Mindali, 2016; Sadowsky and Nelson, 2017).

The importance of studying in advance the potential of such new services and their optimal range of operation has been already underlined in the previous chapter. Such mobility options would be specifically needed in areas where both insufficient transit supply and lower sprawled demand make it difficult to provide mass transit services.

The issue of covering the first/last mile of mass rapid transit is a case in point, being a Many-to-One problem characterized by multiple origin/destination with a low and dispersed demand and a single destination/origin with a concentration of demand (Calabrò et al., 2020). In this case, two main design choices appear, i.e.:

a. The choice between scheduled feeder services (as in conventional public transport) and demand-responsive transit (DRT) services in which routes are dynamically adapted to the users' requests (Koffman, 2004), trying to guarantee maximum flexibility while complying with operating costs.
b. The level of flexibility (in terms of routes, stops and scheduling) of such DRT services.

Both alternatives (fixed and flexible) have their own design questions. A fundamental one for scheduled feeder services is the optimal design of routes and frequency. For flexible services, the design issue to address is what degree of flexibility, from the operator perspective (e.g., routing and dispatching strategies, fleet composition) and the passenger point of view (e.g., user engagement in the booking process) best exploits the trade-off between minimizing the cost of the system and maximizing service quality.

Literature on modelling approaches to study flexible and demand-responsive transport (DRT) services is abundant (Liyanage et al., 2019). In particular, ABM has been largely used thanks to the possibility to simulate complex environments with individual autonomous agents acting and interacting according to their objectives. This is well suited to reproduce DRT services, characterized by real-time user requests and the need to match them with vehicles in an optimal way.

Recently, Di Maria et al. (2018) proposed a modular simulation framework for autonomous mobility on demand and focused on the important issue of optimization strategies using the Manhattan Grid case as a testbed.

Inturri et al. (2019) presented a multi-agent simulation to reproduce a mixed fixed/flexible route transit service with different fleet size and vehicle capacity in the city of Ragusa (Italy), showing an optimal range of operating vehicles that minimizes a total unit cost indicator, accounting both for passenger travel time and operation cost. Giuffrida et al. (2020) extended the results of Inturri et al. (2019), studying the effects of different vehicle assignment and route strategies and comparing its performance with a ride-sharing service provided via low-capacity vehicles.

Some authors have focused on the first and last mile (FMLM) problem of mass rapid transit. Scheltes and Correia (2017) studied the so-called “Automated Last-Mile Transport” via an agent-based simulation model whereby a dispatching algorithm distributes travel requests amongst the available vehicles using a First-In-First-Out (FIFO) sequence and selects a vehicle based on a set of specified control conditions (e.g., travel time to reach a requesting passenger). However, this type of service does not allow shared trips among passengers, which would increase the complexity of the modelling effort. Besides, while solving FMFL issues, it is important to understand which level of flexibility is needed according to demand patterns.

In this chapter we aim at contributing to the current literature in this field by presenting a new ABM to simulate flexible/fixed feeder services with different vehicle fleets and demand patterns, to help solve the FMLM problem of mass rapid transit. The remainder of the chapter is organized as follows. The following section presents the methodology, drawing complete description of the model, outlining the demand-supply interaction and the proposed dispatching algorithm, and describing the performance indicators. In Section 5.3 we apply the model on a real case study, describing the main results. Finally, Section 5.4 concludes the chapter, introducing some future research directions.

5.2 Methodology

The present work is built on and extend the works of Inturri et al. (2019) by allowing for different levels of flexibility, Scheltes and Correia (2017) for the passenger and vehicle dynamics, while allowing for ride sharing, and Calabrò et al. (2020a) by reproducing the operation of a feeder service with optimally designed routes. The model also allows for a more detailed spatial representation of the demand compared to the previous ones, since requests are geocoded to the building scale. The model is specifically designed to compare the performance of the two alternative feeder services, while satisfying the same demand.

5.2.1 Description of the model

The rationale for using ABM is to understand the trade-off between costs and the level of service of feeder services, taking flexibility as a design parameter, while simulating different vehicle fleet capacity and demand patterns. The ABM has been implemented in the NetLogo programming environment (Wilensky, 1999), and takes as reference other previously implemented models (Scheltes and Correia, 2017; Inturri et al. 2019; Calabrò et al, 2020a; Giuffrida et al., 2020). A brief description of the model is provided in the following paragraphs.

Transport network model. The network consists of mandatory stops and optional stops, to encompass both fixed-route feeder (FRF) and demand-responsive feeder (DRF) which routes are built on the real network.

Demand model. The GIS extension of NetLogo is used to map the distribution of socio-demographic data (residents and employees) at a census zone level. A further level of disaggregation is achieved by assigning socio-demographic data to each building proportionally to their surface, whose data were obtained through OpenStreetMap.

The average trip demand rate is based on historical data of the daily distribution of passengers' accessing/egressing the metro station. The service has been simulated for the current demand, but also for higher and lower potential demand, to test the efficiency of the feeder services under different demand rates.

A users' group trip request is generated according to a gravitationally distributed probability from an origin (O) building to the metro station and from the metro station to a destination (D) building, following a Many-to-One demand pattern. The demand model is based on Inturri et al. (2019) and it has been improved through the introduction of an index of attractiveness of the transit mode versus the walking mode to reach the terminal station.

Given a set of n buildings, the trip rate TR_{ij} (where i or j corresponds to the terminal station) is calculated with Equation (1), where TR_i is the generation trip rate from (and to) the building i, proportional to population density and an average trip rate per trip direction (ATR) (simulation variable), calculated with Equation (2), and η_{ij} is the transit index of the attractiveness of the transit mode, which assumes values between 0 and 1, determined for each building i through the exponential function shown in Equation (3).

$$TR_{ij} = TR_i \cdot \eta_{ij} \tag{5.1}$$

$$TR_i = \frac{Pop_i}{\sum_{k=1}^{n} Pop_k} \cdot ATR \tag{5.2}$$

$$\eta_{ij} = 1 - e^{-\frac{\left(d_{ij}-d_{T0}\right)^2}{0.5\, d_{T0}^2}} \tag{5.3}$$

where d_{T0} is the minimum distance from the terminal station to consider the transit service attractive for the users. For distances shorter than d_{T0}, users are assumed to walk directly to the terminal station.

A trip request of a passenger group (with a maximum prefixed size) is generated according to the following rules: (i) from buildings to metro: stochastically generated and Poisson distributed according to the hour trip rate; (ii) from metro to buildings: Poisson distributed, every 10 minutes (which is the headway of the metro service).

Access mode choice. We do not consider private car use, but users have a twofold choice to reach the metro station, i.e., walking and transit. In this respect, we simulate a DRF service with different fleet configurations in comparison with a FRF service while keeping the demand constant.

5.2.2 Fixed-route feeder dynamics

After a trip request is generated, if the distance from the origin to the nearest feeder bus stop (or from the destination, if the origin is the metro station) overcomes a given threshold, the group of travellers assumes the status "rejected". This is because it is assumed that a traveller may decide not to use transit due to excessive access time and will use other modes. Otherwise, the request is confirmed, passengers assume the status "accepted" and move to the stop that allows them to minimize the sum of walking time and on-board time, assuming the status "waiting", while waiting for the feeder service.

If a prefixed maximum waiting time is overcome before a vehicle reaches the stop, each group of travellers gives up and assumes the status "unsatisfied". Otherwise, each passenger boards the vehicle assuming the status "satisfied".

If the overall travel time overcomes a certain desired travel time (given by the vehicle maximum travel time t_{max} multiplied by the index of attractiveness η_{ij}), the passenger assumes the status “delayed”.

Fleet size, vehicle capacity, and speed are set at the beginning of the simulation. Each vehicle is generated at the terminal stop (i.e., the metro station). The FRF vehicle travels along the route until it reaches a stop. Passengers at their destination stop alight and waiting users board the vehicle following the First-Come-First-Served (FCFS) queue rule, but only if the traveller group size is not greater than the available seats, updating vehicle’s available seats.

5.2.3 Demand responsive feeder dynamics

Traveller requests for the DRF service can be served at multiple potential stops either close to origin or destination (according to the vehicles’ availability and schedule). As in the previous case, after a trip request is generated, if the distance from the nearest stop to the origin exceeds a given threshold, the group of travellers assumes the status “rejected”. Otherwise, the request is processed through the dispatching algorithm and the passenger group can be assigned to a predetermined stop and vehicle according to capacity and time constraints as fully explained in the next subsection.

If no vehicle can fulfil these constraints, each traveller of the group assumes the status “rejected”. If accepted, the dynamics of passengers originated at the metro station and those whose origin is one of the buildings follow different rules. In the first case, passengers wait for the assigned vehicle at the metro station, board it, and finally get off the vehicle at the predetermined stop, walking to their destination located at one of the buildings. In the second case, passengers do not go to any stop until the expected time for pick-up, also given the required walking time and an additional “buffer” time (in case the vehicle is earlier than the scheduled arrival time).

Then, the group moves to the assigned stop assuming the status “waiting”, while waiting for the vehicle. If a prefixed maximum waiting time is reached before a vehicle arrives at the stop (e.g., schedule variations and increased travel times of the vehicle due to other following requests), the passenger group gives up and assumes the status “unsatisfied”. Otherwise, each passenger boards the vehicle, alights at the metro station, and assumes the status “satisfied”.

However, if the overall travel time overcomes the desired travel time (given by the vehicle maximum travel time t_{max} multiplied by the index of attractiveness η_{ij}), a passenger assumes the status "delayed". The following flow charts summarize the traveller (Figure 5.1) and vehicle dynamics (Figure 5.2) for the flexible DRT feeder service.

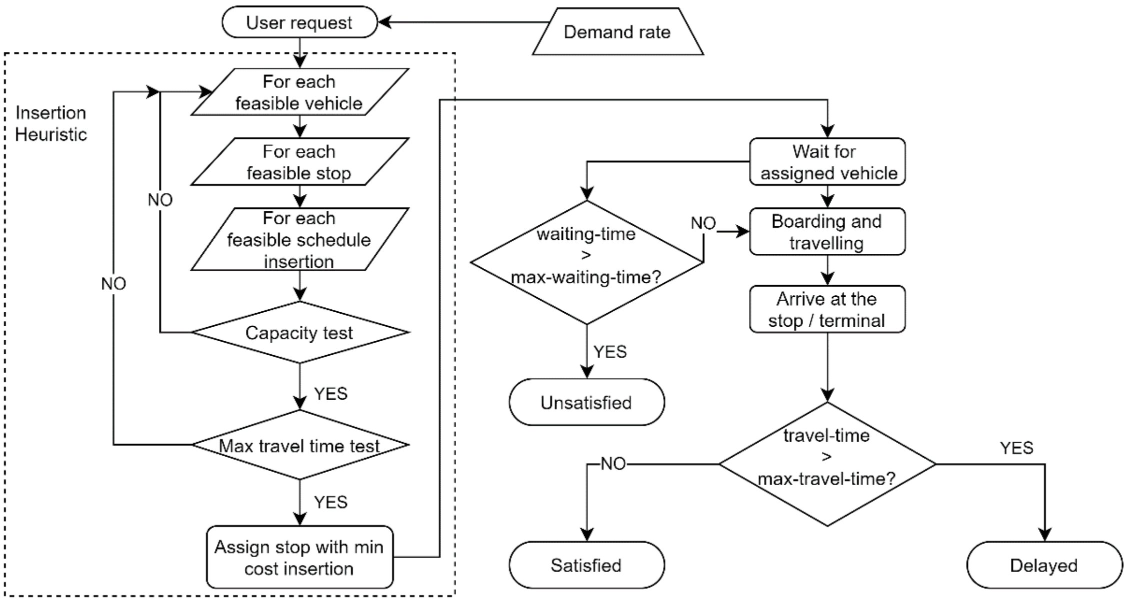


Figure 5.1. DRF traveller dynamics flow chart.

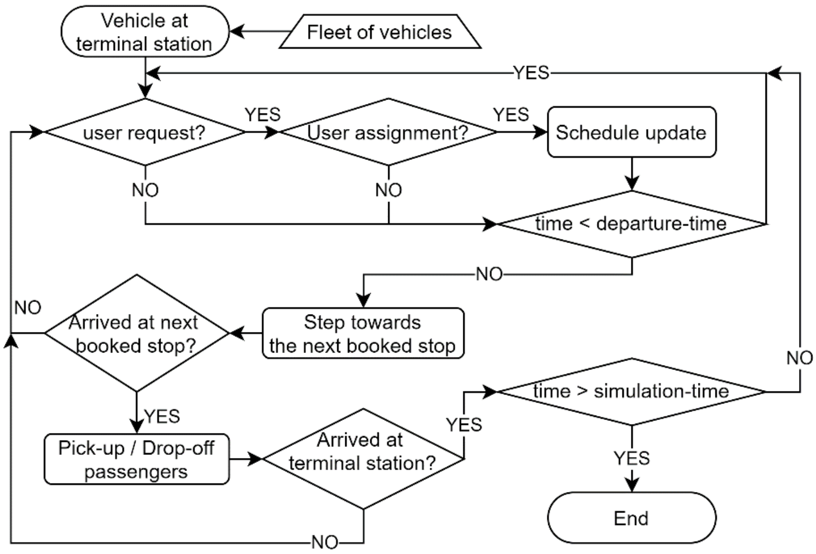


Figure 5.2. DRF traveller dynamics flow chart.

Vehicles of the DRF service start travelling from the metro station at a scheduled departure time across the street network and towards the pre-scheduled mandatory stops (hereafter called waypoints).

Every time that a new request is accepted and assigned to the given vehicle, its schedule is updated with the possible insertion of a new stop to be served between two already scheduled stops. Vehicles drive to pick-up passengers at their origin stop and/or to drop-off passengers at their destination stop.

5.2.4 The dispatching algorithm

As previously said, when a new request is generated for the DRF, an insertion heuristic algorithm is used to determine the feasibility of the insertion, and then the minimum cost insertion of the request in the current schedule of one of the vehicles. Since we deal with a dynamic procedure, vehicles can take new requests in accordance with the maximum travel time t_{max} and/or the maximum capacity constraint according to the FCFS rule and minimizing its cost function.

The main novelty of the procedure lies in the three levels of explorations of feasible solutions. For the insertion of a new stop, the algorithm examines:

i. Each vehicle v of the fleet, considering the list of already scheduled stops, arrival time at each stop and available seats after serving a stop.
ii. Each potential stop s, within a maximum radius of walking distance from/to the origin/destination of the travel request.
iii. Each possible insertion of s between two any subsequent stops belonging to the current schedule of the vehicle v, if s is not already scheduled. The feasibility of each combination of vehicle, stop and insertion location is evaluated by ensuring that it complies with the following constraints: (i) the extra ride time needed to serve stop, also considering the additional time lost during pick-up/drop-off operations, must not be higher than a certain threshold ΔT_{max}, in order not to spend too much travel time in one single detour; (ii) the number of available seats should never be negative.

For every new user request, the best insertion in the schedule is the one that minimizes the cost function (Equation 5.4), which considers the extra waiting and ride times due to the new request insertion:

$$Cost = w_1 \cdot N_{delayed} \cdot \Delta RT + w_2 \cdot N_{UG} \cdot WT_{UG} \tag{5.4}$$

where $N_{delayed}$ is the number of passengers who have to bear an extra ride time ΔRT due to the insertion of the new request, N_{UG} is the number of users who make the new request at time t and need a walking time WT_{UG} to reach the stop or the destination, w_1 and w_2 are weights that regulate the importance of the additional ride time of passengers versus the walking time for the new passenger group. In our application, for a first test, we set both weights equal to 1, leaving for future research the tuning of such parameters.

5.2.5 Performance indicators

The local strategies determining the interaction between passengers and vehicles give rise to global patterns that can be monitored via appropriate performance indicators. They are chosen to capture the different objectives and points of view of the system actors, i.e.: (1) a traveller is interested in reducing the trip cost (distance, travel time, fare); (2) a company providing the service is interested in maximizing the profit, by increasing the number of passengers within a prefixed travelled distance or, conversely, in reducing the amount of travelled distance to serve a prefixed demand; (3) the community is interested in reducing transport-related externalities.

The model can monitor different key performance indicators to compare the two services, related both to the traveller and the operator perspective:

- the total number of transported passengers *NP* (pax).
- the total number of accepted requests *NAP* (pax).
- the total number of satisfied users *PAX* (pax).
- the average passenger travelled distance *APTD* (km).
- the average vehicle load factor *ALF* (pax/vehicle).
- the average passenger travel time T_{pax} (min), in terms of average walking time T_{WK}, average waiting time T_{WT} (min) and average ride time T_{RD} (min).
- the total driven distance *TDD* (km).
- the transport intensity *TI* (km/pax), as the ratio between *TDD* and *NAP*.
- the average traveller travel time T_{trav} (h) (including a penalty time of 60 min for each unsatisfied user).
- the operation cost *OC* (€) (see Equation 4.6 of Section 4.3.1).
- the effectiveness *E* (-) of the service, where $E = PAX / NP$.
- the total unit cost *TUC* (€/pax) (see Equation 4.5 of Section 4.3.1).

The next subsection will illustrate how the model was tested in a real-world case study.

5.3 Case study

The case study focuses on improving the accessibility of the San Nullo metro station (SN) in Catania, a medium-sized city in the south of Italy.

Territorial framework. The station is located in an arterial road that acts as a barrier between two neighbourhoods. In particular, it stands at the outskirt of the northern residential neighbourhood where walking paths are not of great quality, making it difficult for pedestrians to access the station. In such a context, the introduction of a FMLM transit service would help to reduce private car use and increase service coverage. However, it is important to guarantee a good passenger experience, in terms of travel time, and pay attention to the operator's cost.

Besides, the same service strategy and configuration could perform differently, i.e., very well during rush hours but not very well during off-peak hours, so a flexible feeder system able to switch between alternative routing and scheduling strategies in different periods of the day is desirable. We aim to evaluate the best choice between the two operating strategies (FRF vs DRF) under different demand rates and service configurations, identifying their optimal application scopes, through the comparison of passenger-related and cost-related performance indicators (see the next subsection). Figure 5.3 shows the FRF route (in blue) resulting from Calabrò et al. (2020a) and the road network used for the DRF service (in orange).

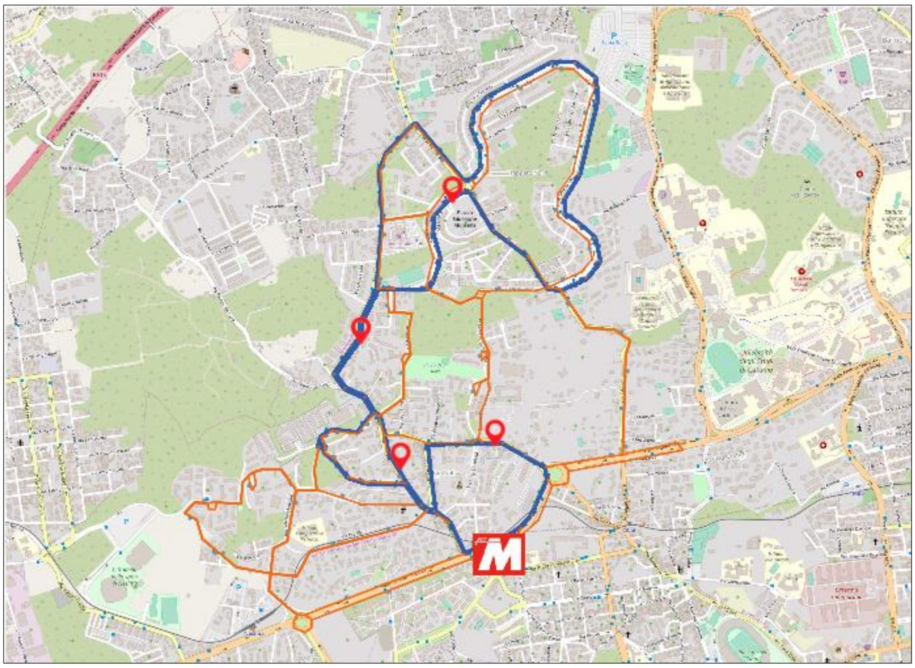


Figure 5.3. Road network for FRF (blue) and DRF (orange) services. Mandatory stops for the DRF are evidenced by red markers.

<u>Input variables and scenario simulation</u>. The main input variables of the system are:

- Service variables, i.e.: the type of service (FRF/DRF), the total simulation time (h), the number of vehicles (n), the vehicle maximum capacity (*cap*, in terms of seats), the average vehicle speed (*S*, in km/h).
- Demand variables, i.e.: the demand rate (*dem_rate* in trips/hour), the maximum number of travellers per request (*max_group*), the maximum waiting time (*mwt* in min).

Scenario simulation considers service operation by combining input values according to Table 5.1. In order to ensure comparability between the two services, we assume the same total capacity (e.g., 45 and 90 seats) for the FRF and the DRF services.

Table 5.1. Values of the input variables in the different scenarios

Type of variable	**Abbreviation**	**Unit**	**Value**	
			FRF	**DRF**
Service variables	n	-	3	3, 5, 6, 10
	cap	-	15, 30	15, 9
	S	km/h	25	25
Demand variables	dem_rate	trips/h	25, 50, 100, 200	25, 50, 100, 200
	max_group	-	3	3
	mwt	s	600	600

5.3.1 Results

The main results of the experiments are reported below. For each scenario, five replications of the simulation were performed, given the stochastic nature of the demand in the model. The model allows monitoring of the parameters and provides graphs and histograms of the main simulation variables during the simulation. In Figure 5.4, the satisfaction plot in terms of satisfied (S), unsatisfied (U) and delayed (D) users, and the average-load-factor plot are reported for a single event in the scenario with 50 pax/h average demand rate, a total capacity of 45 seats, i.e., 3 vehicles of 15 seats for both services.

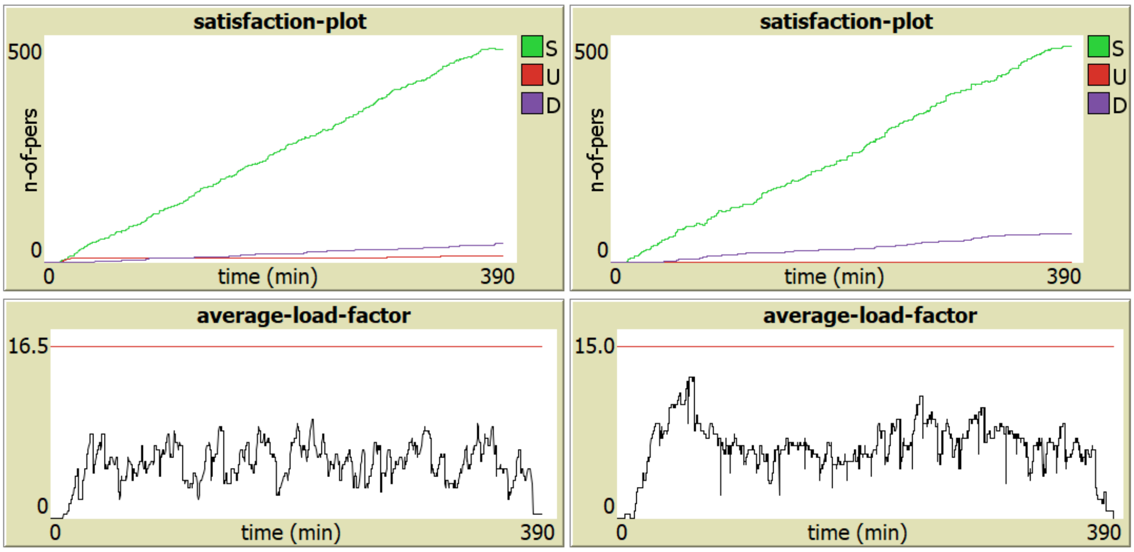


Figure 5.4. Satisfaction plot and average load factor for DRF (left) and FRF (right) for 50 pax/h and 3 vehicles.

Travel time. The experience of travellers is expressed by the travel time that is calculated as the sum of walking, waiting, and ride times. Service configurations lead to variable results according to the demand rate. For the lowest demand rate considered (25 pax/h), the DRF with large fleet and low vehicle capacity (10 vehicles of 9 seats) is the best option. For the highest demand considered (200 pax/h), the fixed feeder is preferable even it implies higher waiting times than the DRF. Results for demand rate of 50 pax/h and 100 pax/h are reported in Figure 5.5.

In the case of 50 pax/h, the FRF result in a lower travel time, while, when it comes to the DRF, it is possible to decrease the travel time only by increasing the vehicle capacity. In particular, the FRF and the DRF with more vehicles (10) can be considered as the best options from the user point of view. More in detail, if the highest weight would be given to waiting and walking time, users would prefer the DRF service. In the case of 100 pax/h, a lower travel time is achievable always with a higher capacity. As in the previous case, the FRF and the DRF with more vehicles (10) should be preferred from the user point of view, even if the willingness to pay for the different times (walking, waiting, riding) should be further investigated since it is context specific. In both scenarios, the ride time weighs more on the DRF performance, due to the various detours required to serve pick-up and drop-off passengers, even though there are considerable savings in waiting and, to a smaller extent, walking time.

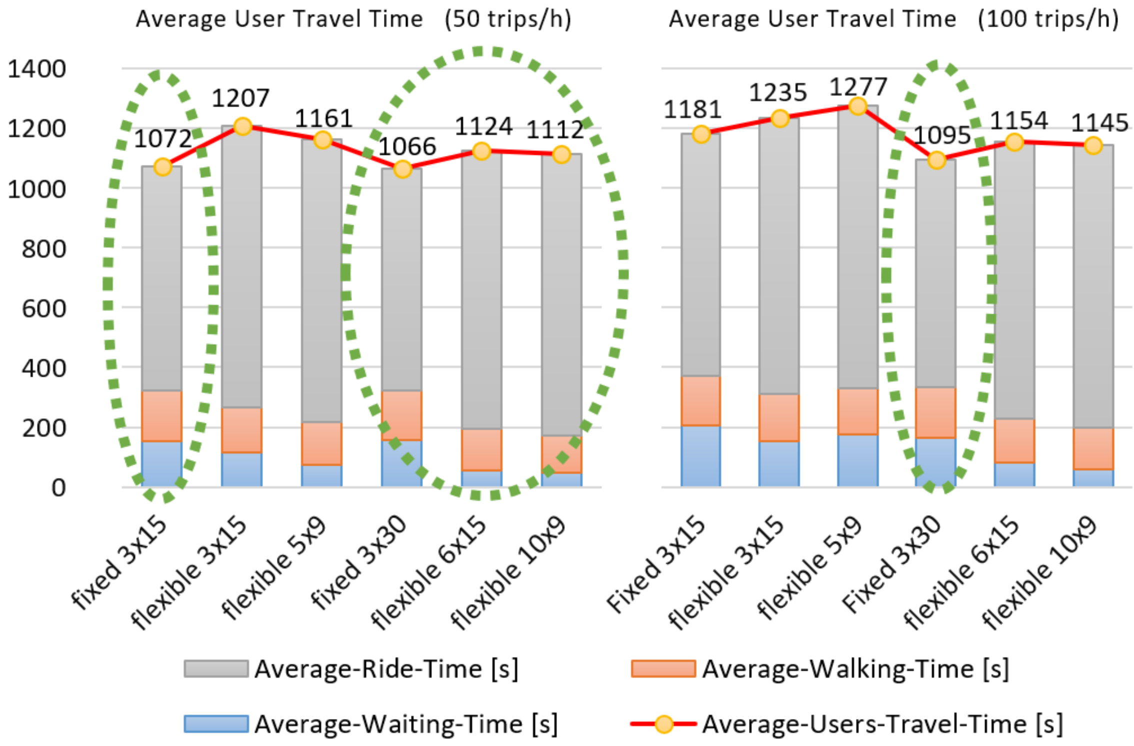


Figure 5.5. Average travel time for demand rates of 50 and 100 pax/h.

TI - E – TUC. We decided to compare the performances of different service configurations using three main indicators, i.e., TI, E and TUC.

TI is the ratio between the total distance travelled by the fleet of vehicles and the total transported passengers. A low TI indicates an efficient service in terms of operation cost per travelled passenger and a low impact on the environment as well.

E, which is the ratio between the number of transported passengers and the number of accepted passenger requests (PAX/NAP), should be high to increase the number of satisfied users compared to the total number of accepted requests.

Finally, TUC should be as low as possible to reduce the total costs of the system (operator and user) and increase the number of satisfied passengers. In this respect, it can be considered as an overall measure of the transport system efficiency.

For low demand rates (25 pax/h), the high capacity DRF service (10x9) is the less convenient for TUC and TI, while E is very high and comparable with the other DRF solutions. For high demand (200 pax/h), the best results in terms of TUC and TI can be achieved by a high-capacity FRF service.

Main results from the scenarios with demand rates of 50 and 100 pax/h are reported in Figure 5.6.

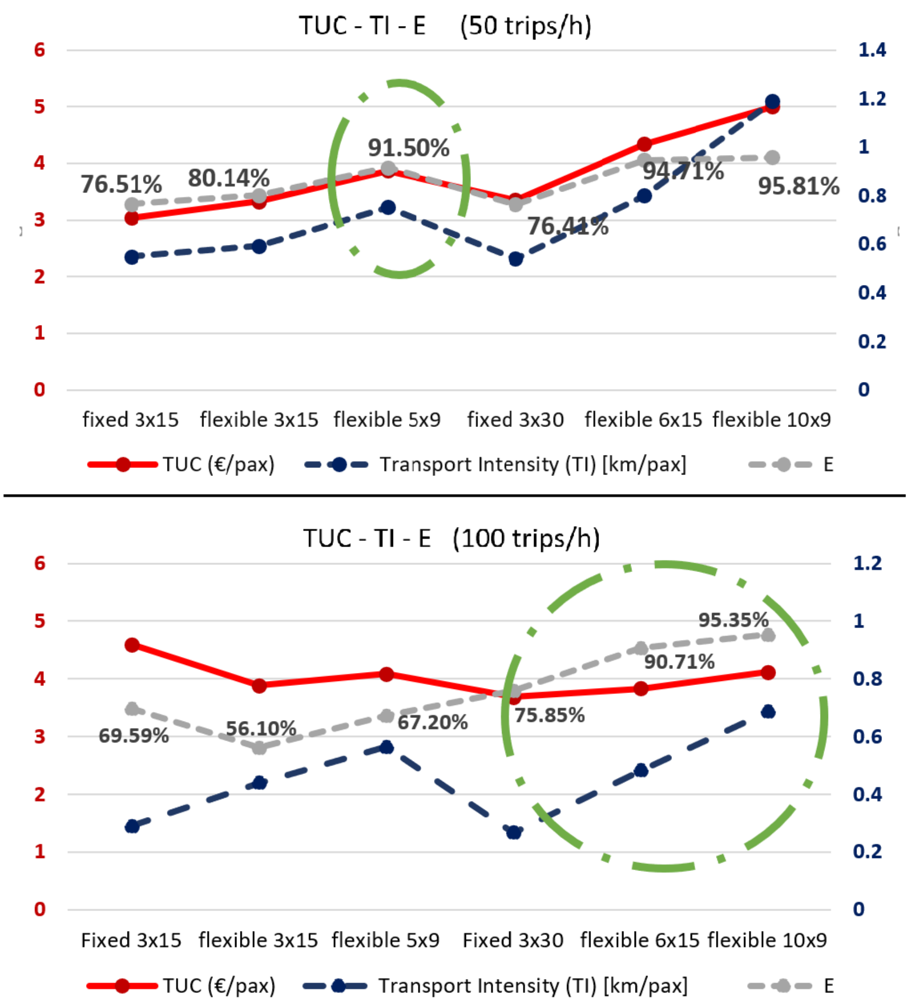


Figure 5.6. TUC, TI and E for 50 pax/h (top) and 100 pax/h (bottom).

In the case of 50 pax/h, the two FRF services are comparable in terms of performance, but they reach approximately 76% of accepted requests. Low capacity in the case of DRF is in general better, with comparable values of TUC. In particular, the FRF would be preferable for the operator since it has lower TI and a good TUC, but with a higher percentage of rejected user requests, while it is possible to cover more than 90% of requests with a DRF service with 5 vehicles of 9 seats.

In the case of 100 pax/h, the two FRF services are also comparable, but allow reaching approximately 70-75% of accepted requests. For 100 pax/h, high capacity is in general better with comparable values of TUC, and this is evident for the DRF with a reduced fleet size, which fails to serve a large percentage of users. As for the previous case, a fixed feeder would be preferable for the operator since it has lower TI and a good TUC, while DRF with 10 vehicles is the one with the highest coverage, but with a high TI. A compromise solution would be the DRF with 6 vehicles and 15 seats, but it would be better evaluated by estimating the extra cost for the system due to each “rejected” passenger, i.e., a user who does not use the feeder service and maybe chooses to use the private car. Once again, the higher driven distance due to the various detours of the DRF service is responsible for the greater TI value compared with fixed feeder, which however is unable to serve a certain percentage of users far from its stops.

5.4 Conclusion

FRF services travel on regular routes at scheduled times, but passengers have to walk to reach the fixed stops and wait for the service. On the other hand, flexible DRF services can pick-up and drop-off passengers wherever and whenever they want, therefore it is expensive to operate, even if passengers are more satisfied. We presented new ABM able to simulate both fixed and flexible mobility services, where the fleet size and capacity, and the demand rates are chosen as input parameters of different simulation scenarios.

The model was tested in a case study with a real network and based on real demand data. First results with different demand rates (from low to high) identify the optimal configuration of DRF to achieve a trade-off among passengers’ convenience, service coverage, operation efficiency, and environmental impacts as well.

In particular, for the case study analysed, the model can tailor the service according to the current demand, where a DRF fleet with 5 vehicles of 9 seats would be suitable for an average demand rate of 50 pax/h, while a fleet of 10 vehicles of 9 seats would fit a demand rate of 100 pax/h. For lower and higher demand rate, a trade-off between coverage and ridership emerges. For higher demand rates, the FRF becomes the best choice even if it implies a lower coverage. Future research should investigate the demand side, in terms of the willingness to pay related to the different components of travel time. Other interesting indicators could be added to better evaluate the services and the related externalities, e.g. CO2 emissions. Moreover, it would be interesting to reproduce different cases with parametric road network topologies and demand (see Chapter 6), to see the impact on travelled distance and level of service.

Another step forward would be to test a multi-station system, with feeder buses serving different metro stations. Finally, the use of autonomous vehicles should be tested, affecting the results of TUC (in terms of operation costs).

In summary, the model can contribute to the development of flexible DRF services, to solve of the coverage/ridership dilemma of rapid transit services, and to understand the impact of land use, road network and demand patterns on the flexible service performances.

CHAPTER 6

Comparing Fixed-Route and Demand Responsive Transit Feeder Services: a Theoretical Agent-Based Model

Feeder transport services are fundamental as first- last-mile connectors of mass rapid transit (MRT). Besides, the rapid spread of new technologies and the shared mobility paradigm gave rise to new mobility on-demand modes that are dynamically able to match demand with service supply. In this context, the new generation of demand-responsive transport (DRT) services can act as on-demand feeders of MRT, but their performance needs to be compared with conventional fixed-route feeders.

This chapter presents an agent-based model able to simulate different feeder services and explore the conditions that make a demand-responsive feeder (DRF) service more or less attractive than a fixed-route fixed-schedule feeder (FRF). The main difference compared to the model presented in the previous chapter is the use of a parametric simulation environment including realistic constraints (e.g., passenger time windows) and parameters (e.g., vehicle capacity) which are usually not considered in analytical models due to high computational complexity. We identified the critical demand density representing a switching point between the two services. Once the demand density is fixed, exploratory scenarios are tested by changing the demand spatial distribution and patterns, service area and service configurations. Main results suggest that the DRF is to be preferred when the demand is spatially concentrated close to the MRT station (e.g., in a TOD[4]-like land-use area) or when station spacing is quite high (e.g., a regional railway service), while FRF performs better when the demand is mainly originated at the MRT station to any other destinations in the service area (e.g., during peak hours). Besides, automated vehicles could play a role in reducing the operator cost if the service is performed with many small vehicles rather than few higher-capacity vehicles, even if this would not imply a major benefit gain for the users.

[4] Transit-Oriented Development

6.1 Introduction

The push towards sustainable mobility and the advances in technology are modifying both transport services and user habits. Besides, the COVID-19 outbreak has heavily hit public transport (PT) and influenced travel behaviour with expected long-lasting impacts, especially at the urban level. In this changing environment, good planning and designing of transport systems becomes increasingly important to shape the future of urban regions (Gkiotsalitis and Cats, 2021). Transport planning cannot ignore new trends and paradigms that are rapidly emerging, oriented towards the concept of sharing assets and services via on-demand services enabled by digital platforms. Mobility as a Service (MaaS) is one good example of such a new concept, being a digital platform enabling multimodal door-to-door trips via a single app that should aim at reducing private car usage (and eventually ownership) (Storme et al., 2020). To achieve this goal, PT is considered to be the backbone of MaaS, complemented by mobility on demand and shared modes (Ambrosino, Nelson and Gini, 2016).

In this context, DRT can play a fundamental role. It came back in the limelight in recent years mainly due to the spread of app-based ride-sharing systems such as Uber, Via, and Lyft (also known as transport network companies).

However, the performance of these systems is often questioned especially concerning the transport demand they are meant to cover, the characteristics of the area, the topology of the road network in which they should operate, and the level of flexibility of the routes they follow. While they are often in competition with PT in the case of densely populated urban areas, they can potentially constitute a good solution to improve the accessibility to mass transit systems in the case of low demand areas by covering the first- last- mile of passenger PT trips (EPRS, 2016). The EU highlights how sparsely populated and underpopulated regions suffer from several structural problems such as lack of transport connections, few job opportunities and inadequate social services (EPRS, 2016). In a recent publication by the European Parliament (Bisaschi et al., 2021), the lack of accessibility to PT is considered as one of the main issues faced by low-density and depopulated areas, with the consequent social exclusion of their inhabitants. The strategic document underlines the role that DRT enabled by novel technologies could have in those regions as a result of its flexibility in meeting the demand. The document also notes that particular attention should be paid to its use in supplementing conventional PT such as train, metro and trams that cannot reach all regions with the same coverage.

The conditions for DRT success have been explored in the literature focusing on some of the key aspects for PT systems viability e.g. the demand patterns making DRT services more attractive than fixed scheduled ones, optimal fleet size and fleet composition. Besides, real case studies (or pilots) have been set up and investigated which are inevitably influenced by the territorial context where they happen leading to results that are often difficult to generalize and thus transfer to other regions (e.g. Mehran et al., 2020; Mishra et al., 2020). Analytical models have been typically used to find demand thresholds and switching points from one service to the other, as in, for example, the work done by Quadrifoglio and Li (2009).

Moreover, while the routing of these systems has long been addressed in the literature (Garaix et al, 2010; Bruni et al., 2014; Ronald et al., 2016)), the recent technological innovations in the trip booking processes have led to the emergence of new problems due to the need to guarantee dispatching of vehicles and requests in real-time. These issues are amplified in the case of shared rides since different users might have different time schedules.

Finally, any analysis of these systems, cannot ignore the technological evolution of the vehicles used. These seem to be oriented towards the conversion to full-electric ones and in the longer-run to full automation (Wang et al., 2019).

Based on these premises, we aim to propose the use of an agent-based model (ABM) to explore the performances of DRT in comparison to a fixed-route service as a feeder to mass transit. We began to address this issue in the previous chapter, by presenting and applying a new ABM to simulate adaptive flexible/fixed feeder services in a low-demand urban area in Catania (Italy). We extend the aforementioned model also to verify current analytical models that aim to distinguish DRT from fixed services (Quadrifoglio and Li, 2009). Moreover, we extend such models introducing a non-uniform demand density and a variable vehicle capacity.

In the process we address several open research issues: (i) the transferability to different contexts by introducing a parametric design model; (ii) the booking process, by taking into account user-based time constraints in the dispatching algorithm; (iii) the possibility to perform the service with automated vehicles, considering the impact they could have especially in terms of the operator cost.

The remainder of the chapter is organized as follows. The following section summarizes the literature on the DRT topic and highlights the current research gaps to be filled in this work. Section 6.3 presents the methodology, drawing an overview of the model with the dispatching algorithm and the selected output indicators. The model is applied to a synthetic parametric application case and results for different scenarios are presented in Section 6.4. Then, we test our novel dispatching procedure on the case study presented in Chapter 5. Finally, conclusions and discussion and future research are outlined in Section 6.5.

6.2 Literature review

DRT planning and design has been addressed in literature mainly since the beginning of the 2000's (e.g., Mageean and Nelson, 2003; Enoch et al., 2004). A fair number of analytical models have been developed to face strategic planning decisions. In 2004, Diana and Dessouky (2004) addressed the dial-a-ride problem with time windows introducing a new regret insertion heuristic able to face a large number of requests and outperform classical heuristics when a high-quality service (narrow time windows) is provided to the users. However, a door-to-door transport policy is assumed, while more trip shareability and travel time savings could be achieved by aggregating requests in virtual stops even if passengers would have to walk a short distance. Later in 2007, Quadrifoglio et al. (2007) proposed an insertion heuristic for flexible services that merge the flexibility of DRT systems with the low-cost operation of fixed-route ones; the proposed service covers a specific geographic zone, with mandatory checkpoints located at major connection points or high-density demand zones. However, the model assumes that customers never reject the insertion proposed by the algorithm, so there is no negotiation phase between the system and the customers. In 2009, Quadrifoglio and Li (2009) proposed a continuous approximation approach to compare the user-related performance of two operating strategies for a feeder bus service: fixed-route and demand-responsive. They provided approximate analytical solutions to estimate the critical demand density that allows a transit agency to switch between the two types of services. Chandra et al. (2013) extended the aforementioned model using a gravity-based accessibility model to evaluate the accessibility impacts for first/last mile transport connectivity in the case of fixed-route transit and DRT, although using the same uniform demand across space. Other issues investigated through analytical models were fleet sizing based on a given quality of service (QoS) for users (Diana et al., 2006), route design (Bruni et al., 2014), the choice between different flexible transit strategies to accommodate a variable demand level (Zheng et al., 2018), and estimating how user and operator costs vary according to the demand density, the service area and the fleet size (Huang et al., 2020).

Analytical models can be considered as design models that provide optimal solutions among infinite alternatives, thanks to the introduction of some approximations and simplifications within the model; such models are in general not capable of reproducing a complex reality without requiring enormous computation times, otherwise, it might be impossible to solve them analytically. In contrast, simulation models are ideal for reproducing the complexity of a system; once the input parameters are set, the simulation model runs a series of operations to the data whose number usually grows linearly with the size of the problem and can then generate several outputs, which can be used as key performance indicators of the system. In this respect, simulation models have been used extensively for making tactical and operational decisions regarding transport services in general and in particular for flexible transport services and last-mile connection (Wang et al., 2019; Shinoda et al., 2003; d'Orey et al., 2012; Bischoff et al, 2017; Cich et al., 2017). Winter et al. (2016) designed, simulated and tested an automated DRT for a campus-train station service; the simulation determined the optimal fleet size for the operation of the service and results showed the importance of adequate vehicle sizes and short vehicle dwell times. Scheltes and Correia (2017) explored the use of automated vehicles as last-mile connection of train trips using an ABM. The ABM incorporated a dispatching algorithm distributing travel requests amongst the available vehicles using a First-come-First-served (FCFS) sequence. However, the type of service that was tested was conceived as being individual, therefore it did not allow for shared trips among the passengers. Araldo, Di Maria et al. (2019) studied the impact of consolidating the demand and limiting the density of waypoints locations through a modular simulation platform, searching for a trade-off between guaranteeing high QoS for the users and providing a high-efficiency system; the model simulated different flexibility levels from door-to-door to bus-like services. Oh et al. (2020b) proposed an agent-based simulation framework to evaluate the performance of an automated demand-responsive transit system as complement and/or substitute to conventional mass transit. Their model suggests advantages in using higher capacity vehicles rather than taxis, resulting in less travelled kilometers for the operator and less congestion. Fielbaum et al. (2021) abandoned the door-to-door scheme showing that significant reductions in the vehicle-hours travelled and in the number of rejections can be achieved by asking travellers to walk a short distance to reach the assigned pick-up/drop-off points. They proposed an insertion heuristic algorithm that encompasses and weights both the travellers and the operator costs. The simulation was performed both on a toy grid network and the real case study of Manhattan. However, they considered neither the integration with PT nor a comparison with other forms of PT. In 2019, Inturri et al. (2019) developed an ABM to compare the performance of a shared DRT with that of a taxi service both for fast-growing cities (Giuffrida et al., 2020) and low demand areas (Inturri et al., 2021). Results showed that DRT shared services are convenient under specific demand patterns for the analyzed case studies.

Literature analysis shows that simulation models are not meant to provide optimal service design; they provide a good description of the performance of the system under specifically designed scenarios. Nevertheless, if the number of combinations between relevant parameters is not great and the simulation run time is not too long it is possible to search the different configurations space in search for solutions that maximize or minimize a certain key performance indicator.

In this chapter, we present an ABM, intending to integrate the benefits of the two methods: the model uses the simulation approach but on an ideal parameterizable environment so that the results are as scalable as possible. Based on the work of Quadrifoglio and Li (2009) and Calabrò et al. (2020b), the new ABM goes beyond a pure analytical model by proposing realistic and real-time dispatching algorithms, but without being tied to a specific simulation network neither to a particular case study, so its results are easily adaptable to other contexts using the main experimental parameters.

6.3 Methodology

Our research focuses on the first and last-mile leg of PT trips, supposing that the PT backbone is a mass rapid transit (MRT) network like rail or BRT. The transit agency might choose between two operational strategies:

a) A fixed-route feeder (FRF) service carried out by buses that pick-up and drop-off passengers at predetermined stops.
b) A demand-responsive feeder (DRF) service, where each vehicle builds customized routes to serve a group of passengers.

6.3.1 Overview of the model

The main components of the simulation model are the service type (FRF or DRF), the geometric features of the service area, the demand model, the supply (vehicle) characteristics, and the simulation duration. The agents of the model are the travellers requesting a ride and the vehicles.

We consider that our feeder service operates in a rectangular area of length L from the terminal station (horizontal direction) and width W (vertical direction). The mobility demand follows a Many-to-One/one-to-many pattern since the focus is on the first/last leg of an entire PT multimodal trip. A trip request can therefore have either origin or destination at the MRT station (the terminal), located at the left side of the service region. The value of W can be considered as the average distance between the terminal and the MRT stations upstream and downstream. Hereinafter, we will refer to the passengers originating at the terminal as *egress passengers*, while the passengers having destination at the terminal will be referred to as *access passengers*.

From a spatial perspective, the base demand density (in $trips/km^2h$) is modelled as a linear function $\lambda = \lambda_0 - m\,x$, where x is the horizontal distance from the terminal (in km), λ_0 is the demand density at $x = 0$ and m is the slope that makes the values decrease. Let us denote with λ_L the as the demand density at $x = L$, with $\bar{\lambda}$ the average demand density (which occurs at $x = L/2$) and with Λ the ratio between λ_L and λ_0 (Figure 6.1).

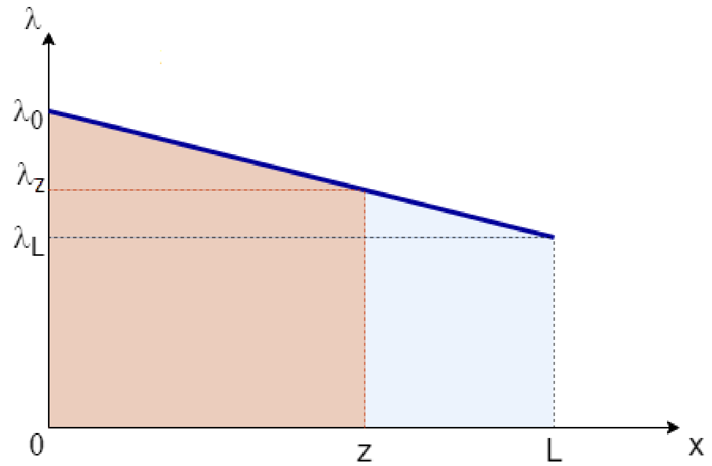


Figure 6.1: A general trend of the demand density along x.

We obtain $m = (\lambda_0 - \lambda_L)/L$, $\bar{\lambda} = (\lambda_0 + \lambda_L)/2$ and $\lambda_L = \Lambda\lambda_0$. We can therefore express the demand function via the two parameters $\bar{\lambda}$ and Λ as follows:

$$\lambda = \frac{2}{1+\Lambda}\,\bar{\lambda}\left(1 - \frac{1-\Lambda}{L}x\right) \tag{6.1}$$

From Equation 6.1 we derive $\lambda_0 = 2\bar{\lambda}/(1+\Lambda)$ and $\lambda_L = 2\Lambda\bar{\lambda}/(1+\Lambda)$. The value $\Lambda = 1$ represents the case where the demand is assumed spatially homogeneous throughout the service region, thus from Equation 6.1 we obtain $\lambda = \bar{\lambda},\ \forall x$. By varying the slope of the demand decay function, we reproduce different types of urban density and land use.

Each trip request i involves a group of a number G_i of travellers. We assume that G_i follows a geometric distribution. If we denote with p_1 the probability that request i consists of a single user ($Pr(G_i = 1)$), therefore the probability that the group is constituted by k users is given by:

$$Pr(G_i = k) = p_k = p_1\,(1 - p_1)^{k-1} \tag{6.2}$$

In our model, we assume that a user chooses between walking and using the feeder service to reach the MRT station. If the Manhattan distance (i.e., the distance measured along axes at right angles) $d_{T,i}$ between the user's origin and the MRT station (access leg) or the MRT station and the destination (egress leg) is lower than a minimum threshold d_0, a user is assumed to walk directly to/from the station. Otherwise, the probability of choosing the feeder service rapidly increases with $d_{T,i}$ and is given by the attractiveness coefficient $\eta \in [0,1]$ we introduced as follows:

$$\eta_i = 1 - e^{-0.5\frac{\left(d_{T,i} - d_0\right)^2}{\gamma^2}} \tag{6.3}$$

where γ is a parameter that rules the increase in attractivity of the feeder service due to the distance from the station: the lower the value of γ, the higher the speed at which η increases with $d_{T,i}$. Figure 6.2 shows a graphical representation of η across the service area.

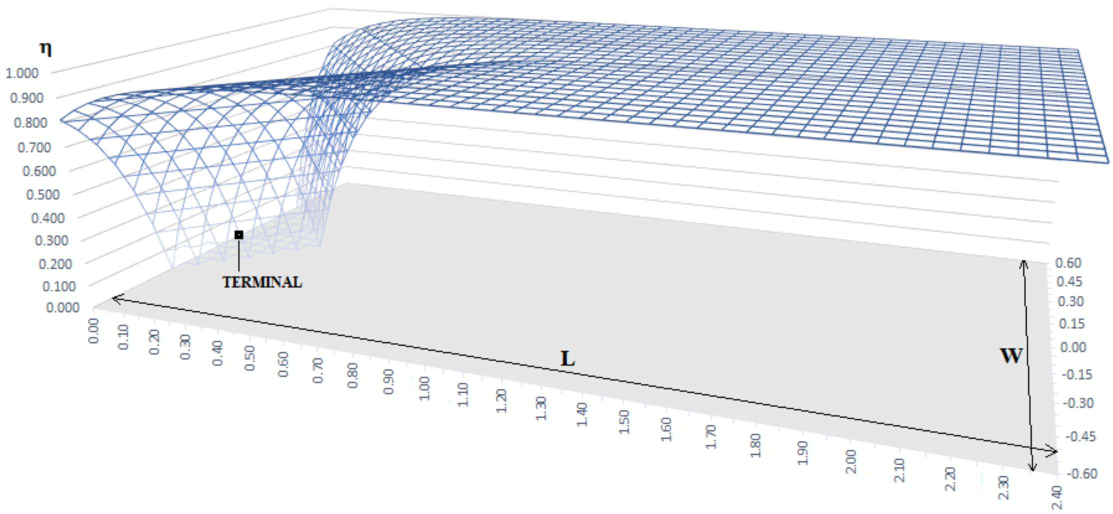


Figure 6.2: Values of the attractiveness coefficient across a service area with L = 2.4 km *and* W = 1.2 km. *In this example, we set* $\gamma = d_0 = 0.3$ km. *From a temporal point of view, the demand follows a Poisson process with* λ *as the rate parameter.*

As in Quadrifoglio and Li (2009), the fraction of trips having the MRT station as the destination is given by the parameter $\alpha \in [0,1]$, so the proportion of users going from the station to a destination inside the service area is $(1-\alpha)$. Although the Poisson process is well suited for the former, the temporal distribution of the egress passengers is strongly related to the schedule of the MRT line. However, assuming that the headway of the MRT line is small enough (i.e., less than 5 minutes) in both directions, we believe that the Poisson distribution with rate parameter λ is a reasonable approximation and should marginally affect the results.

Regarding the supply side, the vehicles (buses, minibuses, vans or automobiles) are defined by three input parameters: the number of vehicles n_V composing the fleet, the cruising speed v and the allowed capacity in available seats in each vehicle Cap.

Finally, the duration of the simulation ST, in which the input parameters are unchanged, should be sufficiently high to ensure that the steady state is reached and that the results are marginally affected by the warm-up period.

6.3.2 The Fixed-Route Feeder

The FRF (Figure 6.3) runs back and forth on a straight line from the MRT station to the farthest bus stop, with spacing d_s (input parameter) between the stops.

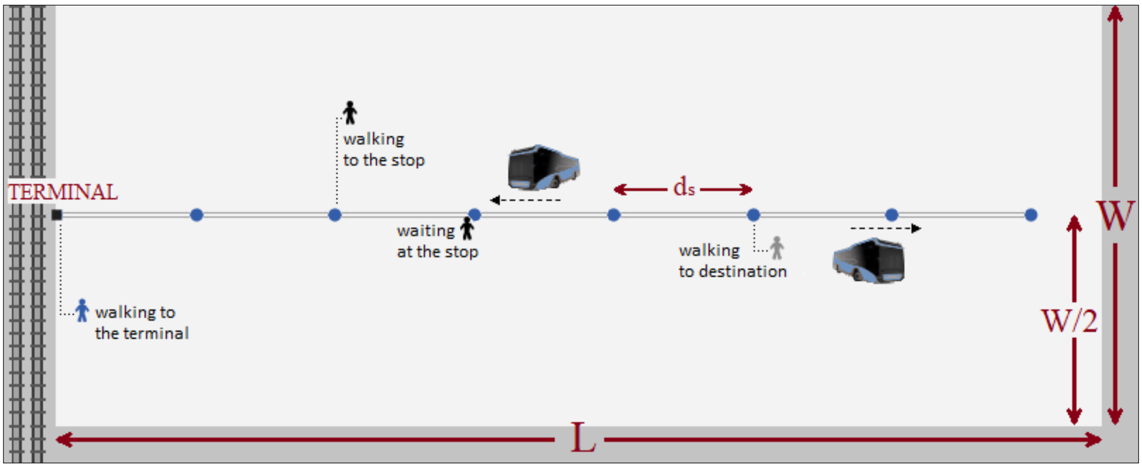


Figure 6.3: Scheme of the FRF service: users in black are about to take the bus, users in grey got off the bus, the user in blue is walking directly to the terminal.

Assuming ideal conditions (no congestion or other disturbance to the service), the maximum cycle time CT_{max}, i.e., the time needed for a vehicle (bus) to complete the round trip, is the sum of two components:

$$CT_{max} = \frac{2\left(L-\frac{d_s}{2}\right)}{v} + \left(\frac{2L}{d_s} - 1\right) \cdot (\tau_s + \tau_p) \tag{6.4}$$

where the first component is the ratio between the length of a complete cycle and the bus cruising speed v, the second one estimates the dwell time at each stop τ_s, including the time of acceleration and braking, and the additional dwell time due to boarding/alighting passengers τ_p.

The headway between two vehicles is given by Equation 6.5:

$$h = \frac{CT_{max} + \tau_t}{n_V} \tag{6.5}$$

where τ_t is the minimum dwell time required at the terminal.

Egress passengers originate at the MRT station, take the bus with the earliest departure time and alight at the desired stop (i.e., the closest one to the destination of the trip) during the first half of the cycle, and walk to their final destination. Access passengers originate in the service area, walk to the closest stop, wait for the first bus headed to the terminal, travel onboard and finally alight at the terminal.

We assume that a traveller is aware of the expected waiting time at the stop thanks to a real-time information system provided by the feeder service. Therefore, if the expected waiting time is above a certain maximum waiting time $t_{w,max}$ (input parameter) traveller i assumes the status "rejected" and walks to the station. Instead, if the overall travel time overcomes the latest drop-off time interval at the station ld_i, the passenger assumes the status of "delayed". This threshold varies according to the different time tolerance of the users, as shown in Equation 6.6:

$$ld_i = \boldsymbol{min}\left\{t_{w,max} + (1 + \delta_i + \gamma_i)\frac{d_{T,i}}{v},\ (1 + \gamma_i)\frac{d_{T,i}}{v_p}\right\} \tag{6.6}$$

Where $\delta_i = d_{T,i} / (L+W/2)$ relates the travellers' distance to the maximum one to reach the terminal (based on the farthest user that could be generated in the service area), and represents the willingness to deviate from the shortest path based on his/her distance to the terminal; $\gamma_i \in [0, 1]$ aims at reproducing travellers' individual willingness to deviate based on the trip purpose, and v_p is the walking speed since Equation 6.6 also takes into account the walking time from origin to destination.

Fleet size, vehicle capacity, and cruising speed are set at the beginning of the simulation. Each vehicle is generated at the terminal stop (i.e., the MRT station). The fixed feeder vehicle travels along the route until it reaches a stop. Passengers at their destination stop alight while waiting travellers board the vehicle (if in the inbound direction), following the FCFS queue rule, and only if the passenger group size is not greater than the available seats.

6.3.3 The Demand Responsive Feeder

The DRF (Figure 6.4) travels along a grid street network, with spacing d_g between the streets. Vehicle routes are dynamically created based on users' requests and each intersection can act as a potential access/egress location for a traveller (virtual stop).

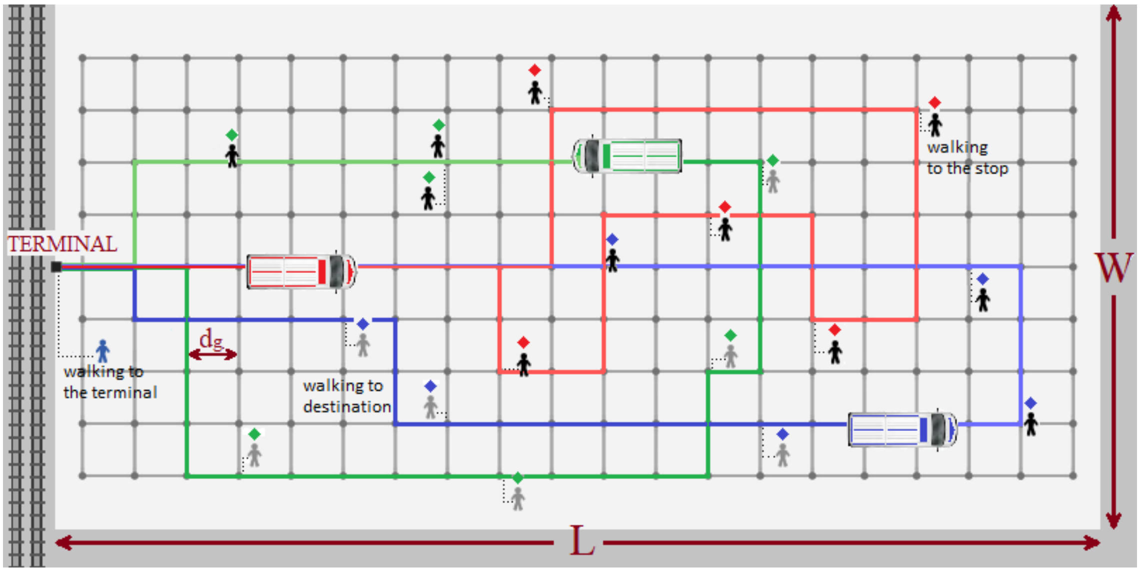


Figure 6.4: Scheme of the DRF service: users in black are waiting for the assigned vehicle, users in grey have left the vehicle and are walking to their destination, the user in blue is walking directly to the terminal.

Unlike in the FRF case, choosing the headway of the DRF is not straightforward. While the flexibility of the DRF implies that vehicles could leave the station when they are full, in the case of a feeder service it could be better to assume a given headway to assure service regularity and synchronization with the MRT. In fact, the length of the full cycle route and the cycle time depend on the expected number of requests n that the vehicle should serve along that route, which in turn depends on the headway, as shown in Equation 6.7:

$$n = n_u \sum_{k=1}^{\infty} \frac{p_k}{k} = \bar{\lambda}\, L\, W\, h \sum_{k=1}^{\infty} \frac{p_k}{k} \tag{6.7}$$

where n_u is the number of users generated during the time interval of duration h. We point out that $n \leq n_u$ because a single trip request can involve k users (see Equation 6.2).

A rough approximation of the expected cycle distance *ECD* for the DRF service, i.e., the length of the round route, can be estimated based on n (Equation 6.7) and is given by the sum of a horizontal component (the vehicle movement from left to the right and vice-versa) and a vertical component (the deviations along the vertical direction to serve the passengers). The former derives from Section 4.2 of Quadrifoglio and Li (2009) and generalizes the result to the non-uniform demand assumption, while the latter derives from Section 4 of Quadrifoglio et al. (2006) and it is improved by a correction coefficient ξ_c (described in the next paragraph), which accounts for the spatial and temporal consolidation of multiple trip requests at the same virtual stop (i.e., served by the same vehicle at the same time).

$$ECD = E[\max\ (x_r) \mid r = 1, \dots, n] + \xi_c \left[\frac{W}{2} + (n-1)\frac{W}{3}\right] \tag{6.8}$$

Now we derive the expected value of the horizontal component of the cycle distance. For a random request r, the probability that $x_r \leq z$, being $0 \leq z \leq L$, is given by the ratio between the red trapezoid and the light-blue trapezoid in Figure 6.1, so that we derive:

$$\Pr(x_r \leq t) = \frac{\frac{\lambda_0+\lambda_z}{2} z}{\frac{\lambda_0+\lambda_L}{2} L} = \frac{\lambda_0+\lambda_z}{\lambda_0+\lambda_L}\frac{z}{L} = \frac{2-(1-\Lambda)\frac{z}{L}}{1+\Lambda}\frac{z}{L} = \frac{2}{1+\Lambda}\frac{z}{L} - \frac{1-\Lambda}{1+\Lambda}\frac{z^2}{L^2} \tag{6.9}$$

where Λ is the ratio between λ_L and λ_0.

The expected value of the horizontal component of the DRF cycle distance can be derived as follows:

$$\begin{aligned} E[\max\ (x_r) \mid r = 1, \dots, n] &= \int_0^L \{1 - \Pr[(\max(x_r) | r = 1, \dots, n) \leq z]\}\, dz \\ &= \int_0^L \{1 - \textstyle\prod_{r=1}^n [\Pr\ (x_r \leq z)]\}\, dz = \int_0^L \left[1 - \left(\frac{2}{1+\Lambda}\frac{z}{L} - \frac{1-\Lambda}{1+\Lambda}\frac{z^2}{L^2}\right)^n\right] dz \end{aligned} \tag{6.10}$$

Imagine dividing the length of the service area L in an adequately large number N_L of segments of length $\Delta z = L\ /\ N_L$ (e.g., $\Delta z \approx 10$ m). Then we can approximate Equation 6.10 as follows:

$$E[\max\ (x_r) \mid r = 1, \dots, n] = \sum_{z=1}^{N_L} \left[1 - \left(\frac{2}{1+\Lambda}\frac{z}{N_L} - \frac{1-\Lambda}{1+\Lambda}\frac{z^2}{N_L^2}\right)^n\right] \Delta z \tag{6.11}$$

In the case of uniform demand density ($\Lambda = 1$), Equation 6.10 becomes:

$$E[\max\ (x_r) \mid r = 1, \dots, n] = \int_0^L \left[1 - \left(\frac{z}{L}\right)^n\right] dt = L\frac{n}{n+1} \tag{6.12}$$

as in Quadrifoglio and Li (2009).

Finally, we can express the expected cycle distance as follows:

$$ECD = \sum_{z=1}^{N_L} \left[1 - \left(\frac{2}{1+\Lambda}\frac{z}{N_L} - \frac{1-\Lambda}{1+\Lambda}\frac{z^2}{N_L^2}\right)^n\right] \frac{L}{N_L} + \xi_c \left[\frac{W}{2} + (n-1)\frac{W}{3}\right] \tag{6.13}$$

Unlike the above-cited works, in our simulation model backtrackings are allowed. In fact, let us imagine that the vehicle serves 3 requests r_i (i = *1, 2, 3)* with coordinates (x_i, y_i) and $x_1 < x_2 < x_3$. With the *no-backtracking policy* of Quadrifoglio and Li (2009) and Quadrifoglio et al. (2006) the requests should have been served in the order [r_1 r_2 r_3], but the dispatching algorithm (discussed in the next paragraph) could find that the order [r_2 r_1 r_3] is a better solution even though it implies a minor backtracking movement.

The headway of the DRF can be ultimately derived from the expected cycle time *ECT* and the fleet size as follows:

$$h = \frac{ECT+\tau_T}{n_V} = \frac{\frac{ECD}{v}+(\xi_c\tau_s+\tau_p)n+\tau_T}{n_V} \tag{6.14}$$

Equation 6.14 is non-linear (as *ECD* depends on h via n (Equation 6.7)), so we choose an initial value h^* of the headway, compute n, *ECD* and *ECT* using Equations 6.7;6.13-6.14, calculate the new value of h and then repeat the process for an adequate number of iterations until convergence. This procedure takes place in the "setup" phase before each simulation starts.

The requests for the DRF service are processed in real-time through a dispatching algorithm that assigns the traveller to a vehicle and a virtual stop (either a pick-up or drop-off location), according to the user time windows and the vehicle available seats, based on the FCFS rule. If no feasible match can be found, the user assumes the status "rejected" and walks directly to the destination. In this way, the penalty due to the rejection is not an arbitrary fixed value, as done e.g., by Inturri et al. (2019) but is directly related to the walking time from the origin to the destination, under the simplification that the rejected user does not have any other modal choices. This implies that rejecting requests involving longer trips plays a significant role in increasing the average user disutility (they will have to walk more to the station or from the station). Figure 6.5 illustrates the dynamics for a DRF passenger, from the trip request moment to the arrival at the destination, while Figure 6.6 shows the vehicle state chart.

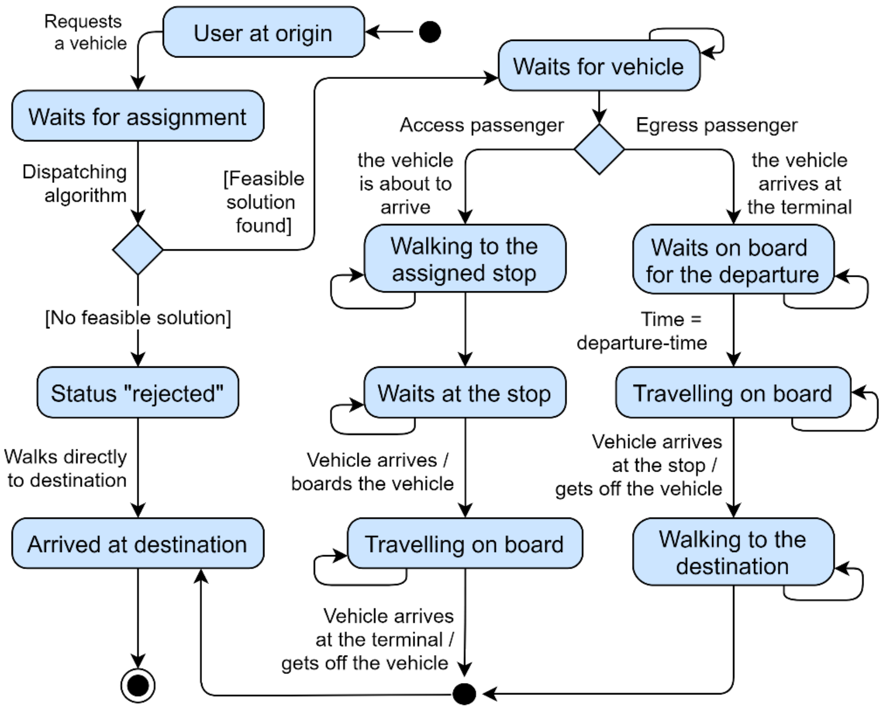


Figure 6.5: Traveller state charts for the DRF feeder service.

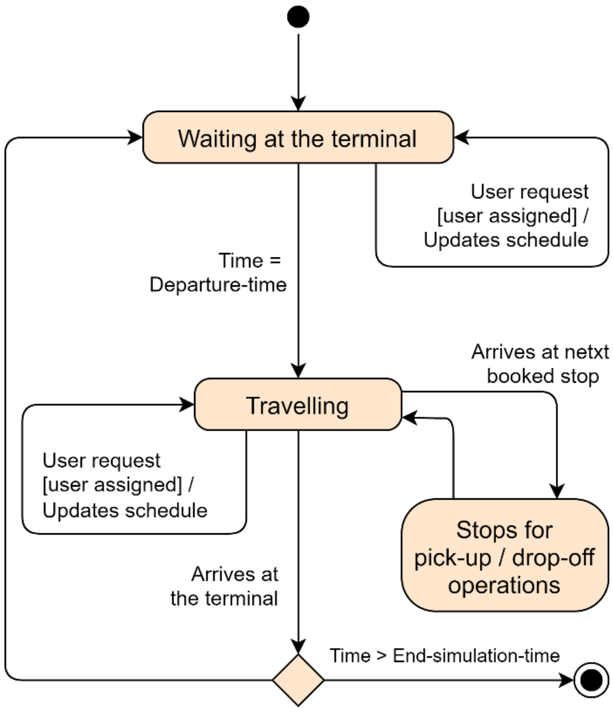


Figure 6.6: Vehicle state chart for the DRF service.

Let us define the set of egress passengers as E and the set of access passengers as A. Egress passengers are assumed to make the trip request as soon as they show up at the terminal, so they are readily available to use the DRF service. Access passengers, instead, need to be given enough time to reach the assigned pick-up location and cannot be served before the earliest pick-up time interval ep_{ij}, which is formulated as follows:

$$ep_{ij} = \begin{cases} 0 & \text{if } i \in E \\ wk_{ij} + \tau_{wk} & \text{if } i \in A \end{cases} \tag{6.15}$$

where wk_{ij} is the walking time needed for passenger i to reach the pick-up location j and τ_{wk} is a fixed extra time to account for the variability of walking time implying the risk of late arrival at the stop.[5]

When making a trip request, users should specify the latest pick-up time interval lp_i they are willing to accept. Once again, this threshold is slightly different for egress and access passengers. The former has to wait for the DRF vehicle at the terminal, so they are supposed to be less tolerant towards the waiting time than the latter, who can generally wait at home or the workplace. The lp_i is computed as follows:

$$lp_i = \begin{cases} t_{w,max} & \text{if } i \in E \\ (1 + \gamma_i)\, t_{w,max} & \text{if } i \in A \end{cases} \tag{6.16}$$

where the time tolerance level is ruled by the coefficient γ_i introduced in Equation 6.6.

The last component of the user time windows is the latest drop-off time interval ld_i, already introduced with the FRF service. This formulation is slightly different, since the term $t_{w,max}$ is replaced by lp_i, as follows:

$$ld_i = min\left\{lp_i + (1 + \delta_i + \gamma_i)\frac{d_{T,i}}{v},\ (1 + \gamma_i)\frac{d_{T,i}}{v_p}\right\} \tag{6.17}$$

6.3.4 The dispatching algorithm for the DRF

Every time a new trip request i (consisting of a user or group of users) occurs, the optimal matching between the demand (users group requesting the trip) and the supply (vehicle fleet) is carried out by the dispatching algorithm (Figure 6.7). Our algorithm follows an insertion heuristic approach, which is widely used in practice to solve transportation scheduling problems, as it is computationally fast, provide very good solutions compared to optimality and can easily handle complicating constraints (Campbell and Savelsbergh, 2004). Our insertion heuristics involve three levels of exploration of the feasible solutions, including:

---

[5] *It can be reasonably set to 1 minute.*

i. The set of routes $r \in R$, i.e., the sequence of already scheduled stops (to be) visited by the vehicles *(0, 1, …, p, p+1, …, m, 0)* where 0 refers to the terminal. Note that each route corresponds to a complete cycle.

ii. The set of possible virtual stops for the new passenger $s \in S$.

iii. The set of feasible insertions of the request in the route schedule between two already scheduled stops $(p, p+1) \in r$.

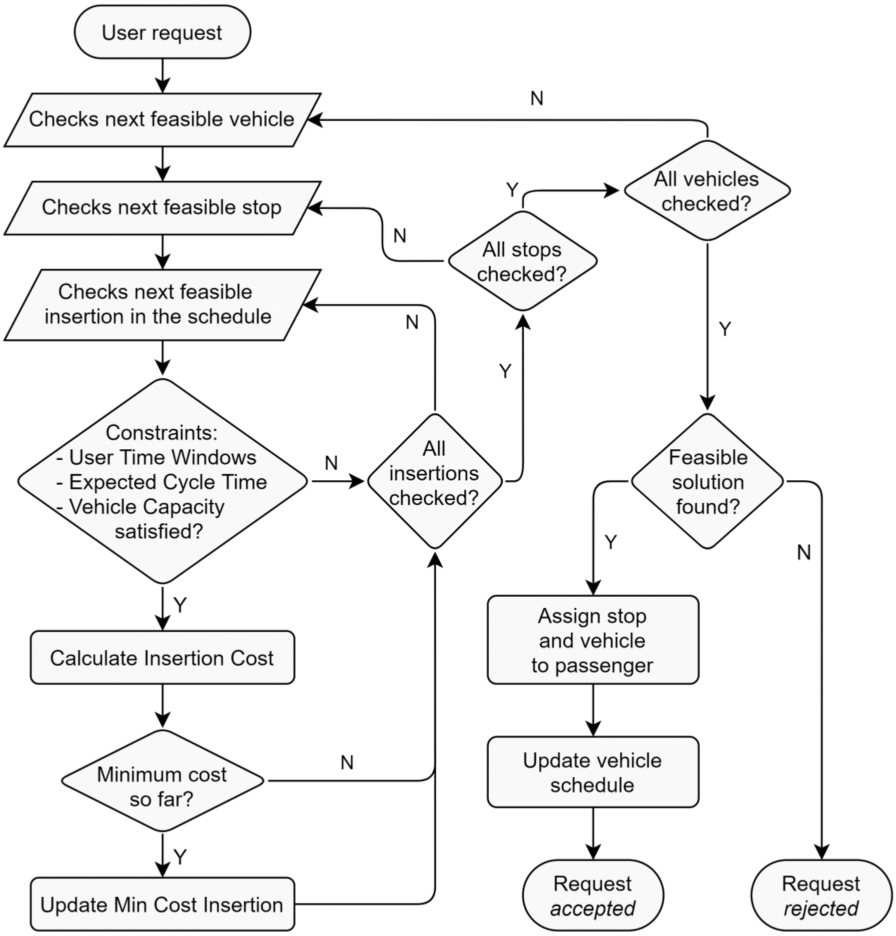


Figure 6.7: Flow chart of the dispatching algorithm for the DRF service.

When determining R for a new trip request i, the algorithm first includes the vehicle routes of the current service cycle. In case no feasible insertion is found, R is updated including the vehicles' route schedule of the next cycle. In this way, the possibility of rejection is reduced.

The algorithm only considers the nearest three intersections to the user's origin (or destination), plus the nearest already scheduled virtual stop, so to limit the vehicles' detour from the original route. This implies a maximum of four possible stops, which is important to limit the length of S, especially in very "dense" street networks, and reduce the computational time. Moreover, through this approach, we keep constant the maximum number of candidates and unbind the time complexity of the algorithm from the granularity of the street network.

In Figure 6.8 the scheme of how passengers can be potentially assigned to the nearest three stops is depicted: the red dotted arrows indicate the stops where a single passenger request can be served, while the green dashed arrows indicate the stops where the consolidation of more than one request can happen. We ignore in this phase, for computational simplicity, the possibility of having a fourth virtual stop which is the nearest one already scheduled but not yet visited by the vehicle.

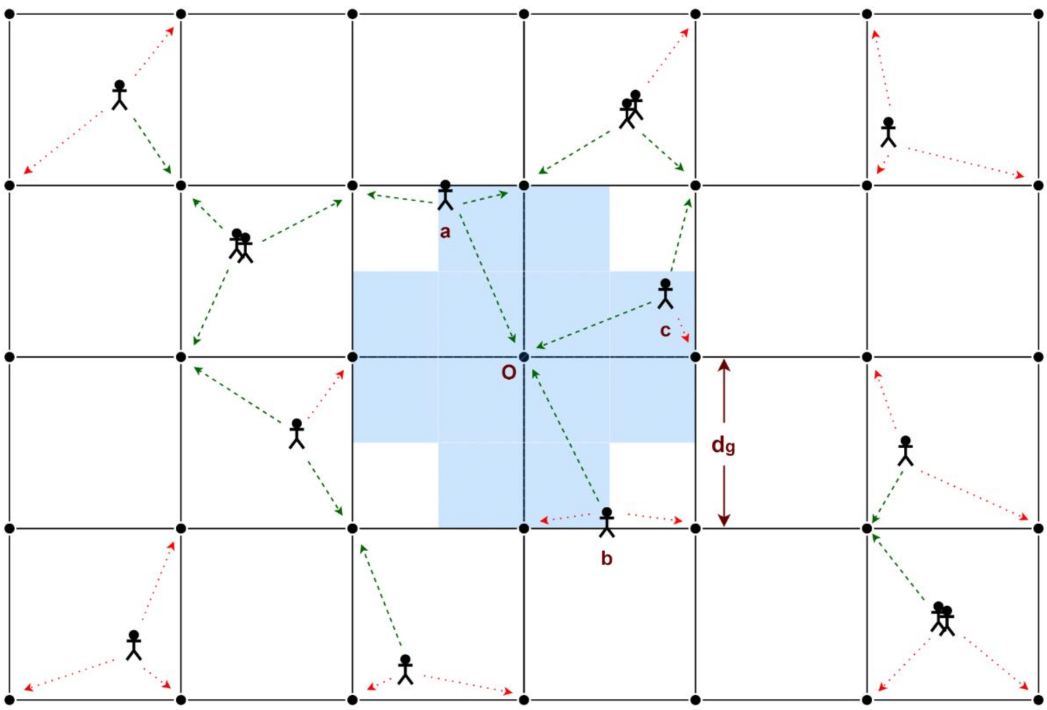


Figure 6.8: Possible options of demand consolidation at virtual stops.

Look at the virtual stop O in Figure 6.8: it is a potential stop for passengers a, b and c, where their trip requests can be consolidated. In general, we can determine an "attractivity area" around each virtual stop (the light blue area around stop O in Figure 6.8) equal to $3d_g^2$. Therefore, considering the whole service region of area LW, the probability for two random requests of being consolidated in the same stop is given by the following equation:

$$p_c = \frac{3\,d_g^2}{L\,W} \tag{6.18}$$

from which we derive the probability of no-consolidation for two requests: $p_{nc} = 1 – p_c$.

Let us consider the expected number of trip requests n that a vehicle should serve along its route (see Equation 6.7). Each new request $i+1$ has a joint probability of no-consolidation with respect to each of the i already scheduled requests, p_{nc}^2. Then, considering the $n–1$ requests after the first one, we can derive[6] the correction coefficient ξ_c which takes into account the decrease of the expected cycle time due to the demand consolidation.

$$\xi_c = \frac{\sum_{i=1}^{n-1} p_{nc}^i}{n-1} = \frac{p_{nc}\,(1-p_{nc}^{n-1})}{(n-1)\,(1-p_{nc})} \tag{6.19}$$

As regards the set of feasible insertions of a new request in the route schedule, for each $r \in R$ and each $s \in S$, the algorithm repeats the procedure of inserting s between p and $p+1$ for each $p = p^*, \ldots, m$, where p^* is the first not yet visited stop of r. verifying that the following constraints are satisfied and computing the cost of the insertion.

The first constraint to be met is that the time t_s at which the vehicle v will stop at s must be consistent with the time windows of the newly arrived user group i, as defined by Equations 6.15-6.17. This can be expressed as follows:

$$\begin{cases} t_0 < t_s \leq ld_i & \text{if } i \in E \\ ep_i \leq t_s \leq lp_i & \text{if } i \in A \end{cases} \tag{6.20}$$

The second constraint is related to the departure time t_0 from the terminal and to the expected cycle time ECT (see Equation 6.14), as shown below:

$$\begin{cases} 0 < t_0 \leq lp_i & \text{if } i \in E \\ t_s < t_0 + ECT \leq ld_i & \text{if } i \in A \end{cases} \tag{6.21}$$

Finally, the third constraint imposes that the number of passengers on board $load_{v,s}$ when vehicle v stops at s must not exceed the vehicle capacity, as expressed by the following relation:

$$load_{v,s} \leq Cap \tag{6.22}$$

If all the constraints are met, the cost function (inspired by Fielbaum et al. (2021)) of inserting s into r between p and $p+1$ is computed as the sum of the cost c_i for the user (group of users) i to be inserted, the additional cost Δc_d for the passengers who are delayed due to the insertion, and the cost Δc_o for the operator due to the detour.

$$cost(r,s,p) = c_i(r,s,p) + \Delta c_d(r,s,p) + \Delta c_o(r,s,p) \tag{6.23}$$

[6] The derivation of ξ_c is based on the sum of the geometric series $\sum_{k=1}^{n} r^k = \frac{r(1-r^n)}{1-r}$

Equations 6.24-6.26 show in detail the three components of the cost function, expressed in terms of time. The cost for the user group is given by the weighted sum of the three components of the travel time minus the ride time related to the shortest path from/to the terminal (since it is not affected by the optimization procedure), multiplied by the number of users in the group G_i (see Equation 6.2) making the request.

$$c_i(r,s,p) = G_i \left(w_{wk}\, t_{wk,i}(s) + w_{wt}\, t_{wt,i}(r,p) + w_{rd} \left(t_{rd,i}(r,s,p) - \frac{d_{T,i}}{v} \right) \right) \tag{6.24}$$

where w_{wk}, w_{wt} and w_{rd} are weighting coefficients related to the walking, waiting and ride time, respectively. The additional costs both depending on the detour time Δt_r caused by the insertion of s in r:

$$\Delta c_d(r,s,p) = \left(w_{rd}\, n_{rd,d}(r,p) + w_{wt}\, n_{wt,d}(r,p) \right) \Delta t_r(r,s,p) \tag{6.25}$$

$$\Delta c_o(r,s,p) = w_o\, \Delta t_r(r,s,p) \tag{6.26}$$

where $n_{rd,d}$ is the number of passengers on board affected by the detour, $n_{wt,d}$ is the number of users who will have to wait an extra Δt_r at the stop due to the schedule update and w_o is the weighting coefficients of the additional cost caused by the detour for the operator, related to the vehicle-kilometres travelled.

After having examined all the feasible solutions (if any), the insertion heuristic chooses the assignment that minimizes the cost function, as shown below:

$$\min_{r \in R; s \in S;\, p \in r} \{ cost\ (r,s,p) \quad \text{s.t. Equations 6.20-6.22} \} \tag{6.27}$$

Regarding the time complexity of the insertion heuristic, considering the “first-level” of exploration of the feasible solution, the size of the set of feasible routes r is linearly related to the fleet size n_V. It is an input parameter, but it should be properly chosen based on the expected number of requests n to serve (imagine having no constraints on n_V). If n is sufficiently high, n_V is linearly related with n. The second level regards the potential stops s to associate with the user request. As already explained, we chose to limit the number of feasible stops to a maximum of four. Finally, in the third level, the number of feasible insertions in the vehicle schedule ($m+1$) is also linearly related with n. Based on these considerations, the time complexity of our insertion heuristic is $O(n^2)$, for each new request. This means that the matching process between passengers and vehicles requires low computational efforts and takes place dynamically, in real-time, even with a high number of requests per unit of time.

6.3.5 Output indicators

The model results can be assessed through different output indicators (Inturri et al., 2019) to compare the two feeder services, FRF and DRF.

Table 6.1. Description of the user-related output indicators.

Indicator	Abbreviation	Description / Formulation
# Users	N_U	Total number of travellers generated during the simulation
% Walking Users	*WLK*	Percentage of users who do not request for the feeder service and walk to/from the terminal
% Accepted Passengers	*ACP*	Percentage of travellers using the feeder service
% Rejected Requests	*REJ*	Percentage of travellers whose trip request was rejected
% Delayed Passengers	*DEL*	Percentage of users who are delayed beyond the time windows
Avg Pre-Trip Time [s]	T_{PT}*	Average time that elapses between when a user makes the trip request (and is accepted) and starts walking to the stop
Avg Walking Time [s]	T_{WK}	Average time for a user of the feeder service walking from the origin to the stop or from the stop to the destination
Avg Waiting Time [s]	T_{WT}	Average time that a traveller waits at the stop for the vehicle
Avg Ride Time [s]	T_{RD}	Average time that a passenger spends on board a vehicle
Avg Total Travel Time [s]	T	Average total travel time, which is the sum of the following components: $T = T_{WK} + T_{WT} + T_{RD}$
Rejection Penalty [s]	T_{rej}	Total walking time related to rejected travellers divided by the total number of passengers.
Avg Time Stretch	*STR*	Average ratio, for all passengers, between T and the sum of the ride and walking time if using the shortest path (no waiting times).
Avg Disutility [min]	U	As defined in Quadrifoglio and Li (2009), it is the weighted sum of T_{WK}, T_{WT} and T_{RD} considering the weighting coefficients of Equation (6.25)

*only for the DRF service

Table 6.2. Description of the operator-related output indicators.

Indicator	Abbreviation	Description / Formulation
Tot. Driven Distance [km]	D	Total distance travelled by the vehicles during the simulation time
Tot. Energy Consumption [kWh]	TEC	Total energy used by the vehicles during the simulation time: TEC = D · EC[7]
Avg Vehicle Occupancy [pax]	AVO	Average number of passengers per vehicle
Transport Intensity [km/pax]	TI	Average distance travelled by the service per transported passenger $TI = D/(NU \cdot ACP)$
Commercial Speed [km/h]	v_c	$v_c = D / ST$
% Stopping Time	ST	Ratio between the total time spent by vehicles at the terminal or the stops, and the simulation time

Another indicator that encompasses both the user and operator point of view is the so-called Total Unit Cost (TUC). It is based on Inturri et al. (2019) and is equal to the sum of the Passenger Unit Cost (PUC) and the Operator Unit Cost (OUC):

$$TUC\ (€/pax) = PUC + OUC \tag{6.28}$$

Where:

$$PUC\left(\frac{€}{pax}\right) = \left[w_{wk}\left(T_{wk} + T_{rej}\right) + w_{wt}\,T_{wt} + w_{rd}T_{rd}\right] \cdot VoT \tag{6.29}$$

$$OUC\ (€/pax) = \frac{D \cdot C_{km} + n_V \cdot ST \cdot C_h}{N_U \cdot ACP \cdot (1 - WLK)} \tag{6.30}$$

With VoT (€/h) the value of time for the travellers, C_{km} (€/veh km) the distance-related operator cost, C_h (€/veh h) the hourly driver cost, and ST (h) the total simulation time. For a first test, VoT was estimed to be 10 €/h, while regarding the operator-related costs we set C_h = 25 €/veh h and C_{km} ranging from 0.5 to 0.1 €/veh km according to the vehicle size.[8] It is worthy of notice that in the case of a service performed by automated vehicles, the second term of the OUC would be equal to zero since no drivers are considered.

---

[7] *Energy consumption (EC): automobile 4 pax 0.18 kWh/km (range 250 km) - minivan 6/8 pax 0.3 kWh/km (range 200 km) - minibus 9 seats (20 pax) 0.6 kWh/km (range 120-150 km) https://www.bluebus.fr/caracteristiques-techniques - electric bus 50 seats 1 kWh/km (range 120 km) https://www.sustainable-bus.com/news/electric-bus-range-focus-on-electricity-consumption-a-sum-up/.*

[8] *Distance-related operator cost ($€_{km}$): automobile 4 pax 0.1 €/veh km - minivan 6/8 pax 0.25 €/veh km - minibus 9 seats (20 pax) 0. 5 €/veh km.*

6.4 Application of the model with different parameters

The model described above is programmed in the *NetLogo* development environment. The operation parameters of the feeder services can be varied, as already depicted in Figure 6.3 and Figure 6.4. This approach provides the advantage of obtaining useful insights for a transit agency operating in different urban contexts and under different demand patterns. Moreover, it is easy to perform a broad range of sensitivity analyses regarding the main input parameters of the system, which are outlined below.

- Geometric: length L and width W of the service area, the distance between stops d_s (for the FRF) and grid street spacing d_g.
- Service: type of service (fixed/flexible), total simulation time ST (h) and headway h (min).
- Supply: number of operating vehicles n_V, vehicle average speed S (km/h) the maximum seat capacity *cap* of a vehicle.
- Demand: average demand density $\bar{\lambda}$ (pax/h), trip direction coefficient α, demand decay coefficient Λ (ratio between demand density at $x = 0$ and at $x = L$), probability of having one user per request p_1, maximum waiting time $t_{wt,max}$ (min); and maximum walking distance $d_{wk,max}$.
- Cost: weighting coefficients related to the passenger w_{wk}, w_{wt}, w_{rd} and the operator w_o.

6.4.1 First set of simulations: finding the critical demand density

We first demonstrate the effectiveness of the simulation model and the proposed insertion heuristic by reproducing the scenarios described by Quadrifoglio and Li (2009). Following this approach, we computed the disutility function U for the FRF and the DRF services under increasing demand levels and keeping supply, cost and geometric parameters fixed. The critical demand density λ_c is calculated as the one that provides an equal passengers' disutility for the two services. The simulation input parameters are: simulation time ST = 8 h, service area dimensions $L = 3.2\ km$ and $W = 0.8\ km$, grid spacing $d_g = 0.1\ km$, trip direction coefficient $\alpha = 0.5$ and demand decay coefficient $\Lambda = 1$ (spatially uniform demand). The input parameters used for the simulations are listed in Table 6.3, where the abbreviation *Sc-A1* refers to a one-vehicle case and *Sc-A2* to a two-vehicles case. To better reproduce the assumptions made by Quadrifoglio and Li (2009), we chose sufficiently high values of $t_{wt,max}$ and *cap* to relax the passenger time windows and to assume an unlimited vehicle capacity, respectively.

Besides, we set up a new scenario (Table 6.3, Sc-Base) able to better exploit the novelty of our methodology and include the “real-world” constraints of the DRF operations (passenger time windows, vehicle capacity, maximum cycle time, etc.) that were not considered in Quadrifoglio and Li (2009). Sc-Base also served as a reference scenario against which to compare other 10 operational scenarios (as shown in the next sub-sections) and thus perform a sensitivity analysis of the most significant input parameters. Each scenario has been replicated 25 times to have a statistic of events.

Table 6.3. Input parameters adopted for the first set of simulated scenarios.

Parameter		**Abbreviation**	**Values**		
			Sc-A1	**Sc-A2**	**Sc-Base**
Stop spacing [m]		d_s	490	490	*425*
Average demand density [pax/ km^{2}h]		$\bar{\lambda}$	5-30	20-40	*5-60*
$\Pr(G_i = k)$ Group probability		p_1	1.0	1.0	*0.8*
Maximum waiting time [min]		$t_{w,max}$	30	30	*10*
Walking speed [m/s]		v_p	0.9	0.9	*1.0*
Vehicle cruising speed [km/h]		v	32.0	32.0	*30.0*
Vehicle capacity [pax]		cap	50	50	*20*
Number of vehicles		nv	1	2	*3*
Headway [min]	(FRF)	h	20	10	*7*
	(DRF)		17-30	10-20	*7*
Cost coefficients [-]		w_{wt}	1	1	*2*
		w_{wk}	3	3	*2*
		w_{rd}	2	2	*1*
		w_o	0	0	*4*

Results of the first set of simulations in terms of average disutility and TUC are reported in Figure 6.10 and Figure 6.12.

Sc-A1 (Figure 6.9) and Sc-A2 (Figure 6.10) find the critical demand density of respectively 15 pax/h-km^2 for the 1-vehicle case and 30 pax/h-km^2 for the 2-vehicle case, a little bit higher than the 12 pax/h-km² and 28 pax/h-km² found by the analytical model of Quadrifoglio and Li (2009). This is probably due to the capability of the simulation model to reproduce a more efficient vehicle dispatching and routing for the DRF service, thus enlarging its range of performance.

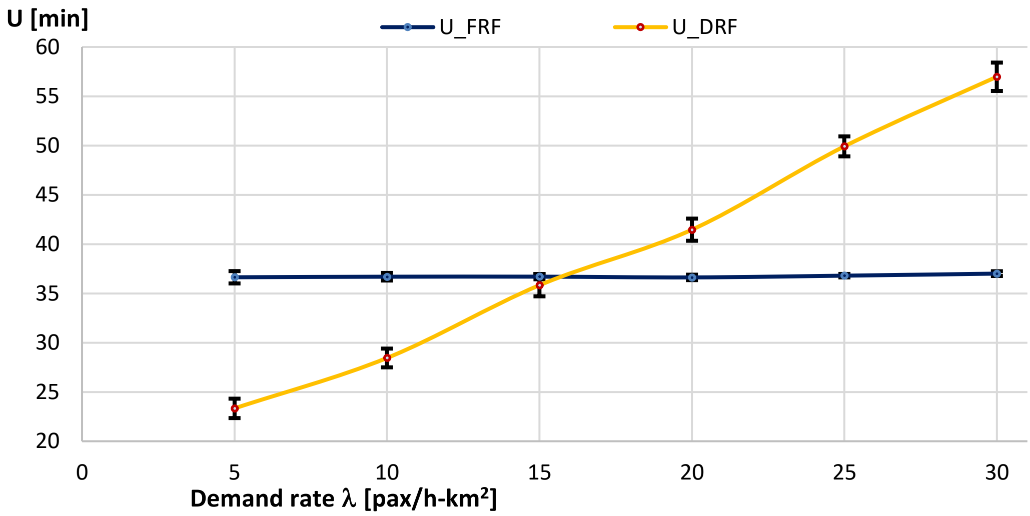


Figure 6.9: Disutility of Sc-A1 (bars showing confidence intervals).

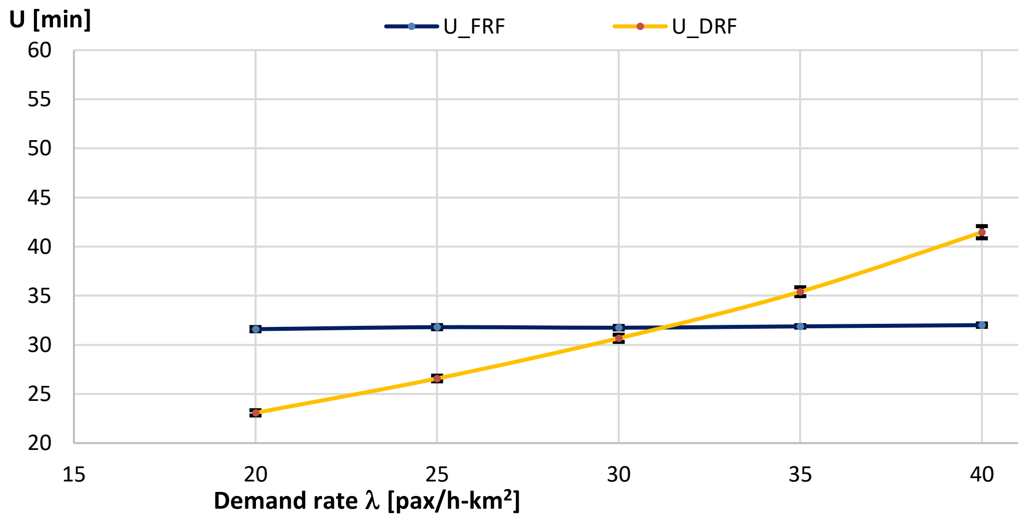


Figure 6.10: Disutility of Sc-A2 (bars showing confidence intervals).

This is even more evident from the results of the Sc-base (Figure 6.11), obtained by removing the constraints corresponding to the analytical model, where the switching point for the passenger convenience shift in the range between 45 and 50 pax/h-km². However, by looking at the TUC (Figure 6.12), which is an indicator of system efficiency including also the operator point of view, the threshold is a bit lower, between 40 and 45 pax/h-km². This is imputable to the fast increase of the DRF supply cost (distance travelled) for a higher demand rate.

As expected, the disutility for the FRF travellers, as opposed to the DRF ones, is not very sensitive to the demand density variation. This can be ascribed to the regularity of the service since the average ride time is almost the same for travellers, the headway (affecting the average waiting time at stops) is constant as well as the walking time, which only depends on W, d_s and v_p. In particular, in Figure 6.11 one can notice a small increase in U with the demand density due to the increasing number of passengers and vehicles causing higher idle times at stops.

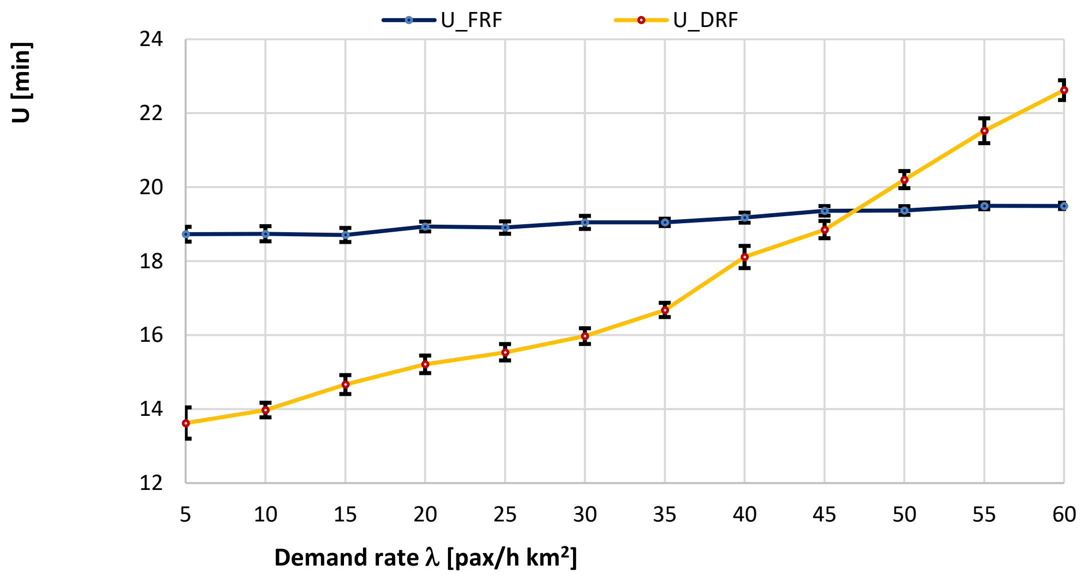


Figure 6.11: Performance of the Sc-Base scenario in terms of User Disutility (bars showing confidence intervals).

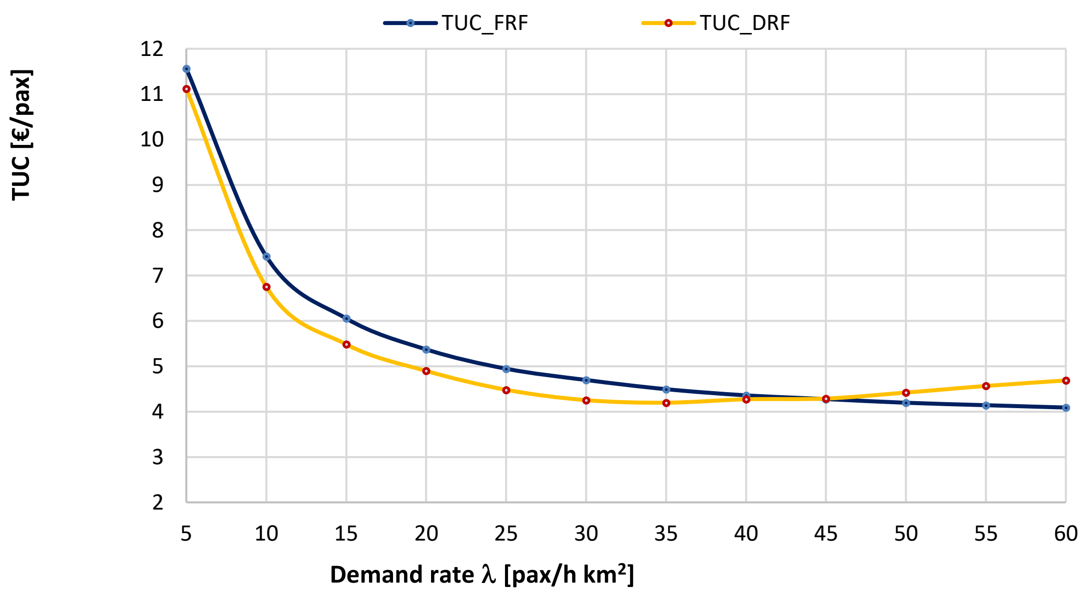


Figure 6.12: Performance of the Sc-Base scenario in terms of TUC.

According to the results obtained, we decided to fix the demand density to the critical value of 40 pax/h-km^2 and perform a second set of simulations to assess the attractiveness of the two services by varying the other model parameters.

6.4.2 Second set of simulations: testing different demand/service configurations

Several scenarios have been defined to reproduce different use cases by varying:

1. The demand spatial distribution (Sc-1): from uniform (Sc-base) to trapezoidal (Sc-1a) or triangular (Sc-1b); the first one might represent a residential area, while the other two might mimic a TOD-like land-use area, where the demand progressively decreases from the MRT station to the outskirt.
2. The demand O/D pattern (Sc-2): from uniform ($\alpha = 0.5$) of the Sc-base, to more concentrated demand patterns to/from the MRT station ($\alpha = 0.9$, Sc-2a; $\alpha = 0.1$, Sc-2b). This allows mimicking different demand configurations according to both land use and temporal demand distribution. Sc-base is more representative of (i) a mixed land-use area with balanced trips from and to the MRT station, or (ii) a mono-functional land-use area during off-peak hours (e.g., a residential area). Both Sc-2a and Sc-2b might represent mono-functional land use areas during peak hours. The former is a residential area during the "morning peak" period or a workplace/service area during the "evening peak" period and the opposite applies to the latter.
3. The area stretching (Sc-3): modifying the ratio L/W from 4 (Sc-base) towards less/more stretched areas (L/W equal to 2 or 3, for Sc-3a and Sc-3b, and 6 for Sc-3c). This represents different MRT services according to the station spacing (from 0.6 km to 1.2 km), e.g., from an urban metro service to a regional railway service.
4. The vehicle capacity (Sc-4): increasing the number of vehicles of smaller capacity, from 3 vehicles of 20 seats (Sc-base) to 5 vehicles of 8 seats (Sc-4a), 10 vehicles of 4 seats (Sc-4b), and 20 vehicles of 2 seats (Sc-4c). This allows testing different types of flexible feeders, from those performed by minibuses or vans to a ride-sharing service with small vehicles (e.g., UberPool).

The ten simulated scenarios are synthetically described in Table 6.4 concerning the Sc-base. All the other input parameters are set equal to those of Sc-base, as reported in Table 6.3.

Table 6.4. Description of the second set of simulated scenarios.

Sc-base	Scenario	a	b	c
MRT station Uniform demand L W	Sc-1	MRT station Trapezoidal demand L W	MRT station Triangular demand L W	-
MRT station $\alpha = 0.5$ O/D point	Sc-2	MRT station $\alpha = 0.9$ O point	MRT station $\alpha = 0.1$ D point	-
MRT station L L/W=4 W	Sc-3	MRT station L L/W=2 W	MRT station L L/W=3 W	MRT station L L/W=6 W
3v x 20s	Sc-4	5v x 8s	10v x 4s	20v x 2s

The main results are reported and commented on the next subsections. Results of the two services will be first separately presented; then, a comparison between them is performed. From the point of view of the demand, we focus on PUC, from the point of view of the operator on OUC, while TUC encompasses both points of view. An overview of scenario results with all the output indicators defined in Table 6.1 and Table 6.2 is presented in Appendix A1.

Fixed-route feeder. Results of the simulations for the FRF in terms of PUC, OUC and U are reported in Figure 6.13.

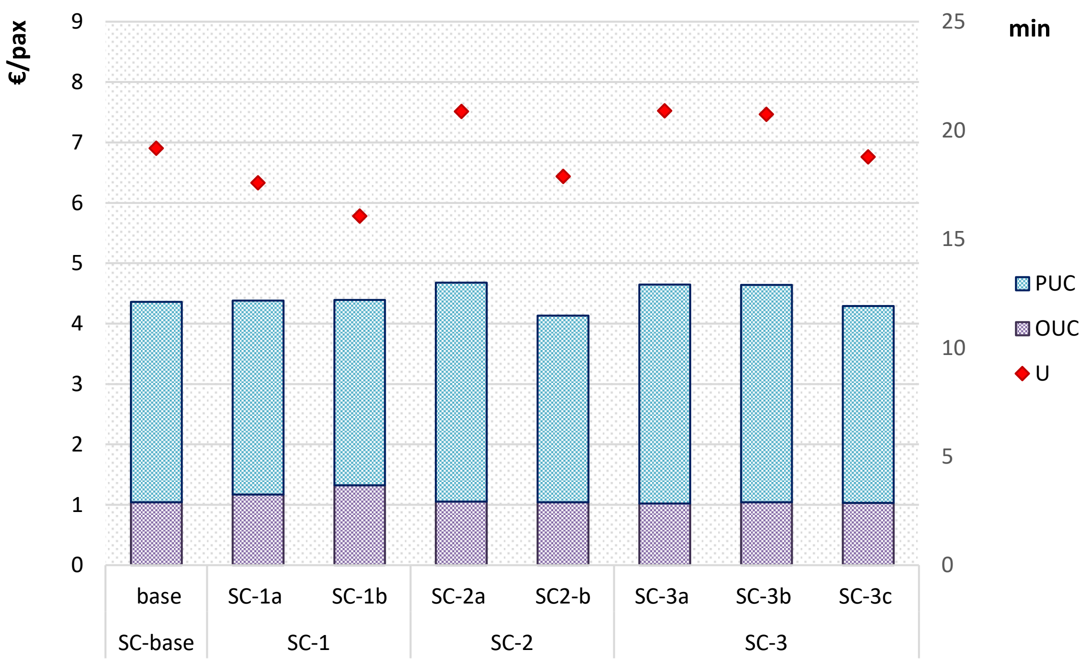


Figure 6.13: Results of the second set of simulations for the FRF.

As expected, the operator cost OUC is quite constant over scenarios Sc-1 and Sc-2 because the fixed service is not affected by the spatio-temporal demand distribution, while Sc3-a provides a smaller operation cost because of the shorter travelled distance.

If we look at demand spatial distribution (Sc-1), when demand is higher near the MRT station, the TUC of the FRF remains almost the same (+1%). We report a reduction of PUC of 3% (SC-1a) and 7% (Sc1-b) if compared to a uniform demand distribution (Sc-base). This is mainly due to shorter passenger ride times since they are more concentrated near the MRT station. On the other hand, the OUC increases, since fewer travellers are using the feeder service.

Interesting results emerge if we consider different demand patterns (Sc-2). In particular, there are two opposite trends: a higher concentration of users directed to the MRT station (Sc-2a) worsens the performance of the FRF in terms of PUC (+9%) and, thus, TUC (+7%), while the opposite occurs when users originate mostly at the MRT station (with multiple destinations) (-7%, -5%). This could be explained in terms of the regularity of the service. In the first case, bus bunching might occur thus worsening the headway regularity and average waiting times. Conversely, the regularity improves in the opposite case (Sc-2b) because passengers are generated mostly at the MRT station and board the FRF with scheduled departure times.

A different configuration of the service area (more or less stretched) leads to different results. In particular, there is a decreasing trend of PUC (and TUC) from less-stretched to more-stretched areas with the best result in the case of Sc-3c. This outcome was expected as well since walking time is more relevant for users than the other time components. In this respect, a more stretched area implies shorter walking times and higher ride time.

<u>Demand-responsive feeder</u>. Figure 6.14 reports the results from DRF simulations, similarly to the FRF case.

Demand spatial distribution and, in particular, a demand concentrated closer to the MRT station (Sc-1) improves the performance of DRF compared to Sc-base, also because of better dispatching of passengers that can be more easily assigned to a few virtual stops. As in the previous case, the decrease in PUC (-2% in Sc-1a and -13% in Sc-1b) are compensated by the increase in OUC (+14% and +28%, respectively).

Interesting and quite different results compared to the FRF case emerge if we consider different demand patterns (Sc-2). In particular, it is possible to see a lower performance of the DRF in Sc-2 with respect to Sc-base. More specifically the worst case occurs in Sc-2b, i.e., when the demand is mainly originated at the MRT station.

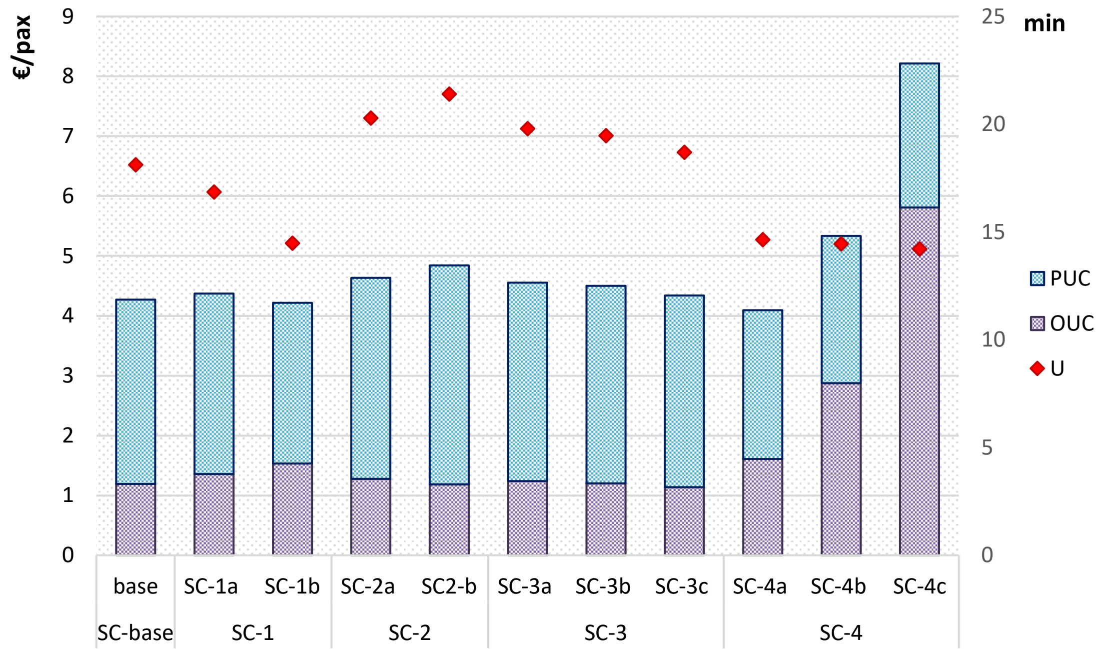


Figure 6.14: Results of the second set of simulations for the DRF.

Three more scenarios are simulated for the DRF case, i.e., by varying the number and capacity of vehicles while keeping the total seat capacity constant (Sc-4). A clear increasing trend of TUC is visible due to higher OUC once the number of vehicles and drivers increase. Interestingly, PUC decreases from Sc-base to Sc-4, but the relative difference between the sub-scenarios is very low. This suggests that it is neither profitable nor beneficial to use a higher number of small vehicles (i.e., cars) to perform the DRF service (Sc4-b, Sc-4c), while vans might represent a good compromise. They would be a suitable solution from the user perspective while implying higher operator cost (+35%) but resulting in a slightly lower TUC (-4%) compared to the minibuses case of Sc-base.

This result confirms the impact of driver costs on the total operator cost and suggests the possibility of looking at automated vehicles (AV) as a potential solution. Figure 6.15 shows a comparison between TUC with and without the driver cost for Sc-base and Sc-4 showing how the difference between the two values increases with the number of vehicles. Clear savings can be obtained in the case of Sc-4c, implying 20 vehicles of 2 seats, i.e., a Uber-like service. However, this option should be analysed more in detail by looking at other important variables like the cost of AVs and the different operation conditions (e.g., road infrastructure, commercial speed, increased road congestion) and the willingness of passengers to use them.

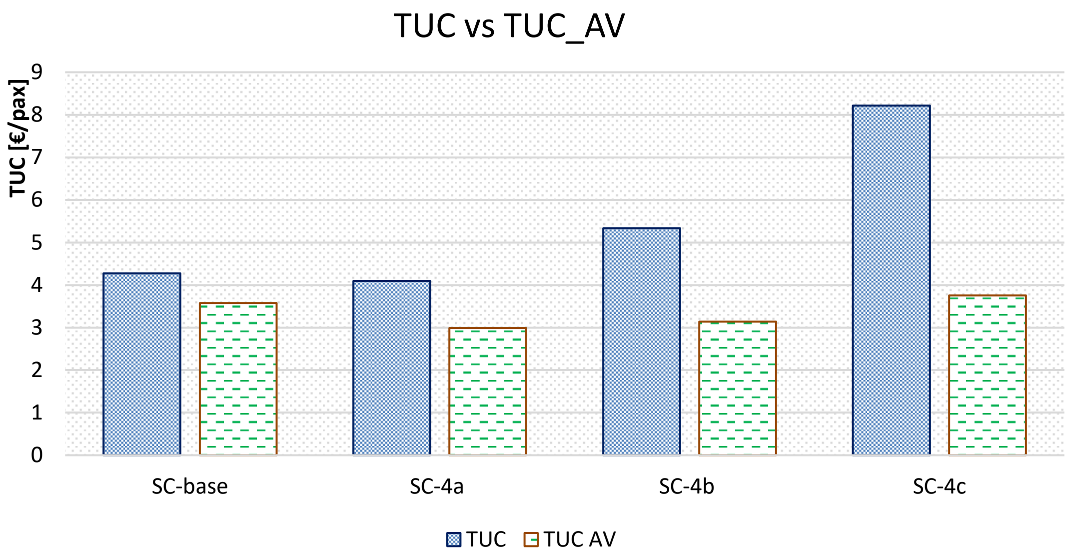


Figure 6.15: Comparison between TUC for traditional and automated vehicles.

DRF vs. FRF

Figure 6.16 shows the comparison between DRF and FRF. This analysis allows finding what service suits better to each particular situation.

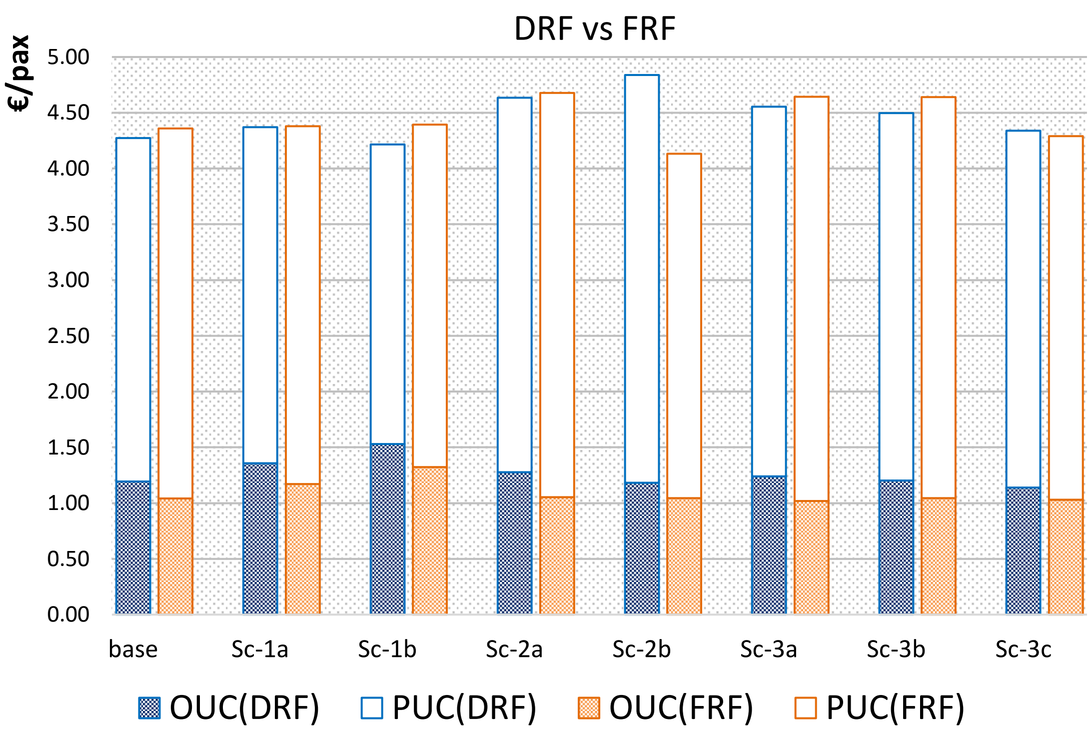


Figure 6.16: Comparison between DRF and FRF in different scenarios.

In general, the DRF performs better in Sc-1b where we have a decreasing demand from the MRT station to the outskirts, thanks to the flexibility of the service which allows a better matching between passengers and vehicles and lower distances driven.

A not very clear trend is obtained by varying the shape of the service area. The two services are almost equivalent in Sc-3a, Sc-3b and Sc3-c where a better performance for the users is compensated by higher costs for the operator (and vice versa). By stretching the area (Sc3-c), thus representing different catchment areas according to the MRT network (in terms of the station and line spacing) the FRF starts becoming more attractive than the DRF.

Demand pattern is a critical issue to consider when planning and designing feeder services (Sc-2). In this respect, simulation results suggest adopting the FRF instead of the DRF when the demand is mainly originated at the MRT station (to multiple destinations) and vice versa, since this would imply smaller TUC, especially in the first case (Sc2-b). This is due to the regularity of the FRF service which is higher compared to the DRF. On the opposite, when the demand is homogeneously distributed in origin/destination from/to the MRT station, a DRF service is to be preferred.

6.5 Conclusions

This chapter presented an agent-based model for the mass rapid transit feeder design. Fixed route and demand-responsive feeder services (FRF and DRF) have been compared to understand their attractiveness under different conditions, by taking into account both operator and user perspectives. The ABM reproduces a parametric synthetic environment where different variables related to demand, service area and service configurations can be easily modified. The first set of scenarios allowed a comparison with the analytical model presented in Quadrifoglio and Li (2009) finding a critical demand density that acts as a switching point between the two services. This was done to demonstrate the effectiveness of the simulation model and the proposed insertion heuristic by reproducing the same scenarios and introducing new realistic constraints, allowing to find a demand density threshold.

The second set of simulations was performed to reproduce different use cases by changing the demand/service configurations while maintaining the same demand density, mimicking realistic scenarios of different transport services and land use areas. Results of ten scenarios show that a demand concentrated near the MRT station improves the performance of both the DRF and FRF services with the DRF to be preferred due to lower total (user and operator) unit costs. Stretching the service area improves the performance of the FRF services because of the lower travellers' walking time, while a non-homogeneous demand pattern may suggest adopting different service configurations (FRF of DRF) along the day. Finally, by increasing the DRF fleet while maintaining the same total capacity (i.e., using smaller vehicles), one can notice that there is no gain in terms of total unit cost since the increase in the operator cost is not compensated by an equivalent decrease of user cost. This could be changed by considering automated vehicles that would drastically reduce the operator cost.

In terms of practical implications, results suggest that the DRF is to be preferred in TOD-like areas characterized by a high negative density gradient from the MRT station (Sc1-b), or in peripheral areas where station spacing is quite high (Sc-base). Vice versa, FRF should be preferred in mono-functional land use areas (e.g., residential or workplaces) during peak periods. Besides, the same transport operator might switch services along the day as the demand pattern changes over time, using an FRF during peak periods and a DRF during off-peak periods. This could be done using the same vehicle fleet and staff, just changing the operation parameters.

In summary, the proposed model can be used to optimize a feeder-trunk transit network, by using variables related to MRT network topology, its line and station spacing, spatial and temporal demand pattern as inputs, and finding the optimal feeder operation as output.

This study comes with some limitations opening future research. First, the theoretical nature of the model, and the application to a synthetic case study, while making it generalizable, does not allow to formulate context-based practical implications. In this respect, the assumptions made on user preferences for different time components, the value of time, and operator cost can influence results. Therefore, the model needs to be tested with more data and use cases. In future research endeavours, it will be interesting to compare the performance of DRF and of parallel FRF services that could be used to serve the service area with the same operation costs, reducing the walking time of users while increasing the headway. Finally, the impact of AV on service performance should be analysed more in detail to further explore the potential of this new technology to provide efficient feeder services.

Appendix A

This appendix reports the summarizing tables related to the FRF (Figure A1) and DRF (Figure A2) simulation outputs, showing their variations compared with the base scenario (SC-base).

OUTPUT	SC-base	SC-1		SC-2		SC-3		
	base	SC-1a	SC-1b	SC-2a	SC2-b	SC-3a	SC-3b	SC-3c
USER-RELATED								
# tot Requests (NR)	820	0%	0%	0%	0%	1%	0%	1%
% Walking Users (WLK)	7.50	135%	260%	11%	2%	32%	16%	3%
% Accepted Requests (PAX)	99.86	0%	0%	-1%	0%	0%	0%	0%
% Rejected Requests (REJ)	0.14	-13%	41%	443%	-96%	-100%	-64%	296%
% Delayed Requests (DEL)	4.78	58%	112%	34%	-17%	117%	86%	-10%
Avg Walking Time (Twk) [min]	4.98	1%	2%	0%	1%	37%	20%	-17%
Avg Waiting Time (Twt) [min]	1.92	-2%	-5%	69%	-76%	-20%	-5%	13%
Avg Ride Time (Trd) [min]	6.10	-11%	-25%	-14%	24%	-17%	-2%	13%
Avg Total Travel Time (T) [min]	12.99	-5%	-11%	4%	0%	3%	6%	1%
Avg Time Stretch (STR)	1.61	2%	5%	4%	-1%	-8%	-3%	7%
Avg Passenger Disutility (Upax)	19.18	-8%	-16%	9%	-7%	9%	8%	-2%
Passenger Unit Cost (PUC) [€/pax]	3.32	-3%	-7%	9%	-7%	9%	8%	-2%
OPERATOR-RELATED								
Total Driven Distance (D) [Km]	381.33	0%	0%	0%	0%	-15%	-7%	-1%
Total Energy Consumption (TEC) [kWh]	228.82	0%	0%	0%	0%	-15%	-7%	-1%
Avg Vehicle Occupancy (AVO)	3.13	-21%	-41%	-14%	23%	-18%	-3%	15%
Transport Intensity (TI) [km/pax]	0.50	13%	28%	1%	0%	-13%	-5%	-2%
Commercial Speed (Vc) [Km/h]	21.73	3%	6%	-3%	6%	-3%	-1%	-1%
% Stopping Time (ST)	46.88	-1%	-1%	-1%	-1%	17%	8%	2%
Operator Unit Cost (OUC) [€/pax]	1.04	12%	27%	1%	0%	-2%	0%	-1%
SYSTEM-RELATED								
Total Unit Cost (TUC) [€/pax]	4.36	0%	1%	7%	-5%	7%	6%	-2%
Total Unit Cost (TUC) [€/pax] AV	3.73	-2%	-4%	8%	-6%	7%	7%	-2%

Figure A1. Summarizing table with FRF simulation outputs (for scenarios 1-3, percentage variations are reported with respect to Sc-base)

OUTPUT	SC-base	SC-1		SC-2		SC-3			SC-4		
	base	SC-1a	SC-1b	SC-2a	SC2-b	SC-3a	SC-3b	SC-3c	SC-4a	SC-4b	SC-4c
USER-RELATED											
# tot Requests (NR)	820	0%	0%	0%	0%	1%	0%	1%	0%	0%	0%
% Walking Users (WLK)	16.17	81%	143%	44%	17%	44%	14%	-5%	-26%	-31%	-23%
% Accepted Requests (PAX)	94.20	-1%	0%	-9%	-3%	-5%	-2%	0%	5%	6%	5%
% Rejected Requests (REJ)	5.80	24%	-3%	141%	53%	80%	30%	-1%	-87%	-96%	-80%
% Delayed Requests (DEL)	0.04	-50%	-44%	-85%	424%	307%	46%	-2%	-100%	-100%	-100%
Avg Pre-trip Time (Tpt) [min]	1.68	-7%	-12%	85%	-78%	-1%	-3%	-10%	-15%	16%	32%
Avg Walking Time (Twk) [min]	2.27	-8%	-14%	-14%	3%	-8%	-2%	-8%	-4%	-6%	-13%
Avg Waiting Time (Twt) [min]	3.08	9%	7%	-26%	47%	14%	12%	11%	-18%	-6%	-5%
Avg Ride Time (Trd) [min]	6.38	-7%	-26%	29%	9%	-1%	6%	7%	-15%	-27%	-36%
Avg Total Travel Time (T) [min]	11.73	-3%	-15%	6%	18%	2%	6%	5%	-14%	-18%	-23%
Avg Time Stretch (STR)	2.00	10%	16%	6%	16%	17%	9%	3%	-12%	-14%	-18%
Avg User Disutility (Utot) [min]	18.11	-7%	-20%	12%	18%	9%	7%	3%	-19%	-20%	-22%
Passenger Unit Cost (PUC) [€/pax]	3.08	-2%	-13%	9%	19%	8%	7%	4%	-19%	-20%	-22%
OPERATOR-RELATED											
Total Driven Distance (D) [Km]	439.44	-14%	-26%	-7%	-15%	-14%	-9%	-9%	46%	118%	280%
Total Energy Consumption (TEC) [kWh]	263.66	-14%	-26%	-7%	-15%	-14%	-9%	-9%	-27%	-35%	14%
Avg Vehicle Occupancy (AVO)	3.70	-14%	-36%	-2%	31%	-2%	5%	10%	-45%	-73%	-87%
Transport Intensity (TI) [km/pax]	0.64	1%	2%	1%	-12%	-7%	-6%	-11%	39%	105%	264%
Commercial Speed (Vc) [Km/h]	23.06	-3%	-3%	-3%	-4%	-4%	-4%	-5%	4%	9%	13%
% Stopping Time (ST)	39.56	21%	39%	12%	21%	20%	12%	12%	19%	52%	64%
Operator Unit Cost (OUC) [€/pax]	1.19	14%	28%	7%	-1%	4%	1%	-5%	35%	141%	387%
SYSTEM-RELATED											
Total Unit Cost (TUC) [€/pax]	4.27	2%	-1%	8%	13%	7%	5%	2%	-4%	25%	92%
Total Unit Cost (TUC) [€/pax] AV	3.57	-1%	-9%	8%	15%	6%	6%	2%	-16%	-12%	5%

Figure A2. Summarizing table with DRF simulation outputs (for scenarios 1-4, percentage variations are reported with respect to Sc-base)

CHAPTER 7

Adaptive Transit Design: Optimizing Fixed and Demand Responsive Multi-modal Transport via Continuous Approximation

In most cities, transit consists of fixed-route transportation only, whence the inherent limited Quality of Service (QoS) for travellers in sub-urban areas and during off-peak periods. On the other hand, completely replacing fixed-route with demand-responsive (DR) transit would imply huge operational cost. It is still unclear how to ingrate DR transportation into current transit systems to take full advantage of it.

In this chapter, we propose a Continuous Approximation model of a transit system that that gets the best from fixed-route and DR transit, adapting to the demand observed in each sub-region of the urban conurbation and time-of-day. The goal is to provide high transportation capacity while guaranteeing high QoS, two objectives that are instead conflicting if only classic fixed-schedule transportation is deployed, as in today's cities. Our model allows to decide, in each area, whether to deploy a fixed-route or a demand-responsive feeder service, and to redesign line frequencies and stop spacing of the main trunk service accordingly. Since such a transit design can adapt to the spatial and temporal variation of the demand, we call it Adaptive Transit. Our numerical results show that *Adaptive Transit* significantly improves user cost, particularly in suburban areas where access time is remarkably reduced, with only a limited increase of agency cost. We believe our methodology can assist in planning future generation transit systems, able to improve urban mobility by appropriately combining fixed and DR transportation.

7.1 The need for Adaptive Transit Network Design

In recent years, urban transportation has witnessed the birth and spread of new demand-responsive and ride-hailing services, mostly provided by private companies (e.g. Uber, Lyft, Via). In most of cities, these “user-centric” services have penalized the conventional public transport (PT), which has basically not evolved in the last decades and still consists in fixed routes and fixed scheduling with some exceptions, or pilots or services for a specific targeted population (elders or handicapped). The detrimental role of ride-hailing and ridesourcing services towards PT can be reversed by transforming them from PT substitutes to PT complement (Sadowsky and Nelson, 2017). In order to pursue this transformation, it is essential to re-think the whole transit network through a multi-modal approach and an integrated design of the various transport modes (conventional and demand-responsive).

As already shown in the previous chapters, conventional PT is inefficient in sparse demand areas: providing a high number of lines with an adequate frequency to ensure an acceptable QoS to travellers would result in low passenger load factors and thus in an excessive agency cost.

This problem is evident in suburbs and is one of the reasons for geographical inequity in modern society (Giuffrida et al., 2017). On the other hand, demand-responsive transportation is not the solution to all mobility needs, as it is not suitable to serve dense demand (Basu et al., 2018), which it would result in tortuous vehicle routes (Araldo, Gao et al., 2019) and thus high operational cost.

Therefore, a combination of fixed-route (FR) and demand-responsive (DR) services is needed to guarantee high capacity for dense demand areas and, at the same time, good QoS in sparse demand areas. For these reasons, in recent years public authorities have launched pilots to experiment with different ways to complement their offers with on-demand services, by subsidizing ride-sharing companies (Sörensen et al., 2021). Moreover, the scientific community has done a big effort in modelling the performance of DR transportation. However, there is no systematic methodology to guide the design of future-generation transit systems integrating FR and DR transport.

Our contribution can be summarized as follows:

- We are the first to propose, to the best of our knowledge, a Continuous Approximation model and an optimization procedure to design a transit system combining both FR and DR transport. We call such a system *Adaptive Transit*. It consists of consists of a Mass Rapid Transit (MRT) system, which is always FR, and a feeder service provided by bus. Depending on the sub-region of the conurbation and the period of the day, the system changes the feeder operation, between FR and DR, in order to adapt to the spatial and temporal variation of the demand density. The optimization procedure decides the deployment parameters of both MRT and feeder (frequencies, stop spacing, etc.).

- Via extensive numerical evaluation, we compare Adaptive Transit with current transit design, where were feeder service is always FR, in multiple scenarios. Our results show that Adaptive Transit may improves the user QoS, in particular during off-peak hours and in suburban areas, while keeping the overall cost (which includes the agency cost) under a reasonable level, even slightly reducing it.

To summarize, the novelty of this work is the joint optimization of both FR and DR transport, integrated in a single system, deciding the overall deployment over space and in different times of day. This chapter is structured as follows: we discuss the related work in Section 2, present the design scheme of transit diagram in Section 3, and compare with classic transit schemes. We then present a Continuous Approximation model of such schemes and the optimization procedure to compute their optimal structure (Section 4). We finally contrast the performance of Adaptive Transit with current schemes in numerical results (Section 5) and conclude (Section 6).

7.2 Related work

During the last decades, transit network design has been studied via several optimization problems, based on different objectives (user and/or operator cost minimization, total welfare maximization, protection of the environment, etc.), parameters and decision variables (network structure, demand patterns, fleet characteristics, headway, route and stop spacing, etc.) and solution methodologies (analytical, heuristics or meta-heuristics). An extensive review on Transit Route Network Design Problems is provided by Farahani et al. (2013).

Only in the last decade, the concept of Mobility as a Service (MaaS) has emerged, which advocates the integration of different modes of transportation into a single multi-modal offer (Smith et al., 2018; Le Pira et al., 2021). We believe that MaaS should not consist in just adding DR services on top of the existing FR ones. On the contrary, to get the most benefits from a multi-modal transit combining the two, it is required to holistically redesign the entire transit and “co-design” FR and DR services. The classic transit design methods are not suitable to this new aim, and new approaches are needed.

In this section, we first briefly motivate our choice of a feeder-trunk structure for the transit, resorting to the literature showing its advantages. We then introduce the work on Continuous Approximation, which is the modelling approach we adopt in this chapter, with a particular focus on studies combining FR and DR operation. We also discuss the work combining FR and DR transport with methods different than CA. We finally state the novelty of our work with respect to the state-of-the-art.

7.2.1 Feeder-trunk transit structure

The so-called “weak demand areas" (i.e., areas with by low residential density and high motorization rate) are the most critical for conventional public transportation, which is unable to ensure at the same time coverage, ridership and cost-efficiency. In these cases, an effective design of FRF or DRF bus lines connecting weak demand areas with mass rapid transit nodes could therefore help to shift passenger’s mode of transport from individual to collective mobility, thus enhancing the accessibility to urban facilities and services.

The advantages of mass transit corridors in the metropolitan transport supply have been shown by Mohaymany and Gholami (2010) and Gschwender et al. (2016). In particular, Gschwender et al. (2016) compared the feeder-trunk scheme against different direct lines structures (where no transfers are required) showing that the first structure performs better when the demand is quite low and dispersed and the distances to travel are high, but also underlining that their results are strongly related to the penalty value assigned to the transfers. Mohaymany and Gholami (2010) demonstrated that feeder lines increase the use of high-capacity mass transit because the travel demand for a more extended area can be satisfied.

7.2.2 Continuous Approximation models in transit-related studies

Addressing transit network design problems at a strategic level with discrete models is often unfeasible when dealing with large-scale instances, which are also not robust to stochasticity and uncertainty of input data (Daganzo, 1987).

To overcome such limitations, Continuous Approximation (CA) models have been proposed. We comply with the literature (Ansari et al., 2018) and define a CA model as the one where demand and supply variables, either as input or as decision variables, are continuous density functions over space. Such models are simple but powerful tools for the strategic stage of a transit plan. The key idea, as reported by Ansari et al. (2018), is to construct an objective function, often including agency- and user-centric costs, based on the integration of localized functions of x,y coordinates, which can be analytically optimized without huge computational efforts. Results obtained via CA provide general insights about the performance of whole transit systems, only dependent on the choice of model's parameters. It must be noted that CA models are very approximated and lack realism. For instance, they cannot include details about transportation topologies, traveller behaviour and vehicle routing. However, CA methodology can provide useful insights in understanding at high level the impact of different design choices on the performance of a transportation network.

Analytical models for demand-responsive transportation under many-to-many demand pattern were proposed in Daganzo (1978) for door-to-door services. CA models considering checkpoints, around which the demand can be clustered, were presented in Daganzo (1984) and Quadrifoglio et al. (2006).

Much further work focused on flexible transit to serve the First Mile/Last Mile (FMLM), in particular to compare the performance of FR and DR feeder therein. On this account, Quadrifoglio and Li (2009) estimated the demand density threshold, for a feeder transit service, below which DR buses are more efficient than FR buses. Edwards and Watkins (2013) expanded the comparison to all types of street networks, transit schedules and passenger demand levels.Recently, Badia and Jenelius (2021) found via CA that electrification and automation will impact the cost structure, so that the situations in which DR will be preferable to FR will extend (although FR will still be irreplaceable in very high demand-density areas).

Note that, while the work mentioned in this subsection only focused on single FMLM areas, we instead aim to devise a design for an entire urban area, consisting in many FMLM sub-regions.

Regarding CA models for metropolitan-scale transit, Daganzo (2010) proposed an analytical model of a FR transit network with a "hybrid" structure, in the sense that it combines the advantage of both the grid (double transit routes coverage in the central area) and the hub-and-spoke (radial routes branching to the periphery) structure. The transit is described by only three decision variables: stop spacing, vehicle headway and ratio between the side of the downtown district and the side of the city boundary. The author found that the more expensive the system's infrastructure, the more it should tilt toward the hub-and-spoke concept. Agency costs are always small compared with user costs; and both decline with the demand density. In all cases, increasing the spatial concentration of stops beyond a critical level tends to increase both the user and agency costs. This result demonstrates how excessive spatial coverage is counterproductive.

This model is reformulated in Badia et al. (2014) and applied to a radial route layout. Among the different outcomes, the authors showed that the radial layout is suitable for a centripetal demand pattern, in which the central area is the major attractor and generator of trips (like we assume in our work). The two aforementioned articles show that a high-performance bus system (i.e., buses running on transit priority corridors) outperforms a rail rapid transit system for a wide range of demand density and coverage areas. However, since the former require quite large streets, it appears unrealistic to imagine that such systems can entirely replace underground transit in the big cities' dense urban fabric.

In Chen et al. (2015), two different city-wide transit structures are compared, showing that the ring-radial layout is more favourable to transit (in terms of costs) than the grid design. However, the demand density is assumed to be spatially uniform over the entire urban area, which is not realistic. Note that none of the aforementioned papers in this subsection consider the simultaneous operation of more than one transit mode, or even the presence of DR transport services, in a multi-modal transit scheme.

Nourbakhsh and Ouyang (2012) proposed a transit network with no fixed route: individual buses cover a tube-shaped predetermined area where to pick up or drop off passengers while sweeping longitudinally back and forth through the tube. These "bus-tubes" form at a microscopic level a structure, although buses operate in a demand-responsive fashion. The optimal structure parameters were obtained via a simple constrained nonlinear optimization problem. The authors showed that under low-to-moderate passenger demand the system incurs lower cost than other conventional counterparts such as the fixed-route transit system and the chartered taxi system. The system is however not suited for high demand. We instead observe that in a big conurbation, the demand can be high or low depending on the geographic sub-region and time-of-day considered and therefore it is not possible to just rely on DR transit. For this reason, we instead keep FR operation at the core of transit and integrate DR to it.

The work discussed so far does not combine feeder in FMLM and trunk MRT, which is instead crucial for our *Adaptive Transit*. We discuss in the next section CA approaches for multimodal transit, which consider such a combination.

7.2.3 Continuous Approximation models for multi-modal transit

CA models have also been applied to multi-modal transit, with FMFL feeder and a trunk, or backbone, which corresponds in our design model to MRT. In these works, the feeder is either FR or DR. The novelty of our work is that we instead let our optimization decide between the two for each distance x from the centre and for each time of day t, based on demand density.

Aldaihani et al. (2004) divided the study areas in a grid, with a FR service along the lines of the grid and a DR service within each sub-region, consisting of a taxi service, serving one passenger at a time. We also divide the entire area in sub-regions, but we let our optimization choose between FRF and DRF therein. Moreover, our DRF is able to serve multiple passengers at a time.

Sivakumaran et al. (2012) propose a CA model to show the benefits of coordinating feeder services and MRT (trunk), but they only consider FRF.

Chen and Nie (2017) studied a grid and a radial network with fixed route transit lines paired with DR lines connecting passengers to the stops of the fixed route. Optimal design is formulated as a mixed integer program. The results show that the paired lines design outperforms the other two systems, one always using FR and the other always DR, under a wide range of scenario configurations. The main limit of Chen and Nie (2007) is that their DR service is forced to run in the entire urban area, everywhere with the same characteristics. e.g., with the same headway. We instead let our optimization problem choose where and when to deploy FRF or DRF (this is captured by our decision variable *F(x)*) and we find that, depending on the time of day, it is optimal to deploy FRF close to the city centre and DRF far from it.

Later, Luo and Nie (2019) compared six distinctive transit systems using the CA approach, most of them already studied in the previously cited works. One finding of Luo and Nie (2019), which we also confirm is that the demand-responsive feeder services tilt the balance of trade-off considerably in the user's favour, at the transit agency's expense. A recent work by Wu et al. (2020) compares a system where feeder is provided by fixed bus with another where feeder is provided by bike sharing.

The systems studied in all the aforementioned work are not adaptive, i.e., they do not take the optimal choice between FR and DR modes, in each region and period of day. Instead, they either use always one or the other. Our work instead shows that it is beneficial to make DR and FR co-exist in the same transit layout.

7.2.4 Other approaches to integrated transit and demand-responsive transportation

Salazar et al. (2020) provided a network flow model of Integrated Autonomous Mobility on Demand (I-AMoD), where a ride sharing service provided with autonomous vehicles is integrated with transit and they are modelled together in a multi-layer graph. They solve a static assignment problem to calculate how the origin-destination matrix demand distributes onto the arcs of such a graph. Their goal is to find optimal pricing, while we aim to optimize the structure of the overall transportation system.

Narayan et al. (2020), Leffler et al. (2021) and Bürstlein et al. (2021) showed in simulation that DR transit can improve accessing line/schedule based public transport system, in terms of level of service and environmental impact. Franco et al. (2020) generated demand for future DR services integrated with fixed transit, based in phone data. Note that none of the mentioned work in this paragraph seeks to find optimal transportation layout for an entire metropolitan area, as we do.

An and Lo (2015) solved the transit network design problem under demand uncertainty trough robust optimization for rapid transit and dial-a-ride services. However, the author did not include the passenger waiting times are not included in the model and determined the travel costs proportional to distance and not to travel time.

Pinto et al. (2020) proposed a model based on dynamic programming and simulation-based assignment. The decision variables they aim to calculate are two: the headways of bus lines and the fleet size S of a taxi-hailing service, where each vehicle can have at most two riders on-board. The main difference of our work lies in the different insights we aim to study: we are interested in studying how the overall transit system can adapt spatially and temporally to the spatio-temporal demand variation, in particular choosing between FRF or DRF bus services. Instead, Pinto et al. (2020) let a mathematical program calculate a single value of S, without letting the agency decide in which regions and at which time of the day such S vehicles should be deployed. This suffers, we believe, from the potential risk to just attract such vehicles in the city centres, where most of the demand is and where fixed transit is already efficient. This would play against our goal of employing DRF in low demand areas and during off-peak hours.

We instead keep the choice of where and when to deploy DRF in the hand of the agency. Moreover, our DRF service is able to serve many users (more than two) at the same time and can act as a ride sharing or minibus service.

The optimization problem of Steiner and Irnich (2020) aimed to “shorten” some bus lines, i.e., eliminate some stops at the beginning and the end of FR lines and replace them with a DR service. We believe that completely removing FR service from periphery of an urban area may worsen, rather than improving, mobility, overall. First, it would disadvantage even further suburban travellers. Second, it would require aggregating the demand close to the centre via a DR service, which may cause congestion and would suffer from limited capacity. We instead let FR service to be deployed up to the extreme periphery of the urban area and adopt DR services as feeder, instead of as a replacement, of FR lines (in our case, the MRT).

Finally, Steiner and Irnich (2020) did not show how the transit service should change configuration over the day to adapt to the time-varying demand pattern.

7.2.5 Positioning of our work

To the best of our knowledge, none of the previous work has tackled the problem of designing *Adaptive Transit*, i.e., to decide how to optimally vary spatially and temporally the layout of transit over an entire urban area, also deciding in which regions and in which time of day (peak / off-peak) FR or DR transit must be deployed. A “variable” layout of this kind allows transit to better adapt to the demand, which is varying over time and space.

To this aim, we do not need to re-invent a model from scratch, but we build upon the previous work discussed in this section, in particular the ones using continuous approximation. We readapt them to our *Adaptive Transit* case as described in the next sections.

7.3 Transit Design Schemes

We focus on the transit system of wide urban and metropolitan areas. They generally show a transition from a central zone to sprawled suburban areas. The former is characterized by dense urban fabric, high population density and presence of numerous “trip attractors” (job places, commercial activities, amenities, etc.). The latter, instead, are often shaped by low residential density and sparse transport demand, both temporally and spatially. The components of transit systems are:

- A Mass Rapid Transit (MRT).
- Possibly, a feeder service, provided by bus, to serve the First Mile and Last Mile (FMLM).

7.3.1 Central and suburban areas

The MRT network is modelled as a ring-radial structure, as in Badia et al. (2014) and Chen et al. (2015), which can be adopted to model several big cities around the world (e.g., Paris, Singapore, Moscow).

As common in Continuous Approximation modelling, we assume the entire area is composed of two parts:

- A central area only served by MRT with double coverage provided by radial and circular rail lines.
- A suburban area covered only by MRT radial lines (and no circular line).

Feeder services can be deployed in the suburban area.

7.3.2 Transit schemes

We discuss three alternative transit schemes, which essentially differ in the way passengers can access MRT:

1. *MRT-only* scheme, in which the access to MRT stations only takes place by walking.
2. *MRT-FRF* scheme, which includes feeder bus lines with fixed routes to increase the accessibility of MRT stations in the suburban area, which can be reached either by walking or using the bus, depending on the distance to/from the station.
3. *Adaptive Transit* scheme, in which the FMLM in the suburban area is still covered by a feeder service, but the feeder can switch between two modes, Fixed-Route Feeder (FRF) and demand-responsive feeder (DRF), choosing optimally between one or the other based on the transport demand density.

Figure 7.1 breaks down the components of the travel time of the passengers using transit. A passenger needs first to access the transit. In case the passenger uses the MRT, she may access it by either walking or using a feeder service (which can be a FRF or a DRF, depending on the scenario). If she uses a feeder service, she has to wait for it and then spend some time in the feeder vehicle before arriving at the MRT stop. Symmetrically, to reach the final destination from the egress MRT station, she needs to walk or use another feeder service. Note that there is no walking time when the DRF is employed, as it is assumed to be a door-to-door service. Waiting time at the MRT station and In-Vehicle Travel Time into the MRT are also depicted.

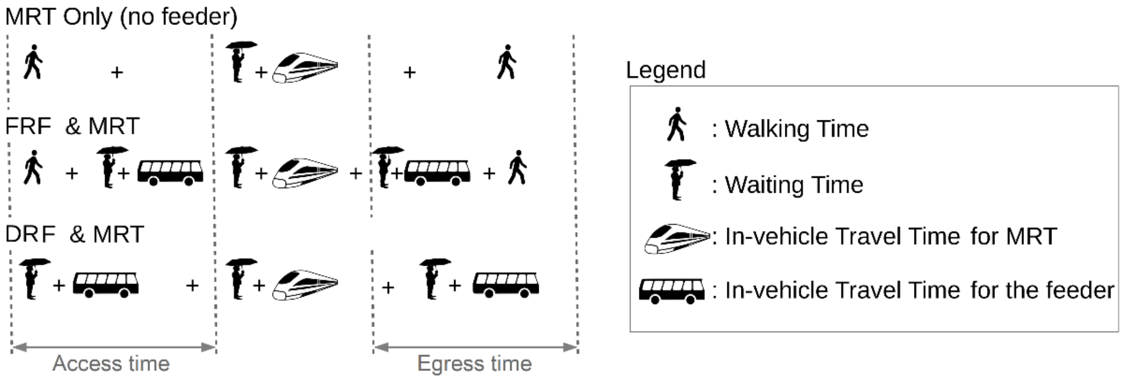


Figure 7.1. Components of the access, egress and waiting time for the MRT-only scheme and when FRF or DRF services are provided.

We clarify that in *Adaptive Transit* the choice of whether to deploy FRF or DRF is not made on-the-fly. On the contrary, we assume that, based on historical observations of the demand density, the authority would plan, for each area, the time periods when FRF or DRF will be operated. Such a plan would be revised only on a seasonal basis.

We use *MRT-only* as a baseline scheme. Its poor cost-efficiency shown later in the numerical results demonstrates that a feeder bus service in the suburban area is, as expected, necessary. The *MRT-FRF* scheme is what it is basically currently deployed in most of cities around the world. The *Adaptive Transit* scheme is the design we propose for future generation transit.

7.4 Continuous Approximation Model

With the Continuous Approximation (CA) approach, an urban conurbation and a transit network are represented with a parametric model, consisting of:

- A set of decision variables, describing the layout of the transit network, i.e., line and stop spacing, headway values, etc.
- A set of input parameters, which are exogenous and describe the scenario, e.g., size of the urban conurbation, demand density.
- A set of constraints, which ensure basic properties like conservation of flows, transit vehicle capacity constraints.

- A cost function, which we want to minimize. It consists of a weighted sum of user-centric and agency-centric terms. It summarizes the performance of transit.

CA models allow to understand the impact of different decision variables on the performance, in an approximated, concise and computationally efficient way. Due to the high level of simplification, results obtained via CA models should not be expected to be exact and realistic and must rather be interpreted as high-level trends, which can guide transit planning considerations. Therefore, we resort to CA modeling to understand the benefits of choosing between a FRF and a DRF service, in order to better adapt to the demand, over time and geographical areas, a concept that we call *Adaptive Transit*. We are not interested in exact results representative of a single specific city. For this reason, CA methodology perfectly fits our needs. Our formulation is mainly based on Chen et al. (2015); Daganzo (2010), both of which, however, do not integrate MRT and feeder. For this reason, we need to introduce some modification, which we will pinpoint in the following pages.

One novelty with respect to the previous work on CA is that we introduce a notion of time-evolution. We need to do so, as we want to evaluate the capacity of our transit design to adapt to change of the intensity of the demand over the day. We therefore day a set $\mathcal{T}$ of time instants $t \in \mathcal{T}$, when the demand and supply of the transit system have given characteristics, and partition the entire day into non-overlapping time intervals starting in instants $t \in \mathcal{T}$, each of duration Δt during which such characteristics remain constant. With slight abuse of notation, we will denote with t a time instant and also the time interval starting at t. Observe that, to keep the mathematical development treatable, we make the simplifying assunmption that the aforementioned time intervals are independent of each other, in the sense that there is no propagation of passenger flows from one interval to the next. This is also equivalent to assuming that the flow starting in the current time interval and terminating in the next is compensated by the flow starting in the previous and terminating in the current.

The notation used in this section is summarized in Appendix B (Table B.1).

7.4.1 Main decision variables

We study a circular metropolitan area of radius R (exogenous input parameter). The transit layout, depicted in Figure 7.2, is organized as described in the previous section. It is described by 10 local decision variables and 4 global decision variables, determining the transit structure.

The local decision variables take a value for each value x of distance (in km) from the centre. They are:

- The angle $\theta_r(x)$, in radiants, between radial MRT lines; based on that, we can also compute the corresponding linear spacing $S_r(x) = \theta_r(x) \cdot x$.
- The spacing $S_c(x)$ between circular MRT lines, defined for $x < r$, where r is the radius of the central area.
- The spacing $s(x)$ between the MRT stations (hereinafter called just "stations") along a radial MRT line.
- The angle $\phi(x)$ between stations on a circular MRT line, defined for $x < r$.
- The headway $H(x)$ on circular and radial MRT lines.
- The headway $h(x)$ of the feeder service (only defined in the suburban area).
- The variable $F(x) \in$ {FRF,DRF,0}, defined for $x > r$, indicating whether at location x a FRF, a DRF or no feeder service is deployed. We introduce the following indicator function $\mathbb{I}_j(x), j \in \{FRF, DRF, 0\}$, which is 1 for the x where $F(x) = j$, and 0 otherwise.
- The variable $d_{FRF}(x)$, defined for $x > r$ and $F(x) =$ FRF, which is the spacing between FRF stops. .
- The variable $d_{0,DRF}(x)$, defined for $x > r$ and $F(x) =$ DRF, which is the value determining the area close to the MRT station where the feeder service does not pick-up or drop-off passengers (see Figure 7.2).
- The number of strips $N_s(x)$, an integer variable defined for $x > r$, in which a FMLM subregion is divided. Each strip is served by a feeder service. In the *MRT-only* scheme, $N_s(x) = 1$.

The variation of such local variables along x can be seen as an approximation of what one could observe in reality. For instance, if radial lines bifurcate with the distance x, we can represent this by reducing the angular spacing $\theta_r(x)$ with x. Accordingly, the headway $H(x)$ would increase with x because less vehicles will travel along a single line. Note that, to limit the number of parameters, we consider the same headway on circular and radial lines $H(x)$. Note that this constraint is only active in the central area ($x < r$), while in the suburbs, where no ring lines are present, $H_r(x)$ is free to take any value.

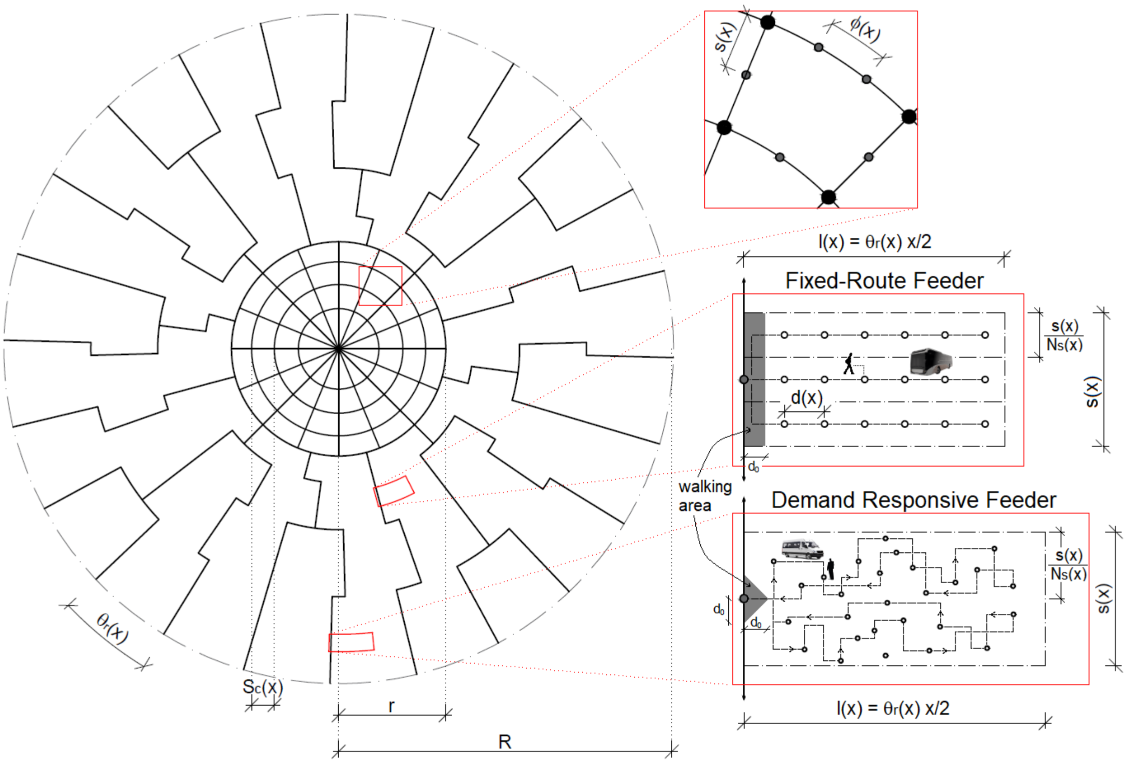


Figure 7.2. Transit network layout (this takes inspiration from Chen et al. (2015) and Quadrifoglio and Li (2009)).

The global decision variables are:

- The radius of the central area r.
- The angle ϕ_B between stations on the outermost circular MRT line, at $x = r$ (city centre's boundary).
- The headway H_B of the MRT line at $x = r$. Observe that we need global variables for ϕ_B and H_B because the outermost circular MRT line serves more trips than all other ring lines. Indeed, the outermost circular MRT line attracts the transfers of the travellers whose origin and destination are in the suburban area (see Figure 7.3b).
- The maximum value Q_0 of the total radial flow of MRT vehicles, which occurs at $x = 0$ (see the next subsection).

Regarding the last decision variable, we define the radial flow $Q(x)$ as the number of MRT vehicles crossing an infinitesimal annulus of radius x in both inward and outward direction.

$$Q(x) = \frac{2\pi}{\theta_r(x)} \cdot \frac{1}{H(x)} \tag{7.1}$$

Therefore, the two local decision variables $\theta_r(x)$, $H(x)$ are interdependent since they determine the radial flow $Q(x)$, which in turn is constrained by the global decision variable $Q_0 \equiv Q(x=0)$, as explained in the next subsection.

7.4.2 Assumptions and constraints

Along each radial MRT line, in the inward direction, we assume a train can depart from any $x \leq R$, but always terminates in the centre ($x = 0$). In the outward direction, a train always departs from the centre and can terminate at any $x \leq R$. This translates to the following constraint:

$$H(x_1) \leq H(x_2), \quad \forall\, x_1, x_2 \mid x_1 < x_2 < R \tag{7.2}$$

By doing this, we allow radial lines to have higher frequency in the city centre (as often observed in real cities). This would occur when some of the trains passing through that line only serve the subset of stops closer to the centre. Also, we prevent the radial flow $Q(x)$ from increasing outward (it would mean that there were vehicles not passing through the city centre, contradicting our assumption) with the following constraint:

$$Q(x_1) > Q(x_2), \quad \forall\, x_1, x_2 \mid x_1 < x_2 < R \tag{7.3}$$

The following vehicle capacity constraint must also be respected:

$$O_j(x) \leq C_{pax,j}, \quad \forall x \tag{7.4}$$

where $O_j(x)$ is the average vehicle occupancy at x and C_j is the vehicle capacity of mode $j \in$ {MRT, FRF, DRF}. $O_j(x)$ is computed in Appendix B.1.

7.4.3 Demand pattern and travel behaviour

We assume that the transit demand density is both temporally and spatially variable and follows the Clark's law (Clark, 1951), i.e., an exponential decline from the centre to the suburbs. Denoting with x the distance (in km) from the city centre, the demand density (measured in trips per hour per km^2) is given by:

$$\rho(x,t) = \rho_0(t) \cdot e^{-\gamma x} \tag{7.5}$$

where $\rho_0(t)$ (pax/km^2 h) is the density of users in the centre and γ (km^{-1}) is the slope (also called density gradient) with which that value decreases as we move away. By

changing $\rho_0(t)$ over the time t of the day, we can capture the temporal variation of the transport demand. In our work we consider that $\rho_0(t)$ varies in a step-wise function, i.e., it remains constant within each time period $t \in \mathcal{T}$.
Note that the demand density decreases towards the outer regions. When we omit t, to simplify notation, it means we are focusing on a single time instant.
With such a model, the city centre, where economic activities are more concentrated, emerges as an attractor and generator of trips from/to the periphery (see Equation B.1). A passenger first accesses the closest transit station (either by walking or via a feeder bus, rides via MRT to the station closest to her destination and finally reaches (either by walking or via a feeder), her destination. Note that, within the MRT, a passenger could change from a radial to a ring line and vice versa. The sequence of such changes obeys classic assumptions in literature (e.g., Badia et al., 2014; Chen et al., 2015) and are calculated in order to minimize the travelled distance via MRT. They are developed in Appendix B.4 and depicted in Figure 7.3.

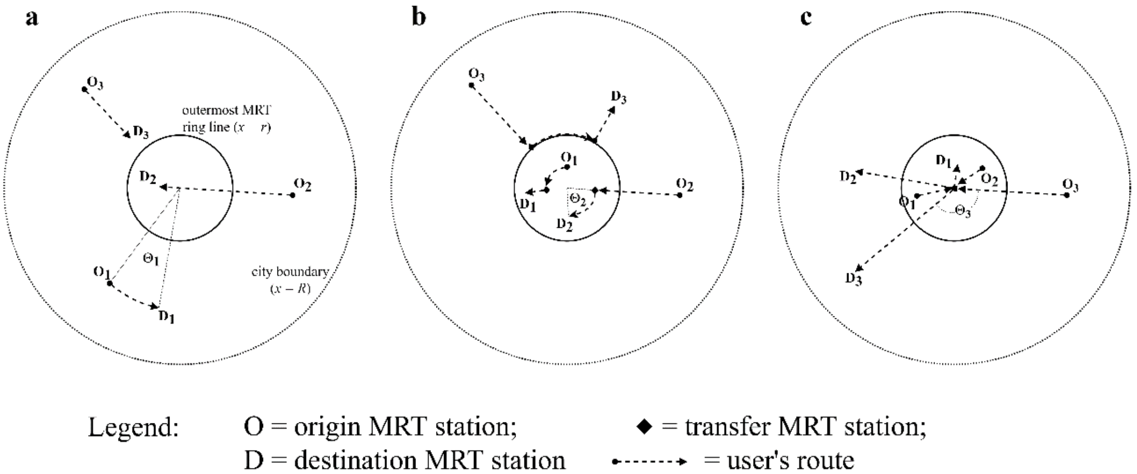


Figure 7.3. User's route choice from the origin MRT station to the destination MRT station (inspired from Chen et al, 2015).

In Figure 7.3a only those trips which do not require transfers between different MRT lines are represented: for such trips, the Origin-Destination (O-D) couple lies within a circular sector of central angle $\Theta_1 \leq \theta_{r,min} = \min_{x\in[0,R]} \theta_r(x)$, which is the minimum angle between radial MRT lines. Moreover, there is a small likelihood (see Appendix B.4) that a trip can be made without using the MRT, when both origin and destination (see O_1 and D_1 in Figure 7.3a) lies near the same MRT station.

In Figure 7.3b we represent trips with angle $\Theta_2 < 2$ [rad] between O and D. In this case, it is easy to see that, in order to minimize the travel distance, a user will travel via both radial and ring lines. Also, when both origin and destination (see O_3

and D_3 in Figure 7.3b) lie in the periphery ($x > r$), this is the only case that requires using the outermost ring line and implies two transfers.

Finally, in Figure 7.3c the cases where the O-D couple has an angle $\Theta_3 > 2$ [rad] are shown: a user will travel by radial line towards the city centre, where she transfers to another radial line to reach her destination.

7.4.4 Feeder services

The suburban area is divided in FMLM sub-areas, each determined by the spacing between the radial lines and the station's spacing along them, as in Figure 7.4. In case of *MRT-only* scheme, passengers can only walk therein. In the other two schemes, instead, each FMLM subarea is associated to a MRT station, and is further divided in a number $N_s(x)$ of strips (as in Guo et al., 2018) served by a feeder service, i.e., a fleet of buses bringing passengers to/from that MRT station. Each FMLM area, forming a ring sector, can be approximated into a rectangle with the following dimensions:

$$\text{length } l(x) = \frac{\theta_r(x)}{2} \cdot x; \quad \text{FMLM subarea width } s(x); \quad \text{strip width } w(x) = s(x)/N_s(x) \tag{7.6}$$

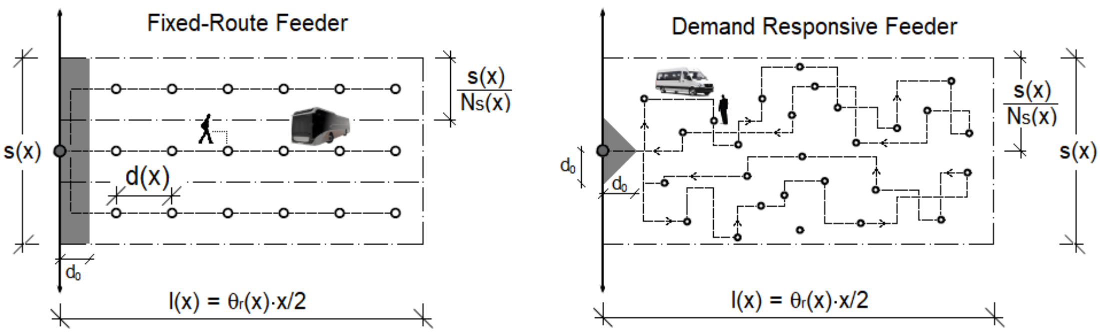


Figure 7.4. FRF and DRF layouts.

Note that, the size of an FMLM subarea depends on the MRT structure (the more the MRT lines and/or the smaller the station spacing, the smaller the FMLM subareas) and determines the total user demand to accommodate.

Fixed-Route Feeder (FRF). The FRF is modelled as a straight route with spacing $d(x)$ between stops and the vehicle moving back and forth between MRT station and the furthest stop, as in Quadrifoglio and Li (2009), so that the length of a complete cycle is given by:

$$CL_{FRF}(x) = 2\left(l(x) + \Delta l(x) - \frac{1}{2}d(x)\right) \tag{7.7}$$

where $\Delta l(x)$ is the average extra vertical distance (see Figure 7.4) which the FRF has to travel due to the different position of the strips with respect to the MRT station they serve, that we approximate to $s(x)/4$ if $N_s(x) > 1$, and 0, otherwise.

We assume that travellers walk on a Cartesian grid, and thus all walking distances are Manhattan distances. Under this assumption, the average walking distance to reach the nearest bus stop is

$$d_{FRF}^{walk}(x) = s(x)/4 + d(x)/4 \tag{7.8}$$

If the origin / destination of a user is in a location close enough to the MRT station (less than threshold $d_{0,FRF} = d(x)/2$), she will prefer to directly walk to / from the MRT station. We call such locations "walking area", represented in grey in Figure 7.4. The fraction of traveller locations in the walking areas is $p_{walk,FRF}(x) = d_{0,FRF}/l(x)$. The time needed to complete a cycle can be calculated as follows:

$$C_{FRF}(x) = \frac{CL_{FRF}(x)}{v_{FRF}} + \tau_s\left(\frac{2l(x)}{d(x)} - 1\right) + \tau_p \cdot n(x) + \tau_T \tag{7.9}$$

where v_{FRF} is the cruising speed of the bus, τ_T is the terminal dwell time, τ_s is the time lost per stop, τ_p is the time lost per passenger due to boarding/alighting operations and it is multiplied by the average number of passengers per vehicle:

$$n(x) = 2\,\rho(x) \cdot w(x) \cdot l(x) \cdot h(x) \cdot \left(1 - p_{walk,FRF}(x)\right) \tag{7.10}$$

being $\rho(x)$ the demand density (in trips/km^{2}h per travel direction).

Demand-Responsive Feeder (DRF). The DRF is assumed to provide a door-to-door service, so passengers do not have to walk to any physical bus stop. We assume that each new request is processed in real-time via an insertion algorithm that aims at minimizing the in-vehicle time experienced by the passengers. We assume a "no-rejection" policy, i.e., the DRF is able to serve all the users' trip requests. The more passengers a single DRF vehicle has to pick-up/drop-off, the longer its route, due to longer detours. Travellers close enough to the MRT station, i.e., at Manhattan distance less than $d_{0,DRF}(x)$ from the station (grey triangular area in Figure 7.4) directly walk to it. Their fraction is given by the ratio between the area of the walking area and the area of the FMLM subarea:

$$p_{walk,DRF}(x) = d_{0,DRF}^2(x)/(l(x) \cdot s(x)) \tag{7.11}$$

The computation of the cycle length (CL_{DRF}) and the cycle time (C_{DRF}) is based on the work of Quadrifoglio et al. (2006) and Quadrifoglio and Li (2009). The cycle length is estimated based on the expected number of passengers per vehicle $n(x)$, which is computed as in Equation 7.10, substituting $p_{walk,FRF}(x)$ with $p_{walk,DRF}(x)$. The cycle length is given by the sum of a horizontal component (the vehicle movement from left to the right and vice-versa) and a vertical component (the deviations along the vertical direction to serve the passengers):

$$CL_{DRF}(x) = 2\, l(x) \frac{n(x)}{n(x)+1} + \frac{w(x)}{3} \cdot n(x) + \frac{w(x)}{2} \tag{7.12}$$

The time needed to complete a cycle can be calculated as follows:

$$C_{DRF}(x) = \frac{CL_{DRF}(x)}{v_{DRF}} + \left(\tau_s + \tau_p\right) \cdot n(x) + \tau_T \tag{7.13}$$

7.4.5 Cost components

The main objective of the present work is to find the optimal transit structure able to integrate fixed and demand-responsive modes. With this aim, we formulate a generalized cost function to be minimized as Badia et al. (2014) and Chen et al. (2015), which combines the disutility for users due to the travel time in its different components (Figure 7.1) and the costs incurred by the transport agency to provide the service (and the related externalities).

As regards the transit users, as usually done in CA work, we do not consider transit fares. This is a reasonable assumption when most of users use monthly passes and since we do not consider a mode choice model in our work.

The quantities involved in the generalized cost are summarized in Table 7.1 and

Table 7.2. They are all converted to money metrics via specific coefficients. They are all density functions over the distance from the centre. We distinguish user-related cost, representing the time spent and the discomfort suffered by the passenger during her trip, both in the FMLM segments and in the MRT segment. We also have agency-related cost, due to capital and operational costs for operating feeder services in the FMLM and the MRT. The detailed computation of all cost components in Appendices B.2–B.5.

Agency's and user's metrics are converted into cost density functions by means of a set of cost coefficients, in order to compute the social cost (per unit of time) as a linear combination of those metrics. We denote with $\mu_{L,MRT}$ or $\mu_{L,FMLM}$ (€/km-h), $\mu_{V,MRT}$ or $\mu_{V,FMLM}$ (€/veh-km) and $\mu_{M,MRT}$ or $\mu_{M,FMLM}$ (€/veh-h) the cost coefficients related to the agency metrics, for MRT and FMLM, respectively. Regarding the MRT, we also consider costs specifically related to the stations through a coefficient μ_{ST} (€/station). The cost components are detailed in Table 7.1 and

Table 7.2, and commented in next sections.

Agency-related costs. The cost components related to the transit agency depend on:

- Infrastructure length L (km), i.e., the construction and maintenance costs. For the MRT we also include an additional unit cost component related to stations.
- Total distance V (veh km/h) travelled by the vehicles per unit of time, i.e., the operational costs.
- Size M (veh) of the vehicle fleet, i.e., the capital and crew costs.

User-related costs. The cost components related to the users of the transit system depend on:

- Walking time A to reach the bus stop or the MRT station or the destination.
- Waiting time W at the bus stop or the MRT station.
- In-vehicle (MRT, FRF or DRF) time T including boarding, riding, dwell and alighting time.
- Transfers: since any possible transfer between different transit lines is an additional disutility, we treat it as a penalty of the extra walking time ΔA.

Similarly, the cost coefficients μ_A, μ_W, μ_T are all equal to the Value of Time (*VoT*) (€/h) associated to walking, waiting and travelling on-board, which have the same value independent on whether they refer to travelling in a feeder or in the MRT.

7.4.6 Optimization problem

To express the cost objective in a concise way, we use i to indicate the type of cost component, $i \in \{L, ST, V, M, A, W, T\}$, emphasizing that FRF and DRF are alternative feeder services. As in Chen et al. (2015), we distinguish:

- Local densities $Y_i(x)$, referred to the MRT, and $y_i(x)$, referred to the FMLM (either FRF or DRF), which vary with the distance from the centre x.
- Global components F_i, which are instead only related to the outermost MRT ring line, i.e., at $x = r$.

We now formulate the optimization problem to minimize it. We separate the decision variables in two sets:

- Global decision variables: $G = \{r;\ Q_0;\ \phi_B;\ H_B\}$,
- Local decision variables (functions of x): $D(x) = \{\theta_r(x);\ S_c(x);\ s(x);\ \phi(x);\ H(x);\ h(x);\ d_j(x);\ N_s(x);\ F(x)\}$.

As in Chen et al. (2015), we constraint the station spacing on the outermost MRT ring line to be the same as the value of the corresponding local variable at $x = r$, so we set:

$$\phi_B = \phi(r) \tag{7.14}$$

For any values for the sets G and $D(x)$, $x \in [0, R]$ at any time interval $t \in \mathcal{T}$, we denote the hourly cost of component Z_i as:

$$Z_i(D,G,t) = \mu_{i,MRT} \cdot \left(F_i(r,\phi_B,H_B,t) + \int_0^R Y_i(D,r,t,x)\,dx\right) + \mu_{i,FMLM} \cdot \int_0^R y_i(D,r,t,x)\,dx \tag{7.15}$$

Observe that all the cost components $y_i(x)$ and $Y_i(x)$ are derived in Appendix B. In the formula above, we emphasize that they depend not only on x, but also on the demand density $\rho_0(x)$ and the decision variables $D(x)$ and G.

Table 7.1. Overview of agency-related and user-related local cost components

		FMLM		**MRT**
User costs	$y_A(x)$	Cost due to walking to/from the feeder bus stop	$Y_A(x)$	Cost due to walking to/from the MRT station
	$y_W(x)$	Cost due to the time to wait for the feeder service	$Y_W(x)$	Cost due to the time to wait for the MRT
	$y_T(x)$	Cost due to the time spent into feeder vehicles	$Y_T(x)$	Cost due to the time spent into MRT
Agency capital costs	$y_L(x)$	Cost for the infrastructure in FMLM, i.e., construction and maintenance	$Y_L(x)$	Cost for the infrastructure of the MRT, i.e., construction and maintenance
	$y_M(x)$	Cost due to the feeder fleet, i.e., vehicles and crew cost	$Y_M(x)$	Cost due to the MRT fleet, i.e., vehicles and crew cost.
			$Y_{ST}(x)$	Cost due to the MRT station density.
Agency operation costs	$y_V(x)$	Cost due to vehicle-distance travelled by feeder vehicles.	$Y_V(x)$	Cost due to vehicle-distance travelled by MRT

Table 7.2. Overview of agency-related and user-related global costs

		MRT
User costs	F_A	Transfer cost, due to users' changing MRT lines at the outermost ring line (Figure 7.3b)
	F_W	Cost due to the time to wait for the MRT at the outermost ring line
	F_T	Cost due to the time spent into MRT along the outermost ring line
Agency capital costs	F_M	Cost due to MRT fleet and crew cost on the outermost ring line
Agency operation costs	F_V	Cost due to the vehicle-distance travelled on the outermost ring line

We distinguish user-related costs and agency-related costs. The latter are further divided into capital costs, which are not dependent on the time of the day, and operation costs, which instead vary with t. The total cost is the sum of the three cost components:

$$Z(D,G,t) = Z_{user}(D(x),G,t) + Z_{cap}(D,G) + Z_{op}(D,G,t) = \sum_{i\in\{A,W,T\}} Z_i(D,G,t) + \sum_{i\in\{L,M\}} Z_i(D,G) + \sum_{i\in\{V\}} Z_i(D,G,t) \quad (7.16)$$

For each of the three schemes of Section 7.3.2 (*MRT-Only*, *MRT-FRF*, *Adaptive Transit*), we first dimension the system in the peak hour. To do so, let t^{peak} be the time interval in which the demand density is the highest, i.e., $t^{peak} = \arg\max_t \rho_0(t)$. We want to find, for every x, the optimal values of local decision function $D^{peak}(x|Q_0)$, which depends on Q_0, and the optimal values of the global variables G^{peak}. We solve the following optimization problem:

$$\{G^{peak}, D^{peak}(x|Q_0)\} = \arg\min_{G,D(x)}\left\{Z\left(D,G,t^{peak}\right) \text{ subject to Equations 7.2-7.4}\right\} \quad (7.17)$$

This optimization allows us to dimension the fleet size and the transit infrastructure needed. Note that, once fixed to satisfy the mobility needs in peak hours, the fleet size and the infrastructure does not change over the day, and the corresponding cost must be supported by the agency, even if there are periods of the day in which they are not fully used. We thus fix the capital cost as:

$$Z_{cap} = \sum_{i\in\{L,M\}} Z_i = \sum_{i\in\{L,M\}} Z_i(D^{peak}(x), G^{peak}) \quad (7.18)$$

We then optimize the system in each time slot independently only minimizing the operational cost:

$$G^*(t), D^*(x,t|Q_0,r) = \arg\min_{G,D(x)}\left\{Z_{op}(D,G,t) + Z_{user}(D,G,t) \quad \text{subject to Equations 7.2-7.4;7.20(a-g)}\right\} \quad (7.19)$$

where we introduced the following constraints, valid when $t \neq t^{peak}$:

a. $r = r^{peak}$

b. $\theta_r(x) = \theta_r^{peak}(x) \quad \forall x$

c. $S_c(x) = S_c^{peak}(x) \quad \forall x < r$

d. $s(x) = s^{peak}(x) \quad \forall x$ (*MRT-only* or *MRT-FRF* scheme)

e. $s(x) = s^{peak}(x) \quad \forall x < r, \quad s(x) \geq s^{peak}(x) \quad \forall x > r$ (*Adaptive Transit* scheme)

f. $d_j(x) = d_j^{peak}(x) \quad \forall x > r, j \in \{FRF, DRF\}$ (*MRT-FRF* scheme)

g. $d_j(x) \geq d_j^{peak}(x) \quad \forall x > r, j \in \{FRF, DRF\}$ (*Adaptive Transit* scheme) (7.20)

Note that the equalities among the constraints of the previous minimization represent the fact that the infrastructure remains, all over the day, the same as the one decided via Equation 7.17. The total cost, over the entire day, is:

$$Z^{24h} = \sum_{t \in \mathcal{T}} \Delta t \cdot \left(Z_{cap} + Z_{op}(D^*(x,t), G^*(t), t) + Z_{user}(D^*(x,t), G^*(t), t) \right) \quad (7.21)$$

We point out that the optimization procedure is done separately for *MRT-only*, *MRT-FRF* and *Adaptive Transit* schemes.

7.4.7 Optimization procedure

We now describe the optimization procedure we execute for each time instant. The optimization of Equation 7.17 is non-convex, so we resort to bi-level optimization to solve it. The lower level subproblem consists, given any value of global variables *r*, Q_0, to solve the following local optimization problem, for all $x \in [0, R]$.

$$D^*(x, t | Q_0, r) = \arg\min_{D(x)} \{ \textstyle\sum_i \mu_i \cdot (Y_i(D, r, t, x) + y_i(D, r, t, x)) \text{ subject to Equations 7.2-7.4;7.20(a-g)} \} \quad (7.22)$$

We use an interior-point algorithm to solve such problem. The higher level subproblem (global optimization) is to determine the set of global variables $G(t)$ that minimizes the total cost Z, which we remember is the sum of global and local costs (Equation 7.15). The following iterative procedure is implemented:

1. Initialize r (sufficiently small) and repeat the global optimization procedure (described in step 3) by increasing r until the total cost found is higher than the average value from the previous 3 iterations. When this occurs, set r^* equal to the value of the third to last iteration.

2. For any value of r, initialize Q_0 and, similarly to the previous point, increase Q_0 until the total cost found is higher than the average value from the previous 3 iterations. When this occurs, set Q_0 equal to Q_0^*, the value of the third to last iteration.
3. Run the global optimization procedure, composed by the following steps:
   a. Run the local optimization (Equation 7.22) in order to find $D^*(x, t|Q_0)$ for all $x \in [0, R]$ (we discretize this interval with a step Δx).
   b. Find $H_B = \arg\min_{\mathrm{H_B}}\{\sum_i \mu_i \cdot F_i(r, \phi_B, H_B, t)$ subject to Equation 7.4}.This problem is simple to solve since it has only one decision variable[9].
   c. Compute the total cost as in Equation 7.16.

Observe that, we first run the optimization procedure explained above in the peak hours. After that, we fix infrastructure length and fleet size. We then repeat the same optimization procedure for all the other time instants, with the additional constraint of keeping infrastructure length and fleet size fixed to the peak hour values.

7.5 Numerical results

We now compare the performance of transit schemes in scenarios representing a large urban conurbation, during peak and off-peak hours.

7.5.1 Scenario parameters

The complete list of the parameter values describing our scenario is reported in Table 7.3. We consider a circular area of radius $R = 25$ km, which corresponds to the size of large metropolitan areas, e.g., the *Greater Paris* region. The value of the demand density $\rho_0(t)$ is estimated from the travel demand data of the regional Household Travel Survey "EGT 2010".[10] We approximately fit Equation 7.5 on residential density data on Paris region.[11] Therefore, an average demand density $\bar{\rho}_0 = 640$ trips/km^2 h (this is the sum of trips departing from and arriving at each km^2) and a slope of $\gamma = 0.12$ km^{-1} are assumed. The shape of the transit demand as a function of the distance from the city centre is represented in Figure 7.5.

[9] Note that $\phi_B = \phi(r)$ (Equation 7.14)
[10] The travel demand of *Greater Paris* consists of 8.3 million PT trips, on average per working day. Since a complete trip needs on average one transfer, we obtain a daily demand of 4.15 million trips made by PT per day. Considering 18 hours of PT operation, we obtain an average demand of 230·10^3 trips/h.
[11] https://www.insee.fr/fr/statistiques

For the sake of simplicity, we divided the day into time intervals $\Delta t = 1$ hour, and assume that $\rho_0(t)$ only takes three values during the day:

- A peak value $\rho_0(t^{peak}) = 2.5 \cdot \bar{\rho}_0 = 1600$ trips/km²h
- An off-peak value $\rho_0(t^{op}) = 0.75 \cdot \bar{\rho}_0 = 480$ trips/km²h
- A low-peak value $\rho_0(t^{lp}) = 0.4 \cdot \bar{\rho}_0 = 256$ trips/km²h

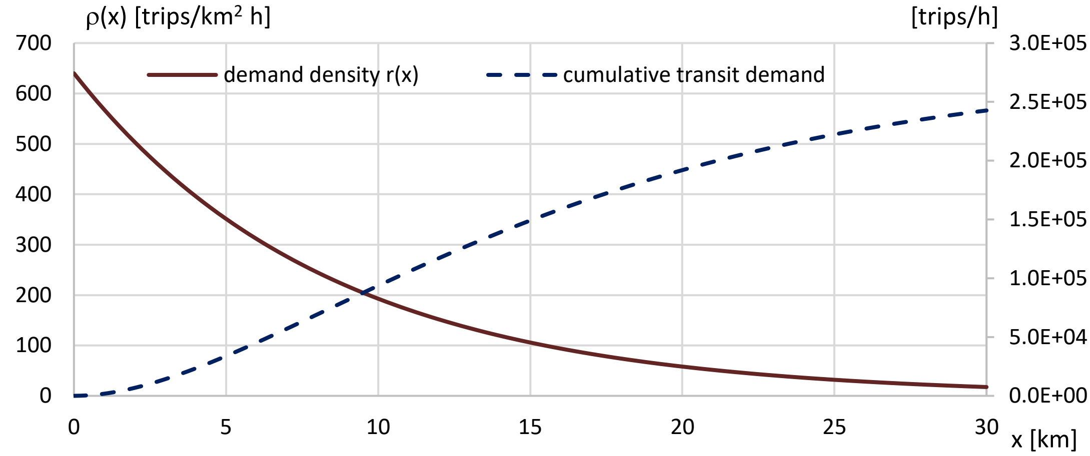


Figure 7.5. Demand density and cumulative transit demand as functions of the distance from the city centre.

Using this scheme, which is represented in Figure 7.6, we obtain a ratio between $\rho_0(t^{op})$ and $\rho_0(t^{peak})$ of 3/10, as suggested by Jara-Díaz et al. (2017).

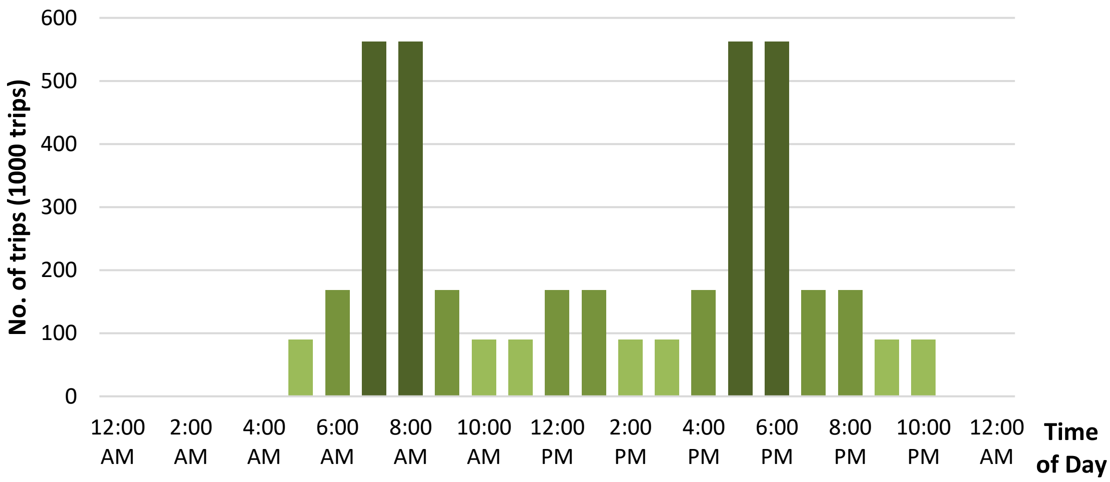


Figure 7.6. Transit demand fluctuation during the day.

Referring to Table 7.3 we assume the cost coefficients of the FRF and the DRF are equal: this is because we assume that the switching between the two feeder services, in the *Adaptive Transit* scheme, occurs maintaining the same vehicles.

Table 7.3. Parameters of the base scenario

Parameter	Value	Name	Reference
R	25 km	Radius of metropolitan area	-
$\rho_0(t^{peak})$	1600 trips/km^{2}h	Demand density (city centre - peak hours)	-
$\rho_0(t^{op})$	480 trips/km^{2}h	Demand density (city centre - off-peak hours)	-
$\rho_0(t^{lp})$	256 trips/km^{2}h	Demand density (city centre - low-peak hours)	-
γ	0.12 km^{-1}	Demand density gradient	-
v_w	4.5 km/h	Walking speed	*Google Maps*
v_{MRT}	60 km/h	MRT cruise speed	Daganzo (2010)
v_{FRF}, v_{DRF}	25 km/h	Feeder cruise speed	Daganzo (2010)
C_{MRT}	1200	MRT vehicle capacity	-
C_{FRF}, C_{DRF}	80	Feeder bus capacity	-
$\tau_{s,MRT}$	45 s	Time lost at MRT station	Daganzo (2010)
$\tau_{s,FRF}, \tau_{s,DRF}$	30 s	Time lost at feeder bus stop	Daganzo (2010)
τ_p	2 s	Time lost per boarding / alighting passenger	-
μ_A, μ_W, μ_T	15, 22.5, 30 €/h	Value of Time related to walking, waiting, travelling	Meunier and Quinet (2015)
ΔA	2 min	Time penalty due to transfers between MRT lines	-
$\mu_{L,MRT}$	600 €/km h ($x < r$); 300 €/km h ($x > r$)	Cost coefficient related to MRT infrastructure	Flyvbjerg et al. (2008)
$\mu_{ST,MRT}$	300 €/km h ($x < r$); 100 €/km h ($x > r$)	Cost coefficient related to MRT stations	Flyvbjerg et al. (2008)
$\mu_{L;FRF}, \mu_{L,DRF}$	10 €/km-h	Cost coefficient related to feeder infrastructure	CERTU (2011)
$\mu_{V,MRT}$	6 €/veh km	Cost coefficient related to MRT vehicle-distance	Daganzo (2010)
$\mu_{V,FRF}, \mu_{V,DRF}$	0.5 €/veh km	Cost coefficient related to feeder vehicle-distance	Cats and Glück (2019)
$\mu_{M,MRT}$	120 €/veh h	Cost coefficient related to MRT fleet size	Daganzo (2010)
$\mu_{M;FRF}, \mu_{M,DRF}$	50 €/veh h	Cost coefficient related to feeder fleet size	Cats and Glück (2019)

Also, we set the cost coefficient $\mu_{L,MRT}$ in the suburban area is half of the one in the central area, since MRT lines do not require extensive tunnelling (which impacts on costs) outside the city centre. For the same reason, the cost coefficient $\mu_{ST,MRT}$ is 3 times larger in the central area.

Finally, to obtain the capital cost coefficients, we use a straight-line amortization assuming 20 years (of 365 days) of useful life for MRT, 12 years for FRF and DRF, considering 18 operating hours per day.

7.5.2 Performance of *Adaptive* design scheme

We now compare the overall cost Z^{24h} obtained with the three transit schemes, the difference in their optimal structure and the impact on user QoS, to show the superiority of our proposed *Adaptive Transit* over today transit design. The results are obtained, for each transit scheme, by applying the optimization procedure on the respective CA model. Such procedure searches for the optimal structure, i.e., the set of values of the decision variables that minimize the social cost.

We remark that such a procedure, which we implement in *MATLAB*, is quite computationally efficient and terminates in about 10-20 minutes on an ordinary laptop for each transit scheme. As a comparison, observe that one single agent-based simulation in Narayan et al. (2020), took 45h on a super-computer. Obviously, we expect that the accuracy of their results is much stronger. However, such huge computation times are not suitable when we do not want to study the performance of a single transit configuration, but we want to find the optimal spatial configuration of the overall transit structure in a large urban area, and we are only interested in high level managerial insights.

Temporal distribution of cost. Figure 7.7 (left) shows the most relevant differences between the performance of classic Fixed transportation vs. our *Adaptive Transit* scheme. Let us partition the time periods into disjoint sets $\mathcal{T} = \mathcal{T}^{peak} \cup \mathcal{T}^{op} \cup \mathcal{T}^{lp}$. We represent the capital cost $Z_{cap} = \sum_{i\in\{L,M\}} Z_i$ (see Equation 7.18), the operational cost $Z_{op} = Z_V$ and the user cost components $Z_{user} = Z_A + Z_W + Z_T$. We recall that the total cost Z^{24h} over the entire day is computed via Equation 7.21. We clearly see that *Adaptive Transit* greatly reduces user cost in all periods of day. This reduction is particularly stronger outside the peak hour, where classic fixed transportation shows more evidently its limitations. The reduction in user cost is mostly achieved thanks to a remarkable reduction of access (walking) time Z_A. Indeed, where and when the demand density is low (suburbs, off-peak), *MRT-only* scheme provides only few stops, to prevent the operational cost to explode, thus requiring users to walk a lot. A similar (although less pronounced) problem occurs in Off- and Low-peak with MRT-FRF scheme. *Adaptive Transit* does not have such an issue, as when and where FRF becomes too disadvantageous for the user, it deploys DRF, which picks up and drops off users at their locations. Observe that this improvement for users requires higher agency costs, both capital cost (fleet Z_M and infrastructure Z_L) and operational (vehicle-km cost Z_V), since the agency needs to deploy more feeder vehicles.

However, such increase of agency cost is worth it, since the overall cost of *Adaptive Transit* outperforms the other schemes. Figure 7.7 shows that our *Adaptive Transit* scheme reduces the overall daily generalized cost (-19.7% and -3.6% than the *MRT-only* and the *MRT-FRF* scheme, respectively). This saving is more pronounced during off-peak hours (7.2% of improvement compared to the *MRT-FRF* scheme).

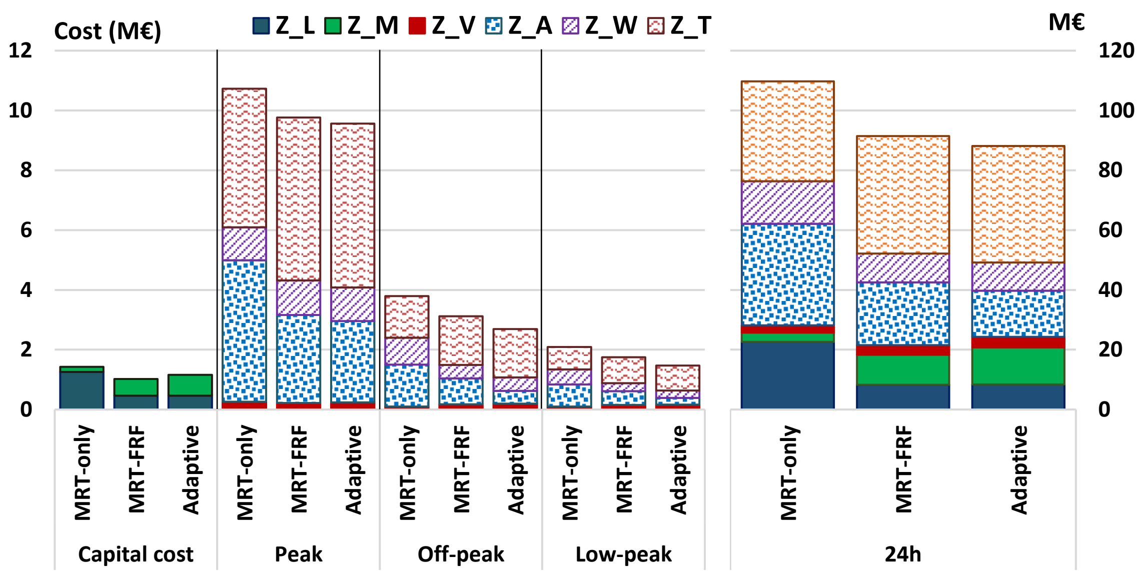


Figure 7.7. Cost of MRT-Only, MRT-FRF and Adaptive Transit scheme across the day.

Spatial distribution of cost. In Figure 7.8, we divide the study area in 3 different zones: the 1st zone for x < 6 km, the second zone for 6 km < x < 15 km, and the third for x > 15 km. Such zones have comparable size to Paris city, *Petite Couronne* and *Grande Couronne*, respectively. With this figure we aim at evaluating how the user cost (in terms of travel time) distributes among the three zones in the different design schemes.

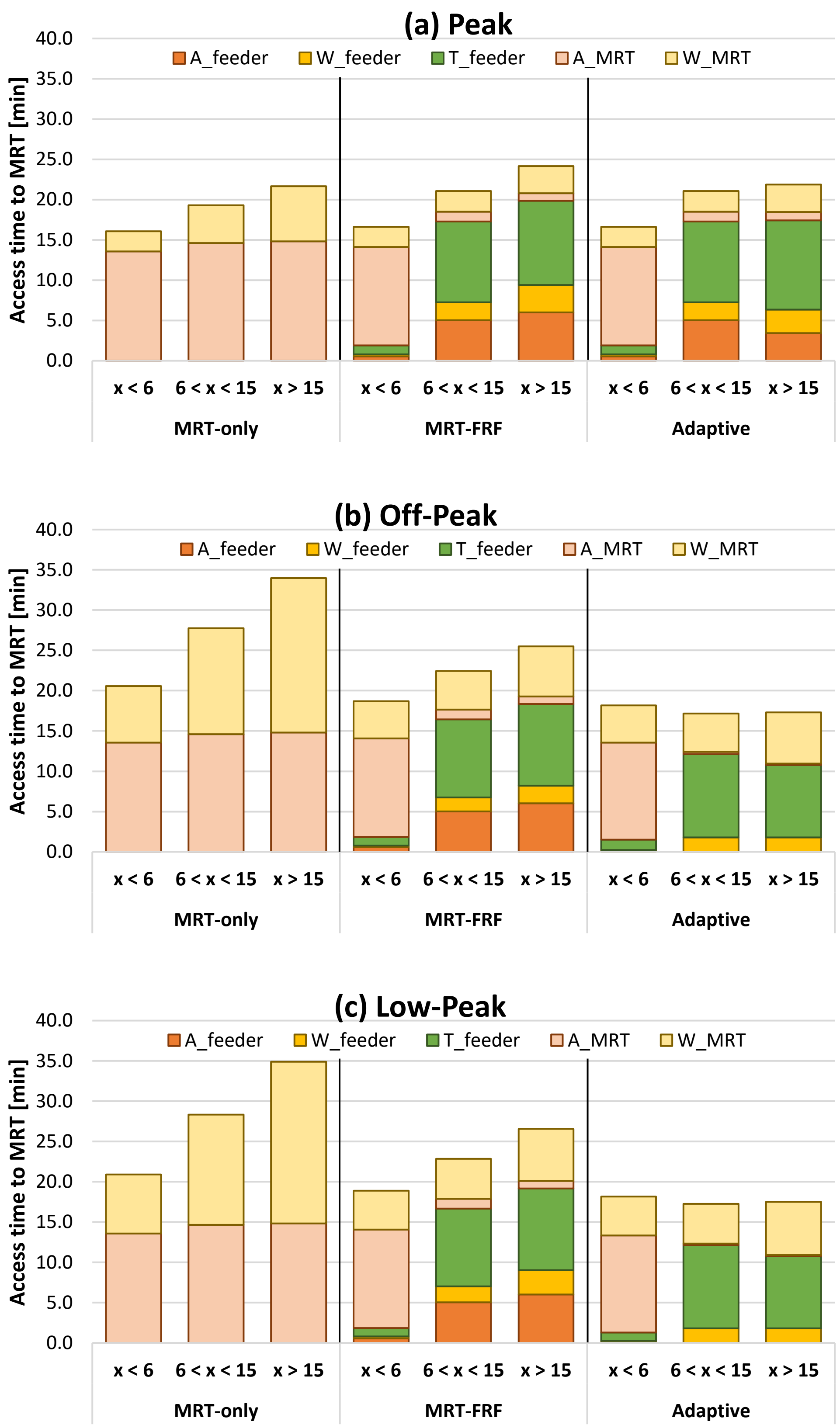


Figure 7.8. Components of the total access time to MRT stations in 3 zones of the study area, for MRT-Only, MRT-FRF and Adaptive Transit scheme, in (a) Peak, (b) Off-peak and (c) Low-Peak hours.

In particular, we represent in each zone the average time (in minutes) suffered by the users of the transit system to access the MRT station. This time is made up by five components, i.e., the walking time A_{feeder} to access/egress the feeder stop (if any), the waiting time W_{feeder} for the feeder service (if any), the in-vehicle time T_{feeder} on the feeder (if any), the walking time A_{MRT} to access/egress the MRT station and the waiting time W_{MRT} at the MRT station.

We observe that the time needed to access the MRT explodes far the from city centre, in particular during off-peak hours, in the *MRT-only* and *MRT-FRF* scheme. In the first case, the discomfort for passengers far from the centre is further exacerbated by high waiting times. Such passengers are the ones that, in reality, would not take transit and would rather use private cars. *Adaptive Transit*, instead, provides a fast connection to MRT stations far from the centre, while keeping the waiting time relatively low. Therefore, *Adaptive Transit* prevents the accessibility to MRT from degrading in suburban areas.

Our *Adaptive* design alleviates the cost suffered from users in the periphery much more than classic designs, by shifting the agency costs toward the outskirt (see the next subsection). Observe instead that classic design suffers from a bias, favouring city centre, in cost distribution: agency invests more in the city centre, so that user cost is minimized there, at the detrimental of suburban population. In order words, classic designs inherently suffer from high inequality. Such an inequality is alleviated with *Adaptive Transit*, which improves user-cost in the suburb, without degrading too much performance in the centre.

7.5.3 Spatial adaptivity of *Adaptive Transit*

Figure 7.9 shows the optimal structure of the three transit schemes with reference to the peak hours and the off-peak hours, explaining the cost results previously discussed. Observe that the three schemes slightly differ in the central area $x < r$, since in any case only MRT is deployed there.

We observe that the optimal value of r for the *MRT-only* scheme (Figure 7.9a-b) is considerably higher than the schemes adopting feeder services in the FMLM. In fact, the outermost MRT ring line should have a 9 km radius vs. the 5.5 km of both *MRT-FRF* (Figure 7.9c-d) and *Adaptive Transit* (Figure 7.9e-f). This ensures a wider double-coverage (ring and radial lines) area, but it would result, as we have already seen, in higher infrastructure-related cost.

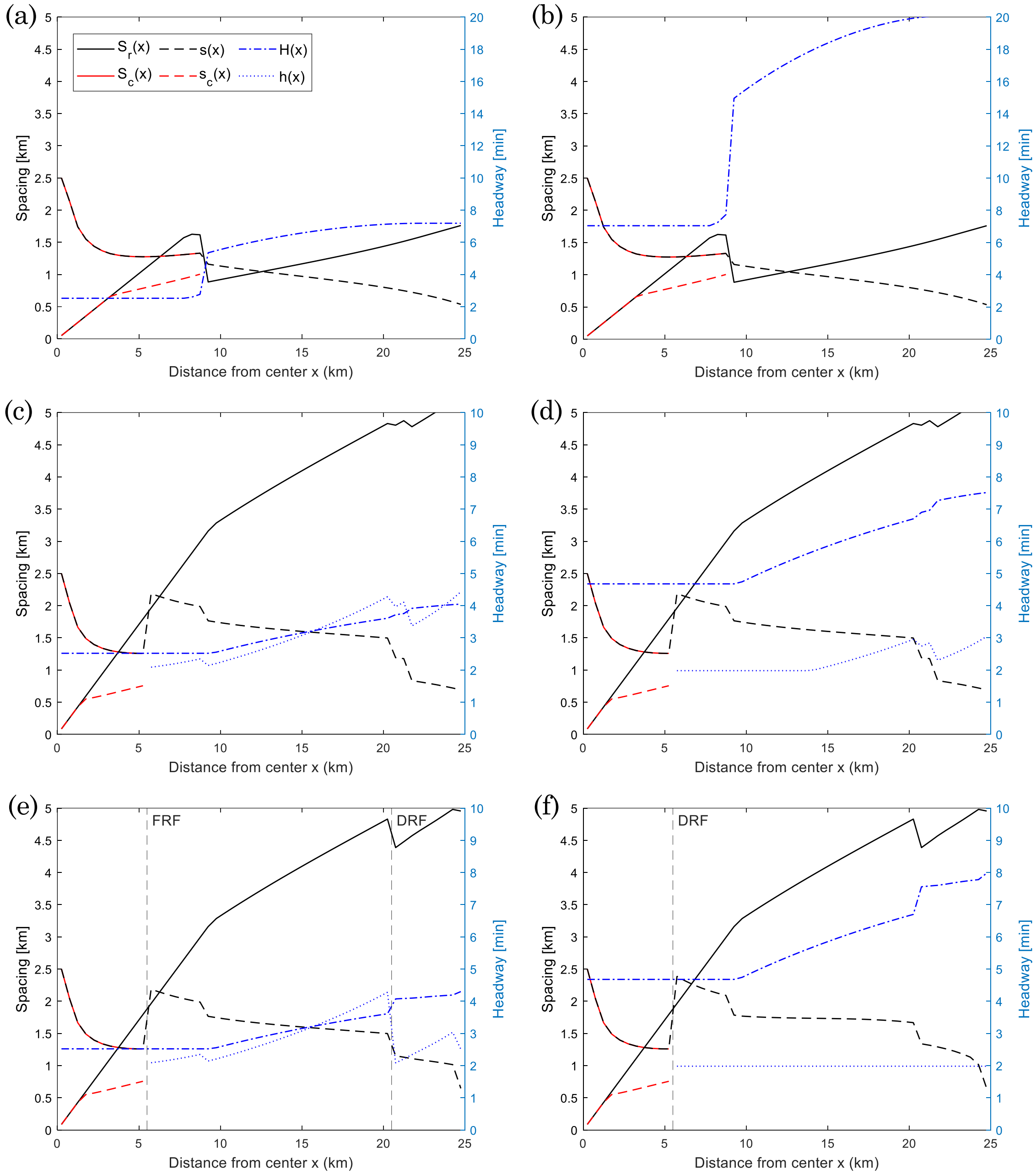


Figure 7.9. Optimal decision variables for the three transit schemes: (a) MRT-only; (b) MRT-FRF; (c) Adaptive. Recall that $S_r(x)$ and $S_c(x)$ are the spacings between MRT lines (radial and circular, respectively), $s(x)$ and $s_c(x)$ are the spacings between MRT stations (radial and circular, respectively), $H(x)$ is the headway of MRT and $h(x)$ is the headway of the feeder service.

The differences between the three schemes are clearly visible in the suburban area. In Figure 7.9b one can see that deploying FRF services allows the transit agency to save on infrastructure cost by increasing the spacing $S_r(x)$ between radial MRT lines and thus to halve the headway of the MRT $H(x)$ with respect to the *MRT-only* scheme. Also, the distance between MRT stations is higher because users can exploit the feeder service instead of walking: increasing the station spacing results in a higher commercial speed on MRT lines and thus in lower in-vehicle times for users. During off-peak, as one can note from Figure 7.9d, $H(x)$ is almost twice higher than the peak value, while $h(x)$ is reduced. Such an improvement in the quality of the feeder service compensates the higher waiting time for the MRT during off-peak and low-peak periods.

Finally, Figure 7.9e and Figure 7.9f show the decision variables derived through the optimization process for peak and off-peak hours for the *Adaptive Transit* scheme. During peak, *Adaptive Transit* prefers to deploy FRF in the close suburbs, where the demand is sufficiently high, and relegates DRF only to the further periphery ($x < 21\ km$), where the feeder service areas are slightly smaller compared to the FRF. Moreover, the DRF requires a lower headway $h(x)$ to better accommodate the demand.

Observe that, for all the schemes the MRT offer is richer where the demand density is high, i.e., the headway $H(x)$ and the radial line spacing $S_r(x)$ are smaller in the suburbs closer to the centre. This is also what we observe in real cities. We also observe that the stop spacing $s(x)$ decreases the further we go from the centre. This trend is more evident in the peak hours, due to the fact that the stop spacing strongly affects the commercial speed of the MRT and, consequently, the total travel time of passengers. Since the demand density is higher close to the city centre, it is more convenient to have a faster service (with less stops) in areas closer to it.

It is important to remark how the structure of our adaptive structure changes over the time and space. It is worth highlighting how, for $x > r$, the low demand density makes *Adaptive Transit* prefer DRF outside the peak. Moreover, the spacing between MRT stations increases going from peak to off-peak and low-peak hours. We pinpoint that such variation should not be interpreted as an unreal modification of the infrastructure, but as an operational strategy consisting of "skipping" some stops in certain moments of the day, i.e., introducing an "express" service not serving all the possible MRT stations, in order to offer faster connections. This shows that Adaptive Transit is able to vary the kind of offered feeder service spatially, to adapt to the geographic demand gradient, and temporally, to adapt to the time-evolution of the demand.

7.5.4 Effects of varying urban area size and value of time

We want to evaluate the impact of varying two parameters, namely the radius R of the city and the VoT for passengers, on the system costs in the Adaptive scheme.

In order to allow for a better comparison between each *R–VoT* combination, costs are expressed per passenger. Not surprisingly, the total cost rises when the study area widens and if the VoT increases. However, it is interesting to evaluate the "gain" in cost that we can achieve via *Adaptive Transit*, with respect to the *MRT-FRF* scheme. From Figure 7.10a, we notice how the higher gain (~5%) can be obtained when the VoT is small (10 €/h), independently of *R*. Increasing the VoT and for larger study areas, the *Adaptive Transit* scheme has still the lower total cost, but the percentage gain decreases.

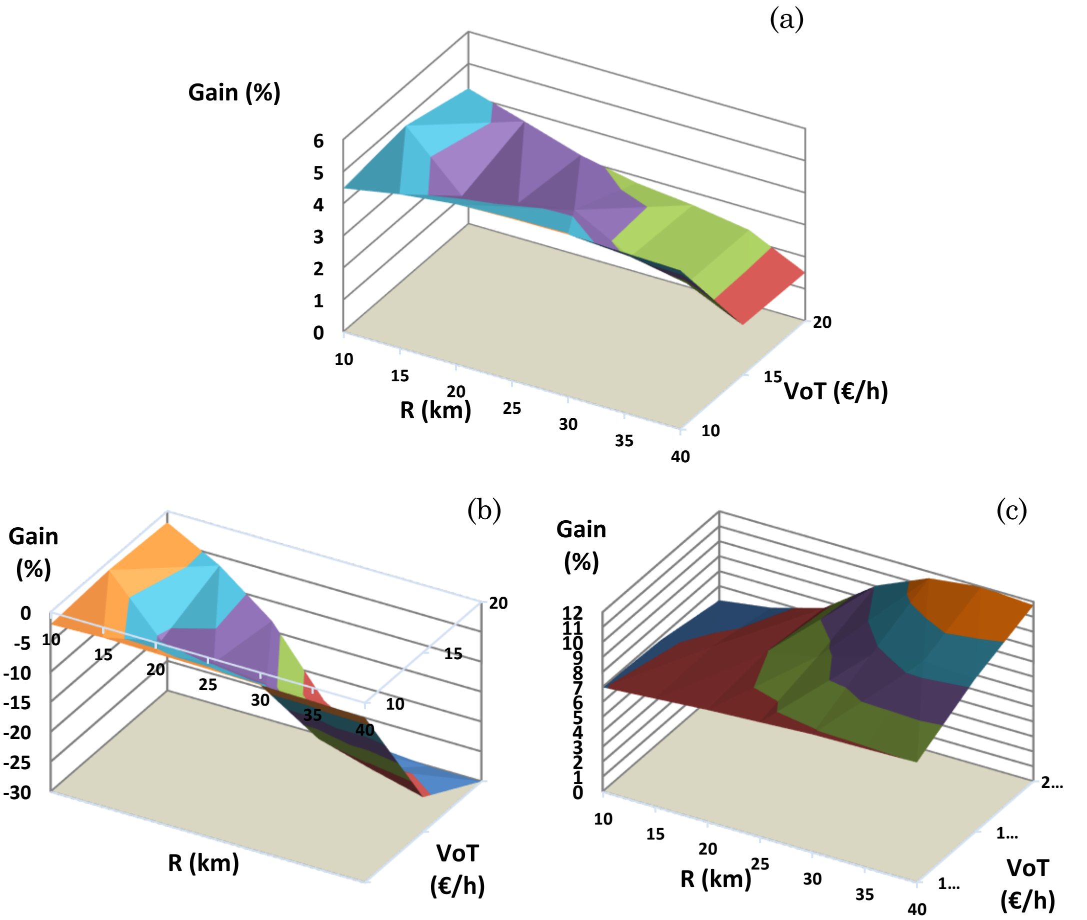


Figure 7.10. Cost improvements (on daily basis) of the Adaptive scheme with respect to the MRT-FRF scheme for different combinations of city size (R) and Value of Time (μ_A), considering: (a) the total cost per transported passenger; (b) the agency-related costs per transported passenger; (c) the user-related costs per transported passenger.

This trend can be better explained if we consider the agency-related (Figure 7.10b) and the user-related costs (Figure 7.10c) separately. Lower VoT (10 €/h) or small *R* (10 km) imply a modest increase of agency-related costs (~2%), but a significant reduction of user-related costs (from 6% to 8%). When considering higher values of VoT and *R*, instead, the optimal configuration of the *Adaptive Transit* leads to a stronger increase in agency-related costs (~30% if VoT = 20 €/h and R > 30 km) than the decrease in user-related costs (~12% if VoT = 20 €/h and R > 30 km). One reason is that *Adaptive Transit*, as already claimed, prevents the accessibility to MRT (and thus travel times) from degrading in suburban areas, by offering a demand-responsive feeder service with high frequency. Ensuring a good QoS results in higher capital and operating costs for the transit agency.

7.6 Conclusion and future research

We have presented the concept of *Adaptive Transit*, which combines fixed-route and demand-responsive transport and alternates between the two services in order to adapt to the spatial and temporal variation of the demand density. We provided a theoretical high-level model of *Adaptive Transit* based on Continuous Approximation, where the demand density and the decision variables defining the transit network are continuous functions across the space and vary over time. Numerical results on such a model show that *Adaptive Transit* tilts the balance of the overall costs in favour of user-centric components, keeping the agency cost at a reasonable level, such that the overall cost (the sum of the two) is improved. An important benefit of *Adaptive Transit* is that it limits the degradation of QoS in suburban areas. In the *MRT-only* Scheme, the line density in the entire study area would be unfeasibly high. *MRT-FRF* and *Adaptive Transit* schemes, instead, can afford a sparser MRT infrastructure by solving the FMLM transportation problem via feeder services, whence the infrastructure savings. An additional advantage, not considered in this work, can be represented by the induced or latent demand, i.e., due to the modal shift from car to PT which would take place thanks to the improved service for passengers.

To summarize, the novelty of this work is that it provides managerial insights on how to holistically optimize the design of a future transit system, able to combine fixed schedule and demand-responsive operations in a single multi-modal service. In future work, we will verify the impact of electrification and automation (Badia and Jenelius, 2021), which deeply modify the agency cost components. Also, intermediate operating strategies between FRF and DRF should be investigated, including the possibility of consolidating the demand for the feeder service in a limited number of stop locations. In conclusion, we believe the concept of *Adaptive Transit* can guide planning agencies in the design of more efficient next-generation transit systems.

Appendix B

The derivation of the cost components and constraints of the Continuous Approximation (CA) model is presented as follows. Table B1 lists the parameters and variables used in the model.

Table B1. Notation of CA model's variables and parameters

Independent variables	
x	radial distance from the centre ($x = 0$) of the urban area
t	time of the day.
Input parameters	
R	Radius of the metropolitan area
$\rho_0(t)$	Demand density in the city centre during the time slot t
γ	Slope of the Clark's law (Equation 7.5)
$v_{MRT}, v_{FRF}, v_{DRF}, v_w$	Cruise speed of MRT, FRF, DRF; walking speed
$C_{MRT}, C_{FRF}, C_{DRF}$	Vehicle capacity
$\tau_{s,MRT}, \tau_{s,FRF}, \tau_{s,DRF}$	Dwell time at MRT stations, FRF and DRF stops
τ_p	Extra dwell time per passenger
μ_A, μ_W, μ_T	*VoT* of the access, walking and in-vehicle time
δ_{tr}	Time penalty due to transfers
$\mu_{L,MRT}, \mu_{L;FRF}, \mu_{L,DRF}$	Cost coefficients related to the infrastructure length
$\mu_{V,MRT}, \mu_{V;FRF}, \mu_{V,DRF}$	Cost coefficients for the distance travelled by the vehicles
$\mu_{M,MRT}, \mu_{M;FRF}, \mu_{M,DRF}$	Cost coefficients related to the fleet size
$N(x)$	Number of passengers in the infinitesimal annulus of radius x, $N(x) = 2\,\rho(x) \cdot (2\pi x)$.
Local decision variables at location x	
$\theta_r(x), S_r(x)$	Angular and linear spacing between radial MRT lines
$S_c(x)$	Spacing between circular MRT lines
$s(x)$	Spacing between stations along a radial MRT line
$\phi(x)$	Angle between stations on a circular MRT line
$H(x)$	Headway on circular and radial MRT lines
$h(x)$	Headway of the feeder service in the suburban area
$d_j(x)$	Spacing between FRF stops if j = FRF, twice the maximum walking distance from the station if j = DRF
$F(x)$	Type of FMLM service, $F(x) \in$ {FRF, DRF, 0}

Derived variables at location x	
$\mathbb{I}_j(x)$	Indicator function, it is 1 if $F(x) = j$, 0 otherwise
$\mathcal{D}(x)$	Local decision functions $\mathcal{D}(x) = \{\theta_r(x), S_c(x), s(x), \phi(x), H(x), h(x), d_j(x), F(x)\}$
$l(x)$	Length of the FMLM rectangle
$CL_j(x)$	Cycle length of the FMLM feeder of type $j \in$ {FRF, DRF}
$C_j(x)$	Cycle time of the FMLM feeder of type $j \in$ {FRF, DRF}
Global decision variables	
r	Radius of the central (double-coverage) area
Q_0	Maximum value of total radial flow of MRT vehicles, $Q_0 = Q(x{=}0)$
H_B	Headway at the outermost ring line $H_B = H(x = r)$
G	Global decision variables $G = (r; Q_0; H_B)$
Output parameters	
$y_{L,j}(x), y_{V,j}(x), y_{M,j}(x)$	Agency local cost at distance x from the centre: infrastructure, vehicle-km and fleet costs, $j \in$ {MRT, FRF, DRF}
$y_{A,j}(x), y_{W,j}(x), y_{T,j}(x)$	Travel time components suffered by users at distance x from the city centre: walking, waiting and in-vehicle time, $j \in$ {MRT, FRF, DRF}
F_L, F_V, F_A, F_W, F_T	Global cost components for agency and users.
$Dem_o(x), Dem_d(x)$	Number of origin or destination trips lying within rings of radii $(x, x + dx)$.
$P_o(x), P_d(x)$	Probability for an origin or destination trip of lying within rings of radii $(x, x + dx)$.
DEM	Total number of trips per hour in the study area.
$v_{cr}(x), v_{cc}(x)$	Commercial speed on radial lines and ring lines at x.
$Q(x)$	Radial flow of MRT vehicles (trains per hour)
$O_j(x)$	Average vehicle occupancy
Z	Global cost

The next sections of the appendix outline the derivation of operational outputs, the agency-related and the user-related cost components. The derivation procedure of the cost components for the MRT takes inspiration from Chen et al. (2015) - Appendix A - except that we consider the same headway on circular and radial lines $H_r(x) = H_c(x) = H(x)$ and treat the spacing between radial $s(x)$ and circular $\phi(x)$ MRT stations as local decision variables, while in Chen et al. (2015) they do not change with x. Also, in this study we consider a symmetric demand pattern, so that $Dem_o(x) = Dem_d(x) = Dem(x) = 2\pi x \cdot \rho(x)$, and we obtain:

$$P(x) = \frac{Dem(x)}{DEM} = \frac{Dem(x)}{\int_0^R Dem(y)dy} \quad \text{(B.1)}$$

B.1. Derivation operational outputs and constraints for the MRT

MRT radial-line and ring-line commercial speed. They are given by the sum, per unit distance, of the cruising time (at speed v_{MRT}) and the time lost at stations due to acceleration and deceleration, including the time spent boarding passengers, as follows:

$v_{cc}(x) = 1/\left(\frac{1}{v_{MRT}} + \frac{\tau_{s,MRT}}{s(x)}\right)$; $v_{cr}(x) = 1/\left(\frac{1}{v_{MRT}} + \frac{\tau_{s,MRT}}{\phi(x)x}\right)$

Commercial speed on the boundary ring. As before, but considering the angle ϕ_B between stations on the outermost ring line, as follows:

$v_{cB} = 1/\left(\frac{1}{v_{MRT}} + \frac{\tau_{s,MRT}}{\phi_B r}\right)$

Vehicle capacity constraint. The expected maximum number of passengers on board a MRT vehicle is constrained to be less than the vehicle's passenger-carrying capacity, i.e., $O_{MRT}(x) \leq C_{pax,MRT}$. To better understand the derivation of the following formulas, the reader can refer to the scheme of Figure 7.3. For ring lines, the total number onboard all MRT vehicles on a ring line at x is: $2\left(P(x)\int_x^R P(y)dy\right) \cdot \frac{2}{\pi}$. The ratio of the average trip length to the length of the ring line is: $x/2\pi x$. The flow of MRT vehicles on that ring is: $2/(S_c(x)H(x))$. Hence:

$O_{MRT,c}(x) = \left(Dem(x)\int_x^R P(y)dy\right) \cdot \frac{2}{\pi} \cdot \frac{(S_c(x)H(x)}{2\pi}$, if $x < r$.

For radial lines, we therefore obtain:

$O_{MRT,r}(x) = \int_0^R \rho(y)dy \cdot \left(\int_x^R P(y)dy \int_0^x P(y)dy \cdot \frac{2}{\pi} + \int_x^R P(y)dy \cdot \left(1 - \frac{2}{pi}\right)\right) \cdot \frac{\theta_r(x)H(x)}{2\pi}$, if $x < r$.

$O_{MRT,r}(x) = \int_0^R \rho(y)dy \cdot \left(\int_x^R P(y)dy\right) \cdot \frac{S_c(x)H(x)}{2\pi}$, if $x > r$.

Vehicle capacity constraint on the boundary ring. Similarly, the vehicle's passenger-carrying capacity constraint for the boundary route is: $O_B \leq C_{pax,MRT}$. Hence:

$O_B = DEM \cdot \left(\int_x^R P(y)dy \cdot \int_x^R P(y)dy\right) \cdot \frac{2}{\pi} \cdot \frac{H_B/2}{2\pi}$.

B.2. Derivation of the agency-related cost components for the MRT

Local cost for the length of MRT lines. Consider the area between two rings of radii x and $x + dx$, which is equal to $2\pi x \cdot dx$. The length of the MRT radial lines within the ring pair is $2\pi/\theta_r(x) \cdot dx$. The length of the ring lines is instead $2\pi x/S_c(x) \cdot dx$. The local cost is given by the length of the MRT lines in the area divided by the area width dx; that is:

$Y_L(x) = \frac{2\pi}{\theta_r(x)} + \frac{2\pi x}{S_c(x)}$ if $x < r$; $Y_L(x) = \frac{2\pi}{\theta_r(x)}$, if $x > r$.

Local cost for the vehicle-distance travelled per hour. It is obtained by multiplying the local cost for the length of ring-lines and radial-lines with their corresponding transit flows, i.e., $1/H(x)$, multiplied by 2 due to the bi-directional travel flows on each line; that is:

$Y_V(x) = \frac{4\pi}{\theta_r(x)H(x)} + \frac{4\pi x}{S_c(x)H(x)}$ if $x < r$; $Y_V(x) = \frac{4\pi}{\theta_r(x)H(x)}$, if $x > r$.

Global cost for the vehicle-distance travelled per hour. Global costs refer only to the outermost (boundary) ring line, hence:

$F_V = \frac{4\pi r}{\mathrm{H_B}}$.

Local cost for the fleet size in the peak hour. It is given by the ratio between the vehicle-distance travelled per hour and the commercial speed, that is:

$Y_M(x) = \frac{4\pi}{\theta_r(x)H(x)v_{cr}(x)} + \frac{4\pi x}{S_c(x)H(x)v_{cc}(x)}$ if $x < r$; $Y_M(x) = \frac{4\pi}{\theta_r(x)H(x)v_{cr}(x)}$, if $x > r$.

Global cost for the fleet size in the peak hour. As before:

$F_M = \frac{4\pi r}{\mathrm{H_B} vcB}$.

B.3. Derivation of the agency-related cost components for the FMLM

Denoting with $sa(x) = 2\pi/(\theta_r(x)/2)$ the number of subareas along the ring at x, the agency local costs for the feeder service are computed as follows. We reasonably assume that the infrastructure cost $y_L(x)$ is the same for FRF and DRF operation and thus we do not need to distinguish between them. Therefore:

$$y_{L,j}(x) = N_s(x) \cdot \frac{CL_{FRF}(x)}{2\ s(x)} \cdot sa(x)$$

We need to do this distinction, instead, for $y_V(x)$ and $y_M(x)$, in which the cycle length $CL_j(x)$ and the cycle time $C_j(x)$ appear, which are different for FRF and DRF (see Equations 7.7; 7.9; 7.12; 7.13). The vehicle-distance travelled per hour is obtained by multiplying the cycle length $CL_j(x)/s(x)$ by their corresponding vehicle frequency $1/h(x)$:

$$y_{V,j}(x) = \sum_{j=\{FRF,DRF\}} N_s(x) \cdot \frac{CL_j(x)}{s(x) \cdot h(x)} \cdot sa(x)$$

The cost of fleet size is derived from the number of vehicles $Cj/(s(x) \cdot h(x))$ needed to ensure the feeder service:

$$y_{M,j}(x) = \sum_{j=\{FRF,DRF\}} N_s(x) \cdot \frac{C_j(x)}{s(x) \cdot h(x)} \cdot sa(x)$$

In the equations above, *sa(x)* multiplies on the left the cost (per distance unit) related to a single service area.

B.4. Derivation of the user-related cost components for the MRT

The cost components related to the users of the transit system depend on:

- Walking time A to reach the bus stop or the MRT station or the destination.
- Waiting time W at the bus stop or the MRT station.
- In-vehicle (MRT, FRF or DRF) time T including boarding, riding, dwell and alighting time.
- Transfers: since any possible transfer between different transit lines is an additional disutility, we treat it as a penalty of the extra walking time ΔA.

To derive the user-related cost components, we consider nine different cases of trips, as depicted in Figure B1.

<u>Local cost for the passenger average access time</u>.

Cases (a) and (g): $Y_A(x) = 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \left(1 - \frac{2}{\pi}\right)$, if $x > r$.

Cases (b) and (h): $Y_A(x) = \begin{cases} 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \left(1 - \frac{2}{\pi}\right), \text{if } x < r \\ 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \left(1 - \frac{2}{\pi}\right), \text{if } x > r. \end{cases}$

Cases (c) and (i): $Y_A(x) = 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \left(1 - \frac{2}{\pi}\right)$, if $x < r$.

Case (d): $Y_A(x) = 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \frac{2}{\pi}$, if $x > r$;

Cases (e): $Y_A(x) = \begin{cases} 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \left(\frac{\phi(x)x}{4} + \frac{S_c}{4}\right)/v_W \cdot \frac{2}{\pi}, \text{if } x < r \\ 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right)/v_W \cdot \frac{2}{\pi}, \text{if } x > r. \end{cases}$

Case (f): $Y_A(x)(x) = 2 \cdot Dem(x) \int_x^r P(y)dy \cdot \left(\frac{\phi(x)x}{4} + \frac{S_c}{4}\right) / v_w \cdot \frac{2}{\pi} + 2 \cdot Dem(x) \int_0^x P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right) / v_w \cdot \frac{2}{\pi}$, if $x < r$;

If we sum up these components (Cases a-i) we obtain:

$$Y_A(x)(x) = \begin{cases} 2 \cdot Dem(x) \cdot \left[\int_0^x P(y)dy \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right) / v_w + \int_x^R P(y)dy \cdot \left(\frac{\phi(x)x}{4} + \frac{S_c}{4}\right) / v_w\right], & \text{if } x < r \\ 2 \cdot Dem(x) \cdot \left(\frac{\theta_r(x)x}{4} + \frac{s}{4}\right) / v_w, & \text{if } x > r. \end{cases}$$

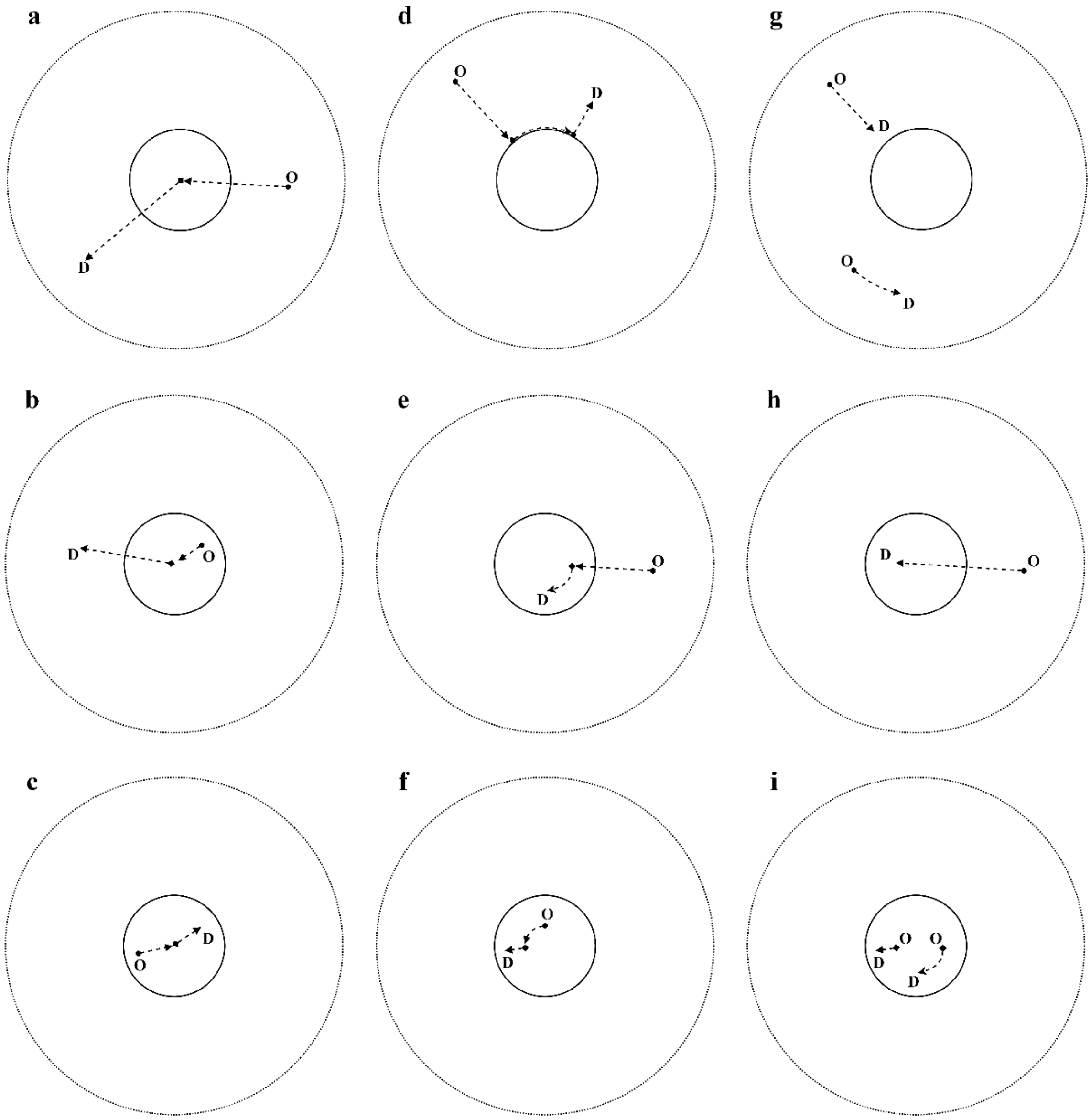


Figure B.1. Different cases of trips for the derivation of user-related cost components.

Global cost for the transfers between MRT lines. During their trips, travellers transfer once (Cases a-c;e-f), twice (case d, where origins and destinations both lie in the periphery and $\Theta < 2$ rad), or do not transfer (cases g-i). The average expected number of transfers is:

$$F_A = 1 + \int_r^R P(y)dy \cdot \int_r^R P(y)dy \cdot \frac{2}{\pi} - \frac{\theta_{r,min}}{\pi}.$$

Local cost for the passenger average waiting time.

Case (a): $Y_W(x) = 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \frac{H(x)}{2} \cdot \left(1 - \frac{2}{\pi}\right)$, if $x > r$.

Case (b): $Y_W(x) = \begin{cases} 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \frac{H(x)}{2} \cdot \left(1 - \frac{2}{\pi}\right), \text{if } x < r \\ 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \frac{H(x)}{2} \cdot \left(1 - \frac{2}{\pi}\right), \text{if } x > r. \end{cases}$

Case (c): $Y_W(x) = 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \frac{H(x)}{2} \cdot \left(1 - \frac{2}{\pi}\right)$, if $x < r$.

Case (d): $Y_W(x) = 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{2}{\pi}$, if $x > r$.

Case (e): $Y_W(x) = \begin{cases} 2 \cdot Dem(x) \int_r^R P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{2}{\pi}, \text{if } x < r \\ 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{2}{\pi}, \text{if } x > r. \end{cases}$

Case (f): $Y_W(x) = 2 \cdot Dem(x) \int_0^r P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{2}{\pi}$, if $x < r$.

The following cases account for those trips which do not require transfers between MRT lines. Therefore, they should be subtracted from the total value of $Y_W(x)$.

Cases (g), (h) and (i): $Y_W(x) = \begin{cases} 2 \cdot Dem(x) \int_0^x P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{\theta_r(x)}{\pi}, \text{if } x < r \\ 2 \cdot Dem(x) \int_{x-s/2}^{x+s/2} P(y)dy \cdot \frac{H(x)}{2} \cdot \frac{\theta_r(x)}{\pi}, \text{if } x > r. \end{cases}$

If we sum up Cases (a-g) and subtract Cases (g-i) we obtain:

$$Y_W(x) = \begin{cases} 2 \cdot Dem(x) \cdot \left[1 - \left(\int_0^x P(y)dy\right) \cdot \frac{\theta_r(x)}{\pi}\right] \cdot \frac{H(x)}{2}, & \text{if } x < r \\ 2 \cdot Dem(x) \cdot \left[1 - \left(\int_0^x P(y)dy + \int_{x-s/2}^{x+s/2} P(y)dy\right) \cdot \frac{\theta_r(x)}{\pi}\right] \cdot \frac{H(x)}{2}, & \text{if } x > r. \end{cases}$$

Global cost for the passenger average waiting time. We only consider the cost that occurs at the outermost ring line, for case (d), excluding the limited portion of trips which do not involve any transfer:

$$F_W = DEM \cdot \int_r^R P(y)dy \cdot \int_r^R P(y)dy \cdot \frac{H_B}{2} \cdot \left(\frac{2}{\pi} - \frac{\theta_{r,min}}{\pi}\right).$$

Local cost for the passenger average in-vehicle time. We consider the annulus $(x, x + dx)$. We obtain the in-vehicle travel time on each annulus weighted by the proportion of users who travel along it and who cross it. Then, the in-vehicle travel time on that annulus by its width dx.

Cases (a), (b) and (c): $Y_T(x) = DEM \cdot 2\int_x^R P(y)dy \cdot \frac{1}{v_{cr}(x)} \cdot \left(1 - \frac{2}{\pi}\right)$, $\forall x$

Case (d): $Y_T(x) = DEM \cdot 2\int_x^R P(y)dy \cdot \frac{1}{v_{cr}(x)} \cdot \frac{2}{\pi}$, if $x > r$.

Cases (e) and (f): $Y_T(x) = DEM \cdot 2\int_x^R P(y)dy \cdot \int_0^x P(y)dy \cdot \frac{1}{v_{cr}(x)} \cdot \frac{2}{\pi} + 2 \cdot Dem(x)\int_x^R P(y)dy \cdot \frac{1}{v_{cc}(x)} \cdot \frac{2}{\pi}$, if $x < r$

The following cases account for those trips which do not require transfers between MRT lines. Therefore, they should be subtracted from the total value of $y_{T,MRT}(x)$.

Case (g): $Y_T(x) = DEM \cdot 2\int_{x+\frac{s}{2}}^{x+\frac{s}{2}} P(y)dy \cdot \int_{x+\frac{s}{2}}^{x+\frac{s}{2}} P(y)dy \cdot \frac{1}{v_{cr}(x)} \cdot \frac{\theta_r(x)}{\pi}$, if $x > r$.

Case (h): $Y_T(x) = DEM \cdot 2\int_x^R P(y)dy \cdot \int_0^x P(y)dy \cdot \frac{1}{v_{cr}(x)} \cdot \frac{\theta_r(x)}{\pi}$, if $x > r$.

If we sum up Cases (a-g) and subtract Cases (g-h) we obtain the local cost $Y_T(x)$.

Global cost for the passenger average in-vehicle time. We consider the in-vehicle travel time along the boundary ring line:

$$F_W = DEM \cdot \int_r^R P(y)dy \cdot \int_r^R P(y)dy \cdot \frac{r}{v_{cB}} \cdot \left(\frac{2}{\pi} - \frac{\theta_{r,min}}{\pi}\right) + \left[DEM \cdot 2\int_x^R P(y)dy \cdot \int_0^x P(y)dy\right] \cdot \frac{1}{v_{cr}(x)} \cdot \frac{2}{\pi} + \left[2 \cdot Dem(x)\int_x^R P(y)dy\right] \cdot \frac{1}{v_{cc}(x)} \cdot \frac{2}{\pi}.$$

B.5. Derivation of the user-related cost components for the FMLM

Let us denote with the walking speed with v_w and the number of passengers in the infinitesimal annulus of radius x with:

$$N(x) = 2\,\rho(x) \cdot (2\pi x)$$

Note that a percentage p_{walk} of passengers does not use the feeder service and that the average walking distance for users residing in the walking area to reach the MRT station in case of DRF is $d_{0,DRF}(x)/3 + d_{0,DRF}(x)/3 = (2/3)\,d_{0,DRF}(x)$. Then, the total time components suffered by users at x are expressed as follows. FRF passengers walk to the closest stop (feeder stop or directly MRT station). According to Equation 7.8, the average walk distance is $s(x)/4 + d(x)/4$. DRF passengers are picked-up and dropped-off in-place, without walking. Only the users within the walking area, i.e., a fraction $p_{walk,DRF}(x)$, walk to/from the MRT station an average distance of $(2/3)\,d_{0,DRF}(x)$. Therefore, the local cost for the feeder access time is:

$$y_{A,j}(x) = \begin{cases} N(x) \cdot \frac{s(x)+d(x)}{4\,v_w} & and \text{ if } j = \text{FRF} \\ N(x) \cdot p_{walk,DRF} \cdot \frac{2/3\,d_{0,DRF}(x)}{4\,v_w} & \text{if } j = \text{DRF} \end{cases}$$

The average waiting time for the feeder service and the in-vehicle time inside the feeder bus are as follows. Note that they are not experienced by passengers directly walking to the MRT station, whence the term $(1-p_{walk,j})$. Assuming that the average waiting time for both feeder services is given by the half of the headway $h(x)$, The local cost for the feeder waiting time is:

$$y_{W,j}(x) = N(x) \cdot \left(1-p_{walk,j}\right) \cdot \frac{h(x)}{2} \quad \text{if } j \in \{FRF, DRF\}$$

The average in-vehicle time for the FRF case is computed as in Quadrifoglio and Li (2009) and is given by $C_{FRF}/4 + (\Delta l(x) + d_{0,FRF}(x)/2)/v_{FRF}$. For the DRF case, the average in-vehicle time is given by $C_{DRF}/4 + \Delta l(x)/v_{DRF}$. The local cost for the feeder in-vehicle time is:

$$y_{T,j}(x) = \begin{cases} N(x) \cdot \left(1-p_{walk,FRF}\right) \cdot \left[\frac{C_{FRF}}{4} + \frac{\left(\Delta l(x) + \frac{d_{0,FRF}(x)}{2}\right)}{v_{FRF}}\right] & \text{if } j = \text{FRF} \\ N(x) \cdot \left(1-p_{walk,DRF}\right) \cdot \left[\frac{C_{DRF}}{4} + \frac{\Delta l(x)}{v_{DRF}}\right] & \text{if } j = \text{DRF} \end{cases}$$

We can summarize the formulas above, for any FMLM feeder type, by using indicator functions.

$$y_A(x) = \sum_{j \in \{FRF, DRF\}} \mathbb{I}_j(x) \cdot y_{A,j}(x) \quad \forall x > r$$

$$y_W(x) = \sum_{j \in \{FRF, DRF\}} \mathbb{I}_j(x) \cdot y_{W,j}(x) \quad \forall x > r$$

$$y_T(x) = \sum_{j \in \{FRF, DRF\}} \mathbb{I}_j(x) \cdot y_{T,j}(x) \quad \forall x > r$$

CHAPTER 8

Conclusions

The focus of this thesis has been on the design and operation of demand responsive transport (DRT) services acting as a feeder towards a transit line with high transport capacity and commercial speed, with the aim of enhancing the transit users' accessibility and make the transport system more efficient. In particular, we have explored the role of DRT in low-demand and suburban areas, where novel design approaches should strengthen the integration between conventional and flexible transit modes and enabling transit operation to adapt to spatial and temporal variations of the mobility demand.

This final chapter summarises and discuss the main findings of the thesis work (Section 8.1) and considers some directions for future research (Section 8.2).

8.1 Main findings and discussion of results

This thesis took its first steps trying to answer the research questions outlined in Chapter 1, regarding the potential of new forms of flexible transit in sparse urban areas and the methodological tools aimed at supporting the design of such services.

The first research question was:

> *How can sprawled urban and suburban areas be effectively served by public transport, thus contrasting the massive use of private vehicles?*

We believe that the optimal design of feeder bus lines would enhance the PT coverage in suburban and low-demand areas, so we proposed a methodology to face this specific problem (Chapter 3) and compared the performance of fixed-route bus versus flexible feeder services (Chapter 5). In Chapter 3 we tackled the problem of designing feeder routes using an ant-colony optimization algorithm and a GIS dataset within an agent-based programming environment, with the objective of maximizing the service coverage in terms of potential demand for the feeder service and according to travel time constraints. The model was applied to different case studies located in Catania (Italy), where a MRT network is currently being developed. Both single-station and multiple-stations routing problems were solved, finding that excessively increasing the route coverage can be deleterious for the travel time experienced by passengers. Moreover, the proposed model can be deployed as tactical decision tool to understand which locations (e.g., potential bus stops) should be served by a conventional (fixed-route) bus line and which areas could be covered by demand-responsive (DR) operations. To this end, a set of comparative simulations of fixed-route and flexible operating strategies for a feeder service was presented in Chapter 5. Simulations focused again on a case study situated in Catania (the *San Nullo* metro station), according to real demand data. Results showed that for an average demand rate of 50 pax/h a fleet of 5 vans (9 seats vehicles) would efficiently serve the demand with a flexible operation, while a fleet of 10 van would fit a demand rate of 100 pax/h, despite the high cost for the operator. In this last case, the transit operator should evaluate the trade-off between QoS and cost, and maybe choose a fixed-route strategy using a reduced number of higher capacity vehicles (e.g., minibuses or buses).

The second research question was of practical interest:

What can be a useful, practical and flexible tool to implement, reproduce and analyse the performance of on-demand transit services with different levels of flexibility?

To answer this question, we developed a handy simulation tool aimed at reproducing the operation of flexible transit services and the interactions between supply and passenger demand. This tool consists in an ABM environment, completely programmable, customizable, and with the possibility of integrating GIS datasets. Specifically, in Chapter 4 we reproduced the operation of an on-demand flexible transit service called MVMANT, experimented in 2019 in Dubai (UAE). Starting from real O-D data we evaluated different routing strategies and fleet characteristics and compared the MVMANT with a ridesharing service with low-capacity vehicles from both the points of view of the transport operator and the community. In Chapter 5, the level of flexibility of the simulated transit service is enhanced, since the requests can be served at multiple potential stops either at origin or destination, according to the vehicles' availability and capacity or time constraints. Every time that a new request occurs, the dispatching algorithm decides if the request can be accepted and assigned to a vehicle, and if the insertion of a new stop between two already scheduled stops is required. Eventually, the ABM presented in Chapter 6 allows to simulate a feeder service which is totally demand-responsive, i.e., where there are no fixed stops to serve, but the vehicle routes are dynamically built and updated according to real-time travellers' requests. Moreover, this ABM is based on a synthetic environment to better generalize and thus transfer the results to different context, and it goes beyond pure analytical models by implementing a dynamic dispatching algorithm and considering practical constraints. To summarize, thanks to the possibility of dynamically interacting with the simulation environment and visualizing routes on the map, and also because of the easy transferability to other contexts, our models can serve as practical and flexible tools for public transport planners and agencies.

The third research question was the following:

How flexible, demand responsive feeder transport services can effectively match the demand with the supply in real-time, trying to maximize shareability, minimize operator cost and contain passenger travel time?

This issue was specifically addressed in Chapters 5 and 6, by proposing dynamic vehicle dispatching algorithms able to satisfy as many trip requests as possible under realistic operating constraints. The algorithms presented in both chapters are based on insertion heuristics involving three levels of exploration of the feasible solutions, i.e., the set of vehicles which can serve the request, the set of locations where the request can be physically served, and the set of feasible insertions between two already scheduled stops. This approach has the advantage of consolidating passenger requests, reducing detours and allowing for a higher shareability of the service. Moreover, it enables the transit operator to vary the flexibility of the operation from almost a door-to-door service to a point- or route-deviation service. The dispatching algorithm proposed in Chapter 6 is more sophisticated than the one outlined in Chapter 5, under two aspects. In fact, the former takes into account the passengers' time windows (both for the access and the egress phase of the trip), as additional constraints, and the operator's additional cost due to each new request, in formulating the objective function. We believe that such dispatching procedure can be applied by real operators in real transit networks, providing good solutions without requiring high computational burden, and it can be improved to deal with wider operating context.

The fourth research question is as follows:

> *When is it more convenient to adopt a fixed-route policy rather than a flexible one in providing feeder services toward mass rapid transit?*

An attempt to answer this question was made in Chapters 6 and 7. First of all, the convenience of a transit service is not absolute, but it depends on the point of view of the actors involved. In fact, a feeder service based on a fixed route covering the most demanded areas and thus ensuring a good ridership can be convenient for a transit operator but not for the community, since many other travellers would be excluded from the PT coverage, or at least they would walk long distances to access the transit system. On the other hand, employing a large fleet of vehicles to serve a sparse but low demand would greatly improve the QoS experienced by passengers, at the expense of operating cost and thus high service fees. The opportunity of switching between different operating strategies, increasing or decreasing the flexibility of routes and schedules, can help face this trade-off problem. In Chapter 6 we presented a methodological approach allowing to find the critical demand density below which a demand responsive feeder service would be more efficient (convenient for passengers and profitable for the operator), given the fleet size and the characteristics (size and shape) of the service area. We found that, for elongated areas (ratio between length and width equal to 4) of about 2.5 km^2 and "symmetric" demand patterns (similar share of travellers accessing and egressing the system) the critical demand density is between 40 and 45 passengers/km^{2}h, when relying on a fleet of 3 small buses. Moreover, when the critical demand value occurs, a DR service performs better when the demand density shows a decreasing trend from the MRT station to the outskirts (e.g., in a transit-oriented development scenario) and when a fleet of 5 vans (i.e., more vehicles with lower capacity) is deployed. Instead, simulation results suggest that the demand pattern is a critical issue to take into account, because a demand mainly originated at (or directed to) the MRT station would make DR operations less efficient. The comparison between conventional and DR feeder services is also addressed by the analytical model presented in Chapter 7, where we did not focus on single local areas around transfer stations, but we propose a design model for the entire urban transit network, which consists of several transfer stations. We showed that a transit system switching from fixed-route to DR feeder operations in off-peak hours can prevent accessibility to MRT from degrading in peripheral and suburban areas. Eventually, the introduction of fully autonomous vehicles in transit operations could radically shift the attention of transit providers towards DRT, due to the opportunity for a significant reduction of operating cost.

Finally, the last research question regarded the entire urban transit network:

> *Where, when and how the layout of the urban transit network should adapt to the spatial and temporal demand variations?*

The analytical design model presented in Chapter 7 was conceived to answer this question. The *Adaptive Transit* we proposed consists of a (fixed) MRT system served by feeder bus routes in suburban regions. Depending on the distance from the city centre and the period of the day, and thus to the demand variations, the system can switch between conventional and DR feeder service, acting on different service features (e.g., MRT and feeder headways, distance between stations, etc.). We modelled the design problem using the continuous approximation technique, which is suited for the strategic phase of transit plans, thanks to the possibility of providing high-level managerial insights on how the transit network should be designed. Numerical results showed that the proposed adaptive approach allows to reduce the social cost, particularly during off-peak hours and in peripheral areas, without exceeding in agency cost. Besides, the outcomes of the model are sensitive to the various input parameters, such as the value of time of travellers, which should be carefully investigated.

To conclude this section, an overview table of the core chapters of the thesis is shown in Table 8.1.

Table 8.1. Summary table of the core chapters of the thesis

	Chapter 3	**Chapter 4**	**Chapter 5**	**Chapter 6**	**Chapter 7**
Goal	Design of optimal feeder bus routes	Comparing flexible routing strategies	Comparison between a fixed-route and a flexible feeder service	Exploring the optimal operating conditions of fixed vs. DRT feeder services	Design of adaptive transit, integrating conventional PT and DRT
Decision level	Tactical	Operational	Operational	Operational and Tactical	Strategic
Methodology	ABM + GIS	ABM + GIS	ABM + GIS	Analytical + ABM	Analytical (CA)
Level of flexibility	Low (fixed route)	Medium/Low (Fixed route + optional routes)	Medium (Fixed stops + optional stops)	Medium/High (demand-responsive stops with consolidation)	High (from fixed-route to door-to-door DRT)
Main performance indicators	Service coverage	Total Unit Cost (TUC); Service coverage; Operator cost.	Passenger travel time; Transport Intensity (TI); TUC	TUC; User disutility.	Generalized transport cost (user + agency costs).
Spatial context	Suburban	Suburban	Suburban	Suburban, periphery	Large urban area
Case study location	Catania (Italy)	Dubai (UAE)	Catania (Italy)	Synthetic	Ideal (based on Paris, France)

8.2 Research directions

The comprehensive objective of this thesis is to contribute to the current methodological framework for the design and operation of flexible transit services, with specific reference to low-demand suburban areas. The recent technological innovations have led to the introduction of innovative forms of DRT services, which are currently (almost always) operated by private companies and often in competition with PT authorities.

The transport planning process should involve policy-makers, transit agencies and on-demand transport providers to ensure the optimal integration between the various components of the shared mobility ecosystem. In order to achieve this, a more detailed analysis of cost and profit for the DRT operator and the transit agency (in case it is the provider of both conventional and flexible services) should be addressed in future research.

Another step forward would be to consider a broader range of mobility options in our models, introducing discrete choice models (Le Pira, Marcucci et al., 2017; Fournier et al., 2018; Oh et al., 2021) hence the possibility for travellers of choosing between alternative modes of transport (e.g., PT, ridesharing services, private car, micromobility, etc.) and accounting for more complex and realistic trip chains.

In this thesis we chose to focus on the first and last mile leg of travellers' journeys, often modelling the demand patterns according to a many-to-one scheme, where a transfer station attracts the ridership. In future research, our ABMs would be applied to a broader scope than just the local confined areas around MRT stations, thus dealing with a many-to-many demand pattern.

Also, an attractive issue which is currently poorly addressed in literature is the evaluation of dynamic pricing schemes (Lei et al., 2019) for flexible transit, which would help the service providers improve system performance, maximizing the profit according to the spatial and temporal variation in demand. Dynamic pricing strategies can be included in more sophisticated ABMs, where negotiation mechanisms between the vehicle fleet and the users and among the vehicles themselves would be simulated.

Future research would evaluate the impact of vehicle electrification and automation, under different maturity stage of this technologies, on the operations of conventional or flexible transit services. Once again, automated feeder transit solutions in suburban areas would be of particular interest (Badia and Jenelius, 2021), since they are promising in the opportunity of reducing operating costs.

The current research direction is envisioning near-future transport systems where conventional PT and DRT services are integrated in the broader context of Mobility as a Service (MaaS), where a tailored mobility package is offered to the users, by integrating different modes of transport in one online platform, using geospatial data generation tools, multimodal transport ticketing, and an integrated e-payment platform (Jittrapirom et al., 2017). Despite literature on the topic is still at an early stage, it is perceivable that the MaaS paradigm opens new attractive scenarios for demand-adaptive (collective and individual) mobility. Besides, MaaS would increase the overall attractiveness of the transit systems, especially for low-demand areas where the PT operations are often expensive and ineffective.

Finally, this thesis work proposes a methodological framework to support the design of flexible transit services in different stages of the transport planning process. We believe that this study represents the initial step of a wider research on the ever-changing era of the shared mobility.

Bibliography

A

Adnan, M., Pereira, F. C., Azevedo, C. M. L., Basak, K., Lovric, M., Raveau, S., ... and Ben-Akiva, M. (2016, January). Simmobility: A multi-scale integrated agent-based simulation platform. In 95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record.

Aldaihani, M. M., Quadrifoglio, L., Dessouky, M. M., and Hall, R. (2004). Network design for a grid hybrid transit service. Transportation Research Part A 38(7), 511-530, ISSN 0965-8564,URL http://dx.doi.org/https://doi.org/10.1016/j.tra.2004.05.001.

Almasi, M. H., Sadollah, A., Oh, Y., Kim, D.-K., and Kang, S. (2018). Optimal Coordination Strategy for an Integrated Multimodal Transit Feeder Network Design Considering Multiple Objectives. Sustainability, 1-28.

Alonso-González, M. J., Liu, T., Cats, O., Van Oort, N., and Hoogendoorn, S. (2018). The potential of demand-responsive transport as a complement to public transport: An assessment framework and an empirical evaluation. Transportation Research Record, 2672(8), 879-889.

Alonso-González, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S., and Hoogendoorn, S. (2020). Value of time and reliability for urban pooled on-demand services. Transportation Research Part C: Emerging Technologies, 115, 102621.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., and Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proceedings of the National Academy of Sciences, 114(3), 462-467.

Ambrosino, G., Nelson, J. D., and Gini, S. (2016). The pivotal role of Public Transport in designing the integration of mobility services and in operating MaaS offer: the concept of Shared Mobility Centre and the experience of Arezzo. In REAL CORP 2016–SMART ME UP! Proceedings of 21st International Conference on Urban Planning, Regional Development and Information Society (pp. 737-747). CORP–Competence Center of Urban and Regional Planning.

Ambrosino, G., Nelson, J. D., Boero, M., and Pettinelli, I. (2016). Enabling intermodal urban transport through complementary services: From Flexible Mobility Services to the Shared Use Mobility Agency: Workshop 4. Developing inter-modal transport systems. Research in Transportation Economics, 59, 179-184.

An, K., and Lo, H. K. (2015). Robust transit network design with stochastic demand considering development density. Transportation Research Part B: Methodological, 81, 737-754.

Anand, N., van Duin, R., and Tavasszy, L. (2019). Carbon credits and urban freight consolidation: An experiment using agent based simulation. Research in Transportation Economics, 100797.

Ansari, S., Başdere, M., Li, X., Ouyang, Y., and Smilowitz, K. (2018). Advancements in continuous approximation models for logistics and transportation systems: 1996–2016. Transportation Research Part B: Methodological, 107, 229-252.

Araldo, A., Di Maria, A., Di Stefano, A., and Morana, G. (2019, October). On the Importance of demand Consolidation in Mobility on Demand. In 2019 IEEE/ACM 23rd International Symposium on Distributed Simulation and Real Time Applications (DS-RT) (pp. 1-8). IEEE.

Araldo, A., Gao, S., Seshadri, R., Azevedo, C. L., Ghafourian, H., Sui, Y., ... and Ben-Akiva, M. (2019). System-Level Optimization of Multi-Modal Transportation Networks for Energy Efficiency using Personalized Incentives: Formulation, Implementation, and Performance. Transportation Research Record, 2673(12), 425-438.

Armendáriz, M., Burguillo, J. C., Peleteiro-Ramallo, A., Arnould, G., and Khadraoui, D. (2011, June). Carpooling: A Multi-Agent Simulation In Netlogo. In ECMS (pp. 61-67).

Atasoy, B., Ikeda, T., Song, X., and Ben-Akiva, M. E. (2015). The concept and impact analysis of a flexible mobility on demand system. Transportation Research Part C: Emerging Technologies, 56, 373-392.

Axhausen, K. W., Horni, A., and Nagel, K. (2016). The multi-agent transport simulation MATSim (p. 618). Ubiquity Press.

Azevedo, C. L., Marczuk, K., Raveau, S., Soh, H., Adnan, M., Basak, K., ... and Ben-Akiva, M. (2016). Microsimulation of demand and supply of autonomous mobility on demand. Transportation Research Record, 2564(1), 21-30.

B

Baaj, M. H., and Mahmassani, H. S. (1995). Hybrid route generation heuristic algorithm for the design of transit networks. Transportation Research Part C: Emerging Technologies, 3(1), 31-50.

Badia, H. (2020). Comparison of bus network structures in face of urban dispersion for a ring-radial city. Networks and Spatial Economics, 20(1), 233-271.

Badia H, Estrada M, Robusté F (2014) Competitive transit network design in cities with radial street patterns. Transportation Research Part B 59:161–181.

Badia, H., and Jenelius, E. (2020). Feeder transit services in different development stages of automated buses: comparing fixed routes versus door-to-door trips. Transportation Research Procedia, 47, 521-528.

Badia, H., and Jenelius, E. (2021). Design and operation of feeder systems in the era of automated and electric buses. Transportation Research Part A: Policy and Practice, 152, 146-172.

Baldacci, R., Mingozzi, A., Roberti, R. (2012). Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints. European Journal of Operational Research, 218(1), 1-6.

Banister, D. (2008). The sustainable mobility paradigm. Transport policy, 15(2), 73-80.

Basu, R., Araldo, A., Akkinepally, A. P., ... and Ben-Akiva, M. (2018). Automated mobility-on-demand vs. mass transit: a multi-modal activity-driven agent-based simulation approach. Transportation Research Record, 2672(8), 608-618.

Bettencourt, L. M. (2015). Cities as complex systems. Modeling complex systems for public policies, 217-236.

Biazzo, I., Monechi, B., and Loreto, V. (2019). General scores for accessibility and inequality measures in urban areas. Royal Society open science, 6(8), 190979.

Bisaschi, L., Romano, F., Carlberg, M. et al. (2021). Research for TRAN Committee – Transport infrastructure in low-density and depopulating areas. European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.

Bischoff, J., Maciejewski, M., and Nagel, K. (2017, October). City-wide shared taxis: A simulation study in Berlin. In 2017 IEEE 20th international conference on intelligent transportation systems (ITSC) (pp. 275-280). IEEE.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the national academy of sciences, 99(suppl. 3), 7280-7287.

Borowska-Stefanska, M., and Wisniewski, S. (2017). Vehicle Routing Problem as urban public transport optimization tool. Computer Assisted Methods in Engineering and Science, 23(4), 213-229.

Brake, J., Nelson, J. D., and Wright, S. (2004). Demand responsive transport: towards the emergence of a new market segment. Journal of Transport Geography, 12(4), 323-337.

Bruni, M. E., Guerriero, F., and Beraldi, P. (2014). Designing robust routes for demand-responsive transport systems. Transportation research part E: logistics and transportation review, 70, 1-16.

Bürstlein, J., López, D., and Farooq, B. (2021). Exploring first-mile on-demand transit solutions for North American suburbia: A case study of Markham, Canada. Transportation Research Part A: Policy and Practice, 153, 261-283.

C

Calabrò, G., Inturri, G., Le Pira, M., Pluchino, A., and Ignaccolo, M. (2020a). Bridging the gap between weak-demand areas and public transport using an ant-colony simulation-based optimization. Transportation Research Procedia 45C, 234–241.

Calabrò, G., Correia, G., Giuffrida, N., Ignaccolo, M., Inturri, G., and Le Pira, M. (2020b). Comparing the performance of demand responsive and schedule-based feeder services of mass rapid transit: an agent-based simulation approach. In 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS) (pp. 280-285). IEEE.

Campbell, A. M., and Savelsbergh, M. (2004). Efficient insertion heuristics for vehicle routing and scheduling problems. Transportation science, 38(3), 369-378.

Cascetta, E. (2009). Transportation systems analysis: models and applications (Vol. 29). Springer Science and Business Media.

Cats, O., and Glück S (2019) Frequency and vehicle capacity determination using a dynamic transit assignment model. Transportation Research Record 2673(3):574–585.

Ceder, A., and Wilson, N. H. (1986). Bus network design. Transportation Research Part B: Methodological, 20(4), 331-344.

Ceder, A. (2001). Operational objective functions in designing public transport routes. Journal of advanced transportation, 35(2), 125-144.

Ceder, A. (2013). Integrated smart feeder/shuttle transit service: simulation of new routing strategies. Journal of advanced transportation 47, 595–618.

CERTU (2011). Les coûts des transports collectifs urbains en site propre – chiffres clefs – principaux paramètres.

Chan, N. D., and Shaheen, S. A. (2012). Ridesharing in North America: Past, present, and future. Transport reviews, 32(1), 93-112.

Chandra, S., Bari, M. E., Devarasetty, P. C., and Vadali, S. (2013). Accessibility evaluations of feeder transit services. Transportation Research Part A: Policy and Practice, 52, 47-63.

Chandra, S., and Quadrifoglio, L. (2013). A model for estimating the optimal cycle length of demand responsive feeder transit services. Transportation Research Part B: Methodological, 51, 1-16.

Chang, S. K., and Schonfeld, P. M. (1991). Optimization models for comparing conventional and subscription bus feeder services. Transportation Science, 25(4), 281-298.

Chen, H., Gu, W., Cassidy, M. J., and Daganzo, C. F. (2015). Optimal transit service atop ring-radial and grid street networks: A continuum approximation design method and comparisons. Transportation Research Part B 81, 755–774.

Chen, P. W., and Nie, Y. M. (2017). Analysis of an idealized system of demand adaptive paired-line hybrid transit. Transportation Research Part B 102, 38–54.

Chen, P. W., and Nie, Y. M. (2018). Optimal design of demand adaptive paired-line hybrid transit: Case of radial route structure. Transportation Research Part E: Logistics and Transportation Review, 110, 71-89.

Ciaffi, F., Cipriani, E., and Petrelli, M. (2012). Feeder bus network design problem: a new metaheuristic procedure and real size applications. Procedia: Social and Behavioral Sciences 54, 798-807.

Cich, G., Knapen, L., Maciejewski, M., Bellemans, T., and Janssens, D. (2017). Modeling demand responsive transport using SARL and MATSim. Procedia Computer Science, 109, 1074-1079.

Cisterna, C., Giorgione, G., and Viti, F. (2021). Explorative analysis of potential MaaS customers: an agent-based scenario. Procedia Computer Science, 184, 629-634.

Clark, C. (1951). Urban population densities. Journal of the Royal Statistical Society. Series A (General) 114(4), 490–496.

Coffman Jr, E. G., Jelenkovic, P., and Poonen, B. (1999). Reservation probabilities. Adv. Perform. Anal, 13, 129–158.

Cohen-Blankshtain, G., and Rotem-Mindali, O. (2016). Key research themes on ICT and sustainable urban mobility. International Journal of Sustainable Transportation, 10(1), 9-17.

Cordeau, J. F., Laporte, G., Savelsbergh, M. W., Vigo, D. (2007). Vehicle routing. Handbooks in operations research and management science, 14, 367-428.

Cordeau, J.-F., and Laporte, G. (2007). The dial-a-ride problem: models and algorithms. Annals of Operations Research, 29-46.

Cordeau, J.-F., Gendreau, M., Hertz, A., Laporte, G., and Sormany, J.-S. (2005). New Heuristics for the Vehicle Routing Problem. In A. langevin, and D. Riopel, Logistics Systems: Design and Optimization (p. 279-297). New-York: Springer.

Currie, G., and Fournier, N. (2020). Why most DRT/Micro-Transits fail–What the survivors tell us about progress. Research in Transportation Economics, 83, 100895.

D

Daganzo, C. F. (1978). An approximate analytic model of many-to-many demand responsive transportation systems. Transportation Research 12(5):325–333.

Daganzo, C. F. (1984). Checkpoint dial-a-ride systems. Transportation Research Part B 18(4-5):315–327.

Daganzo, C. F. (1987). Increasing model precision can reduce accuracy. Transportation Science, 21(2), 100-105.

Daganzo, C. F. (2010). Structure of competitive transit networks. Transportation Research Part B: Methodological, 44(4), 434-446.

Daganzo, C. F., and Ouyang, Y. (2019). A general model of demand-responsive transportation services: From taxi to ridesharing to dial-a-ride. Transportation Research Part B: Methodological, 126, 213-224.

Dantzig, G. B., and Ramser, J. H. (1959). The truck dispatching problem. Management science, 6(1), 80-91.

Davison, L., Enoch, M., Ryley, T., Quddus, M., and Wang, C. (2014). A survey of demand responsive transport in Great Britain. Transport Policy, 31, 47-54.

Di Maria, A., Araldo, A., Morana, G., Di Stefano, A. (2018). "AMoDSim: An Efficient and Modular Simulation Framework for Autonomous Mobility on Demand", Internet of Vehicles (IoV).

Diana, M., and Dessouky, M. M. (2004). A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. Transportation Research Part B: Methodological, 38(6), 539-557.

Diana, M., Dessouky, M. M., and Xia, N. (2006). A model for the fleet sizing of demand responsive transportation services with time windows. Transportation Research Part B: Methodological, 40(8), 651-666.

d'Orey, P. M., Fernandes, R., and Ferreira, M. (2012, September). Empirical evaluation of a dynamic and distributed taxi-sharing system. In 2012 15th International IEEE Conference on Intelligent Transportation Systems (pp. 140-146). IEEE.

Dorigo, M., and Di Caro, G. (1999). The ant colony optimization meta-heuristic. In D. Corne, and M. Dorigo, New Ideas in Optimization (p. 11–32). McGraw-Hill.

Dorigo, M., and Gambardella, L. M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Transactions on evolutionary computation, 1(1), 53-66.

Dorigo, M., and Socha, K. (2006). An Introduction to Ant Colony Optimization. In Approximation Algorithms and Metaheuristics, (p. 1-25). Artificielle, IRIDIA - Institut de Recherches Interdisciplinaires et de Développements en Intelligence, Université Libre de Bruxelles: T. F. Gonzalez.

Dorigo, M., and Stützle, T. (2004). Ant colony optimization. Cambridge, MA: MIT Press.

E

Edwards, D., and Watkin,s K. (2013). Comparing fixed-route and demand-responsive feeder transit systems in real world settings. Transportation research record 2352(1), 128–135.

EEA. (2006). Urban sprawl in Europe: The ignored challenge. Copenhagen: EEA, European Environmental Agency.

Enoch, M., Potter, S., Parkhurst, G., and Smith, M. (2004). Intermode: Innovations in demand responsive transport. https://www.researchgate.net/profile/Graham-Parkhurst/publication/37183508_INTERMODE_innovations_in_demand_responsive_transport_final_report/links/5616ac2008ae1a8880031d6a/INTERMODE-innovations-in-demand-responsive-transport-final-report.pdf

EPRS (2016). Sparsely populated and under– populated areas. https://www.europarl.europa.eu/RegData/etudes/BRIE/2016/586632/EPRS_BRI(2016)586632_EN.pdf

Errico, F., Crainic, T. G., Malucelli, F., and Nonato, M. (2013). A survey on planning semi-flexible transit systems: Methodological issues and a unifying framework. Transportation Research Part C: Emerging Technologies, 36, 324-338.
Estrada, M., Roca-Riu, M., Badia, H., Robusté, F., and Daganzo, C. F. (2011). Design and implementation of efficient transit networks: procedure, case study and validity test. Procedia-Social and Behavioral Sciences, 17, 113-135.

Ettema, D. (2015). Complexity methods applied to transport planning. Modeling Complex Systems for Public Policies, 279.

F

Fagnant, D. J., and Kockelman, K. M. (2014). The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transportation Research Part C: Emerging Technologies, 40, 1-13.

Fagnant, D. J., and Kockelman, K. M. (2018). Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas. Transportation, 45(1), 143-158.

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., and Rashidi, H. (2013). A review of urban transportation network design problems. European Journal of Operational Research, 229(2), 281-302.

Fielbaum, A., Bai, X., and Alonso-Mora, J. (2021). On-demand ridesharing with optimized pick-up and drop-off walking locations. Transportation research part C: emerging technologies, 126, 103061.

Flyvbjerg, B., Bruzelius, N., and Van Wee, B. (2008). Comparison of capital costs per route-kilometre in urban rail. European journal of transport and instrastructure research EJTIR, 8 (1).

Fournier, N., Chen, S., De Lima, I. H. V., Needell, Z., Deliali, A., Araldo, A., ... and Ben-Akiva, M. (2018, November). Integrated simulation of activity-based demand and multi-modal dynamic supply for energy assessment. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2277-2282). IEEE.

Franco, P., Johnston, R., and McCormick, E. (2020). Demand responsive transport: Generation of activity patterns from mobile phone network data to support the operation of new mobility services. Transportation Research Part A: Policy and Practice, 131, 244-266.

Fu, L. (2002). Planning and design of flex-route transit services. Transportation Research Record, 1791(1), 59-66.

G

Gambardella, L. M., Taillard, E., and Agazzi, G. (1999). MACSVRPTW: A multiple ant colony system for vehicle routing problems with time windows. In New Ideas in Optimization (p. 63–76). London, UK: McGraw-Hill.

Garaix, T., Artigues, C., Feillet, D., and Josselin, D. (2010). Vehicle routing problems with alternative paths: An application to on-demand transportation. European Journal of Operational Research, 204(1), 62-75.

Geurs, K. T., and Van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. Journal of Transport geography, 12(2), 127-140.

Giuffrida, N., Ignaccolo, M., Inturri, G., Rofè, Y., and Calabrò, G. (2017). Investigating the correlation between transportation social need and accessibility: the case of Catania. Transportation Research Procedia, 27, 816-823.

Giuffrida, N., Le Pira, M., Inturri, G., Ignaccolo, M., Calabrò, G., … and Pluchino, A. (2020). On-demand flexible transit in fast-growing cities: The case of dubai. Sustainability 12(11), 4455.

Gkiotsalitis, K., and Cats, O. (2021). Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. Transport Reviews, 41(3), 374-392.

Gschwender, A., Jara-Díaz, S., Bravo, C. (2016). Feeder-trunk or direct lines? economies of density, transfer costs and transit structure in an urban context. Transportation Research Part A 88, 209–222.

Guihaire, V., and Hao, J. K. (2008). Transit network design and scheduling: A global review. Transportation Research Part A: Policy and Practice, 42(10), 1251-1273.

Gudmundsson, H., Marsden, G., and Josias, Z. (2016). Sustainable transportation: Indicators, frameworks, and performance management. Springer: Berlin/Heidelberg, Germany. ISSN 2192-4333.

Guo, Q. W., Chow, J. Y., and Schonfeld, P. (2018). Stochastic dynamic switching in fixed and flexible transit services as market entry-exit real options. Transportation Research Part C: Emerging Technologies, 94, 288-306.

H

Hansen, W. G. (1959). How accessibility shapes land use. Journal of the American Institute of planners, 25(2), 73-76.

Ho, S. C., Szeto, W. Y., Kuo, Y. H., Leung, J. M., Petering, M., and Tou, T. W. (2018). A survey of dial-a-ride problems: Literature review and recent developments. Transportation Research Part B: Methodological, 111, 395-421.

Holden, E., Banister, D., Gössling, S., Gilpin, G., and Linnerud, K. (2020). Grand Narratives for sustainable mobility: A conceptual review. Energy Research and Social Science, 65, 101454.

Huang, D., Tong, W., Wang, L., and Yang, X. (2020). An Analytical Model for the Many-to-One Demand Responsive Transit Systems. Sustainability, 12(1), 298.

I

Ibarra-Rojas, O. J., Delgado, F., Giesen, R., and Muñoz, J. C. (2015). Planning, operation, and control of bus transport systems: A literature review. Transportation Research Part B: Methodological, 77, 38-75.

IEA (2018), CO2 Emissions from Fuel Combustion 2018, IEA, Paris, https://doi.org/10.1787/co2_fuel-2018-en.

Ignaccolo, M., Inturri, G., Le Pira, M., Caprì, S., and Mancuso, V. (2016). Evaluating the role of land use and transport policies in reducing the transport energy dependence of a city. Research in Transportation Economics, 55, 60-66.

Ingram, D. R. (1971). The concept of accessibility: a search for an operational form. Regional studies, 5(2), 101-107.

Inturri, G., Giuffrida, N., Ignaccolo, M., Le Pira, M., Pluchino, A., Rapisarda, A., and D'Angelo, R. (2021). Taxi vs. demand responsive shared transport systems: An agent-based simulation approach. Transport Policy, 103, 116-126.

Inturri, G., Le Pira, M., Giuffrida, N., Ignaccolo, M., Pluchino, A., Rapisarda, A., and D'Angelo, R. (2019). Multi-agent simulation for planning and designing new shared mobility services. Research in Transportation Economics, 73, 34-44.

J

Jara-Díaz, S., Fielbaum, A., and Gschwender, A. (2017) Optimal fleet size, frequencies and vehicle capacities considering peak and off-peak periods in public transport. Transportation Research Part A 106(June), 65–74, ISSN 09658564, URL http://dx.doi.org/10.1016/j.tra.2017.09.005.

Jennings, N., and Wooldridge, M. (1998). Applications of Intelligent Agents. In N. Jennings, and M. (. Wooldridge, Agent Technology (p. 3-28). Berlin, Heidelberg: Springer.

Jittrapirom, P., Caiati, V., Feneri, A. M., Ebrahimigharehbaghi, S., González, M. J. A., and Narayan, J. (2017). Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges. Urban Planning, 2(2), 13.

K

Kepaptsoglou, K., and Karlaftis, M. (2009). Transit Route Network Design Problem: Review. Journal of Transportation Engineering 135 (8), 491–505.

Koffman, D. (2004). Operational experiences with flexible transit services (No. 53). Transportation Research Board.

Kuah, G. K., and Perl, J. (1989). The feeder-bus network-design problem. Journal of the Operational Research Society, 40(8), 751-767.

Kuan, S., Ong, H. L., and Ng, K. (2006). Solving the feeder bus network design problem by genetic algorithms and ant colony optimization. Advances in Engineering Software 37, 351–359.

Kwan, M. P. (1998). Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. Geographical analysis, 30(3), 191-216.

L

La Greca, P., Barbarossa, L., Ignaccolo, M., Inturri, G., and Martinico, F. (2011). The density dilemma. A proposal for introducing smart growth principles in a sprawling settlement within Catania Metropolitan Area. Cities, 28(6), 527-535.

Leffler, D., Burghout, W., Jenelius, E., and Cats, O. (2021). Simulation of fixed versus on-demand station-based feeder operations. Transportation Research Part C: Emerging Technologies, 132, 103401.

Le Pira, M., Inturri, G., Ignaccolo, M., Pluchino, A., and Rapisarda, A. (2017). Finding shared decisions in stakeholder networks: an agent-based approach. Physica A: Statistical Mechanics and its Applications, 466, 277-287.

Le Pira, M., Marcucci, E., Gatta, V., Inturri, G., Ignaccolo, M., and Pluchino, A. (2017). Integrating discrete choice models and agent-based models for ex-ante evaluation of stakeholder policy acceptability in urban freight transport. Research in transportation economics 64, 13-25.

Le Pira, M., Tavasszy, L., Correia, G., Ignaccolo, M., and Inturri, G. (2021). Opportunities for integration between Mobility as a Service (MaaS) and freight transport: a conceptual model. Sustainable Cities and Society, 103212.

Levine, J., and Garb, Y. (2002). Congestion pricing's conditional promise: promotion of accessibility or mobility?. Transport Policy, 9(3), 179-188.

Li, X., Hu, S., Fan, W., and Deng, K. (2018). Modeling an enhanced ridesharing system with meet points and time windows. PloS one, 13(5), e0195927.

Li, X., and Quadrifoglio, L. (2010). Feeder transit services: choosing between fixed and demand responsive policy. Transportation Research Part C: Emerging Technologies, 18(5), 770-780.

Liang, X., de Almeida Correia, G. H., and Van Arem, B. (2016). Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. Transportation Research Part E: Logistics and Transportation Review, 93, 115-129.

Lin, C., Choy, K. L., Ho, G. T., Chung, S. H., and Lam, H. Y. (2014). Survey of green vehicle routing problem: past and future trends. Expert systems with applications, 41(4), 1118-1138.

Liyanage, S., Dia, H., Abduljabbar, R., and Bagloee, S. A. (2019). Flexible Mobility On-Demand: An Environmental Scan. Sustainability, 11(5), 1262.

Lu Y, Adnan M, Basak K, Pereira FC, Carrion C, Saber VH, Loganathan H, and Ben-Akiva ME (2015) Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model. TRB.

Lu, X., Yu, J., Yang, X., Pan, S., and Zou, N. (2015). Flexible feeder transit route design to enhance service accessibility in urban area. Journal of Advanced Transportation 50, 507-521.

Lucas, K. (2006). Providing transport for social inclusion within a framework for environmental justice in the UK. Transportation Research Part A: Policy and Practice, 40(10), 801-809.

Luo, S., and Nie, Y. M. (2019) Impact of ride-pooling on the nature of transit network design. Transportation Research Part B 129, 175–192.

Luo, S., and Nie, Y. M. (2020). Paired-line hybrid transit design considering spatial heterogeneity. Transportation Research Part B: Methodological, 132, 320-339.

M

Ma, X., Yuan, Y., Van Oort, N., and Hoogendoorn, S. (2020). Bike-sharing systems' impact on modal shift: A case study in Delft, the Netherlands. Journal of Cleaner Production, 259, 120846.

Machado, C. A. S., de Salles Hue, N. P. M., Berssaneti, F. T., and Quintanilha, J. A. (2018). An overview of shared mobility. Sustainability, 10(12), 4342.

Mageean, J., and Nelson, J. D. (2003). The evaluation of demand responsive transport services in Europe. Journal of Transport Geography, 11(4), 255-270.

Marcucci, E., Le Pira, M., Carrocci, C. S., Gatta, V., and Pieralice, E. (2017a). Connected shared mobility for passengers and freight: Investigating the potential of crowdshipping in urban areas. In 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) (pp. 839-843). IEEE.

Marcucci, E., Le Pira, M., Gatta, V., Ignaccolo, M., Inturri, G., and Pluchino, A. (2017b). Simulating participatory urban freight transport policy-making: Accounting for heterogeneous stakeholders' preferences and interaction effects. Transportation Research Part E 103, 69-86.

Martínez, L. M., and Eiró, T. (2012). An optimization procedure to design a minibus feeder service: an application to the sintra rail line. Procedia-Social and Behavioral Sciences, 54, 525-536.

Martinez, L. M., Correia, G. H., and Viegas, J. M. (2015). An agent-based simulation model to assess the impacts of introducing a shared-taxi system: an application to Lisbon (Portugal). Journal of Advanced Transportation, 49(3), 475-495.

Martinez, L. M., and Viegas, J. M. (2017). Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of Lisbon, Portugal. International Journal of Transportation Science and Technology, 6(1), 13-27.

Mehran, B., Yang, Y., and Mishra, S. (2020). Analytical models for comparing operational costs of regular bus and semi-flexible transit services. Public Transport, 12(1), 147-169.

Mendes, R. S., Miranda, D. S., Wanner, E. F., Sarubbi, J. F., and Martins, F. V. (2016). Multiobjective Approach to the Vehicle Routing Problem with Demand Responsive Transport. IEEE Congress on Evolutionary Computation (CEC), (p. 3761-3768). Vancouver, Canada.

Meunier, D., and Quinet, E. (2015). Value of time estimations in cost benefit analysis: the French experience. Transportation Research Procedia, 8, 62-71.

Meyer, J., Becker, H., Bösch, P. M., and Axhausen, K. W. (2017). Autonomous vehicles: The next jump in accessibilities?. Research in transportation economics, 62, 80-91.

M'hamdi, A., and Nemiche, M. (2018). Bottom-up and top-down approaches to simulate complex social phenomena. International Journal of Applied Evolutionary Computation (IJAEC), 9(2), 1-16.

Mishra, S., Mehran, B., and Sahu, P. K. (2020). Assessment of delivery models for semi-flexible transit operation in low-demand conditions. Transport Policy, 99, 275-287.

Mohaymany, A. S., and Gholami, A. (2010). Multimodal feeder network design problem: ant colony optimization approach. Journal of Transportation Engineering, 136(4), 323-331.

Montemanni, R., Gambardella, L., Rizzoli, A.-E., and Donati, A. (2005). Ant Colony System for a Dynamic Vehicle Routing Problem. Journal of Combinatorial Optimization, 327-343.

Mourad, A., Puchinger, J., and Chu, C. (2019). A survey of models and algorithms for optimizing shared mobility. Transportation Research Part B: Methodological, 123, 323-346.

N

Narayan, J., Cats, O., van Oort, N., and Hoogendoorn, S. (2020). Integrated route choice and assignment model for fixed and flexible public transport systems. Transportation Research Part C: Emerging Technologies, 115, 102631.

Newell, G. F. (1973). Scheduling, location, transportation, and continuum mechanics: some simple approximations to optimization problems. SIAM Journal on Applied Mathematics, 25(3), 346-360.

Newman, P., and Kenworthy, J. (2006). Urban design to reduce automobile dependence. Opolis, 2(1).

Nourbakhsh, S. M., Ouyang, Y. (2012). A structured flexible transit system for low demand areas. Transportation Research Part B 46(1), 204–216.

O

Oh, S., Lentzakis, A. F., Seshadri, R., and Ben-Akiva, M. (2021). Impacts of Automated Mobility-on-Demand on traffic dynamics, energy and emissions: A case study of Singapore. Simulation Modelling Practice and Theory, 110, 102327.

Oh, S., Seshadri, R., Azevedo, C. L., Kumar, N., Basak, K., and Ben-Akiva, M. (2020a). Assessing the impacts of automated mobility-on-demand through agent-based simulation: A study of singapore. Transportation Research Part A 138, 367–388.

Oh, S., Seshadri, R., Le, D. T., Zegras, P. C., and Ben-Akiva, M. E. (2020b). Evaluating automated demand responsive transit using microsimulation. IEEE Access, 8, 82551-82561.

Ouyang, Y., Nourbakhsh, S. M., and Cassidy, M. J. (2014). Continuum approximation approach to bus network design under spatially heterogeneous demand. Transportation Research Part B: Methodological, 68, 333-344.

P

Pan, S., Yu, J., Yang, X., Liu, Y., and Zou, N. (2015). Designing a flexible feeder transit system serving irregularly shaped and gated communities: Determining service area and feeder route planning. Journal of Urban Planning and Development, 141(3), 04014028.

Papanikolaou, A., Basbas, S., Mintsis, G., and Taxiltaris, C. (2017). A methodological framework for assessing the success of Demand Responsive Transport (DRT) services. Transportation Research Procedia, 24, 393-400.

Papanikolaou, A., and Basbas, S. (2021). Analytical models for comparing Demand Responsive Transport with bus services in low demand interurban areas. Transportation Letters, 13(4), 255-262.

Park, C., Lee, J., and Sohn, S. Y. (2019). Recommendation of feeder bus routes using neural network embedding-based optimization. Transportation Research Part A: Policy and Practice, 126, 329-341.

Pinto, H. K., Hyland, M. F., Mahmassani, H. S., and Verbas, I. Ö. (2020). Joint design of multimodal transit networks and shared autonomous mobility fleets. Transportation Research Part C: Emerging Technologies, 113, 2-20.

Pluchino, A., Rapisarda, A., and Garofalo, C. (2010). The Peter Principle revisited: A computational study. Physica A, 389, 467–472.

Q

Qiu, F., Shen, J., Zhang, X., and An, C. (2015). Demi-flexible operating policies to promote the performance of public transit in low-demand areas. Transportation Research Part A: Policy and Practice, 80, 215-230.

Quadrifoglio, L., Dessouky, M. M., and Ordóñez, F. (2008). A simulation study of demand responsive transit system design. Transportation Research Part A: Policy and Practice, 42(4), 718-737.

Quadrifoglio, L., Dessouky, M. M., and Palmer, K. (2007). An insertion heuristic for scheduling mobility allowance shuttle transit (MAST) services. Journal of Scheduling, 10(1), 25-40.

Quadrifoglio, L., Hall, R. W., and Dessouky, M. M. (2006). Performance and design of mobility allowance shuttle transit services: bounds on the maximum longitudinal velocity. Transportation science, 40(3), 351-363.

Quadrifoglio, L., and Li, X. (2009). A methodology to derive the critical demand density for designing and operating feeder transit services. Transportation Research Part B: Methodological, 43(10), 922-935.

R

Ronald, N., Thompson, R., and Winter, S. (2015). Simulating demand-responsive transportation: A review of agent-based approaches. Transport Reviews, 35(4), 404-421.

Ronald, N., Yang, J., and Thompson, R. G. (2016). Exploring co-modality using on-demand transport systems. Transportation Research Procedia, 12, 203-212.

Russel, S. J., and Norvig, P. (1995). Intelligent agents. In S. J. Russel, and P. Norvig, Artificial Intelligence: A Modern Approach (p. 31-52). Englewood Cliffs, N.J.: Prentice-Hall.

S

Sadowsky, N., and Nelson, E. (2017). The Impact of Ride-Hailing Services on Public Transportation Use: A Discontinuity Regression Analysis. Economics Department Working Paper Series. 13. https://digitalcommons.bowdoin.edu/econpapers/13

Sakai, T., Alho, A. R., Bhavathrathan, B. K., Dalla Chiara, G., Gopalakrishnan, R., Jing, P., ... and Ben-Akiva, M. (2020). SimMobility Freight: An agent-based urban freight simulator for evaluating logistics solutions. Transportation Research Part E: Logistics and Transportation Review, 141, 102017.

Salazar, M., Lanzetti, N., Rossi, F., Schiffer, M., and Pavone, M. (2019). Intermodal autonomous mobility-on-demand. IEEE Transactions on Intelligent Transportation Systems, 21(9), 3946-3960.

Sayarshad, H. R., and Gao, H. O. (2020). Optimizing dynamic switching between fixed and flexible transit services with an idle-vehicle relocation strategy and reductions in emissions. Transportation Research Part A: Policy and Practice, 135, 198-214.

Scheltes, A., and Correia, G. H. A. (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(1), 28-41.

Schlüter, J., Bossert, A., Rössy, P., and Kersting, M. (2021). Impact assessment of autonomous demand responsive transport as a link between urban and rural areas. Research in Transportation Business & Management, 39, 100613.

Shaheen, S., Cohen, A., Yelchuru, B., Sarkhili, S., and Hamilton, B. A. (2017). Mobility on demand operational concept report. U.S. Department of Transportation.

Shaheen, S., and Chan, N. (2016). Mobility and the sharing economy: Potential to facilitate the first-and last-mile public transit connections. Built Environment, 42(4), 573-588.

Shen, J., Yang, S., Gao, X., and Qiu, F. (2017). Vehicle routing and scheduling of demand-responsive connector with on-demand stations. Advances in Mechanical Engineering, 9(6), 1687814017706433.

Shen, Y., Zhang, H., and 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.

Shi, H., and Gao, M. (2020). Analysis of a Flexible Transit Network in a Radial Street Pattern. Journal of Advanced Transportation, 2020.

Shinoda, K., Noda, I., Ohta, M., Kumada, Y., and Nakashima, H. (2003, August). Is dial-a-ride bus reasonable in large scale towns? Evaluation of usability of dial-a-ride systems by simulation. In International Workshop on Multi-Agents for Mass User Support (pp. 105-119). Springer, Berlin, Heidelberg.

Sivakumaran, K., Li, Y., Cassidy, M. J., and Madanat, S. (2012). Cost-saving properties of schedule coordination in a simple trunk-and-feeder transit system. Transportation Research Part A: Policy and Practice, 46(1), 131-139.

Smith, G., Sochor, J., and Karlsson, I. M. (2018). Mobility as a Service: Development scenarios and implications for public transport. Research in Transportation Economics, 69, 592-599.

Sörensen, L., Bossert, A., Jokinen, J. P., and Schlüter, J. (2021). How much flexibility does rural public transport need?–Implications from a fully flexible DRT system. Transport Policy, 100, 5-20.

Steiner, K., and Irnich, S. (2020). Strategic planning for integrated mobility-on-demand and urban public bus networks. Transportation Science, 54(6), 1616-1639.

Stiglic, M., Agatz, N., Savelsbergh, M., and Gradisar, M. (2018). Enhancing urban mobility: Integrating ride-sharing and public transit. Computers and Operations Research, 90, 12-21.

Storme, T., De Vos, J., De Paepe, L., and Witlox, F. (2020). Limitations to the car-substitution effect of MaaS. Findings from a Belgian pilot study. Transportation Research Part A: Policy and Practice, 131, 196-205.

Stützle, T., and Hoos, H. H. (2000). MAX–MIN ant system. Future generation computer systems, 16(8), 889-914.

T

Tak, S., Woo, S., Park, S., and Kim, S. (2021). The City-Wide Impacts of the Interactions between Shared Autonomous Vehicle-Based Mobility Services and the Public Transportation System. Sustainability, 13(12), 6725.

Tavasszy, L. A. (2019). Predicting the effects of logistics innovations on freight systems: Directions for research. Transport Policy.

Teodorović, D. (2008). Swarm intelligence systems for transportation engineering: Principles and applications. Transportation Research Part C: Emerging Technologies, 16(6), 651-667.

Tisue, S., and Wilensky, U. (2004). NetLogo: A Simple Environment for Modeling Complexity. International Conference on Complex Systems. Boston.

Torkjazi, M., and Huynh, N. (2019). Effectiveness of dynamic insertion scheduling strategy for demand-responsive paratransit vehicles using agent-based simulation. Sustainability, 11(19), 5391.

Toth, P., and Vigo, D. (Eds.). (2002). The vehicle routing problem. Society for Industrial and Applied Mathematics.

V

Van Wee, B., and Geurs, K. (2011). Discussing equity and social exclusion in accessibility evaluations. European journal of transport and infrastructure research, 11(4).

Vaughan, R. (1986). Optimum polar networks for an urban bus system with a many-to-many travel demand. Transportation Research Part B: Methodological, 20(3), 215-224.

W

Walker, J. (2012). Ridership or Coverage? The Challenge of Service Allocation. In Human Transit (pp. 117-134). Island Press, Washington, DC.

Wang, S., Correia, G. H. D. A., and Lin, H. X. (2019). Exploring the performance of different on-demand transit services provided by a fleet of shared automated vehicles: An agent-based model. Journal of Advanced Transportation, 2019.

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer Based Modeling. Northwestern University, Evanston, IL. In: http://ccl.northwestern.edu/netlogo/.

Winter, K., Cats, O., Correia, G. H. D. A., and Van Arem, B. (2016). Designing an automated demand-responsive transport system: Fleet size and performance analysis for a campus–train station service. Transportation Research Record, 2542(1), 75-83.

X

Xiong, J., Guan, W., Song, L., Huang, A., and Shao, C. (2013). Optimal routing design of a community shuttle for metro stations. Journal of Transportation Engineering, 139(12), 1211-1223.

W

Wang, H. (2019). Routing and scheduling for a last-mile transportation system. Transportation Science, 53(1), 131-147.

Wang, S., Correia, G. H. D. A., and Lin, H. X. (2019). Exploring the performance of different on-demand transit services provided by a fleet of shared automated vehicles: An agent-based model. Journal of Advanced Transportation, 2019.

Wang, L., Zeng, L., Ma, W., and Guo, Y. (2021). Integrating Passenger Incentives to Optimize Routing for Demand-Responsive Customized Bus Systems. IEEE Access, 9, 21507-21521.

Wooldridge, M. (2009). An introduction to multiagent systems. John wiley & sons.

World Health Organization. (2018). Global Status Report on Road Safety 2018.

Wu, L., Gu, W., Fan, W., and Cassidy, M. J. (2020). Optimal design of transit networks fed by shared bikes. Transportation research part B: methodological, 131, 63-83.

Z

Zhang, J. (2021). Agent-based optimizing match between passenger demand and service supply for urban rail transit network with NetLogo. IEEE Access, 9, 32064-32080.

Zhang, T., Tian, W. X., Zhang, Y. J., and Liu, S. X. (2008). Improved ant colony system for VRPSPD with maximum distance constraint. Systems engineering-theory and Practice, 28(1), 132-140.

Zhang, L., Yang, H., Wu, D., and Wang, D. (2014). Solving a discrete multimodal transportation network design problem. Transportation Research Part C: Emerging Technologies, 49, 73-86.

Zhang, W., Jenelius, E., and Badia, H. (2019). Efficiency of semi-autonomous and fully autonomous bus services in trunk-and-branches networks. Journal of Advanced Transportation, 2019.

Zheng, Y., Li, W., and Qiu, F. (2018). A methodology for choosing between route deviation and point deviation policies for flexible transit services. Journal of Advanced Transportation, 2018.

Zhu, Z., Guo, X., Zeng, J., and Zhang, S. (2017). Route design model of feeder bus service for urban rail transit stations. Mathematical Problems in Engineering, 2017.

Annex

Here we present two additional papers produced during the doctoral research and focusing on the optimization (Annex A) or simulation (Annex B) of logistic operations.

Annex A presents a model for the solution of a Capacitated Vehicle Routing Problem (CVRP) using a novel Ant Colony Optimization (ACO) algorithm, developed and implemented in agent-based modelling environment, with an application to a real case study regarding a freight transport and logistics company in South Italy. This work resulted in the following journal article:

> *Calabrò, G., Torrisi, V., Inturri, G., and Ignaccolo, M. (2020). "Improving inbound logistic planning for large-scale real-world routing problems: a novel ant-colony simulation-based optimization". European Transport Research Review, vol. 12, no. 1, pp. 1-11.*

Annex B presents a new agent-based model to explore different scenarios of e-commerce urban deliveries, comparing home deliveries with consolidation-based strategies, reproducing operation under different demand patterns and including the possible matching of customer systematic trips and collection/delivery points with small detour from the scheduled trip. This work has led to the following conference proceeding:

> *Calabrò, G., Le Pira, M., Giuffrida, N., Fazio, M., Inturri, G., and Ignaccolo, M. (forthcoming). "Modelling the dynamics of fragmented vs. consolidated last-mile e-commerce deliveries via an agent-based model". Transportation Research Procedia (presented at 24th EURO Working Group on Transportation Meeting, EWGT 2021, 8-10 September 2021, Aveiro, Portugal).*

Calabrò et al. European Transport Research Review (2020) 12:21
https://doi.org/10.1186/s12544-020-00409-7

European Transport Research Review

ORIGINAL PAPER Open Access

Improving inbound logistic planning for large-scale real-world routing problems: a novel ant-colony simulation-based optimization

Giovanni Calabrò[1], Vincenza Torrisi[1*], Giuseppe Inturri[2] and Matteo Ignaccolo[1]

Abstract

This paper presents the first results of an agent-based model aimed at solving a Capacitated Vehicle Routing Problem (CVRP) for inbound logistics using a novel Ant Colony Optimization (ACO) algorithm, developed and implemented in the *NetLogo* multi-agent modelling environment. The proposed methodology has been applied to the case study of a freight transport and logistics company in South Italy in order to find an optimal set of routes able to transport palletized fruit and vegetables from different farms to the main depot, while minimizing the total distance travelled by trucks. Different scenarios have been analysed and compared with real data provided by the company, by using a set of key performance indicators including the load factor and the number of vehicles used. First results highlight the validity of the method to reduce cost and scheduling and provide useful suggestions for large-size operations of a freight transport service.

Keywords: Ant Colony optimization, Vehicle routing problem, Multi-agent simulation, logistics

1 Introduction

Logistics is the set of services and activities that allow goods to be carried from the place of origin in which they are available to the destinations where they are required. Transport helps to connect the sources of raw materials, production centres and markets, generating an increase in the value of goods sufficiently to justify the transport cost incurred. The first component of the logistics system is inbound logistic, which deals with the management of incoming materials, so it has to do with the purchases and supplies of raw materials, components or semi-finished products arriving from upstream suppliers of the logistics network. Among the activities of order management, collection, storage, internal handling and transport of goods, the latter often represents the main cost item. Therefore, a transport company that is able to provide an efficient and timely service achieves a competitive advantage in the increasingly competitive national and international markets. By improving route assignments to the vehicles of the fleet, it is possible to obtain significant time and cost savings.

However, even big companies often plan loading and distribution operations based on their empirical experience, without optimization methods able to minimize driving distance, avoid space waste inside the transport vehicles or at worse infeasible loading [1]. Since the relatively recent development of computer tools, a huge amount of scientific literature has been produced with the aim of optimizing delivery and/or pickup operations for a fleet of vehicles serving a set of customers and subject to side constraints. This gave rise to a whole class of problems sharing the generic name of *Vehicle Routing Problem* (VRP). The original version of the VRP was proposed by Dantzig and Ramser [8] under the definition of Truck

* Correspondence: vtorrisi@dica.unict.it; enza.torrisi@hotmail.it
[1]Department of Civil Engineering and Architecture, University of Catania, Catania, Italy
Full list of author information is available at the end of the article

Dispatching Problem, which dealt with the calculation of optimal routes for a fleet of trucks for petrol deliveries. This issue, in turn, may be considered as a generalization of the Traveling-Salesman Problem (TSP), consisting in finding the shorter route (or, in general terms, the lowest cost path) connecting all vertex of a graph, starting and finishing at a specified vertex after having visited each other vertex exactly once. Thanks to its numerous practical implications (especially in logistics but also in passenger transport), several variants of the basic problem have been put forward in recent years. One of the most studied members of the VRP family is the Capacitated Vehicle Routing Problem (CVRP), in which a fleet of identical vehicles has to be optimally routed from a central depot to supply a set of geographically dispersed customers with known demands [2]. Although CRVPs are not so "hard" to deal with as problems with pickups and deliveries and/or time-windows, when we deal with large-scale instances, it is fundamental to reduce the computational demand by acting both on the optimization algorithm and on the network topology, which is precisely the point on which this paper is focused.

The work is organized as follows: Section 1 introduced the research topic, highlighting the applicability of the Ant Colony Optimization metaheuristic to solve freight transport problems; Section 2 presents a brief literature review on VRP instances and their resolution approaches, with reference to the research contribution; Section 3 describes in detail the methodology adopted and the algorithm implemented in the multi-agent simulation environment; Section 4 presents the application of the model to a real case study; Section 5 shows and discusses the experimental results; finally, Section 6 concludes the work, providing some considerations for further research.

1.1 The use of ant Colony optimization to solve the VRP

It is well known that VRP, in its various specifications, is a non-determistic polynomial-time hard problem (NP-hard problem) which is not easily addressed with exact algorithms, since the computational time grows exponentially with problem size (with the increase in the scale of logistics and distribution this time would be extremely high). Therefore, a feasible option consists in formulating heuristic and metaheuristic algorithms, conceived so as to generate solutions that are as close as possible to the optimal one.

Ant Colony Optimization (ACO) algorithms are derived from an analogy with ants which lay a volatile substance called "pheromone" on their trail when foraging for food. In this family of metaheuristics, by extension, a certain number of simple artificial agents cooperate to build good solutions to hard combinatorial optimization problems via low-level based communications [9]. Iteration after iteration, more pheromone is deposited on the more frequented trails and this brings out a learning mechanism: when constructing a VRP solution, the probability of selecting a certain move is higher if this move has previously led to a better solution in previous iterations. Therefore, the "auto catalytic" nature of the process leads to the convergence towards good near-optimal solutions. A detailed explanation of the algorithm proposed in the present work will be provided in section 3.2.

In general, ACO is conceived to find the minimum cost paths within a network, so it presents several applications to routing and scheduling problems and is of particular interest in transport problems [4, 17]. Besides, thanks to its easy applicability to dynamic problems, where the topology of the characteristics of the network changes during the simulation, ACO algorithms are able to perform better than other metaheuristics. The excellent performances of ACO in solving such optimization problems are highlighted by the works of Catay [6] and Carabetti et al. [5], which applied the ACO approach to a series of benchmark problems finding results that were comparable and in some cases better than those available from the literature.

2 Literature review and research contribution

An extensive review on VRP instances exists [7, 21] and numerous variations of the basic problem in real-world applications have been addressed, including supply chain and freight transport issues [14], public transport [23], street cleaning, urban solid waste collection [3], school bus routing [12] and other instances.

Zhang et al. [26] investigate the reverse logistics vehicle routing problem with a single depot, simultaneous distribution and collection of the goods by a homogeneous fleet of vehicles under the restrictions of maximum capacities and maximum distance. They proposed an Ant Colony System (ACS) approach in which the vehicle residual loading capacity is introduced into the heuristic function to consider the dynamic fluctuation of vehicle load. Xiao et al. [22] extended the classical CVRP introducing the objective of minimizing fuel consumption, assumed as load dependent function, and using a simulated annealing algorithm to solve the problem. Lin et al. [15] addressed the recent trend of the environmental sensitivity in the supply chain management through a survey of green vehicle routing problems. Schneider et al. [18] introduced the electric VRP with time windows with the possibility for vehicles of recharging at any of the available stations, exploiting a hybrid heuristic that combines neighbourhood search and tabu search. Wang et al. [24] propose a modified ACO algorithm integrated with other savings algorithms in order to solve the CVRP allowing ants to go in and out the depots more than once until they have visited all customers, aiming at simplifying the procedure of constructing feasible solutions. Martin et al. [16] developed a multi-agent framework for scheduling and routing problems where agents use different metaheuristics and cooperate by sharing partial solutions during the search, giving rise to a reinforcement learning and pattern matching process. Hannan et al. [11] address the routing and scheduling optimization problem in waste collection by using a modified particle swarm optimization algorithm in a CVRP model, with the objective of minimizing travel distance, collected waste and tightness. Song et al. [19] propose a multi-objective approach to solve a CVRP with Time Windows and two-dimensional loading constraints, making use of mixed integer linear programming and a generalised variable neighbourhood search algorithm.

This paper contributes to the current literature by proposing a new agent-based modelling framework for the optimized planning of truck routes in large-scale inbound operations. This work provides a twofold contribution, methodological, since the ACO has been applied to a specific real transport network, allowing to vary and verify the incidence of some specific parameters related to truck freight (e.g. maximum working time, average speed, truck capacity, etc.) in each simulation, and operational, considering that the proposed model stands as a useful practical optimization tool able to support logistics operators in the route planning phase of their service.

3 Methodology

The problem addressed in the present work is to plan and design an optimal set of routes for the collection of goods through a new methodological approach. Its main components are described as follows:

- a single depot (D) where collection, groupage and distribution activities are centralized;
- a set of clients (farms) clustered in client nodes (N) spread on the study area, with a pick-up demand (P_j);
- a fleet of vehicles (V) with the same capacity (Q), able to collect goods from the farms to the depot;
- the road network and the set of possible links (i, j ∈ L) between different client nodes.

Starting from the assumption of no congestion on the road network (considering the suburban and rural context) and thus assuming a constant average speed for all vehicles, the objective of the optimization process is to minimize the Total Distance Travelled (TDT) by the vehicles of the fleet (minimization of the operational costs), while taking into account a work hours constraint. Decision variables, objective function and constraints will be explained in section 3.1. The optimal assignment of vehicles to routes is achieved by using an ACO algorithm that will be described in detail in section 3.2. Simulations have been carried out in NetLogo [25], a multi-agent programming and modelling environment which allows to model and simulate complex systems and allowing the visualization of their parameters in real time.

3.1 Model conceptualization

The agent-based model is structured on a double-layer network. The base layer spatially reconstructs the real road network, while the upper layer reproduces the directed graph of the possible connections (links) between the different client nodes and between the depot and the all the nodes.

The first step is the reconstruction of the road network in the NetLogo workspace and the localization of client nodes, to each of which an array of indivisible loading units from each farm belonging to the client-node itself is associated. This particular procedure has been designed to reduce the number of nodes and links of the graph by creating clusters of neighbouring farms since when the problem deals with hundreds of customers (large-scale VRP) it is computationally demanding and difficult to tackle (Fig. 1).

Starting from the base layer of the road network, the upper layer is created by connecting each client node with the depot and with a certain number of other client nodes. This number varies from node to node once a maximum road distance (*d-max*) to other nodes and a minimum number of links (*n-links*) to the nearest nodes are selected. The road distance (d_{ij}) is an attribute of the link connecting two client nodes; it is calculated for each link through a shortest path algorithm before the simulation starts.

The CVRP model formulated as follows:

$$\text{Minimize } TDT = \sum_{i\in N}\sum_{j\in N}\sum_{v\in V} d_{ij}\cdot x_{ij,v} \tag{1}$$

$$\text{Subject to } 0 \le \sum_{i\in N}\sum_{j\in N} P_j\cdot x_{ij,v} \le Q \qquad v\in V \tag{2}$$

$$\sum_{i\in N}\sum_{j\in N}\left(t_{ij} + n_f\cdot t_s\right)\cdot x_{ij,v} \le TT_{max} \qquad v\in V \tag{3}$$

Eq. (1) is the objective function which minimizes the TDT, $x_{ij,v}$ is a binary variable equal to 1, if vehicle v travels along the link *(i, j)*, or 0, otherwise. Eq. (2) ensures that the pick-up load of each vehicle never exceeds its capacity. Eq. (3) imposes a maximum travel time TT_{max} (work hours constraints, generally equal to 8 h) for each vehicle, defining t_s as the service-time that the

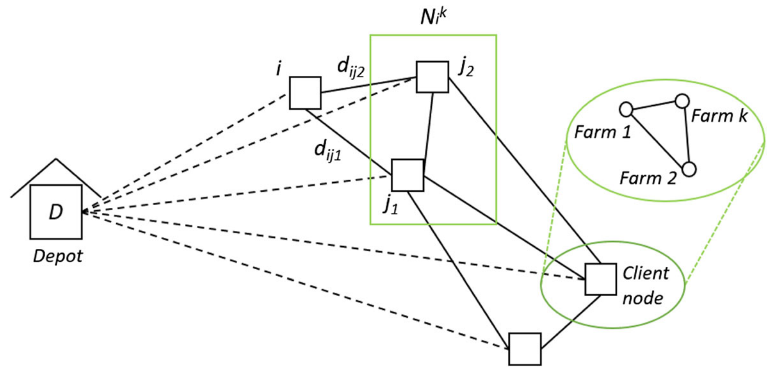


Fig. 1 "Conceptual" schematization of the depot, client nodes and possible connections

vehicle takes in each farm to for loading operations (about 15 min), t_{ij} as the travel time along the link *(i, j)* and n_f as the number of farms of the client-node *j* served by vehicle *v*.

It should be specified that each farm must be visited only once (indivisibility of pickup loads), considering a constant service-time, and each vehicle has a certain capacity that cannot be exceeded considering any pickup operations (however, since a client node could have an overall demand exceeding the vehicle capacity, it may be served by more than one vehicle of the fleet, as opposed to classical VRP instances).

3.2 The ant Colony optimization algorithm

It is well known that CVRPs are NP-hard problems in the field of operations research. Moreover, the proposed model aims at addressing large-scale instances, where the residual orders of client nodes changes dynamically during the simulation, so they are practically impossible to solve using exact methods. In last decades, a huge number of heuristic procedures have been developed in order to find good suboptimal solutions with acceptable computational efforts. Among them, metaheuristics take inspiration from natural optimization mechanisms, translating them into specific algorithms. In particular, ACO algorithms [10] derives from the social behaviour of some ant species which are capable to find the shortest paths between their nest and a food source. This ability arises because ants can exploit a sort of communication based only on pheromone trails, a volatile chemical substance deposited on the ground by ants. Artificial ant colonies, despite being very simple organisms, can form systems able to perform highly complex tasks and jointly solve optimization problems by dynamically interacting with each other.

The algorithm implemented in the present model derives from MAX-MIN Ant System [20], which improved the first member of the ACO family, named Ant System, originally applied to the resolution of the Travelling Salesman Problem. Simulations are based on an iterative optimization process that ends after a given number of generations (*g)* of a chosen number of colonies (*m)* made of a specified number of ants. This process leads to the quality improvement of the final solution comes from the comprehensive exploitation of three different information components, iteration after iteration: simulated artificial ants build routes by considering a) the pheromone trail, b) the "visibility" and c) the residual capacity. The first component is updated for each link when a new generation *g* of colonies is launched. The last two component are included in the heuristic function, which structure is shown in Eq. (4). The visibility is given by the reciprocal of the distance related to a link and represents the fixed information available a priori. As concern the residual capacity, when ants explore their neighbourhood (client nodes linked with the actual node), the feasible combination of orders from the farms belonging to the next client node has to be recalculated every time. So, if a client node consists of *n* farms, each one with a given pick-up demand (number of loading units), ant *k* at iteration *t* investigates the combination of all pick-up demands $p_{j,h}$ that can be served without exceeding the residual loading capacity (orders combination list).

$$\eta_{ij}(t) = \frac{1}{d_{ij}} \cdot \sum_{h=1}^{n} p_{j,h}(t) \tag{4}$$

Where d_{ij} is the road distance between node *i* and node *j* and $p_{j,h}$ *(t)* is $p_{j,h}$ if it belongs to the orders combination list, zero otherwise. However, there are situations in which, although the remaining capacity makes possible to satisfy another pick-up demand, it may be preferable to return to the depot in order not to further lengthen the distances covered. In consequence, we took into account a heuristic function related to the links from client node *i* to the depot *D*, given by the reciprocal of the product between the distance d_{iD} and the residual loading capacity at iteration t. Since the total numbers of orders, represented by standard loading unit, is usually much higher than the single vehicle capacity, multiple routes must be found, each one served by one vehicle. So, every solution must be built by an ant colony, able to jointly minimize the total distance travelled without violating the constraints of maximum capacity and maximum working time. Step by step, each ant of the colony applies a random proportional rule to decide the next farm to go. Therefore, the probability with which ant k, currently at farm i, chooses to go to farm j is given by Eq. (5):

$$p_{ij}^{k}(t) = \frac{\left[\tau_{ij}(g)\right]^{\alpha} \cdot \left[\eta_{ij}(t)\right]^{\beta}}{\sum_{l \in N_i^k} \left[\tau_{il}(g)\right]^{\alpha} \cdot \left[\eta_{il}(t)\right]^{\beta}} \quad \text{if } j \in N_i^k \tag{5}$$

where and α and β are calibration parameters that control the relative importance of the pheromone trail τ_{ij} versus the heuristic information η_{ij}, N_i^k is the feasible neighbourhood.

When an ant of the colony reaches the loading capacity (or its feasible neighbourhood is empty) it comes back to the depot and the next ant of the same colony starts its tour. When pick-up demands are all satisfied, the following colony of ants is allowed to explore other possible solutions, until the given number m of colonies is reached.

Once all the m colonies have found their solution, only the "best" colony (i.e. the one that founds the solution that minimises the total distance travelled) is allowed to reinforce the pheromone trail, to better exploit the best of the m solutions found, using the following updating rule (Eq. (6)):

$$\tau_{ij}(g+1) = (1-\rho) \cdot \tau_{ij}(g) + \Delta\tau_{ij}^{best}(g) \tag{6}$$

where ρ is the evaporation rate, ranging from 0 to 1, $\Delta\tau_{ij}^{best}$ is the amount of pheromone deposited on link *(i, j)* used by the best colony at generation g, which is given by Eq. (7):

$$\Delta\tau_{ij}^{best}(g) = Q \cdot \left(\frac{E^{best}(g)}{E^{global-best}}\right)^2 \tag{7}$$

E represents the value of the objective function (i.e. the reciprocal of the total distance travelled by vehicles) and Q is the diffusion rate, which is greater than zero. Finally, when the maximum number of generations is reached, the simulation stops and outputs the results. The whole process described so far is showed through a flow chart in Fig. 2.

4 Case study

4.1 Territorial framework

The described methodology is applied to the case study of *Gali Group*, a freight transport and logistics company, located in Ispica (Sicily), on the eastern end of the province of Ragusa, bounding Siracusa's district. Ispica is 33 km from Ragusa (Fig. 3); it has an area of about 110 km^2 with a population density of 143,54 inhabitants/ km^2 [13]. Its economy is primarily agricultural boasting major outputs of early fruit, tomato, vegetables and carob – for which Ispica is Italian's biggest producer and exporter. Industry has developed in recent decades, particularly the agriculture-related businesses. Thus, the main industrial activities are those involved in processing and marketing the agricultural products.

In this context, the *Gali Group* company is recognized as a landmark for the activity of pick-up and delivery of horticultural products from Sicily to the central-northern Italy and in some cases also abroad. Its activity is based on road transport. The company offers the possibility to request the pick-up order by clients up to 5.00 p.m.; and only after this time the pick-up activity is carried out. This way of working arises from the concept that logistics operators can decide the routes when they have an almost complete awareness of the orders. This clearly affects the subsequent phases of the logistics process. Hence, this work analyses the upstream of the process by addressing to the pick-up procedure, in order to provide an optimized route planning in terms of times and costs. The study area is represented by the catchment area of *Gali Group*, as shown in Fig. 3 (on the right side) and the data analysis and the algorithm implementation are provided below.

4.2 Data analysis and algorithm implementation

Data analysis is referred to a period between May 2018 and March 2019. The initial basis of data for this study is essentially represented by the analysis of 3 days with a maximum flow of goods. For all these days, incoming orders have been collected, with the following information:

- number, code and time of arrival of the order;
- name and pick-up zone of the customer/provider company corresponding to each order;
- number and type of loading units.

On average, about 90 orders have been registered for each day with more than 1400 loading units. In addition, the operating program of the logistics operators concerning the procedure of pick-up of goods has been recorded for each day, i.e. total distance travelled (TDT), number of vehicles (NV) and load factor (LF).

Once the study area has been identified, coinciding with the catchment area of *Gali Group* characterized by the various provider companies, the first step to start

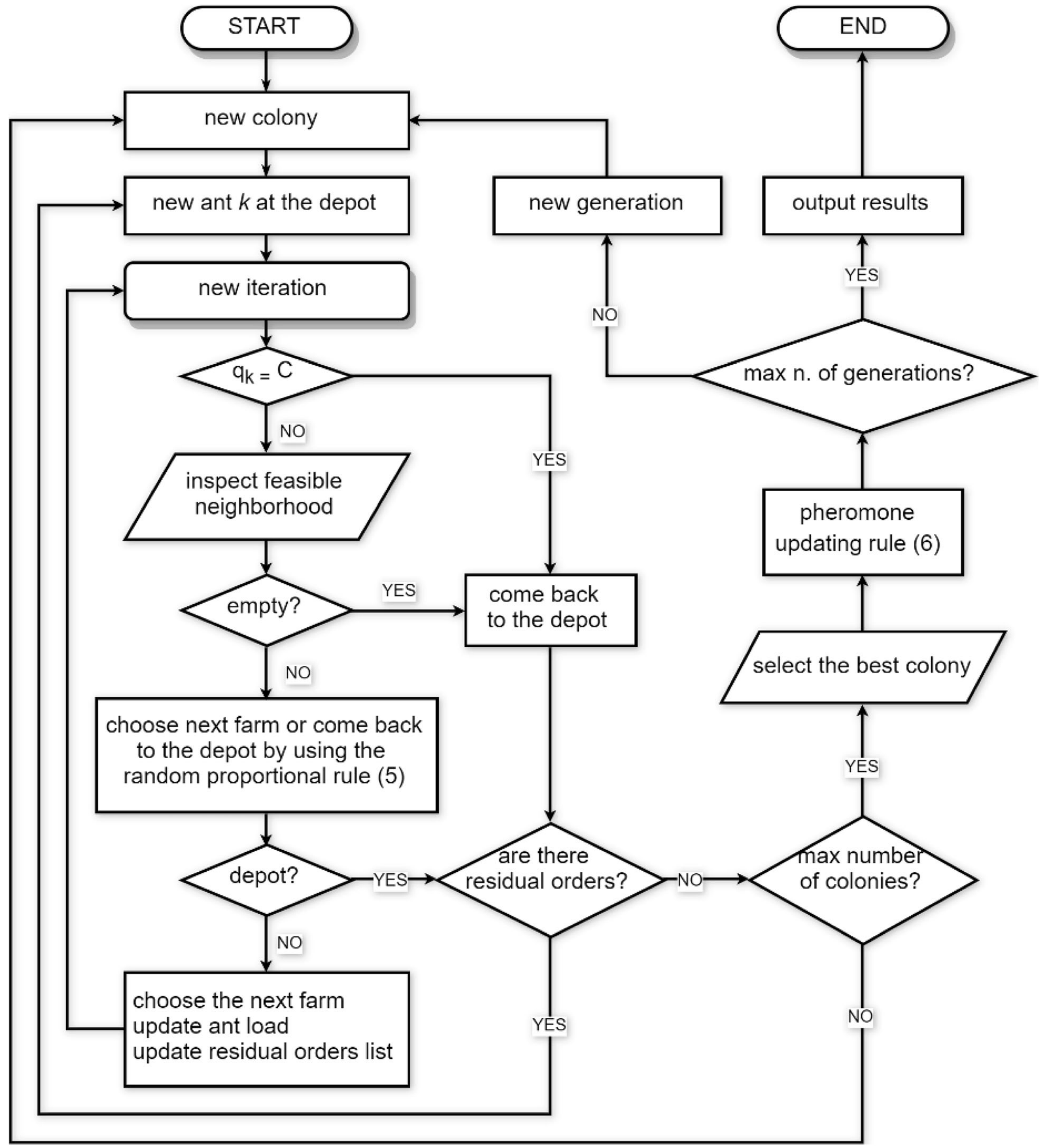


Fig. 2 Flow chart of the ACO algorithm

simulations in NetLogo consisted in the definition and construction of the road network graph. It has been drawn using OpenStreetMap as a basis and it is characterized by a double-layer network (as stated in section 3.1). Figure 4 shows the reconstruction of the real road network (on the left side) and the directed graph of the possible connections between the different client nodes and between the depot and all the nodes (on the right side), reproduced by links.

In the proposed model, the clients have been considered by creating clusters of neighbouring farms clustered in client nodes, with the corresponding number of loading units organized in an array for each node. Therefore, 60 client nodes or farms have been positioned.

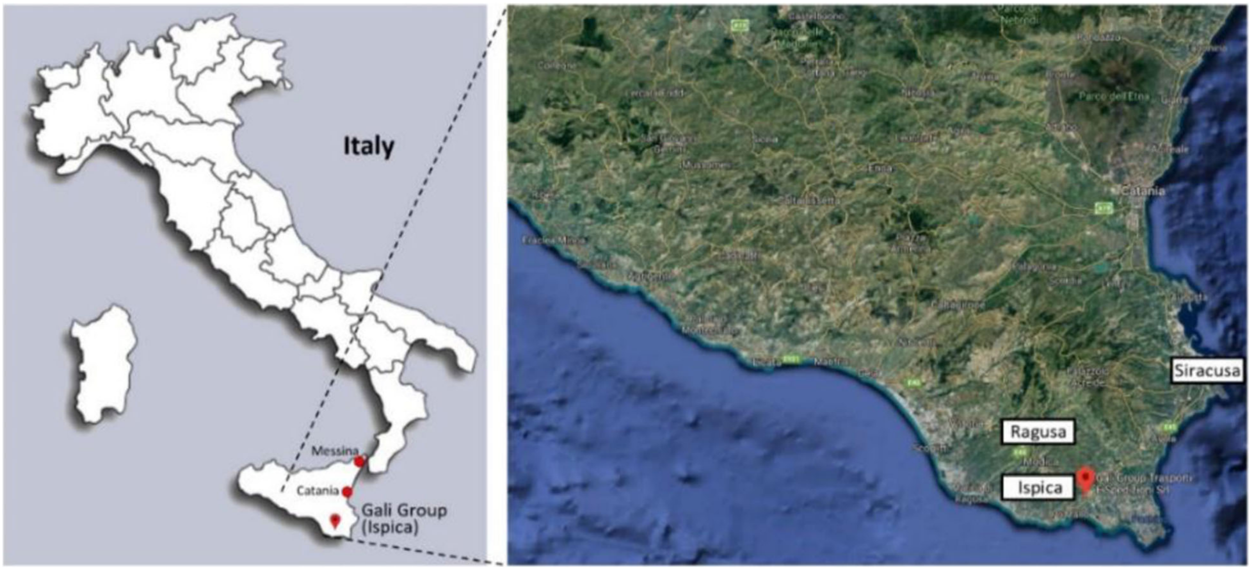


Fig. 3 Satellite Images of the study area (https://satellites.pro/)

5 Results and discussion

Several simulations have been performed to test the model and results are shown in Table 1. Since the ants apply to each node a probability-based choice criterion, an initial level of pheromone concentration is assigned to each link of the network. Moreover, the maximum number of generations g_{max}, the number of colonies for each generation m, the vehicles capacity C (i.e. maximum number of units), the diffusion rate Q, the evaporation rate ρ, exponents α and β have been fixed as input parameters. They have been chosen after several tests which resulted in better computational times and model outcomes. The only two variable parameters have been constituted by the maximum road distance *d-max* to other client nodes and the minimum number of links *n-links* to the nearest nodes. Three combinations of *d-max* and *n-links* have been considered. The first one, 10–5 is characterized by a low number of connections and from the nearest ones considering the reduced radius. In the second combination 50–10, both the radius and the number of minimum connections have been increased. The third one, 50–30 has an equal radius value while the number of minimum connections is increased.

The results of these three sets of simulations highlight that the shorter distances have been obtained for the second combination 50–10. This proves that an increase in exploration possibilities does not correspond to a better solution found by the algorithm (in this case a decrease of TDT).

To demonstrate the effectiveness of the model, Fig. 5 shows the convergence curve of the objective function obtained in one of the simulations for the day 1. The x-coordinate denotes the number of generations and the y-coordinate denotes the corresponding TDT. It can be

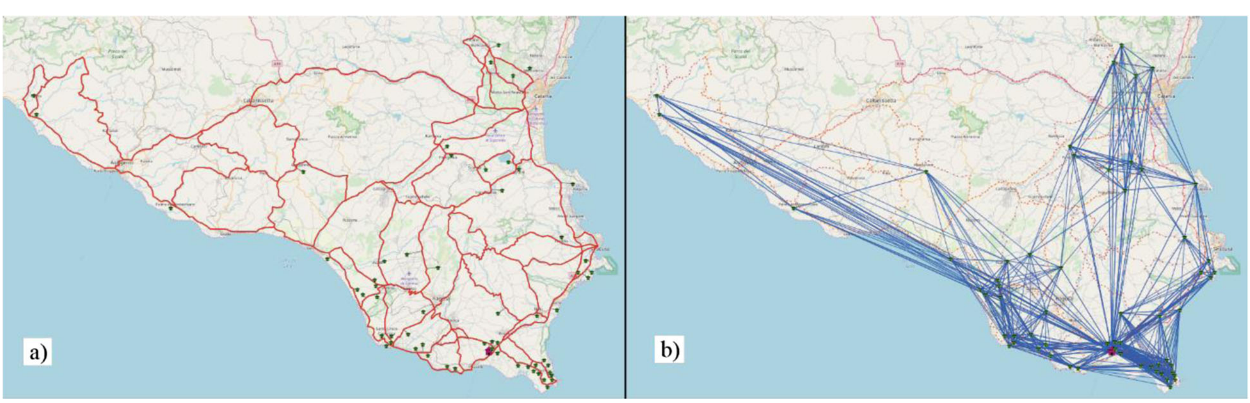


Fig. 4 Road network graph implemented in NetLogo; **a** Base layer; **b** Upper layer

Table 1 Input parameters set and simulations results (day 1)

N. Sim	d-max	n-links	g_{max}	m	C	Initial ph. val.	Q	ρ	α	β	TDT [km]	LF	NV	Avg. TDT
1	10	5	150	25	33	2.0	0.1	0.05	2.0	1.0	4951.8	0.922	39	
2											4919.3	0.922	39	
3											5018.5	0.897	40	4976
4											5006.7	0.926	39	
5											4983.1	0.903	40	
1	**50**	**10**	150	25	33	2.0	0.1	0.05	2.0	1.0	4857.5	0.923	39	
2											4867.1	0.923	39	
3											4901.2	0.927	39	**4880**
4											4846.7	0.959	38	
5											4928.5	0.920	39	
1	50	30	150	25	33	2.0	0.1	0.05	2.0	1.0	4920.8	0.952	38	
2											4956.7	0.945	38	
3											4912.6	0.954	39	**4915**
4											4866.3	0.935	38	
5											4916.4	0.958	39	

seen that the convergence TDT within 150 iterations is increasing almost constantly. It remains fixed at a value of approximately 4900 between 95 and 130 iteration, and finally converge to the best solution.

For each analysed day, considering the second combination of *d-max* and *n-links* (50–10), several simulations have been performed. Scheduled data (provided by the company) with reference to the total number of travelled kilometers (calculated considering the optimal minimum path for the trip of each vehicle), the number of vehicles and the load factors have been taken into consideration to make a comparison with data derived from simulations. Table 2 shows the aggregated results of simulations deriving from the elaboration of collected data during the three analysed days. It is noticeable that the total number of travelled kilometres (TDT) deriving from the simulation is significantly lower than the scheduled one provided by the company. This outcome is more emphasized in the case of the second and third days, during which the scheduled number of travelled km and vehicles is greater. Moreover, the load factors of simulated vehicles are higher than the scheduled ones, consequently leading to a fewer number of vehicles to carry out the procedure of pick-up of goods. Only the Average Distance Travelled (ADT) by vehicles is higher than the scheduled one for "Day 1", but this outcome is due to the minor number of vehicles resulting from simulations.

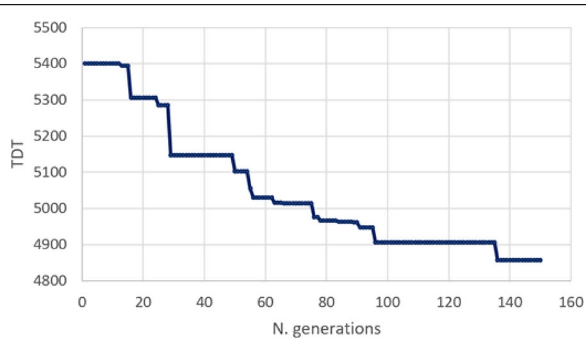


Fig. 5 Convergence curve of the objective function

These findings are much more evident from a graphical point of view in Fig. 6 which show the comparison between the daily programme of the logistics company and data obtained through some simulations, related to the number of travelled kilometres and the load factor for each vehicle and for all days.

For most vehicles, the number of travelled kilometres deriving from simulations is consistent with the scheduled one. Then, despite for the latest vehicles this number appears higher in the case of simulation (Fig. 6 on the left side). This is largely justified by the fact that the load factor of vehicles is higher (Fig. 6 on the right side).

Figure 7 shows the aggregated results for all 3 days. Comparing the average results obtained from the simulations and scheduled data, it is demonstrated how the model can optimize the routes for the collection of goods. This optimization is configured not only in a reduction of the travelled kilometres and a higher load factor of vehicles (as stated before), but also in a lower simulated number of vehicles (e.g. vehicles' number = 47 for the second day) compared to the one scheduled (i.e. vehicles' number = 60 for the second day) (see Fig. 7 on the right).

Table 2 Aggregated results of simulations for the analysed days

	Day 1		Day 2		Day 3	
	Simulated	Scheduled	Simulated	Scheduled	Simulated	Scheduled
TDT [km]	4880	5280	5340	7952	5690	7380
LF	0.930	0.748	0.962	0.796	0.959	0.789
NV	39	44	45	60	48	58
ADT [km]	125.1	120.0	118.7	132.5	118.5	127.2
Min TDT [km]	4846.7		5287.9		5954.0	

As seen, the implementation of the ACO in NetLogo, besides representing an optimization tool, it allows the logistics company to optimize resources through the model results of routing, in terms of travelled km, load factor and number of vehicles. Moreover, the proposed study has an important advantage represented by the fact that NetLogo gives the possibility of graphically representing networks and displaying the best routes, configuring NetLogo as an operational tool for the company from a practical point of view. Furthermore, the

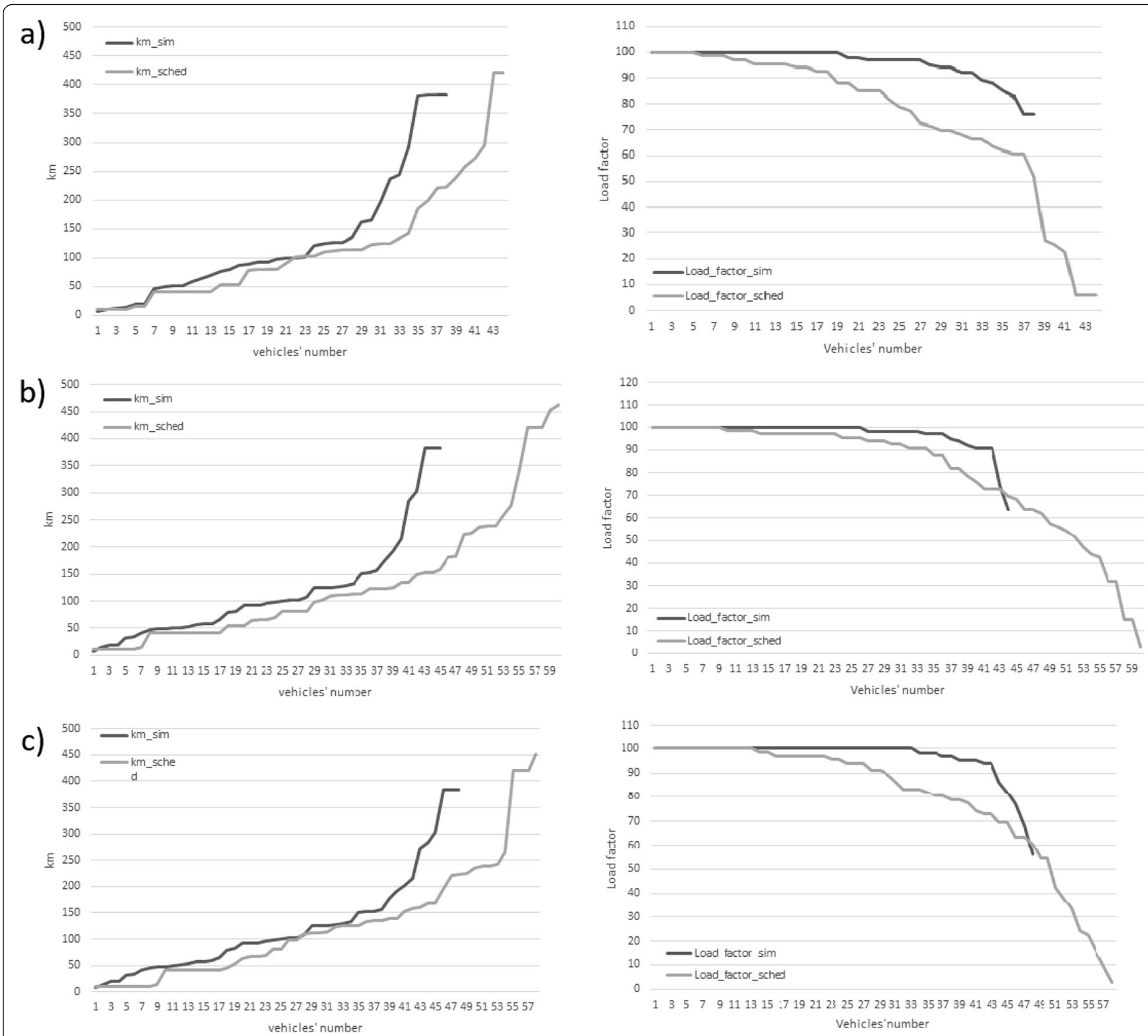


Fig. 6 Scheduled and simulated travelled kilometres and load factors; **a** first day; **b** second day; **c** third day

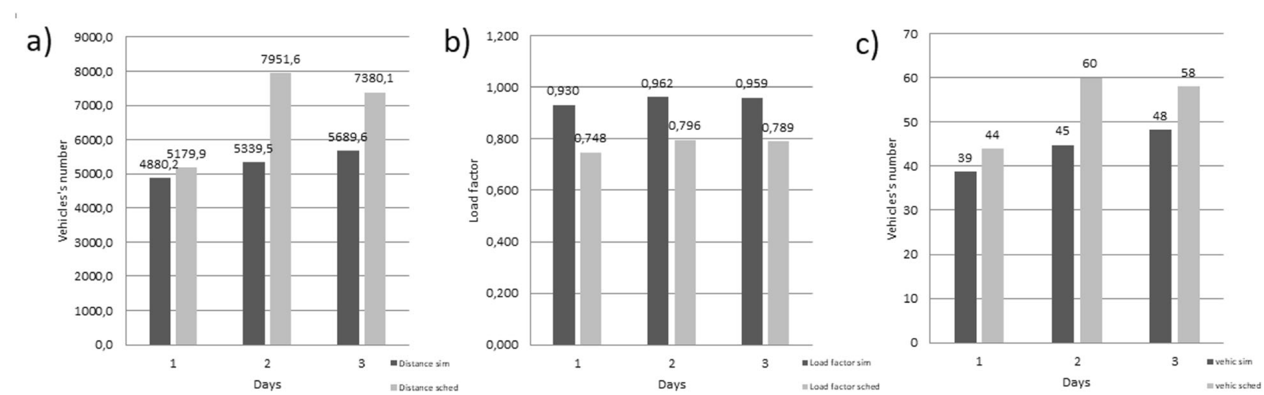


Fig. 7 Comparison between simulated and scheduled (**a**) total travel distance and (**b**) load factors (**c**) number of vehicles

element of absolute originality of the work consists in having applied the multi-agent modelling environment NetLogo to solve optimization problems in the planning of transport operations. The simulations based on parameters of Table 1, after an initial calibration, could be replicated in other contexts, allowing benchmark analyses to be performed between different contexts. Currently, the model does not consider the congestion level of each link belonging to the transport network, but the computations are based on the travelled distance. Furthermore, the time-wasters due to the pick-up of goods at each farm are considered in an equivalent manner in the probabilistic choice made by the model in each iteration (i.e. when the number of array components of each pick-up zone varies). This is because more farms are clustered in a single pick-up zone to simplify the model and to decrease computational demand. These aspects could be further investigated in future research.

6 Conclusions

This study proposes a new methodological approach of CVRP to optimize inbound logistics in a large-scale problem. The model has been tested using input data provided by the logistics company *Gali Group*, located in the Sicilia region in southern of Italy. Data have been acquired for a significant period and 3 days with a maximum flow of goods have been considered for comparisons. From the analysis of data, the catchment area of *Gali Group* has been identified and the road network graph has been constructed in NetLogo to start simulations. The CVRP using ACO algorithm has been implemented for the identification of an optimal set of routes for the collection of goods, by using an objective function coinciding with the travel distance and a maximum working time of 8 h/day as a constrain. In this way, the proposed model is able to support the logistics operators during the route planning phase, optimizing the operations related to their service: in fact, comparing the obtained results from simulations and the scheduled ones, it is evident a significant reduction in the traveled distances by vehicles, as well as in the number of vehicles itself, compared to those planned by the company, with a corresponding higher load factor. Therefore, this study lays the basis for a deeper analysis in order to investigate the logistics process in its overall perspective. Further analysis will be carried out in future research, in order to obtain more information on the operation of the logistics company through interviews to planners and drivers (e.g. departure and arrival times of vehicles from-to the depot; groupage and delivery processes), identifying algorithm improvements (although in the initial phase, the model is providing very interesting results) and paving the way for a well-thought-out decision support service of an optimized logistics freight transport.

Acknowledgements
Our sincere thanks are expressed to those who participated in the study. This work has been partially financed by the University of Catania within the project "Piano della Ricerca Dipartimentale 2016-2018" of the Department of Civil Engineering and Architecture and the project "Piano per la Ricerca 2016-2018 - Linea di intervento 2" of the Department of Electric, Electronic and Computer Engineering. This study was also supported by the MIUR (Ministry of Education, Universities and Research [Italy]) through a project entitled WEAKI TRANSIT: WEAK-demand areas Innovative TRANsport Shared services for Italian Towns (Project code: 20174ARRHT; CUP Code: E44I17000050001), financed with the PRIN 2017 (Research Projects of National Relevance) programme. We authorize the MIUR to reproduce and distribute reprints for Governmental purposes, notwithstanding any copyright notations thereon. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the MIUR.

Authors' contributions
VT and GC designed the methodological approach. GC implemented the model. VT ran all computations and analyses. GI and MI participated in the literature review and interpretation of the results. The author(s) read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
This study was approved by the Ryerson University Research Ethics Board, No. 2014–250-2.

Competing interests
The authors declared that they have no competing interests.

Author details
[1]Department of Civil Engineering and Architecture, University of Catania, Catania, Italy. [2]Department of Electrical electronic and computer engineering, University of Catania, Catania, Italy.

Received: 11 July 2019 Accepted: 8 March 2020
Published online: 14 April 2020

References

1. Aprile, D., Egeblad, J., Aravelli, A. C., Pisinger, D., & Lisi, S. (2007). Logistics optimization: Vehicle routing with loading constraints. In *ICPR –19, the development of collaborative production and Service Systems in Emergent Economies, 19th international conference on production research, Valparaiso, CL, Jul 29 - Aug 2, 2007*.
2. Baldacci, R., Mingozzi, A., & Roberti, R. (2012). Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints. *European Journal of Operational Research, 218*(1), 1–6.
3. Benjamin, A. M., & Beasley, J. E. (2010). Metaheuristics for the waste collection vehicle routing problem with time windows, driver rest period and multiple disposal facilities. *Computers & Operations Research, 37*(12), 2270–2280.
4. Calabrò, G., Inturri, G., Le Pira, M., Pluchino, A., & Ignaccolo, M. (2020). Bridging the gap between weak-demand areas and public transport using an ant-colony simulation-based optimization. *Transportation Research Procedia*, *45*, 234–241.
5. Carabetti, E. G., de Souza, S. R., Fraga, M. C. P., & Gama, P. H. A. (2010). An application of the ant colony system metaheuristic to the vehicle routing problem with pickup and delivery and time windows. In *2010 eleventh Brazilian symposium on neural networks*. 2010 (pp. 176–181). Sao Paulo: IEEE.
6. Catay, B. (2009). Ant colony optimization and its application to the vehicle routing problem with pickups and deliveries. In R. Chiong & S. Dhakal (Eds.), *Natural intelligence for scheduling, Planning and packing problems* (pp. 219–244). Berlin Heidelberg: Springer.
7. Cordeau, J. F., Laporte, G., Savelsbergh, M. W., & Vigo, D. (2007). *Vehicle routing. Handbooks in operations research and management science* (Vol. 14, pp. 367–428).
8. Dantzig, G. B., & Ramser, R. H. (1959). The truck dispatching problem. *Management Science, 6*, 80–91.
9. Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on evolutionary computation, 1*(1), 53–66.
10. Dorigo, M., & Stützle, T. (2003). The ant Colony optimization Metaheuristic: Algorithms, Applications, and Advances. https://doi.org/10.1007/0-306-48056-5_9.
11. Hannan, M. A., Akhtar, M., Begum, R. A., Basri, H., Hussain, A., & Scavino, E. (2018). Capacitated vehicle-routing problem model for scheduled solid waste collection and route optimization using PSO algorithm. *Waste management, 71*, 31–41 https://doi.org/10.1016/j.wasman.2017.10.019.
12. Huo, L., Yan, G., Fan, B., Wang, H., & Gao, W. (2014). School bus routing problem based on ant colony optimization algorithm. In *IEEE transportation electrification conference and expo, ITEC Asia-Pacific 2014 - conference proceedings* https://doi.org/10.1109/ITEC-AP.2014.6940973.
13. ISTAT (2017). Italian National Institute of Statistics, Population data. http://demo.istat.it/pop2017/index.html
14. Lee, Y. H., Jung, J. W., & Lee, K. M. (2006). Vehicle routing scheduling for cross-docking in the supply chain. *Computers & Industrial Engineering, 51*(2), 247–256.
15. Lin, C., Choy, K. L., Ho, G. T., Chung, S. H., & Lam, H. Y. (2014). Survey of green vehicle routing problem: Past and future trends. *Expert systems with applications, 41*(4), 1118–1138.
16. Martin, S., Ouelhadj, D., Beullens, P., Ozcan, E., Juan, A. A., & Burke, E. K. (2016). A multi-agent based cooperative approach to scheduling and routing. *European Journal of Operational Research, 254*(1), 169–178.
17. Neogi, A. G., Mounika, S., Kalyani, S., & SA, Y. (2018). A comprehensive study of vehicle routing problem with time windows using ant Colony optimization techniques. *International Journal of Engineering & Technology, 7*(2.32), 80–85.
18. Schneider, M., Stenger, A., & Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science, 48*(4), 500–520.
19. Song, X., Jones, D., Asgari, N., & Pigden, T. (2019). Multi-objective vehicle routing and loading with time window constraints: A real-life application-routing and loading. *Annals of Operations Research*, 1–27. https://doi.org/10.1007/s10479-019-03205-2.
20. Stützle, T., & Hoos, H. H. (2000). MAX–MIN ant system. *Future Generation Computer Systems, 16*(8), 889–914.
21. Toth, P., & Vigo, D. (2014). Vehicle Routing: Problems, Methods, and Applications. In *Society for Industrial and Applied Mathematics and the Mathematical Optimization Society* https://doi.org/doi:10.1137/1.9781611973594.
22. Xiao, Y., Zhao, Q., Kaku, I., & Xu, Y. (2012). Development of a fuel consumption optimization model for the capacitated vehicle routing problem. *Computers & Operations Research, 39*(7), 1419–1431.
23. Xiong, J., Guan, W., Song, L., Huang, A., & Shao, C. (2013). Optimal routing design of a community shuttle for metro stations. *Journal of Transportation Engineering, 139*(12), 1211–1223.
24. Wang, X., Choi, T. M., Liu, H., & Yue, X. (2016). Novel ant colony optimization methods for simplifying solution construction in vehicle routing problems. *IEEE Transactions on Intelligent Transportation Systems, 17*(11), 3132–3141.
25. Wilensky, U. (1999). *NetLogo. Center for Connected Learning and Computer Based Modeling*. Evanston: Northwestern University http://ccl.northwestern.edu/netlogo/.
26. Zhang, T., Tian, W. X., Zhang, Y. J., & Liu, S. X. (2008). Improved ant Colony system for VRPSPD with maximum distance constraint. *Systems Engineering - Theory & Practice* https://doi.org/10.1016/s1874-8651(09)60008-9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

24th EURO Working Group on Transportation Meeting, EWGT 2021, 8-10 September 2021, Aveiro, Portugal

Modelling the dynamics of fragmented vs. consolidated last-mile e-commerce deliveries via an agent-based model

Giovanni Calabrò[a]*, Michela Le Pira[a], Nadia Giuffrida[b], Martina Fazio[c], Giuseppe Inturri[d], Matteo Ignaccolo[a]

[a]*Department of Civil Engineering and Architecture, University of Catania, Via S. Sofia 64, Catania, 95125, Italy*
[b]*Spatial Dynamics Lab, University College Dublin, UCD Richview Campus, D04 V1W8, Belfield, Dublin, Ireland*
[c]*Department of Physics and Astronomy, University of Catania, Via S. Sofia 64, Catania, 95125, Italy*
[d]*Department of Electric, Electronic and Computer Science Engineering, University of Catania, Via S. Sofia 64, Catania, 95125, Italy*

Abstract

This paper proposes a new agent-based model (ABM) to explore different scenarios of e-commerce urban deliveries, comparing door-to-door deliveries with consolidation-based strategies. The ABM reproduces operation under different demand patterns and include the possible matching of customer systematic trips and collection/delivery points with small detour from the scheduled trip. Several variables of the model can be changed in a parametric simulation environment, allowing to infer the level of convenience of consolidation strategies for the different actors involved. The model provides indicators able to take into account customer and logistics operator perspectives, and the impact of the service on the community. Results can give useful information to understand how to manage growing on-demand urban deliveries and to measure the impact of freight transport on city sustainability.

Peer-review under responsibility of the scientific committee of the 24th EURO Working Group on Transportation Meeting.
Keywords: parcel deliveries; e-commerce; consolidation; parcel lockers; city logistics

1. Introduction

E-commerce is a constantly growing phenomenon, becoming an increasingly common buyers' practice, also pushed by digital technologies spread. The COVID-19 outbreak has contributed to a deep change of users' habits, bringing people ever closer to the world of online shopping and increasing the related demand.

Fragmented door-to-door deliveries may not be sufficient to satisfy this rising demand nor sustainable as they can lead to different transport and delivery issues such as missed deliveries, low load factors of vehicles, increase in travel distances and related GHG emissions (Lachapelle et al., 2018). Moreover, e-commerce-related deliveries generate additional traffic in cities that hardly compensate the reduction in individual shopping trips (World Economic Forum, 2020).

This is becoming one of the big concerns of policy-makers engaged in preventing the negative impact of freight urban deliveries (Le Pira et al., 2017; Allen et al. 2018). Besides, customers can count on a faster and more reliable service compared to the recent past. In this respect, same-day and instant deliveries are two fast-growing logistics segments, although they are still not the most diffused. Deliveries should be arranged by considering demanding customers, city constraints and last-mile logistics costs, which account for most of the supply chain costs (Gevaers et al., 2011). Innovative logistics solutions could be explored to this purpose, by appropriately taking into account: (i) the impact of logistics activities on city sustainability, e.g. congestion and emissions in urban areas; (ii) the quality of service for customers, e.g. in terms of the time elapsing between the order/purchasing and the parcel delivery (lead

* Corresponding author. Tel.: +39-095-7382211
E-mail address: giovanni.calabro@unict.it

Peer-review under responsibility of the scientific committee of the 24th EURO Working Group on Transportation Meeting.

time) that should be minimized; (iii) logistics costs, e.g. the number of vehicles needed for the deliveries, which depends on the demand and the possibility of consolidation of goods, expressed in terms of average load factor of parcel couriers vehicles and vehicle kilometres travelled (VKT).

In this paper, we present a new agent-based model (ABM) built to compare the performance of two parcel delivery strategies, namely Home Delivery and Collection-and-Delivery Point (CDP), by varying customer demand patterns, vehicle fleet capacity and the spatial density of CDPs. The first option is the most common in last-mile delivery, even if it the logistics operator must manage a fragmented delivery process, a more complex routing and higher VKT (Iwan et al., 2016). The second option includes selected locations (e.g. parcel lockers, shops) where customers can pick-up their parcels. This allows consolidating parcels, solving the expensive problem of failed deliveries and potentially reducing the impact of freight deliveries on urban traffic (Lachapelle et al., 2018). The potential of CDPs is usually related to their closeness to origin/destination of customer trips, mainly activities, services and residences. The novelty of the ABM proposed here is the simultaneous dynamic simulation of customer and freight movements with the possibility for customers to pick up their parcels in CDP along their daily trip path, considering different trip purposes. Combining passenger mobility and last mile freight deliveries through a consolidation-based approach has the potential of increasing the overall efficiency of the transport, thus turning in to one of the most promising strategies for sustainable freight transport planning. Next section will frame our research in the context of the recent research endeavours on this topic.

2. Fragmented vs. consolidated deliveries

Literature on this topic has recently started to emerge. Schnieder and West (2020) introduced the concept of Time-Area requirements, i.e. the product of the space usage and the time needed for the delivery process (including both couriers and customers movements), to evaluate the two delivery options, deriving some useful insights for policy-makers to reduce the externalities of urban deliveries. Mitrea et al. (2020) investigated the feasibility of parcel lockers and the attitude of users towards this alternative delivery solution. Authors collected useful information through a survey about the willingness to use parcel lockers instead of the traditional deliveries and the most important features that parcel must have. Regarding locations, users seem to prefer parcel lockers close to their daily origin/destination (e.g. home, workplace) so to integrate the collection of goods in their daily routine. Besides, about two third of the e-consumers are willing to deviate up to 10 minutes to collect their parcels. Van Duin et al. (2020) used three different methods for evaluating the suitability of parcel lockers in different scenarios: (i) Cost Effectiveness Analysis to calculate the delivery cost for each scenario; (ii) Multi-Criteria Analysis to identify the most important features for each alternative (e.g. safety, comfort); (iii) simulations to understand how different scenario works according to parameter set. Results show that the parcel locker solution can provide great benefits but it requires an appropriate analysis for the identification of their optimal location. There are some very recent studies that dealt with the issue of e-commerce delivery through ABM. Le Pira et al. (2020) proposed an approach based on the integration of discrete-choice models and ABM able to simulate the propensity of users on choosing between three different e-grocery alternatives: home delivery, click-and-pick and the traditional way of shopping. Sakai et al. (2020) implemented an ABM framework, validated with real-world user’s survey, to model the general e-commerce delivery demand. Another related approach is the work of Palanca et al. (2021), where they simulated different urban mobility and delivery solutions. Alves et al. (2019) analysed different scenarios by changing the number of delivery lockers. The study integrated the use of ABM and GIS tool and main results show that lockers can reduce the missed delivery phenomenon, generate additional monetary revenue for the supermarkets that contain them, and improve the cost effectiveness for curriers. Literature so far focused on demand analysis for different delivery strategies or configurations of deliveries based on consolidation. Our paper contributes to the literature in this field by building an ABM able to dynamically reproduce different delivery strategies and passenger movements. This is done in a parametric simulation environment to study the overall impact of the use of CDPs on the different actors involved, and specifically on the couriers routing, on the delivery cost and on the discomfort for customers, according to different demand patterns. Next section enters into the details of the methodology used to perform such analysis.

3. Methodology

The aim of the ABM is to identify the trade-off between the operator cost, the quality of service for customers and the environmental impact, considering design parameters such as the spatial density of CDPs. The ABM is implemented in the NetLogo programming environment (Wilensky, 1999) and includes four types of agents, namely customers, delivery vehicles (couriers), parcels and CDPs. A huge advantage of using ABMs is that each agent is programmed to carry out series of tasks without being necessarily aware of the status of the other agents in the system and one can study the emergence of the collective behaviour, sometimes with unexpected outcomes. The main input parameters of our simulation model concern the geometric features of the service area, the demand (customers) characterization, the supply (couriers) characteristics, and the simulation duration. We assume that delivery operations take place in a rectangular area of length L and width W. Couriers are defined by the capacity C_V (m^3), the cruise speed v_c (km/h) and the average energy consumption E_V (kWh/km) assuming they are fully electric. Operation costs are related to the drivers' wage (€$_{dr}$) and vehicle cost by distance (€$_{dist}$), not considering in the present work the fixed cost of vehicles. Couriers depart from a depot r distance away from the service area and travel along a grid street network (of spacing d_g). Each intersection (node) of the network could act as a potential delivery location (stop-node) where couriers stop to serve a home delivery request or a CDP. The customer demand is assumed homogeneous throughout the area, according to a demand density λ (*customers /km*^{2}*-day*). Customer purchase orders can be delivered either at home or in a CDP, which for simplicity is assumed to be constituted by a certain number of parcel lockers. The latter serves multiple customers, who have to go and pick-up their parcels, and are characterized by the input parameter "density of CDPs" δ (CDP/km^2) . Fig. 1 presents a schematic representation of the deliveries using the two options.

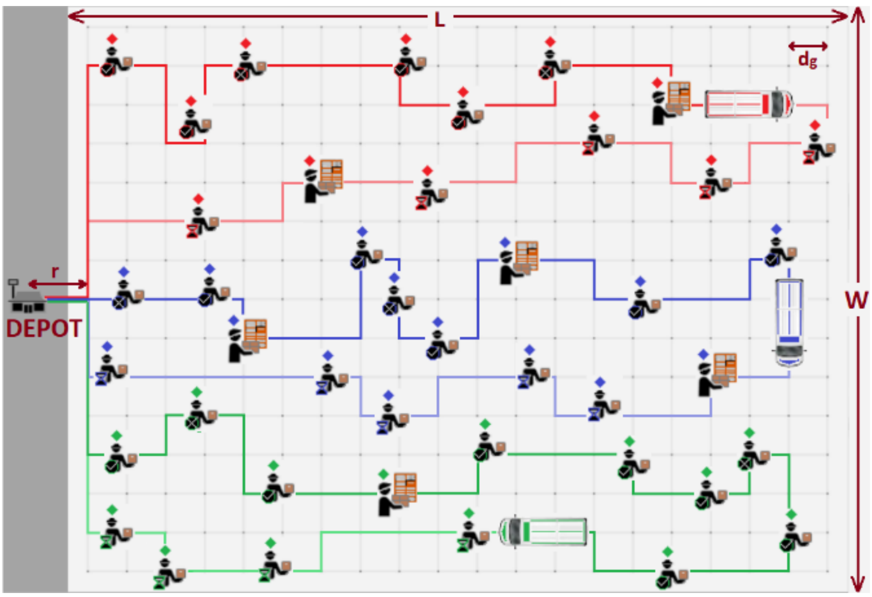


Fig. 1. Scheme of the delivery operations (Home Delivery and CDP) carried out by a fleet of couriers along the grid street network. Each courier's route is depicted in different couleurs.

The total number of deliveries in the simulated day, assuming one parcel per customer, is determined as follows:

$$N_C = \lambda \cdot L \cdot W \cdot (1 + K_R) \cdot (1 + K_F) \quad (1)$$

where K_R and K_F are coefficients accounting for the percentage of returned parcels and the percentage of failed deliveries. The latter should be calibrated based on the results of test simulations, since it depends on the previous days' delivery operations. Each customer is characterized by a different "activity profile" (as synthesized in Table 1).

Customers have random origin (O) and destination (D) of their daily activities within the service region. The distance d_{OD} has thresholds based on the activity profile. There are also customers not moving from home, except for picking up (or returning) their parcel. Customers choose between four mode of transport for their trips (“*car”*, “*bike”*, “*PT”* (public transport) and “*walk”*) and the probability P_{ij} for customer *i* to choose transport mode *j* for his trip is given by a logit model: $P_{ij} = e^{Vij} / \sum_j e^{Vij}$, where V_{ij} is the utility function associated with the alternative *j*. V_{ij} depends on the travel time t_{OD} from origin to destination, which in turn depends on the distance d_{OD}, and the alternative-specific attributes (see Table 2). In the present application, the parameters of the utility functions are taken from Cascetta (1998). We reasonably assume that customers consider as “target” CDP (possible alternative to home delivery) the one minimizing the detour time to reach it, but not exceeding the maximum detour $t_{det,max}$ (a key input parameter to be further investigated). If *j = “walk”*, the customer *i* can only choose the nearest CDP to the origin. If *j = “PT”*, one can also choose a CDP near to the destination. Finally, if *j = “car”* or *j = “bike”*, one can also choose a CDP near the O-D route (note that, given the same detour time, the detour distances are greater since car speed and bike speed are greater than walking speed).
Regarding customers’ requests, we consider four different parcel sizes (Table 3), the percentage of which is hypothesized considering a standard courier capacity of 8 m^3 and an average number of parcels (per fully loaded vehicle) of about 150 parcels (Llorca and Moeckel, 2020). The characterization of each customer (activity profile, mode of transport, type of parcel and delivery option) is made in the “setup” phase before the simulation starts.

Table 1. Input parameters related to the activity profile of customers.

Activity profile	**Home-Work (Employee)**	**Home-Work (Self-employed)**	**Home-Shopping**	**No displacement**
O-D Distance [km]	> 0.2	> 0.2	< 5.0	0
Start travel time	7:00 ÷ 8:30	6:00 ÷ 9:00	9:00 ÷ 17:00	-
Travel time duration [hour]	6 ÷ 10	8 ÷ 10	1 ÷ 3	-

Table 2. Input parameters related to the mode of transport.

Mode of transport	**Car**	**Bike**	**PT**	**Walk**
Avg. Speed [km/h]	30	15	15	5
Utility (V)	$\beta_{t,car}\, t_{OD,car} + \beta_{c,car}\, d_{OD,car} + \beta_{0,car}$	$\beta_{t,bike}\, t_{OD,bike} + \beta_{0,bike}$	$\beta_{t,PT}\, t_{OD,PT} + \beta_{0,PT}$	$\beta_{t,}\, t_{OD,walk}$

Table 3. Input parameters related to parcels.

Parcel size	**Small (S)**	**Medium (M)**	**Large (L)**	**Extra-Large (XL)**
Percentage	40 %	40 %	15 %	5 %
Size [m^3]	0.005 m3 (0.4x0.25x0.05)	0.04 m3 (0.5x0.4x0.2)	0.12 m3 (0.8x0.5x0.3)	0.36 m3 (1x0.6x0.6)
# reserved lockers in a CDP	20	26	4	0

In order to limit the variables affecting the model, we assume that the probability for customer *i* of choosing the delivery at a CDP *k* only depend on an impedance function η_{ik} which is the ratio between the detour time to reach *k* and $t_{det,max}$, as below:

$$P_{ik} = (1 - \eta_{ik}) \cdot x_{CDP} = \left(1 - \frac{t_{i,k}}{t_{det,max}}\right) \cdot x_{CDP} \tag{2}$$

where x_{CDP} is a Boolean variable which assumes the value 1 if there are available lockers for the parcel and 0 otherwise. Note that, for simplicity, the detour time only includes the one-way time needed to reach the CDP, not the round-trip. It may happen that the customer choosing home delivery is not at home when the courier arrives: in this case a failed delivery occurs, if nobody else is available (a family member, a neighbour, a doorman, etc.). The probability of having an alternative person to receive the delivery is a calibration parameter, which we set assuming a percentage of failed deliveries of about 25% (Van Duin et al., 2016) when only the home delivery option is available.

The vehicle routing problem (VRP) is solved in the "setup" phase, as follows. We first determinate the minimum fleet size n_V by dividing the total volume of the parcel to be delivered $V_{del,tot}$ by the courier capacity C. Then we divide the service region into n_V rectangular sub-regions (this approximation is justified by the homogeneous spatial demand distribution), assigning the stop-nodes to serve the and estimate the travel time (including the idle time with the parked vehicle during delivery operations). In real-world scenarios, determining the optimal route assignments to the vehicles of the fleet allows the operator to obtain significant time and cost savings, so the VRP should be addressed by *ad-hoc* optimization algorithms (Calabrò et al., 2020). Since we propose a synthetic case study with an idealized grid network, we follow the guidelines suggested by Daganzo (2004) and applied in the analytical model of Quadrifoglio and Li (2009), assuming that each courier travels through the upper half of the sub-region in a no-backtracking policy (e.g. left-to-right), and then through the bottom half in a no-backtracking policy in the opposite direction (e.g. right-to-left). This pattern can be observed in Fig. 1. We consider an additional time lost at each stop τ_s = 30 s, including the time of acceleration and deceleration, and average delivery times per parcel τ_p (home delivery) and τ_{cdp} (CDP).

The behaviour of the couriers and the behaviour of the customers are represented by two state charts in Fig. 2a and Fig. 2b, respectively. Every agent goes through several states, each one is activated when a given event occurs and/or if a set of conditions are met (square brackets), causing the transitions (the black arrows) between different states.

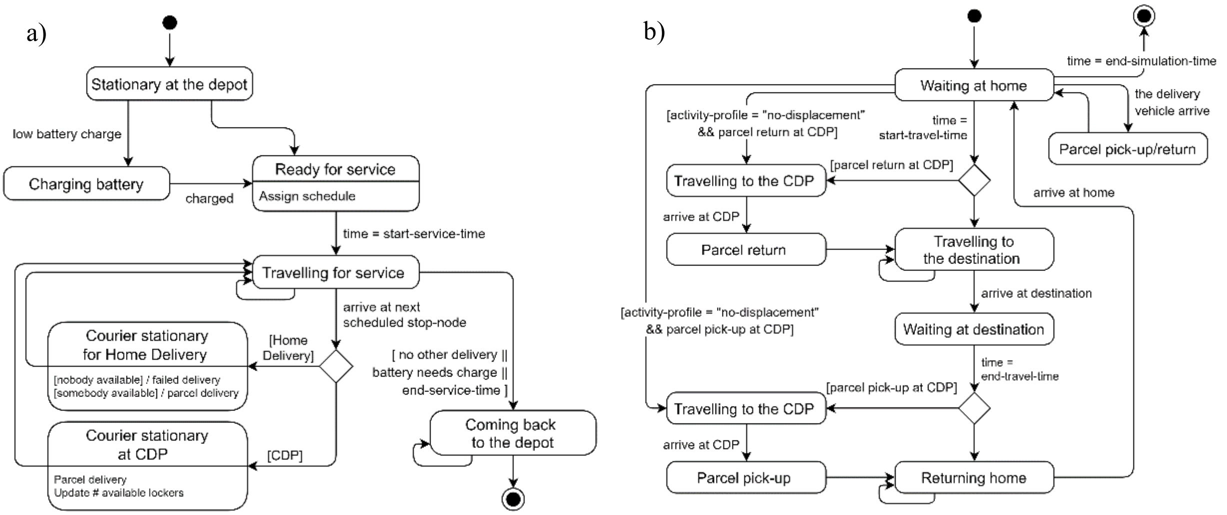


Fig. 2. (a) Courier state chart; (b) Customer state chart.

4. Preliminary results and discussion

The ABM allows monitoring of different key performance indicators, related to both customers and operators' points of view (Calabrò et al., 2021). Different scenarios can be simulated by varying the number and location of parcel lockers, courier fleet size, and customer demand patterns, allowing to infer about the attractiveness of different delivery options from the point of view of customers and logistics operators, and by evaluating the overall impact of deliveries on city sustainability.

For a first test of the model, we simulate customers' activities during an entire working day and delivery operations during 8 work hours (from 9 A.M. to 6 P.M., including a 1-hour lunch break). We applied the explained methodology to a synthetic case study that we called "Virtual-CT". It mimics a medium-sized city (30 km^2) with a grid-like network and the freight depot in the city outskirt. Operator-related costs are taken from Llorca and Moeckel (2020). We performed two sets of simulations, by varying (1) the density of CDP and (2) customer willingness to deviate from their usual trip to reach a CDP (i.e. the max detour). The inputs of the different scenarios are reported in Table 4:

Table 4. Scenario input parameters setup.

GEOMETRIC PARAMETERS				DEMAND PARAMETERS				SUPPLY PARAMETERS							
L	W	r	δ	λ	K_R	v_w	$t_{det,max}$	n_V	C_V	v_c	τ_s	τ_p	τ_{cdp}	$€_{dr}$	$€_{dist}$
[km]	[km]	[km]	**[CDP/km²]**	[cust/km²]		[km/h]	**[min]**		[m³]	[km/h]	[s]	[s]	[s]	[eur/h]	[eur/km]
6	5	1	**0-0.5-1-2-5-10**	25	0.15	5	**6 - 10**	*5-7* *	8	25	30	120	30	26.9	0.77

*The number of couriers is the minimum to satisfy the number of requests. It can vary in different simulation runs

Eleven scenarios are simulated. In particular, scenario 0 only considers home deliveries and scenarios 1-5 are related to increasing density of CDP and are repeated twice by increasing the max detour from 6 to 10 min to account for more flexibility of consumers. In particular, scenario 1 considers a density of 0.5 CDP/km^2, scenario 2 of 1 CDP/km^2, scenario 3 of 2 CDP/km^2, scenario 4 of 5 CDP/km^2 and scenario 5 of 10 CDP/km^2. Each scenario has been run five times to test result fluctuations and have a statistics of events. Main results are reported in the following figures (Fig. 3-4-5).

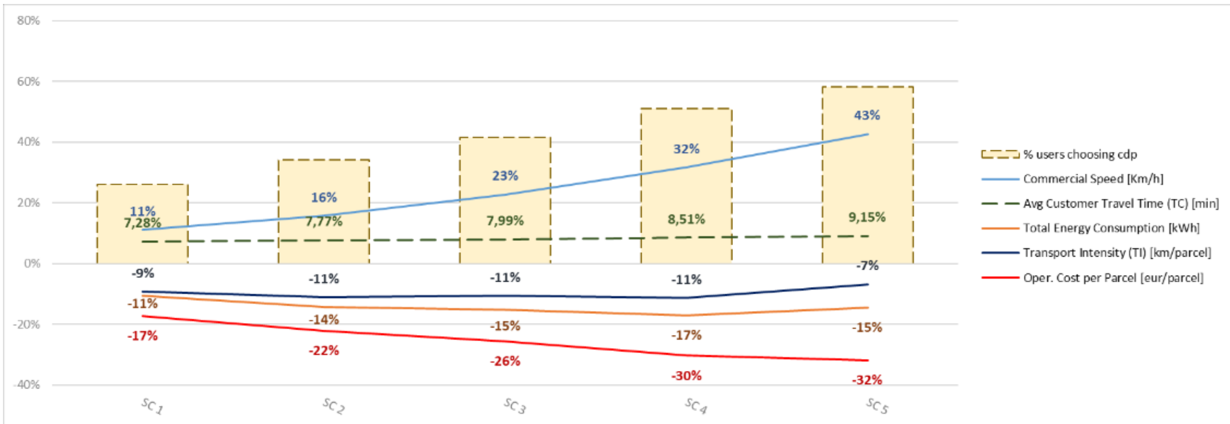


Fig. 3. Variation of customer- and operator-related outputs with respect to scenario 0 (home delivery) when increasing the CDP density (max detour time = 6 min).

Fig. 3 compares scenarios in terms of variation of customer- and operator-related outputs. In general, while increasing the density of the CDP, the percentage of customers choosing them goes up from 26% to 58%. This occurs at the expense of a small increase in their average travel time due to the detour needed, which goes from 7% to 9% (from 16 to 18 minutes) from scenario 1 to 5. From the point of view of the operator, this implies fewer costs per parcel, and higher commercial speeds due to shorter stopping times (in line with Allen et al., 2018). However, the transport intensity, which is linked to the distance travelled per parcel, and the total energy consumption do not show a clear decreasing trend. In particular, one can notice a critical value of CDP density (5 CDP/km2) after which these indicators start to get worse with respect to the previous scenarios. This suggests that, for the specific area of analysis and simulation parameters, a further increase in the CDP density would not be so beneficial in terms of distance travelled per parcel and environmental impact, meaning that there is not a meaningful consolidation effect.

If we look at failed deliveries (Fig. 4), there is a general improvement once the density of CDP increases. They drop from 22% in the scenario totally based on home deliveries (scenario 0) to 9 or 7% in the scenario with a density of CDP equal to 10 CDP/km2 (scenario 5, respectively considering 6 or 10 minutes detour).

Finally, when comparing scenario results obtained by increasing customer willingness to deviate from their original trip from 6 to 10 minutes (Fig. 5), one can see a small increase in the average customer travel time, and a higher benefit in terms of transport intensity, especially for CDP density equal to 2 CDP/km2.

To sum up, scenarios based on consolidation are to be preferred from the point of view of logistics efficiency and negative impact, even if they imply a small discomfort for the customers. The best solution from both the customer and the operator point of view (i.e. in terms of transport intensity) should be to provide a delivery service using CDP with a density from 2 to 5 CDP/km2. A mechanism of incentives for customers choosing CDP could be considered to compensate the increase in travel time due to the detour to reach the CDP.

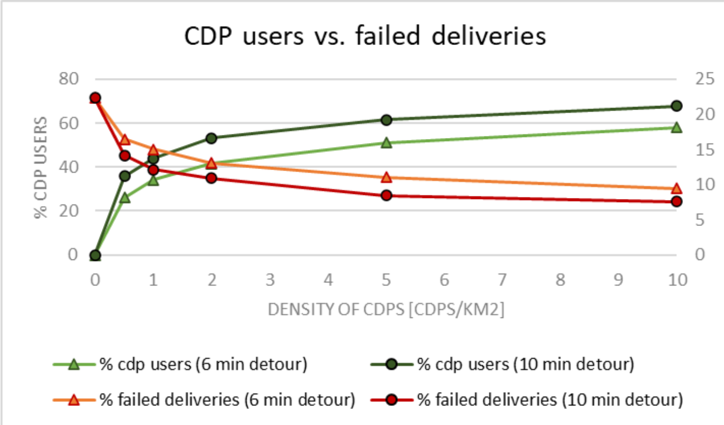


Fig. 4. %CDP users with respect to failed deliveries in the different scenarios.

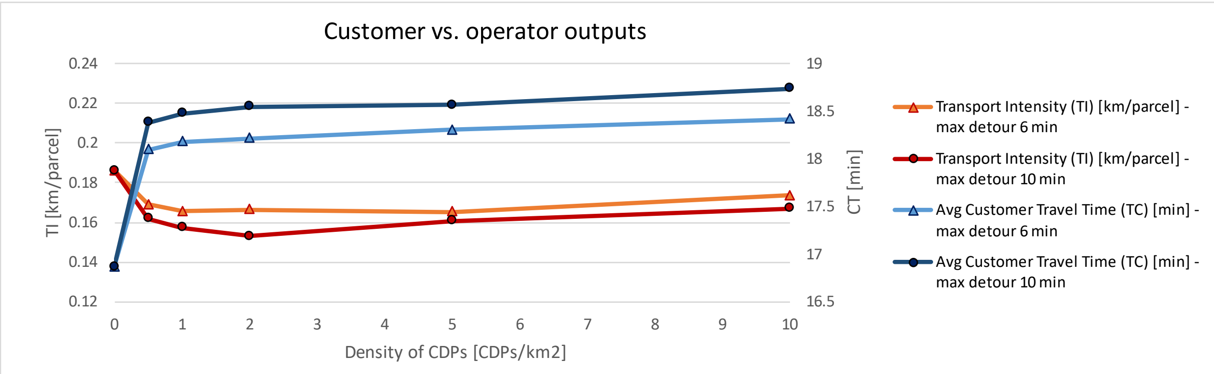


Fig. 5. Customer vs. operator outputs in terms of transport intensity (TI) and customer travel time (CTT) in the different scenarios.

5. Conclusions

This paper presented the first results of a new agent-based able to reproduce the complexity of on-demand last-mile parcel deliveries. These are expected to grow in the near future and are becoming a big concern for policy-makers. The model reproduces a parametric environment and is able to dynamically simulate freight deliveries and customer movements. For a first test, it is used to reproduce scenarios based on delivery consolidation via CDP.

Main results suggest that a trade-off between freight vehicle travelled distance, customer satisfaction and logistics costs can be found while proposing a solution to last mile parcel deliveries based on consolidation via parcel lockers. In particular, while increasing the density of CDP, there are more opportunities for reducing the operator costs and improving the logistics efficiency at the expense of a small discomfort for users in terms of travel time to pick up the parcel at the CDP. It is also possible to find an optimal range of CDP density implying the best results in terms of delivery impact (travelled distance per parcel and total energy consumption).

In conclusion, the proposed methodology can be used as a support tool to understand how to plan e-commerce deliveries by considering efficient solutions both from sustainability and the customer satisfaction point of view. Future research steps will focus on more tests (e.g. by simulating more than one day or different demand patterns) and to make it more realistic by reproducing real case studies with demand data.

Acknowledgements

The work is partially supported by the project "WEAKI-TRANSIT: WEAK-demand areas Innovative TRANsport Shared services for Italian Towns" (unique project code: E44I17000050001) under the programme "PRIN 2017" and by the project of M. Le Pira "AIM Linea di Attività 3 – Mobilità sostenibile: Trasporti" (unique project code CUP E66C18001390007) under the programme "PON Ricerca e Innovazione 2014-2020 – Fondo Sociale Europeo, Azione 1.2 "Attrazione e mobilità internazionale dei ricercatori"".

References

Allen, J., Piecyk, M., Piotrowska, M., McLeod, F., Cherrett, T., ... and Austwick, M., 2018. Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of London. Transportation Research Part D: Transport and Environment, 61, 325-338.

Alves, R., da Silva Lima, R., Custódio de Sena, D., Ferreira de Pinho, A., & Holguín-Veras, J. (2019). Agent-based simulation model for evaluating urban freight policy to e-commerce. *Sustainability*, *11*(15), 4020.

Calabrò, G., Torrisi, V., Inturri, G., Ignaccolo, M., 2020. Improving inbound logistic planning for large-scale real-world routing problems: a novel ant-colony simulation-based optimization. European Transport Research Review 12,1-11.

Calabrò, G., Correia, G., Giuffrida, N., Ignaccolo, M., Inturri, G., & Le Pira, M., 2020. Comparing the performance of demand responsive and schedule-based feeder services of mass rapid transit: an agent-based simulation approach. In 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS) (pp. 280-285). IEEE

Daganzo, C.F., 2004. Logistics Systems Analysis. Springer-Verlag, Heidelberg, Germany.

Gevaers, R., Van de Voorde, E., and Vanelslander, T., 2011. Characteristics and typology of last-mile logistics from an innovation perspective in an urban context. City Distribution and Urban Freight Transport: Multiple Perspectives, Edward Elgar Publishing, 56-71.

Iwan, S., Kijewska, K., & Lemke, J.,2016. Analysis of parcel lockers' efficiency as the last mile delivery solution–the results of the research in Poland. *Transportation Research Procedia*, *12*, 644-655.

Lachapelle, U., Burke, M., Brotherton, A., & Leung, A., 2018. Parcel locker systems in a car dominant city: Location, characterisation and potential impacts on city planning and consumer travel access. *Journal of Transport Geography*, *71*, 1-14.

Le Pira, M., Marcucci, E., Gatta, V., Inturri, G., Ignaccolo, M., & Pluchino, A., 2017. Integrating discrete choice models and agent-based models for ex-ante evaluation of stakeholder policy acceptability in urban freight transport. Research in transportation economics, 64, 13-25.

Le Pira, M., Marcucci, E., Gatta, V., Pluchino, A., Fazio, M., Inturri, G., & Ignaccolo, M., 2020. Simulating urban freight flows in e-grocery scenarios accounting for consumer heterogeneous preferences. In *2020 Forum on Integrated and Sustainable Transportation Systems (FISTS)* (pp. 286-291). IEEE.

Llorca, C. and Moeckel, R., 2020. Study of cargo bikes for parcel deliveries under different supply, demand and spatial conditions. In 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS) (pp. 39-44). IEEE.

Mitrea, I. A., Zenezini, G., De Marco, A., Ottaviani, F. M., Delmastro, T., & Botta, C., 2020. Estimating e-Consumers' Attitude Towards Parcel Locker Usage. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1731-1736). IEEE.

Palanca, J., Terrasa, A., Rodriguez, S., Carrascosa, C., & Julian, V., 2021. An agent-based simulation framework for the study of urban delivery. *Neurocomputing*, *423*, 679-688.

Quadrifoglio, L., and Li, X., 2009. A methodology to derive the critical demand density for designing and operating feeder transit services. Transportation Research Part B: Methodological, 43(10), 922-935.

Sakai, T., Hara, Y., Seshadri, R., Alho, A., Hasnine, M. S., Jing, P., ... & Ben-Akiva, M., 2020. E-Commerce Delivery Demand Modeling Framework for An Agent-Based Simulation Platform. *arXiv preprint arXiv:2010.14375*.

Schnieder, M., and West, A. A., 2020. Comparison of Time-Area Requirements of Parcel Lockers vs. Home Delivery: A Cyber-Physical System of Last Mile Delivery. In 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS), IEEE, 298-303.

Van Duin, J. H. R., Wiegmans, B. W., van Arem, B., & van Amstel, Y., 2020. From home delivery to parcel lockers: A case study in Amsterdam. *Transportation Research Procedia*, *46*, 37-44.

Wilensky, U., 1999. NetLogo. Center for Connected Learning and Computer Based Modeling. Northwestern University, Evanston, IL. In: http://ccl.northwestern.edu/netlogo/.

World Economic Forum, January 2020, The Future of the Last-Mile Ecosystem: Transition Roadmaps for Public- and Private-Sector Players.