# Integrated timetabling and vehicle scheduling of an intermodal urban transit network
## A distributionally robust optimization approach

Xia, Dongyang; Ma, Jihui; Sharif Azadeh, Sh

**DOI**
10.1016/j.trc.2024.104610

**Publication date**
2024

**Document Version**
Final published version

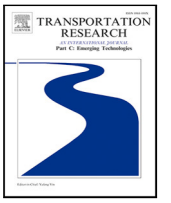
**Published in**
Transportation Research Part C: Emerging Technologies

Transportation Research Part C 162 (2024) 104610

Contents lists available at ScienceDirect

Transportation Research Part C

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



# Integrated timetabling and vehicle scheduling of an intermodal urban transit network: A distributionally robust optimization approach☆

Dongyang Xia [a], Jihui Ma [a], Sh. Sharif Azadeh [b,*]

[a] Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China
[b] Department of Transport & Planning, Delft University of Technology, Netherlands

ARTICLE INFO

*Keywords:*
Inter-modal transit system
Modular vehicles
Dispatching plan
Stop-skipping strategy
Uncertain and time-dependent passenger demand
Adaptive large neighborhood search

ABSTRACT

Integrating emerging shared mobility with traditional fixed-line public transport is a promising solution to the mismatch between supply and demand in urban transportation systems. The advent of modular vehicles (MVs) provides opportunities for more flexible and seamless intermodal transit. The MVs, which have been implemented, are comprised of automated modular units (MUs), and can dynamically change the number of MUs comprising them at different times and stops. However, this innovative intermodal urban transit brings with it a new level of dynamism and uncertainty. In this paper, we study the problem of jointly optimizing the timetable and the vehicle schedule within an intermodal urban transit network utilizing MVs within the context of distributionally robust optimization (DRO), which allows MVs to dynamically (de)couple at each stop and permits flexible circulations of MUs across different transportation modes. We propose a DRO formulation to explore the trade-off between operators and passengers, with the objective of minimizing the worst-case expectation of the weighted sum of passengers' and operating costs. Furthermore, to address the computational intractability of the proposed DRO model, we design a discrepancy-based ambiguity set to reformulate it into a mixed-integer linear programming model. In order to obtain high-quality solutionss of realistic instances, we develop a customized decomposition-based algorithm. Extensive numerical experiments demonstrate the effectiveness of the proposed approach. The computational results of real-world case studies based on the operational data of Beijing Bus Line illustrate that the proposed integrated timetabling and vehicle scheduling method reduces the expected value of passengers' and operating costs by about 6% in comparison with the practical timetable and fixed-capacity vehicles typically used in the Beijing bus system.

## 1. Introduction

Virtually every major city is experiencing rapid population growth. United Nations (2019) reports that, the world's population is expected to increase to 9.7 billion in 2050. This increase provides an opportunity to reshape urban mobility, making a flexible, effective and integrated transportation system ever more desirable. On the other hand, with cities around the world moving towards

☆ This article belongs to the Virtual Special Issue on VSI: Coordinated Multimodal.

* Corresponding author.
*E-mail address:* s.sharifazadeh@tudelft.nl (S. Sharif Azadeh).

https://doi.org/10.1016/j.trc.2024.104610
Received 23 September 2023; Received in revised form 1 April 2024; Accepted 2 April 2024
Available online 8 April 2024

**Fig. 1.** Modular vehicles developed by the Next Future Transportation Inc. (NEXT Future Transportation Inc., 2019).

low car diet in urban areas due to mainly pollution, energy, and spatial shortage, there is an urgent need to design and operationalize seamless urban mobility services that can integrate new shared mobility solutions with fixed-line public transport.

In this paper, we study the timetabling and vehicle scheduling problem within an inter-modal urban transit system. We design a tailored distributionally robust optimization framework. Our study focuses on two transportation modes: Fixed-Line and Schedule (FLS) and Demand Responsive Transport (DRT) services. DRT services are operated similarly to ride-hailing services in which passengers are picked up and dropped off from a pre-defined set of stops, referred to as stop-to-stop services (Narayan et al., 2020). Both FLS and DRT modes utilize modular vehicles (MVs) composed of modular units (MUs). Each MV can be flexibly coupled and decoupled at each stop, thus dynamically adjusting their capacity at different times and stops—a concept referred to as *dynamic capacity allocation*. The *formation* of an MV is defined as the number of MUs contained within it. In practice, the MVs developed by the Next Future Transportation Inc., as illustrated in Fig. 1, already have the functionality to support the aforementioned operations and have been tested in Dubai (NEXT Future Transportation Inc., 2023). In their MVs, each MU is capable of docking with other MUs at speed on public roads and can be unlocked smoothly, and these operations can be performed relying solely on MU's own device (NEXT Future Transportation Inc., 2019).

The MUs employed in the intermodal urban transit network exhibit the capability to circulate between different modes. Specifically, these units can be operated within the DRT system or allocated to FLS trips. Therefore, the dispatch center of the FLS system also serves as the controller for the DRT system. This innovative, highly flexible intermodal transit system provides effective and economical services to passengers, enabling them to reach destinations through FLS or DRT services at public transportation prices. Simultaneously, operators can reduce operating costs by adopting dynamic-capacity allocations and the flexible vehicle scheduling in response to the time-dependent passenger demand.

However, these DRT services introduce a new level of dynamism and uncertainty to the currently existing transportation networks. Passenger demand in the intermodal transit system is frequently influenced by multiple uncertainty sources, such as weather conditions and specific dates. Accordingly, it is crucial to take uncertainties into account in the planning phase to ensure high-quality services. During the last decade, there has been a growing body of research on the integrated FLS and DRT system. Nevertheless, current studies mainly assume that the demand is deterministic and (or) static, and they often limit their scope to optimizing DRT services while fixing FLS trips. Rarely do they consider flexible vehicle circulations between different transportation modes. The possible reasons may involve: (1) The dynamics of passenger boarding, alighting, and on-boarding are tightly coupled between different transportation modes in an intermodal transit system, which raises highly nonlinear characteristics for integrated optimization. (2) Constructing a set of stochastic scenarios is a typical method to model the demand uncertainty, but it would dramatically increase the dimensionality of variables and make finding solutions more challenging. (3) More importantly, existing research has already shown the computational intensities in solely solving the vehicle scheduling problem for the DRT system or the timetabling and dynamic-capacity allocations for a single FLS line (e.g., Xia et al. 2023). Consequently, the difficulties are intuitive if we further consider an integrated framework of timetabling and vehicle scheduling of an intermodal urban transit network under demand uncertainty while the capacity can dynamically change.

In addition, the most common approach to capture the demand uncertainty in transportation systems is the stochastic programming (SP) method, in which stochastic scenarios with deterministic probability distributions are employed (e.g., Wu et al. 2019, Sadrani et al. 2022, Ma et al. 2023). However, obtaining precise probability distributions for the various scenarios in real-world operations remains a challenge. Recently, the distributionally robust optimization (DRO) method has gained significant attention as an emerging approach to address the demand uncertainty. Within a DRO framework, the distribution of uncertain parameters is assumed to be partially known and follows a probability distribution that resides within an ambiguity set (Rahimian and Mehrotra, 2022). While it is widely acknowledged that the DRO method mitigates the over-conservatism inherent in robust optimization and enhances the robustness of optimal solutions relative to the SP approach, one limitation is that the DRO formulations are difficult to be solved for large-scale problems due to considering a set of probability distributions.

To overcome the aforementioned issues, this paper presents a distributionally robust timetabling and vehicle scheduling (TTVSP) framework of an integrated FLS and DRT system. The proposed methodology employs the DRO method to model the time-dependent

and uncertain passenger demand, as well as the uncertainty of the possibility of stochastic scenarios used to portray the randomness of passenger demand. The goal of this paper is to generate a global-optimal robust timetable of FLS trips, routing and departure times of DRT services, vehicle schedules of this inter-modal transportation system, and gain insight into the trade-off between the passengers' and operating costs. To do so, we formalize the TTVSP for an integrated FLS and DRT system with MVs, which, in contrast to the traditional timetabling problem in the FLS system using fixed-capacity vehicles, has not received much attention in the literature. Next, we formulate an SP model and introduce the first DRO model with the objective of minimizing the expected value of passengers' and operating costs. We allow for a flexible vehicle circulations within an intermodal urban transit system. Furthermore, to efficiently solve real-case instances based on the Beijing bus line, we develop a hybrid algorithm combining the tailored Adaptive Large Neighborhood Search (ALNS) algorithm with a mixed-integer linear programming (MILP) solver. The small realistic case study and real-world instances based on the operating data of Beijing Bus Line 468 are constructed to verify our formulations and solution methodologies.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature and summarizes our contributions. Section 3 provides a description of the problem. In Section 4, an SP model and a DRO model are formulated for the TTVSP. Then, we design an effective solution algorithm combining the tailored ALNS and an MILP solver in Section 5. Computational results are presented in Section 6, followed by conclusions and future research directions in Section 7.

## 2. Literature review

There is a large amount of literature on the timetabling problem for bus networks (e.g., Ibarra-Rojas et al. 2014, Dai et al. 2020, Gkiotsalitis and Van Berkum 2020, Chen and An 2021, Zhang et al. 2021, Cortés et al. 2023). For the sake of brevity, our literature review will focus on the research related to the vehicle scheduling problem of DRT systems, timetabling and dynamic-capacity allocation of FLS systems, and the integrated optimization of FLS and DRT services in a multi-modal transportation system.

### 2.1. Vehicle scheduling of DRT systems

Over the past decade, many studies have been conducted to investigate the tactical and operational planning of DRT systems in terms of vehicle scheduling (e.g., Narayan et al. 2020, Chen et al. 2021a, Galarza Montenegro et al. 2021, Ma et al. 2023). For example, Chen et al. (2021b) addressed the routing problem of customized buses considering the time-varying passenger demand. The authors formulated a bi-objective optimization model to minimize passengers' and operating costs and designed an adaptive variable neighborhood search algorithm. The results show that, in comparison with FLS trips, the customized bus services with flexible routes can provide a higher level of quality and are an economically efficient travel options for passengers. Sharif Azadeh et al. (2022a) tackled the choice-driven dial-a-ride problem considering private and shared demand-responsive mobility services. They proposed an MILP formulation incorporating customer choices and developed a ALNS algorithm to solve the model. Fu and Chow (2023) addressed the dial-a-ride problem with MVs, where the capacity of each MV assigned to serve requests varies at different times and locations, and passengers can en-route transfers during vehicle platooning. The results show that using MVs can lead to significant savings in vehicles' travel costs, passengers' service time, and total costs against existing mobility-on-demand services.

In summary, the above literature contributes to the vehicle scheduling for the DRT system considering time-varying passenger demand or static requests, minimizing operating costs or (and) passengers' costs. Nevertheless, these works consider only the DRT system, with the assumption that the timetables, fleet sizes, and vehicle schedules of the FLS system are pre-given.

### 2.2. Timetabling and dynamic-capacity allocations of FLS systems

With the boom in modular vehicles, a more recent idea to emerge in the literature is to use their flexibility to dynamically adjust capacity on an existing bus line to match time-varying passenger demand as closely as possible, conserving capacity and road resources. Notable works include Chen et al. (2020), Shi and Li (2021), Tian et al. (2022), Khan et al. (2023). The objective is mostly to minimize the weighted sum of passengers' and operating costs. The main differences include whether the dynamics and the uncertainty of passenger demand, whether MUs can be decoupled and coupled at each station, and whether the re-balancing of decoupled MUs is considered.

Optimizing timetables and capacity allocations with time-dependent passenger demand is one research area that is especially relevant to our study. In this domain, Chen et al. (2019) formulated an MILP model for the joint design problem of the dispatch headways and vehicle capacities for shuttle systems, with the objective of minimizing passenger waiting costs and vehicle energy costs. Liu et al. (2023) proposed a mathematical model for integrated timetabling, bus formation adjusting, and vehicle scheduling on a bus loop line, aiming to minimize the total passengers' waiting cost, operational costs, and penalty costs of (de)coupling operations. Xia et al. (2023) formulated SP and DRO models considering time-dependent and uncertain passenger demand on a bus line. Computational results indicate that adopting dynamic-capacity allocations of MVs reduces operational costs, and the DRO method outperforms RO and SP approaches in in-sample and out-of-sample performance, respectively. In addition, as far as we know, Tian et al. (2023) is the only paper that takes into account re-balancing the MUs that are decoupled at downstream stops to upstream stops. Nevertheless, the above research mainly focuses on timetabling and scheduling of MVs within a single-mode transit system, while none of those studies consider synergistic utilization of these highly flexible vehicles in an intermodal urban transit network.

### 2.3. Integrated optimization of FLS and DRT services in a multi-modal transit system

In recent years, the body of literature concerning the integrated FLS and DRT system has been expanding. Several studies have explored with a particular emphasis on last-mile service design (e.g., Guo et al. 2018, Wang 2019, Steiner and Irnich 2020). Some studies examined how mixed fleets should be switched between different transportation services at different times (e.g., Kim and Schonfeld 2014). Some research sought to determine the optimal network design for an integrated fixed-line and demand-responsive transit system (e.g., Sharif Azadeh et al. 2022b).

For instance, considering the heterogeneous demand, Luo and Nie (2020) explored the design of demand-adaptive paired-line hybrid transit system consists of both FLS and DRT services, aiming to minimize both agency and user costs. The decisions include the number and headway of circular lines inside and outside the central business district. Wang et al. (2022) investigated the design problem of the multi-modal transportation system characterized by static passenger arrival rates. This study integrated ride-hailing, bike-sharing services, and fixed-route buses, modeling them within a grid network. The authors addressed this problem in the view of a central designer, focusing on decision variables related to the numbers of shared bikes and ride-hailing vehicles, density of available ride-hailing vehicles, mileage fee of the ride-hailing service, distance between two nearby bus stops, and bus headway. However, these two aforementioned studies assume that formations of vehicles are fixed, and don not take the flexible circulation of vehicles among various transportation modes into account. Lin et al. (2023) proposed an integrated co-modality system that combines public transit with last-mile logistics services through the use of MVs, and thoroughly discussed the prospects for future applications. They concluded that such a flexible transit system has the potential to profoundly reshape urban mobility. Gu and Chen (2023) proposed an equilibrium mode choice formulation to model the long-term modal split in multi-modal transportation networks, determining the numbers of passengers choosing customized bus services, private car mode and conventional transit mode. This study focuses on the travel choice behavior of passengers in a multi-modal transportation system, considering the timetabling and vehicle scheduling as out of scope. More recently, based on the assumption that passenger arrivals are uniformly distributed, Basciftci and Van Hentenryck (2023) proposed a bi-level optimization model for the network design problem of on-demand multi-modal transit systems (ODMTSs), supposing that buses are operated between hubs and on-demand shuttle services are used to serve passengers between stops. However, in that study, it is assumed that fleet sizes are fixed and the vehicles used to serve on-demand services and fixed-line bus trips are different and cannot be flexibly shared between different transportation modes. To sum up, even though the aforementioned studies share several characteristics with our proposed integrated optimization framework, none of them explore the integration of timetabling and vehicle scheduling within an intermodal network, particularly in the context of time-varying and uncertain passenger demand. Moreover, the flexible circulation of vehicles across different transportation modes has not been incorporated into their mathematical models.

Table 1 presents a comparison of our study with closely related literature. The most closely associated literature is Xia et al. (2023), which focuses only on FLS services on a single bus line while leaving the vehicle scheduling aspect out of scope. It uses a parameter with two-dimensional properties of time and space to portray the uncertainty of the passenger demand, and introduces a Wasserstein distance-based ambiguity set. However, such a two-dimensional parameter introduces a significant computational complexity. In this current paper, we expand the research scope and introduce novel methodologies to integrated timetabling and vehicle scheduling of an innovative intermodal transit network. Specifically, (i) the introduced approach allows operators to manage decoupled MUs and enables flexible vehicle circulations across different transportation modes. (ii) We implement a skip-stop strategy to enhance the operational efficiency to tackle real-world challenges. (iii) Our proposed method adopts a *passenger group modeling* approach, offering a more precise characterizations of complex passengers' dynamics in the intermodal transportation system and better tracking of key performance indicators related to passengers' movements. (iv) We build stochastic scenarios to ensure that the model scales reasonably well while portraying the uncertainty of passenger flows, which can be easily handled using the designed decomposition solution method. Besides, we employ a discrepancy-based ambiguity set which can depict the uncertainty of the occurring probability of scenarios.

### 2.4. Paper contributions

In this study, we aim to construct a distributionally robust optimization model for the integrated timetabling and vehicle scheduling problem, and develop an effective decomposition-based algorithm. The four key contributions of this paper are summarized as follows:

- We explore the timetabling and vehicle scheduling within an innovative intermodal transit system that integrates DRT services with FLS trips. Vehicles utilized in this integrated system are emerging MVs. Thanks to the operational flexibility inherent in MVs, the forthcoming inter-modal systems could enable autonomous MUs to serve DRT services, which may be decoupled from the MVs allocated to FLS trips. Moreover, the MVs allocated to DRT services have the capacity for real-time integration with the MUs servicing traditional FLS trips, thereby augmenting both the continuity and efficacy of the intermodal transportation network.
- We formulate an SP model for the studied problem based on the assumption that the possibility distribution of stochastic parameters is fully known in advance. Furthermore, we propose a mixed-integer nonlinear programming (MINLP) formulation within the context of distributionally robust optimization method considering uncertainties of both the time-dependent demand and the possibility distribution of scenarios. The objective is to minimize the worst-case expectation of the weighted sum of passengers' and operating costs, where passengers' costs include waiting and in-vehicle costs. Moreover, we demonstrate that the proposed models can be equivalently reformulated into MILP formulations.

**Table 1**
Summary of relevant recent studies. Abbreviations: Intermodal transit network = ITN, FLS = Fixed-Line and Schedule, DRT = Demand-Responsive Transport, DP = Dynamic programming algorithm, AVNS = Adaptive variable neighborhood search algorithm, RH = Rolling horizon method, ALNS = Adaptive large neighborhood search algorithm, BD = Benders decomposition method, ADMM = Alternating direction method of multipliers, IB = Iterative balancing algorithm, STILNS = Steiner tree-inspired local neighborhood search algorithm.

| Publications | Transportation mode | Passenger demand | Timetabling | Flexible formations of vehicles | Flexible circulations of vehicles across modes | Solution method |
|---|---|---|---|---|---|---|
| Chen et al. (2019) | FL | Time-dependent | Yes | At terminus | No | DP |
| Luo and Nie (2020) | ITN | Static | Yes | No | No | Monte Carlo simulation |
| Leffler et al. (2021) | DRT | Static | No | No | No | Simulation |
| Chen et al. (2021b) | DRT | Time-varying | No | No | No | AVNS |
| Sharif Azadeh et al. (2022b) | ITN | Time-varying | No | No | No | ALNS |
| Wang et al. (2022) | ITN | Static | Yes | No | No | DIRECT algorithm |
| Ma et al. (2023) | DRT | Time-dependent and uncertain | No | No | No | RH embedded ALNS |
| Basciftci and Van Hentenryck (2023) | ITN | Static | No | No | No | BD |
| Tian et al. (2023) | FL | Time-dependent | Yes | At each stop | No | Two-step heuristic |
| Liu et al. (2023) | FL | Time-dependent | Yes | At certain stop and terminus | No | ADMM |
| Gu and Chen (2023) | ITN | Static | No | No | No | IB |
| Xia et al. (2023) | FL | Time-dependent and uncertain | Yes | At each stop | No | Integer L-shaped |
| Fu and Chow (2023) | DRT | Time-varying | Yes | At each location | No | STILNS |
| **This paper** | **ITN** | **Time-dependent and uncertain** | **Yes** | **At each stop and across modes** | **Yes** | **ALNS combined with GUROBI** |

• Our proposed formulations, extended from Xia et al. (2023), simultaneously consider strict capacity constraints and constraints for flexible vehicle circulations within an intermodal transit network. Different from the single-modal vehicle scheduling problem studied, e.g., in Liu et al. (2023) and Chen et al. (2021b), we propose formulations for flexibly and dynamically schedule homogeneous MUs across different transportation modes. Our formulations allow each MU to be decoupled from the MV of which it is a part at each stop, and provide three alternatives for each decoupled MU: rebalancing to upstream stops, being stored in depots attached to each stop with strict capacity limitations, or being allocated to serve the other transportation mode.

• We design a customized hybrid algorithm to generate high-quality solutions for real-world instances within the reasonable computational time. First, we present a decomposition method to divide the proposed MILP formulation into two subproblems. Subsequently, we develop an effective algorithm combining an ALNS algorithm and an MILP solver, which embeds the constrained compass search and simulated annealing algorithms. To evaluate the effectiveness of our proposed methodologies, we conduct two sets of numerical experiments: one is based on a small-scale case study and the other utilizes practical operating data of the Beijing bus line. The experimental results show that our methodologies offer remarkable improvements compared with the state-of-the-art solver and the practical plan used on the Beijing bus line.

## 3. Problem description

In this section, we provide details on the dynamics of the proposed intermodal network. We provide illustrative examples to provide further clarity.

### 3.1. Problem definition

Our study considers the distributionally robust timetabling and vehicle scheduling (TTVSP) for an integrated FLS and DRT system utilizing MVs. We take an automated bus line containing $|\mathcal{U}|$ stops into account. These stops are denoted as $1, 2, \ldots, |\mathcal{U}|$ and are indexed by $u$ or $v$. The physical connection between stops $u$ and $u+1$ is defined as segment $u$. The study time horizon $[0, T]$ is discretized into a finite number of time intervals $\mathcal{T} = \{t \mid 1, 2, \ldots, |\mathcal{T}|\}$, each of duration $\Delta$. When an FLS trip or a DRT service departs from stop $u$ at time $t$ and docks at both stops $u$ and $u+1$, the time-varying running time of trips and DRT services on segment $u$ is denoted as $\hat{r}_{u(u+1)t}$. In the studied intermodal urban transit system, both FLS and DRT services utilize MVs composed of MUs, where MUs can be used interchangeably for both FLS and DRT services. We denote the set of the number of MUs that can be contained in an MV as $\mathcal{M} = \{m \mid 1, 2, \ldots, |\mathcal{M}|\}$, and the capacity of each MU as $C$. Each operated MV has the flexibility to dynamically adjust its formation by coupling and decoupling MUs at each stop.

**Time-dependent and uncertain passenger demand.** It is widely acknowledged that passenger demand in public transportation networks exhibits time-dependent and inherent uncertainty. First, to model the uncertainty of passengers in practical operating environments, we construct a set of stochastic scenarios using historical data. Specifically, let $\mathcal{W} = \{1, 2, \ldots, w, \ldots, |\mathcal{W}|\}$ denote the set of stochastic scenarios, where the probability of each scenario $w \in \mathcal{W}$ is represented by $\rho(w)$. Then, we adopt the time-dependent origin–destination (OD) matrices in each scenario. Each element in these matrices represents a passenger group, denoted

as $p$, characterized by five attributes: $(n_p, O_p, D_p, AT_p, w)$. Here, $n_p$ represents the total of number of passengers in group $p$. In this group, passengers originate from stop $O_p$ at time $AT_p$ and travel to destination stop $D_p$ in scenario $w$. We denote the set of all passenger groups in scenario $w \in \mathcal{W}$ as $\mathcal{P}_w = \{p \mid 1, 2, \ldots, |\mathcal{P}_w|\}$.

**FLS services.** During the study time horizon, a total of $|\mathcal{I}|$ FLS trips are scheduled to operate from stops 1 to $|\mathcal{U}|$, of which the set is denoted as $\mathcal{I} = \{i \mid 1, 2, \ldots, |\mathcal{I}|\}$. Each FLS trip is performed by an MV comprising various MUs at different times and stops. It is noteworthy that each trip is not required to dwell at every stop. These trips are permitted to skip certain stops and employ variable formations to improve the matching of capacity and passenger demand with the uneven spatial and temporal distribution. The other benefit of employing the skip-stop strategy is that it contributes to a reduction in the running time on segments because the it eliminates the need for the acceleration process for starting or the deceleration process to stop.

In the context of FLS services, the decisions involve finding robust timetables across multiple scenarios and optimal vehicle schedules within each individual scenario, all subject to constraints such as the minimum headway, bounded shifting time and dwell times, and strict capacity limitations. The main reason of this bifurcation is that the timetable is released to passengers and is intended to be easily memorizable, thus necessitating robustness, while vehicle scheduling is more operationally focused and can be flexibly adjusted on a daily basis.

**DRT services.** The services provided by the DRT system is the ride-hailing services. Each DRT service is also executed by an MV comprising MUs. In the context of DRT services, decisions encompass when to depart from which stop, whether and when a DRT service arrives at a certain stop, the formation, and the vehicle scheduling of each service in each individual scenario. To ease the notation, we introduce a binary variable $\eta_{uvt}(w)$ for all $u, v \in \mathcal{U}, v > u, t \in \mathcal{T}, w \in \mathcal{W}$ to simultaneously represent whether a DRT service departs from stop $u$ at time interval $t$ in scenario $w$, and whether it arrives at stop $v$. Besides, we define the other binary variable $\pi_{uvtm}(w)$ to indicate whether the MV composed by $m$ MUs is allocated to execute the aforementioned DRT service.

**Flexible circulations of MUs between FLS and DRT services.** Thanks to the high degree of flexibility provided by MVs, in this paper we assume that each MU can flexibly serve these two transportation modes. Specifically, prior to operations, MUs are stored at the large depots at stops 1 and $|\mathcal{U}|$ as well as depots with a capacity $\hat{N}_u$ belongs to each stop $u \in \mathcal{U}\backslash\{1, |\mathcal{U}|\}$. During operations, MUs at depots can form MVs to execute FLS trips and DRT services. All operated MVs can be decoupled and coupled at each stop. The decoupled MUs at stop $u \in \mathcal{U}\backslash\{1, |\mathcal{U}|\}$ have three potential destinations: (1) they can be stored in the strictly capacity-constrained depot attached to this stop, i.e., storage, (2) to return empty to a preceding stop $v \in \mathcal{U}, v < u$ for storage, i.e., rebalancing, or immediate coupling to an MV operating at that stop, or (3) to directly perform a DRT service with a route covers stops $u$ and $v \in \mathcal{U}, v > u$, i.e., reallocation. The coupled MUs at stop $u$ may originate from one of following three sources. (1) They may be decoupled at a downstream stop $v > u$ and rebalanced to this stop; (2) they may come directly from the depot at this stop; or (3) they may have been decoupled from an MV previously assigned to a DRT service ending at this stop.

**Passengers' and operators' objectives.** For passengers, it is assumed that they always choose the first available MV that is capable of completing their journey, irrespective of the particular mode of transportation. To be specific, upon arrival at their originating stops, passengers wait either for an FLS trip that would stop at both their origin and destination stops, or a DRT service whose route includes their origin and destination stops. Passengers then board the first of the aforementioned services – be it FLS or DRT – that arrives at their origins. From the passenger's point of view, the expectation of the weighted sum of the waiting time and the in-vehicle time of all passenger groups among all scenarios $p \in \mathcal{P}_w, w \in \mathcal{W}$ during the journey needs to be minimized.

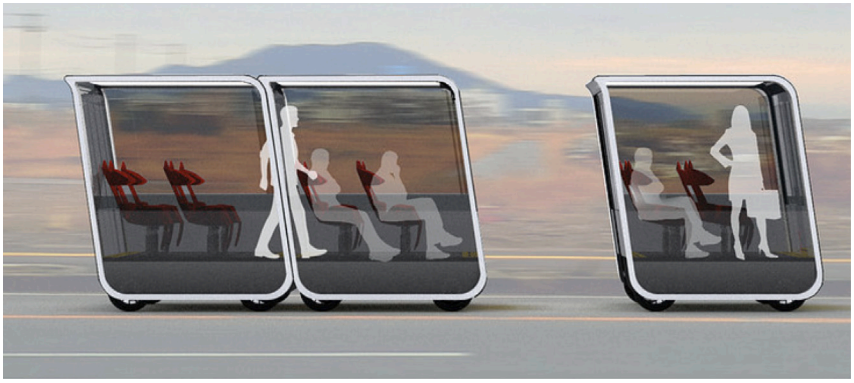
On the other hand, the goal of operators is to minimize operating costs. Operating costs include the usage of MUs to execute FSL trips and DRT services, as well as the rebalancing process of the decoupled MUs run empty back to upstream stops. Following Dai et al. (2020), operating costs of an FLS trip and a DRT service are related to the formation of an MV assigned to this trip or service on the segment between stops, i,e, the parameter $\varphi_{uvm}$ is defined in this study to represent operating costs of using an MV comprising $m$ MUs on the segment between stops $u$ and $v$.

Furthermore, we adopt the following assumptions to formulate the described problem. Firstly, it is assumed that both stops 1 and $|\mathcal{U}|$ are equipped with large depots, while each intermediate stop $u \in \mathcal{U}\backslash\{1, |\mathcal{U}|\}$ has its own depot with strict capacity limits for storing idle MUs. The second assumption is that the number of MUs are sufficient at depots so that there are always available MUs for dispatching, in accordance with Chen et al. (2020). Thirdly, following Xia et al. (2023), the time required for coupling and decoupling MUs is assumed to be negligible. Lastly, we assume that demand is aggregated.

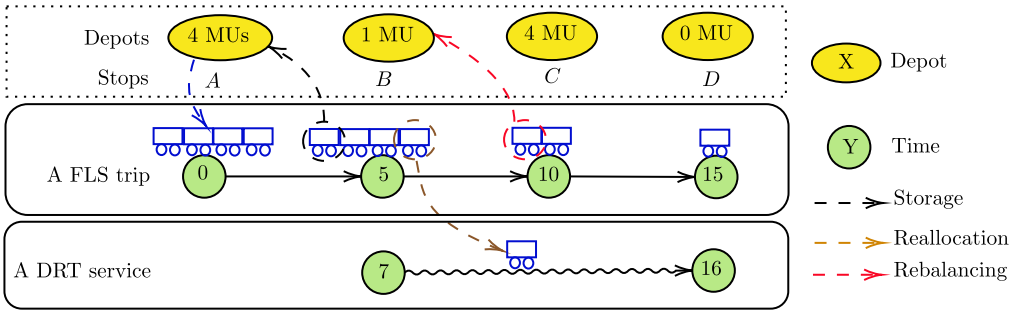
### 3.2. Problem illustrations

In this subsection, we illustrate the flexibility of circulations with respect to MUs and the TTVSP for an integrated FLS and DRT system utilizing MVs.

Fig. 2(a) showcases MUs engineered by Next Future Transportation Inc., providing a foundation for the flexible scheduling under investigation. Fig. 2(b) depicts the flexible circulations of these MUs, assuming a capacity of four MUs at each depot belonging to a stop. Initially, an MV, consisting of four MUs, departs from stop A and performs an FLS trip, which originates from the depot belonging to this stop. Upon reaching stop B, this MV decouples into two individual MUs and one MV with two MUs. The MV composed with two MUs continues the FLS trip, one MU is stored at stop B's depot, and the remaining MU is reallocated for a DRT service departing from stop B in the seventh time interval. The MV on the FLS trip takes five time intervals to reach stop C, where it decouples into two separate MUs. Due to the lack of available space in the depot affiliated with stop C, one MU is rebalanced back to the depot belonging to stop B, while the other proceeds to perform the FLS trip to stop D. Besides, the MU designated for DRT service directly reaches stop D after nine time intervals.



(a) Modular units with the capability of flexible circulations




(b) Flexible circulations of modular units

**Fig. 2.** An illustration of the modular units developed by Next Future Transportation Inc. (NEXT Future Transportation Inc., 2019) and the investigated flexible circulations.

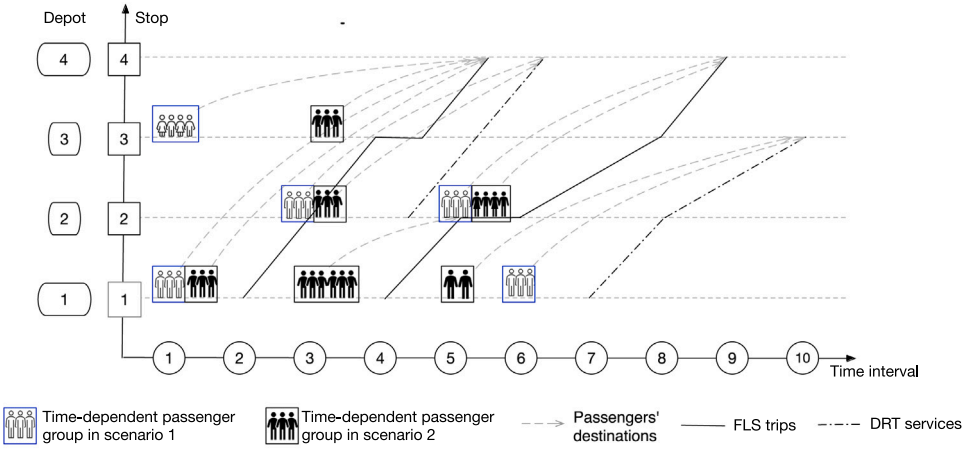
Next, an example is given to illustrate the problem explored in this study. As shown in Fig. 3(a), two scenarios with uncertain and time-dependent passenger groups are taken into consideration. For the sake of simplifying this illustration, each MU is limited to carrying a maximum of two passengers. Moreover, the running time on each segment is time-varying, e.g., for departures from stop 2 at the 3rd and 6th time intervals, the respective running times between stops 2 and 3 are 1 and 2 time intervals.

Fig. 3(a) depicts the robust timetable across two scenarios. Additionally, for passenger demand in scenarios 1 and 2, the vehicle schedules under this timetable are presented in Fig. 3(b) and (c), respectively. Let us take the first FSL trip and the second DRT service as examples. In the first scenario, there are three passengers who arrive at stop 1 at time interval 1 with a destination of stop 4. Consequently, the formation of the MV allocated to the first FLS trip is 2 MUs. According to the robust timetable, this trip skips stops 2 and docks at stop 3, where four waiting passengers are waiting. Therefore, two additional MUs stored in the depot attached to stop 3 are coupled to this MV, increasing its capacity to eight passengers when it runs from stops 3 to 4. Similarly, in the second scenario, three passengers arrive at stop 1 at time interval 1 and intend to reach stop 4. The MV assigned to the first FLS trip also consists of two MUs. This trip likewise skips stop 2 and arrives at stop 3, where there are three waiting passengers. In this case, a single additional MU is coupled to this MV.

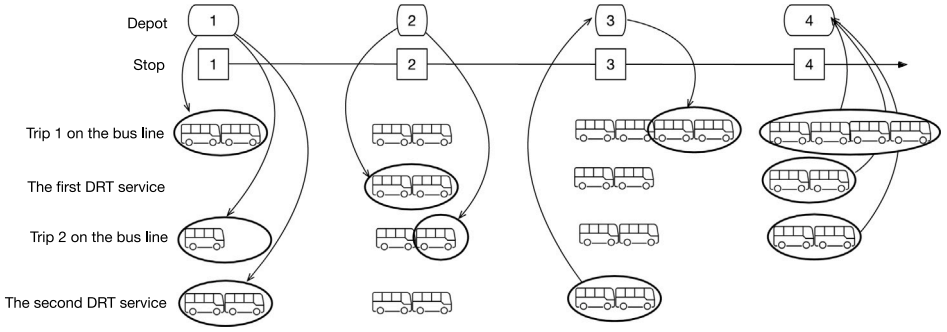
Regarding the second DRT service, there are discrepancies in the number of waiting passenger at stop 1 across the two scenarios: three passengers in the first scenario while two in the second. Consequently, the MV allocated to this DRT service consists of 2 MUs in the first scenario and a single MU in the second. In the first scenario, both MUs are sourced from the depot attached to stop 1. In contrast, the MU used in the second scenario comes from the MU decoupled from the MV which is initially assigned to the second FLS trip. Specifically, this MU is decoupled at stop 2, travels empty back to stop 1, and is subsequently utilized for the second DRT service.

## 4. Mathematical formulations

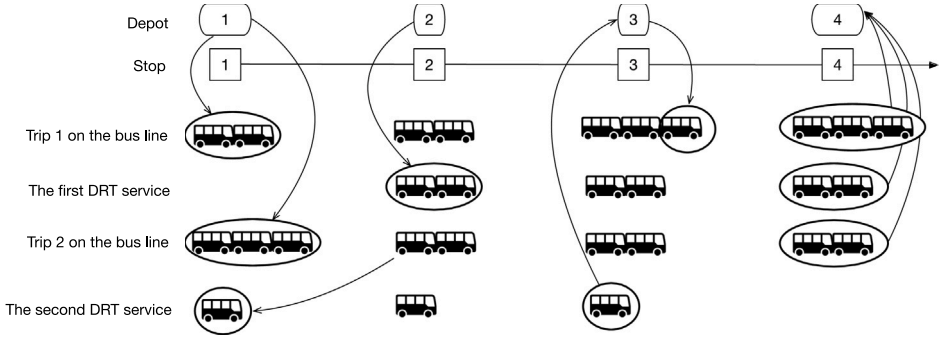
In order to solve the TTVSP, we first formulate the problem as a stochastic programming model which contains the majority of the notations, decision variables and constraints involved in the DRO version in Section 4.1. Subsequently, we propose a DRO model based on the assumption that the probability distribution information of the random parameters is partially known and further reformulate it into an MILP model in Section 4.2.




**(a)** Robust timetable.




**(b)** Vehicle circulations in scenario 1.




**(c)** Vehicle circulations in scenario 2.

**Fig. 3.** Problem illustration.

### *4.1. A stochastic programming formulation for the TTRDCP*

In this subsection, we intend to formulate an SP model by considering the time-dependent passenger demand and running times on segments, where the possibility distribution of stochastic parameters is fully known in advance.

**(1) Notations of sets, parameters, dependent variables and decision variables**

For the sake of modeling, we list all notations of sets, parameters and dependent variables used throughout this paper in Tables 7 and 8 in Appendix A. In addition, we define the following decision variables related to the timetabling of FLS trips, the vehicle scheduling of FLS trips and DRT services:

- $\iota_{i1}$: Shifting time of FLS trip $i$ at the first stop.
- $s_{iu}$: Dwell time of FLS trip $i$ at stop $u$.
- $y_{iu}$: A binary variable indicating whether FLS trip $i$ docks at stop $u$. If yes, $y_{iu} = 1$.
- $z_{iut}$: A binary variable representing whether FLS trip $i$ has departed from stop $u$. If yes, $z_{iut} = 0$.
- $e_{iut}$: A binary variable representing whether FLS trip $i$ has arrived at stop $u$. If yes, $e_{iut} = 0$.
- $x_{iu}(w)$: The number of MUs composed in the MV allocated to FLS trip $i$ serving the section between stops $u$ and $u+1$ in scenario $w$.
- $q_{ium}(w)$: If the number of MUs in the MV assigned to FLS trip $i$ at stop $u$ is $m$ under scenario $w$, $q_{ium}(w) = 1$; $q_{ium}(w) = 0$, otherwise.
- $\eta_{uvt}(w)$: If a DRT service is dispatched for stop $u$ to stop $v$ at time interval $t$ under scenario $w$, $\eta_{uvt}(w) = 1$; $\eta_{uvt}(w) = 0$, otherwise.
- $\pi_{uvtm}(w)$: If the MV composed by $m$ MUs is allocated a DRT service from stop $u$ to stop $v$ at time interval $t$ in scenario $w$, $\pi_{uvtm}(w) = 1$; $\pi_{uvtm}(w) = 0$, otherwise.

**(2) The objective function**

In this study, we aim to take into account both operators' and passengers' points of view. Specifically, we adopt operating costs as the evaluation index to obtain a reasonable operational plan from the operators' perspective. With respect to passengers, traveling costs including both waiting and in-vehicle costs is an important metric for assessing the quality of transportation services. Subsequently, we would introduce the deduction of the aforementioned objectives.

To be precise, traveling costs of passengers in each scenario can be formulated as follows:

$$Z_{pa}(w) = \varphi^T \cdot \Delta \cdot \sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}_w} (\xi_1 \cdot WT_p + \xi_2 \cdot TP_p) \cdot n_p, \tag{1}$$

where $\varphi^T$ represents the equivalent monetary value per unit of passengers' traveling time, $\Delta$ denotes the duration of each time unit, $\xi_1$ and $\xi_2$ are the weighting coefficients of waiting time $WT_p$ and in-vehicle time $TP_p$ with respect to passenger group $p$.

Secondly, operating costs of operators in each scenario comprise the costs of FLS trips, DRT services, and rebalancing. This can be formally expressed as follows:

$$Z_{op}(w) = \sum_{i\in\mathcal{L}} \sum_{u\in\mathcal{U}\backslash|\mathcal{U}|} \sum_{m\in\mathcal{M}} \varphi_{u(u+1)m} \cdot q_{ium}(w) + \sum_{u\in\mathcal{U}} \sum_{v\neq u, v\in\mathcal{U}} \sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} \varphi_{uvm} \cdot \pi_{uvtm}(w), \tag{2}$$

where $\varphi_{uvm}$ represents operating costs of using an MV comprising $m$ MUs on the segment between stops $u$ and $v$. In particular, if $v < u$, then $\varphi_{uvm}$ indicates operating costs associated with rebalancing MUs by sending the MUs decoupled at stop $u$ back to a previous stop $v$; if $v > u$, then $\varphi_{uvm}$ indicates costs of operating DRT services. The overall costs of operating MVs to complete all FLS trips are shown in the first term in Eq. (2), and the total costs of both operating DRT services and rebalancing MUs are calculated through the second term.

To facilitate the trade-offs between operators and passengers, the aforementioned objective functions are reformulated as a single one through the weighting coefficients $\zeta_{pa}$ and $\zeta_{op}$, shown as below:

$$OBJ_w(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \zeta_{pa} \cdot Z_{pa}(w) + \zeta_{op} \cdot Z_{op}(w). \tag{3}$$

Based on the assumption that the distribution of the possibility of each scenario is completely known, the expected value of the weighted sum of operating and passengers' costs can be formulated as an expected-value functional $\mathbb{E}_{\boldsymbol{\rho}}[\cdot]$:

$$\mathbb{E}_{\boldsymbol{\rho}}[\mathbf{OBJ}(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi})] = \sum_{w\in\mathcal{W}} \rho(w) OBJ_w(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}), \tag{4}$$

where $\boldsymbol{\rho} = (\rho(w_1), \rho(w_2), \ldots, \rho(w_{|\mathcal{W}|}))^T$, $\rho(w) \geq 0$ is the probability of scenario $w$ and $\sum_{w\in\mathcal{W}} \rho(w) = 1$.

**(3) Timetabling constraints related to FLS trips**

In the context of FLS trips, the timetable specifies when each trip arrives at and departs from each stop, whether it stops along with the dwell time. Establishing a robust timetable across various stochastic scenarios greatly enhances the scheduling operability. Specifically, the dynamics of FLS trips can be modeled through a system of linear constraints, as detailed below:

$$a_{iu} = \begin{cases} a_{i1} + \iota_{i1}, & \text{if } u = 1 \\ d_{i(u-1)} + r_{iu-1}, & \text{if } u \in \mathcal{U}\backslash\{1\} \end{cases} \qquad \forall i \in \mathcal{I}. \tag{5}$$

$$d_{iu} = a_{iu} + s_{iu} \qquad \forall i \in \mathcal{I}, u \in \mathcal{U}. \tag{6}$$

$$\underline{\iota} \leq \iota_{i1} \leq \bar{\iota} \qquad \forall i \in \mathcal{I}. \tag{7}$$

$$\underline{s}_{iu} y_{iu} \leq s_{iu} \leq \bar{s}_{iu} y_{iu} \qquad \forall i \in \mathcal{I}, u \in \mathcal{U}. \tag{8}$$

$$e_{iu(t+1)} \leq e_{iut} \qquad \forall i \in \mathcal{I}, u \in \mathcal{U}, t \in \mathcal{T}\backslash\{|\mathcal{T}|\}. \tag{9}$$

$$z_{iu(t+1)} \leq z_{iut} \qquad \forall i \in \mathcal{I}, u \in \mathcal{U}, t \in \mathcal{T}\backslash\{|\mathcal{T}|\}. \tag{10}$$

$$g_{iut} = \begin{cases} 1 - e_{iut} & \text{if } t = 1 \\ e_{iu(t-1)} - e_{iut} & \text{if } t \in \mathcal{T}\backslash\{1\} \end{cases} \quad \forall i \in \mathcal{I}, \forall u \in \mathcal{U}. \tag{11}$$

$$h_{iut} = \begin{cases} 1 - z_{iut} & \text{if } t = 1 \\ z_{iu(t-1)} - z_{iut} & \text{if } t \in \mathcal{T}\backslash\{1\} \end{cases} \quad \forall i \in \mathcal{I}, \forall u \in \mathcal{U}, \tag{12}$$

$$a_{iu} = \Delta(1 + \sum_{t \in \mathcal{T}} e_{iut}) \quad \forall i \in \mathcal{I}, u \in \mathcal{U}. \tag{13}$$

$$d_{iu} = \Delta(1 + \sum_{t \in \mathcal{T}} z_{iut}) \quad \forall i \in \mathcal{I}, u \in \mathcal{U}. \tag{14}$$

$$r_{iu} = R_{iu} - r_{iu}^{acc}(1 - y_{iu}) - r_{i(u+1)}^{dec}(1 - y_{i(u+1)}) \quad \forall i \in \mathcal{I}, u \in \mathcal{U}\backslash\{|\mathcal{U}|\}. \tag{15}$$

$$R_{iu} = \sum_{t \in \mathcal{T}} \hat{r}_{u(u+1)t} \cdot h_{iut} \quad \forall i \in \mathcal{I}, u \in \mathcal{U}\backslash\{|\mathcal{U}|\}. \tag{16}$$

$$d_{(i+1)u} - d_{iu} \geq \underline{h} \quad \forall i \in \mathcal{I}\backslash\{|\mathcal{I}|\}, u \in \mathcal{U}. \tag{17}$$

$$y_{i1} + y_{i|\mathcal{U}|} = 2 \quad \forall i \in \mathcal{I}. \tag{18}$$

$$e_{iut}, z_{iut} \in \{0, 1\} \quad \forall i \in \mathcal{I}, u \in \mathcal{U}, t \in \mathcal{T}. \tag{19}$$

$$y_{iu} \in \{0, 1\} \quad \forall i \in \mathcal{I}, u \in \mathcal{U}. \tag{20}$$

Constraints (5) and (6) are proposed to capture the arrival and departure times of the MV assigned to trip $i$ at stop $u$, respectively. Constraints (7) ensure that the shifting time at the first stop of each trip remains within realistic upper and lower thresholds. The decision variable $s_{iu}$, which represents the dwell time of the MV assigned to trip $i$ at stop $u$ is intrinsically linked with the decision variable $y_{iu}$, that is, whether this stop is skipped or not. Constraints (8) ensure that if stop $u$ is skipped by trip $i$, then $y_{iu} = 0$ and $s_{iu} = 0$; otherwise, $y_{iu} = 1$ and the value of $s_{iu}$ is adjusted within the given upper and lower limitations. Motivated by Niu and Zhou (2013), we introduce four types of binary indicators to facilitate the linear modeling of the coupling between movements of MVs assigned to trips and passenger dynamics. We propose the FLS trip arrival and departure indicators, $e_{iut}$ and $z_{iut}$, respectively, to represent whether an MV assigned to trip $i$ arrives at or departs from stop $u$ at time interval $t$. Furthermore, we define $g_{iut} = 1$ (or $h_{iut} = 1$) to precisely indicate the arrival (or departure) of the FLS trip at time $t$, with a value of 0 otherwise. Constraints (9) and (10) are formulated to ensure the non-increasing property of these two binary variables. Constraints (11) and (12) track the exact time intervals of arrival and departure times, respectively. For clarity, we provide a straightforward illustration of the four variables through Fig. 4. Constraints (13) and (14) are imposed to track the real-valued arrival and departure times. Constraints (15) ensure the running time from stops $u$ to $u+1$ of trip $i$ is closely coupled with the its stopping pattern. To be specific, on the segment between stops $u$ and $u+1$, if trip $i$ skips either stop $u$ or $u+1$, the acceleration or deceleration phase when departing from stop $u$ or arriving at stop $u+1$ is replaced by the cruise phase, thereby yielding a reduction in the running time on this segment. Constraints (16) establish a relationship between the time-varying running time on the segment from stops $u$ to $u+1$ at time $t$ (hereinafter denoted as $\hat{r}_{u(u+1)t}$) and the running time of the MV assigned to trip $i$ on the same segment. Constraints (17) guarantee that the departure headway between two consecutive trips meets a minimum restriction in order to ensure safe operations. Constraints (18) ensure that the MV assigned to each trip has to dock at the first and last stops. Constraints (19) and (20) define the domain of binary decision variables related to the timetabling of FLS trips.

**(4) Coupling constraints among passenger mobility, vehicle movements and vehicle scheduling within the intermodal urban transit network**

In practice, a passenger waiting at a stop would choose the first available MV capable of completing their journey, irrespective of the specific mode of transportation. Additionally, the routing and departure times of DRT services are tightly linked with passenger dynamics. To model the couplings among passenger mobility, inter-modal vehicle movements, and vehicle scheduling within the intermodal urban transit network, we adopt the constraints (21)–(33).

$$M(\vartheta_{pi}^{\mathrm{FLS}}(w) - 1) \leq d_{iu} \cdot y_{iu} \cdot y_{iv} - AT_p \leq M \cdot \vartheta_{pi}^{\mathrm{FLS}}(w) - \epsilon \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, v = D_p. \tag{21}$$

$$M(\vartheta_{puvt}^{\mathrm{DRT}}(w) - 1) \leq t \cdot \eta_{uvt}(w) - AT_p \leq M \cdot \vartheta_{puvt}^{\mathrm{DRT}}(w) - \epsilon \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p. \tag{22}$$

$$\varsigma_{pi}^{\mathrm{FLS}}(w) = (d_{iu} - AT_p)\vartheta_{pi}^{\mathrm{FLS}}(w) + M(1 - \vartheta_{pi}^{\mathrm{FLS}}(w)) \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p. \tag{23}$$

$$\varsigma_{puvt}^{\mathrm{DRT}}(w) = (t \cdot \eta_{uvt}(w) - AT_p)\vartheta_{puvt}^{\mathrm{DRT}}(w) + M(1 - \vartheta_{puvt}^{\mathrm{DRT}}(w)) \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p. \tag{24}$$

$$WT_p = \min\{\varsigma_{p1}^{\mathrm{FLS}}(w), \varsigma_{p2}^{\mathrm{FLS}}(w), \ldots, \varsigma_{p|\mathcal{L}|}^{\mathrm{FLS}}(w), \varsigma_{puv1}^{\mathrm{DRT}}(w), \ldots, \varsigma_{puv|\mathcal{T}|}^{\mathrm{DRT}}(w)\} \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p, v = D_p. \tag{25}$$

$$TP_p = \sum_{i \in \mathcal{I}} \varphi_{pi}^{\mathrm{FLS}}(w)(a_{iv} - d_{iu}) + \sum_{t \in \mathcal{T}} \hat{r}_{uvt}\varphi_{puvt}^{\mathrm{DRT}}(w) \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p. \tag{26}$$

$$\sum_{i \in \mathcal{I}} \varphi_{pi}^{\mathrm{FLS}}(w) + \sum_{t \in \mathcal{T}} \varphi_{puvt}^{\mathrm{DRT}}(w) = 1 \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p, v = D_p. \tag{27}$$

$$\varsigma_{pi}^{\mathrm{FLS}}(w) \geq WT_p - M \cdot \varphi_{pi}^{\mathrm{FLS}}(w) + \epsilon_3 \quad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}. \tag{28}$$

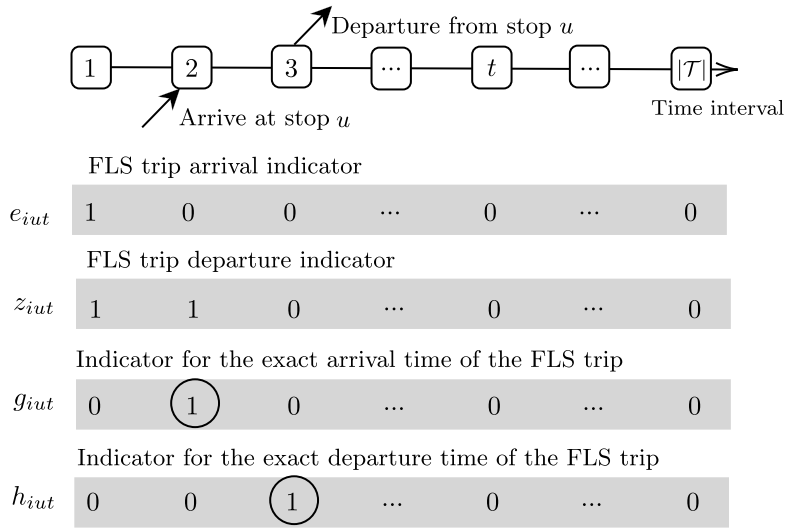



**Fig. 4.** Illustration of 0-1 binary variables related to the timetabling of FLS trips.

$$\varsigma_{pi}^{\mathrm{FLS}}(w) \le WT_p + M(1-\varphi_{pi}^{\mathrm{FLS}}(w)) \qquad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}. \tag{29}$$

$$\varsigma_{puvt}^{\mathrm{DRT}}(w) \ge WT_p - M \cdot \varphi_{puvt}^{\mathrm{DRT}}(w) + \epsilon_4 \qquad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u, v \in \mathcal{U}, v > u, t \in \mathcal{T}. \tag{30}$$

$$\varsigma_{puvt}^{\mathrm{DRT}}(w) \le WT_p + M(1-\varphi_{puvt}^{\mathrm{DRT}}(w)) \qquad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u, v \in \mathcal{U}, v > u, t \in \mathcal{T}. \tag{31}$$

$$\varphi_{pi}^{\mathrm{FLS}}(w) \in \{0,1\} \qquad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{I}. \tag{32}$$

$$\varphi_{puvt}^{\mathrm{DRT}}(w) \in \{0,1\} \qquad \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u, v \in \mathcal{U}, v > u, t \in \mathcal{T}. \tag{33}$$

Constraints (21)–(31) ensure that a passenger boards the MV when the first MV assigned to FLS trips or DRT services that can transport them to their destinations arrives. For an MV allocated to FLS trip $i$, constraints (21) specify two necessary conditions for determining whether it has the potential to serve passenger group $p$: (1) the departure time of this trip, denoted as $d_{iu}$, must precede the arrival time of passenger group $p$ (denoted as $AT_p$), and (2) this trip must stop at both the origin stop $u$ and destination stop $v$ of passenger group $p$. Here, the variable $\vartheta_{pi}^{\mathrm{FLS}}(w) = 1$ if the MV assigned to FLS trip $i$ can potentially transport passenger group $p$ in scenario $w$, and 0 otherwise. For an MV allocated to a DRT service departing from stop $u$ at time $t$, constraints (22) impose analogous conditions to ensure that this service may transport passenger group $p$ only if the departure time of this service is earlier than $AT_p$ and the route of this service covers stops $u$ and $v$. The variable $\vartheta_{puvt}^{\mathrm{DRT}}(w)$ equals 1 if a DRT service, starting operations at time $t$ and serving stops $u$ and $v$, meet the aforementioned conditions for transporting passenger group $p$ in scenario $w$, and 0 otherwise. Constraints (23) and (24) are formulated to calculate the waiting time of passenger group $p$ prior to boarding an MV assigned to a FLS trip or a DRT service, respectively. Constraints (25) ensure that the actual waiting time of passenger group $p$ is equal to the minimum of their waiting time across all feasible options of FLS trips or DRT services. Constraints (26) is formulated to calculate the in-vehicle time experienced by passenger group $p$, which is dependent on the used transportation mode. Here, variables $\varphi_{pi}^{\mathrm{FLS}}(w) = 1$ or $\varphi_{puvt}^{\mathrm{DRT}}(w) = 1$ if the passenger group $p$ boards an MV assigned to a FLS trip or a DRT service in scenario $w$, and 0 otherwise. The decision variable $\eta_{uvt}(w) = 1$ if the DRT service with a route that covers stops $u$ and $v$ is $m$ in scenario $w$ departs from stop $u$ at time $t$, and it is set to 0 otherwise. Constraints (27) ensure the passenger group $p$ is served through one mode. Constraints (28)–(31) are established to track the specific MV assigned to a FLS trip or a DRT service that passenger $p$ boards. Lastly, constraints (32) and (33) provide the range of the decision variables.

Next, considering that capacity of such inter-modal transportation system needs to be temporally and spatially matched to passenger demand, we develop the following constraints (34)–(39). These constraints capture various processes of passenger mobility, including waiting at stops, boarding and alighting MVs assigned to FLS trips and DRT services, the number of in-vehicle passengers, and dynamic-capacity allocations subject to the strict capacity limitations. The detailed formulations reads as follows:

$$b_{iu}^{\mathrm{FLS}}(w) = \sum_{p \in \mathcal{P}_w : O_p = u} n_p \cdot \varphi_{pi}^{\mathrm{FLS}}(w) \qquad \forall i \in \mathcal{I}, u \in \mathcal{U}, w \in \mathcal{W}. \tag{34}$$

$$b_{uvt}^{\mathrm{DRT}}(w) = \sum_{p \in \mathcal{P}_w : O_p = u, D_P = v} n_p \cdot \varphi_{puvt}^{\mathrm{DRT}}(w) \qquad \forall u, v \in \mathcal{U}, v > u, t \in \mathcal{T}, w \in \mathcal{W}. \tag{35}$$

$$o_{iu}^{\mathrm{FLS}}(w)=\begin{cases} b_{iu}^{\mathrm{FLS}}(w) & \text{if } u=1 \\ o_{i(u-1)}^{\mathrm{FLS}}(w)-l_{iu}^{\mathrm{FLS}}(w)+b_{iu}^{\mathrm{FLS}}(w) & \text{if } u\in\mathcal{U}\backslash\{1,|\mathcal{U}|\} \\ 0 & \text{if } u=|\mathcal{U}| \end{cases} \quad \forall i\in\mathcal{I},w\in\mathcal{W}. \tag{36}$$

$$l_{iu}^{\mathrm{FLS}}(w)=\sum_{p\in\mathcal{P}_{iw}:D_p=u} n_p\cdot\varphi_{pi}^{\mathrm{FLS}}(w) \quad \forall i\in\mathcal{I},u\in\mathcal{U},w\in\mathcal{W}. \tag{37}$$

$$o_{iu}^{\mathrm{FLS}}(w)\le C\cdot x_{iu}(w) \quad \forall i\in\mathcal{I},u\in\mathcal{U},w\in\mathcal{W}. \tag{38}$$

$$b_{uvt}^{\mathrm{DRT}}(w)\le C\cdot\sum_{m\in\mathcal{M}} m\cdot\pi_{uvtm}(w) \quad \forall u,v\in\mathcal{U},v>u,t\in\mathcal{T},w\in\mathcal{W}. \tag{39}$$

To be specific, constraints (34) and (35) calculate the number of passengers served by different transportation modes at each stop in each scenario, respectively. Constraints (36) is formulated to model the number of in-vehicle passengers on the MV assigned to FLS trip $i$ when it departs from stop $u$ in scenario $w$. Constraints (37) are used to compute the number of passengers alighting the MV assigned to FLS trip $i$ at stop $u$ in scenario $w$. Constraints (38) limit the maximum allowed number of in-vehicle passengers in the MV assigned to FLS trip $i$ in each scenario, where the maximum allowable number being related to the formation of that MV (denoted as $x_{iu}(w)$). Lastly, constraints (39) ensure that the number of passengers loaded on a DRT service does not exceed its capacity in each scenario. Here, the decision variable $\pi_{uvtm}(w)=1$ if the number of MUs composing a DRT service with a beginning time of $t$ and a route that covers stops $u$ and $v$ is $m$ in scenario $w$, and 0 otherwise.

**(5) Constraints related to flexible vehicle circulations between different transportation modes**

In this study, it is assumed that an MV assigned to each FLS trip and each DRT service can be decoupled and coupled at each stop. The decoupled MUs have three destinations: (1) to be stored in a strictly capacity-constrained depot attached to that stop, (2) to return empty to one of the preceding stops and be stored in its depot, or to be directly coupled to an MV operating to that stop, or perform a DRT service at that stop, and (3) to perform a DRT service at this stop directly. Such highly flexible vehicle circulations both within the same transportation mode and between different modes can be formulated as follows:

$$N_{ut}(w)=\begin{cases} N_u+AV_{ut}(w)-DV_{ut}(w) & \text{if } t=1 \\ N_{u(t-1)}(w)+AV_{ut}(w)-DV_{ut}(w) & \text{if } t\in\mathcal{T}\backslash\{1\} \end{cases} \quad \forall w\in\mathcal{W},2\le u\le|\mathcal{U}|-1,t\in\mathcal{T}. \tag{40}$$

$$AV_{ut}(w)=\sum_{i\in\mathcal{L}} x_{i(u-1)}(w)g_{iut}+\sum_{v\neq u,v\in\mathcal{U}}\sum_{m\in\mathcal{M}} m\pi_{vu(t-\hat{r}_{uvt})m}(w) \quad \forall w\in\mathcal{W},2\le u\le|\mathcal{U}|-1,t\in\mathcal{T}. \tag{41}$$

$$DV_{ut}(w)=\sum_{i\in\mathcal{L}} x_{iu}(w)h_{iut}+\sum_{v\neq u,v\in\mathcal{U}}\sum_{m\in\mathcal{M}} m\pi_{uvtm}(w) \quad \forall w\in\mathcal{W},2\le u\le|\mathcal{U}|-1,t\in\mathcal{T}. \tag{42}$$

$$\sum_{m\in\mathcal{M}} q_{ium}(w)=1 \quad \forall w\in\mathcal{W},i\in\mathcal{I},u\in\mathcal{U}. \tag{43}$$

$$x_{iu}(w)=\sum_{m\in\mathcal{M}} m\cdot q_{ium}(w) \quad \forall w\in\mathcal{W},i\in\mathcal{I},u\in\mathcal{U}. \tag{44}$$

$$\eta_{uvt}(w)=\sum_{m\in\mathcal{M}}\pi_{uvtm}(w) \quad \forall w\in\mathcal{W},u,v\in\mathcal{U},v\neq u,t\in\mathcal{T}. \tag{45}$$

$$0\le N_{ut}(w)\le\hat{N}_u \quad \forall w\in\mathcal{W},u\in\mathcal{U},t\in\mathcal{T}. \tag{46}$$

$$x_{iu}(w)\in\{0,1\} \quad \forall i\in\mathcal{I},w\in\mathcal{W},u\in\mathcal{U},t\in\mathcal{T}. \tag{47}$$

$$q_{ium}(w)\in\{0,1\} \quad \forall i\in\mathcal{I},w\in\mathcal{W},u\in\mathcal{U},m\in\mathcal{M}. \tag{48}$$

$$\eta_{uvt}(w)\in\{0,1\} \quad \forall w\in\mathcal{W},u,v\in\mathcal{U},v\neq u,t\in\mathcal{T}. \tag{49}$$

Constraints (40) are proposed to compute the number of MUs stored in each depot attached to each stop at time $t$, where $N_u$, $AV_{ut}(w)$, $DV_{ut}(w)$, and $R_{ut}(w)$ represent the initial storage quantity, the inflow of MUs at time $t$, the outflows at time $t$, and the number of MUs stored in this depot at time $t$, respectively. Constraints (41) and (42) track the number of MUs arriving at and departing from the depot attached to stop $u$ at time $t$ during the study period, which are highly linked with the timetable of FLS trips and the dynamics of DRT services. Constraints (43) ensure that the formation of the MV assigned to each trip is unique in each scenario. Constraints (44) are used to calculate the number of MUs comprising the MV assigned to FLS trip $i$ at stop $u$. Constraints (45) are imposed to determine whether the DRT service with a route covering stops $u$ and $v$ is dispatched at time $t$. Constraints (46) ensure that the available MUs stored at the depot attached to stop $u\in\mathcal{U}$ is greater than or equal to 0 and not exceeding the capacity limitation (denoted as $\hat{N}_u$) at each time interval. Constraints (47) and (48) formulted the domin of decision variables $\mathbf{x}$ and $\mathbf{q}$. Constraints (49) provide the domain of the decision variable $\eta_{uvt}(w)$, which equals to 1 if a DRT service serving stops $u$ and $v$ in scenario $w$ departs from stop $u$ at time $t$, and 0 otherwise.

With the aforementioned analyses, the entire SP formulation for the TTVSP of an intermodal urban transit network, which is denoted as the model SP-TTVSP, reads as the following MINLP model:

$$\begin{cases} \min & \displaystyle\sum_{w\in\mathcal{W}}\rho(w)OBJ_w(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi}) \\ \text{s.t.} & \text{Constraints (1)–(49).} \end{cases} \tag{50}$$

**Remark 1.** Constraints (21), (23)–(26), (41)–(42) in the model (50) are all nonlinear. To derive a formulation with stronger mathematical properties, these constraints are linearized in Appendix B. Based on these linear constraints, an equivalent MILP model (51) is proposed, which is expressed as follows:

$$[\mathrm{SP-TTVSP}] \quad \begin{cases} \min & \sum_{w\in\mathcal{W}} \rho(w) OBJ_w(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}) \\ \text{s.t.} & \text{Constraints (1)–(20), (22), (27)–(40), (43)–(49), (58)–(62).} \end{cases} \tag{51}$$

### 4.2. A distributionally robust formulation with probability uncertainty

On the basis of the notations and constraints discussed in the previous sections, we formally present a distributionally robust model for the TTVSP in Section 4.2.1. This DRO model, which is not computationally tractable, takes into consideration uncertain probability distributions related to various scenarios. Subsequently, in Section 4.2.2, we design a discrepancy-based ambiguity set and reformulate the proposed DRO formulation as an MILP model to overcome its computational intractability.

#### 4.2.1. The DRO model

Considering the uncertainty inherent in the probability distributions of scenarios in practice, we are now ready to present a DRO formulation in which optimal solutions are evaluated under the worst-case expectation, as shown below:

$$\begin{cases} \min\limits_{\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi}} & \left\{ \max\limits_{\rho\in\mathscr{P}} \mathbb{E}_{\rho}\left[ \mathbf{OBJ}(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}) \right] \right\} \\ \text{s.t.} & \text{Constraints (1)–(3), (5)–(49),} \end{cases} \tag{52}$$

where $\mathscr{P}$ denotes the ambiguity set containing all possible probability distributions, and $\max_{\rho\in\mathscr{P}}[\cdot]$ represents the worst-case evaluation with respect to a family of probability distributions of the uncertain parameters.

**Remark 2.** If $\mathscr{P}$ contains only the true distribution of the uncertain vector $\boldsymbol{\rho}$, the DRO model (52) reduces to the SP model. On the other hand, if $\mathscr{P}$ contains all probability distributions on the support of the random vector $\boldsymbol{\rho}$, the DRO model (52) turns to the RO model.

As pointed out in Rahimian and Mehrotra (2022), Problem (52) is a semi-infinite program (SIP) which cannot be solved directly with numerical methods. Hence, a key step of the solution method is to handle the semi-infinite qualifier (i.e., $\forall \boldsymbol{\rho} \in \mathscr{P}$). A promising method in literature to reformulate the computationally intractable terms into forms that are amenable to existing optimization techniques is to design a proper ambiguity set and employ the dual method. In addition, the complexity and tractability of such SIPs and their subsequent reformulations are contingent upon the geometric characteristics and intrinsic properties of the ambiguity set. Next, we detail the designed ambiguity set and the reformulation of Problem (52).

#### 4.2.2. The ambiguity set and the reformulation of the DRO model

Consider a neighborhood of the nominal probability distribution by allowing certain perturbations around it is a natural way to hedge against the distributional ambiguity because it is easy to have a nominal estimate of the probability distribution in practice. Hence, in this study, we use the *discrepancy-based ambiguity set*, which has the following form:

$$\mathscr{P} = \left\{ \boldsymbol{\rho} = \boldsymbol{\rho}_0 + \boldsymbol{\varrho}\boldsymbol{\varepsilon} \,\middle|\, \mathbf{e}^T \boldsymbol{\varrho}\boldsymbol{\varepsilon} = 0, \boldsymbol{\rho}_0 + \boldsymbol{\varrho}\boldsymbol{\varepsilon} \geq 0, \|\boldsymbol{\varepsilon}\|_1 \leq 1 \right\}, \tag{53}$$

where $\boldsymbol{\rho}$ denote the unknown true probability distribution and $\boldsymbol{\rho}_0$ represents the known nominal probability distribution, i.e., the most likely probability distribution in the estimation. Besides, $\boldsymbol{\varrho}$ is the scaling matrix, $\boldsymbol{\varepsilon}$ is the perturbation vector with respect to the nominal probability distribution, and $\|\boldsymbol{\varepsilon}\|_1$ equals to $\sum_{w\in\mathcal{W}} |\varepsilon(w)|$. In addition, the conditions $\mathbf{e}^T\boldsymbol{\varrho}\boldsymbol{\varepsilon} = 0$ is proposed to guarantee the sum of the probability $\boldsymbol{\rho}$ equals to 1; and $\boldsymbol{\rho}_0 + \boldsymbol{\varrho}\boldsymbol{\varepsilon} \geq 0$ ensures the nonnegative property of the probability distribution.

**Proposition 1.** *Under the proposed discrepancy-based ambiguity set $\mathscr{P}$ and the linear constraints deduced in Appendix B, the computationally tractable form of the DRO model (52), i.e., the equivalent MILP model, can be expressed as follows:*

$$[\mathrm{DRO-TTVSP}] \quad \begin{cases} \min\limits_{\boldsymbol{\theta},\upsilon,\zeta,\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi}} & \mathbf{OBJ}(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi})^T \boldsymbol{\rho}_0 + \boldsymbol{\theta}^T \boldsymbol{\rho}_0 + \upsilon \\ \text{s.t.} & \|\boldsymbol{\varrho}^T \mathbf{OBJ}(\boldsymbol{\iota}, \mathbf{s}, \mathbf{y}, \mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}) + \boldsymbol{\varrho}^T \boldsymbol{\theta} + \boldsymbol{\varrho}^T e\zeta\|_\infty \leq \upsilon, \\ & \boldsymbol{\theta} \geq 0, \upsilon \geq 0, \\ & \text{constraints (1)–(3), (5)–(20), (22), (27)–(40),} \\ & \qquad \text{(43)–(49), (58)–(62),} \end{cases} \tag{54}$$

*where* $(\boldsymbol{\theta}, \upsilon, \zeta) \in R^{|\mathcal{W}|} \times R \times R$.

**Proof.** See Appendix C. □

**Lemma 1.** *When $\rho = \rho_0$, that is, the value of each element in $\epsilon$ equals to 0, the discrepancy-based ambiguity set reduces to the singleton that only includes the nominal probability distribution, and thus, the DRO model (54) defined over the set $\mathscr{P}$ reduces to the stochastic programming model.*

## 5. Solution methodologies

It worth noting that the proposed linear programs, i.e., SP model (51) and DRO model (54), possess an exponential number of variables, which results in a significant computational intensity for solving real-world instances. For example, we attempted to directly use GUROBI to solve a case study based on the data of the Beijing Bus Line 468, but after 4 h of computation, no feasible solution was found. Motivated by these observations, to find high-quality solutions within an acceptable computing time, this section focuses on the design of a tailored decomposition method and an effective algorithm. To this end, we first decompose the proposed models into two subproblems and propose the solution framework that combine a tailored Adaptive Large Neighborhood Search (ALNS) algorithm with an MILP solver in Section 5.1. Subsequently, we introduce the detailed procedure of the developed ALNS algorithm in Section 5.2.

### 5.1. Decomposition of models and design of the solution framework

In the proposed models, both robust timetables and scenario-related vehicle schedules are jointly optimized. To solve models effectively, a fundamental idea is to break the tight coupling between these two problems. Note that the timetabling problem involves a substantial quantity of binary decision variables and is coupled with the other problem, thereby posing computational challenges. For instance, as the problem size grows, the variables $z_{iut}$ and $e_{iut}$ related to timetables would increase exponentially. In light of this complexity, we discern that the original problem (54) can be divided into two subproblems. The first subproblem determines binary variables $\mathbf{y}, \mathbf{e}, \mathbf{z}$ as well as integer variables $\boldsymbol{\iota}$, $\mathbf{s}$, which are all related to the robust timetabling problem considering the skip-stop strategy. The second one optimizes the vehicle scheduling within the intermodal urban transit system given $\mathbf{y}, \mathbf{e}, \mathbf{z}, \boldsymbol{\iota}, \mathbf{s}$ as input parameters. Clearly, when treating timetabling variables as parameters, the couplings—both between timetables and vehicle schedules, and between timetables and passenger movements—are degraded in the second subproblem. As a result, the second subproblem becomes an MILP model with a relatively smaller scale of mixed-integer variables and constraints.

For clarity reasons, we next express the second subproblem as follows.

$$\left[\mathrm{P}(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}})\right] \quad \begin{cases} \min\limits_{\boldsymbol{\theta}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}, \upsilon, \zeta} & \mathbf{OBJ}(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi})^T \boldsymbol{\rho}_0 + \boldsymbol{\theta}^T \boldsymbol{\rho}_0 + \upsilon \\ \text{s.t.} & \|\boldsymbol{\rho}^T \mathbf{OBJ}(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}}, \mathbf{x}, \mathbf{q}, \boldsymbol{\eta}, \boldsymbol{\pi}) + \boldsymbol{\rho}^T \boldsymbol{\theta} + \boldsymbol{\rho}^T e\zeta\|_\infty \leq \upsilon, \\ & \boldsymbol{\theta} \geq 0, \upsilon \geq 0, \\ & \text{constraints (22), (27)–(40), (43)–(49), (58)–(62).} \end{cases} \tag{55}$$

As variables related to timetabling problem are fixed as parameters $\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}}$, the second subproblem $\mathrm{P}(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}})$ can be efficiently solved to optimality. Given this, we design an iterative solution framework to solve the original problem (54) effectively. Specifically, the solution framework starts with heuristically finding a good timetable of FLS trips at each iteration while concurrently generating and dynamically updating a pool of candidate DRT services. This timetable and the DRT pool subsequently serve as input parameters for the second subproblem. Then, an MILP solver is adopted to find the optimal vehicle schedules within the integrated FLS and DRT system. The optimal objective value at each iteration functions as an evaluation metric to guide the subsequent searching process in the next iteration. With this solution method, it is expected that the studied integrated optimization problem can be approximately solved with a rapid computational speed. For clarity, the overall flowchart of our developed solution framework, which is denoted as ALNS + GUROBI, is given in Fig. 5. To quickly find an approximate optimal solution, we design a problem-based ALNS algorithm for the timetabling process. The goals of this algorithm are two-fold: (1) to search the possible neighbors of candidate solution $(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}})$; and (2) to dynamically update the pool of DRT services through removing redundant options and adding new DRT services with potential to reduce the objective values. The specific procedures of the customized ALNS algorithm are detailed in the following discussions.

### 5.2. A tailored ALNS algorithm

In this section, we aim to introduce the detailed techniques of the tailored ALNS algorithm for solving the subproblem related to timetabling, including the initial solution generation, destroy operators, repair operators, adaptive searching strategy, and termination criteria. The destroy operators are use to explore search directions that possibly improve the performance of the solution, and repair operators aim to fix infeasible solutions. In addition, to avoid falling into local optima, we incorporate the simulated annealing into the ALNS framework to change the search neighborhood.

**(1) Initial solution generation.** Finding an appropriate initial solution quickly as the beginning point of the search is critical in designing an effective ANLS algorithm. In our problem-based ALNS algorithm, we adopt a timetable for FLS trips with uniform headway as the initial solution, which is commonly employed by bus operators in practice. Furthermore, it is crucial to draw attention that the number of decision variables associated with the DRT services, i.e., $\pi_{uvtm}(w)\ \forall i \in \mathcal{I}, u, v \in \mathcal{U}, v > u, t \in \mathcal{T}, w \in \mathcal{W}$, is $\frac{|\mathcal{U}|\times(|\mathcal{U}|-1)\times|\mathcal{T}|\times|\mathcal{M}|\times|\mathcal{W}|}{2}$, is large in real-case instances. This exponential growth in dimensionality critically impedes solution efficiency. Besides, after conducting extensive numerical experiments, we find that the number of operated DRT services is limited in

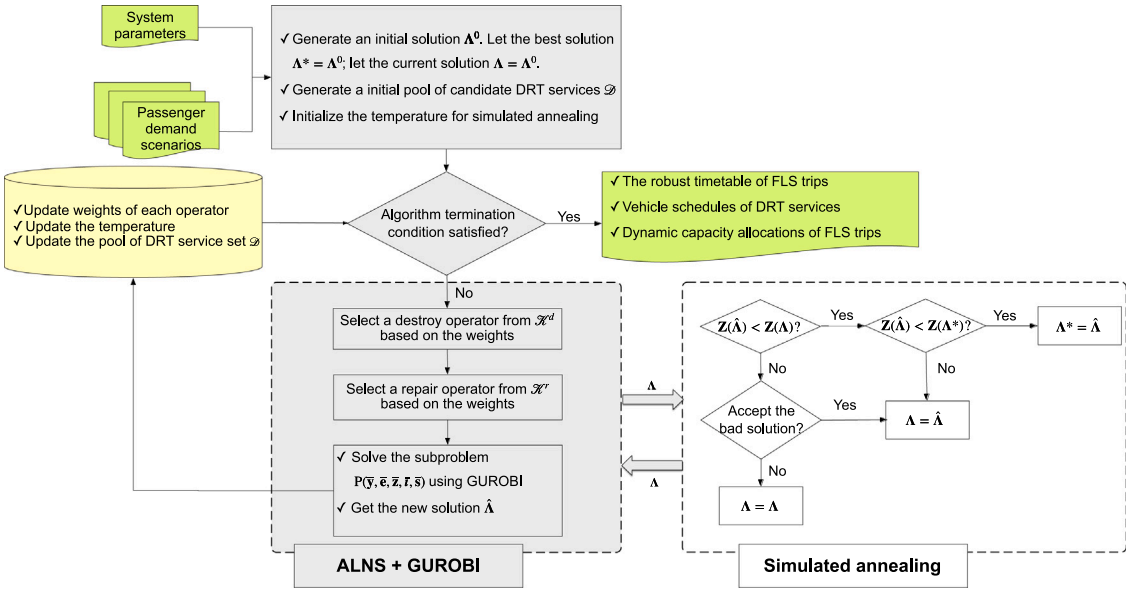



**Fig. 5.** Overall flowchart of the problem-based algorithm combining ALNS and GUROBI.

the optimal solution. Consequently, at initialization, we only incorporate a limited subset of potential DRT services into a DRT pool $D$. This subset is dynamically updated in subsequent iterations by introducing new and removing redundant DRT services through the following destroy operators.

**(2) Destroy operators related to timetabling of FLS trips.** We design the following four destroy operators to generate candidate timetables of FLS trips. The goal of utilizing destroy operators is to expand the neighborhood of the candidate solution, thus facilitating the escape from the local optimum.

**Destroy operator 1: Greedy search for the shifting plan of FLS trips at the first stop.** This operator aims to find good shifting plans with lower objective function values by adopting the Constrained Compass Search Algorithm (CCSA), as outlined in Zhang et al. (2021). The core task of each iteration of the CCSA process is to find the steepest decreasing direction from the feasible ones. In this operator, the values of all other decision variables remain unchanged except for the shifting decision. The values of the shifting decision are adjusted by the CCSA in an iterative manner. Specifically, the values of decision variables $\mathbf{e}, \mathbf{s}, \mathbf{y}, \mathbf{z}$ are held constant during each CCSA iteration, while the values of the shifting decision $\boldsymbol{\iota}$ are changed. The current solution is then evaluated by solving $\mathrm{P}(\bar{\mathbf{y}}, \bar{\mathbf{e}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\iota}}, \bar{\mathbf{s}})$ with an MILP solver. For clarify, an algorithmic description of the CCSA is provided in Algorithm 1.

**Algorithm 1** The Constrained Compass Search Algorithm (CCSA)

**Require:** Feasible solution $\boldsymbol{\iota}_0$, an initial step-size $v_0$, a step-size compressing constant $v$. Let $v_n$ be the step-size of the $n$th iteration. Let $\mathbf{e}_m$ be the standard basis vector in $\mathbb{R}^N$ where the $m$-th element is 0 and all other elements are 0. Let $n = 1$.

**while** $v_n > 0$ **do**
  Generate the set of candidate solutions $\Theta = \{\boldsymbol{\iota}_n, \boldsymbol{\iota}^i_{n+1} | \boldsymbol{\iota}^i_{n+1} = \boldsymbol{\iota}_n + \mathbf{e}_i, \forall i \in \mathcal{I}\}$.
  Calculate the objective value $f(\hat{\boldsymbol{\iota}})$ for each $\hat{\boldsymbol{\iota}} \in \Theta$. Denote $\boldsymbol{\iota}^* = \arg\min_{\hat{\boldsymbol{\iota}} \in \Theta} f(\hat{\boldsymbol{\iota}})$.
  **if** $f(\boldsymbol{\iota}^*) < f(\boldsymbol{\iota}_n)$ **then**
    Let $\boldsymbol{\iota}_{n+1} = \boldsymbol{\iota}^*$, and let $v_{n+1} = v_n$.
  **else**
    Let $\boldsymbol{\iota}_{n+1} = \boldsymbol{\iota}_n$, and let $v_{n+1} = v_n - v$.
  **End if**
  Let $n = n + 1$.
**end while**
**return** $\boldsymbol{\iota}_n$.

**Destroy operator 2: Greedy search for the skip-stop strategy of FLS trips.** This operator is designed to identify good skip-stop strategies with lower objective function values by employing the CCSA. The specific procedures are the same as those for the destroy operator 1, except that the values of the decision variable related to the skip-stop strategy are adjusted in the destroy operator 2, i.e., $\mathbf{y}$.

**Destroy operator 3: Randomly adjust dwell times of FLS trips.** This operator randomly selects a trip and modifies its dwell time at a particular stop. Specifically, for the selected trip and stop, the dwell time is changed by adding an integer number $tt$ drawn from the interval $\{-\Delta, \Delta\}$. If the adjusted dwell time satisfies constraints (8), then update this dwell time, i.e., $s_{iu} \leftarrow s_{iu} + tt$; otherwise, repeat the above procedure until a compliant dwell time modification is achieved.

**Destroy operator 4: Randomly adjust the stop-skip strategy of FLS trips.** This operator aims to eliminate the skip-stop strategy for a trip at a specific stop. More precisely, this operator randomly selects a $y_{iu}$ with a value equal to 0 in the current solution and reassigns it to 1.

**(3) Destroy operators associated with updating the pool of DRT services.** To update the DRT pool $D$, we introduce the following two destroy operators to add new DRT services, and develop a pruning strategy to remove infrequently operated DRT services.

**Destroy operator 5: Randomly add new DRT services.** This operator generates new DRT services. To elaborate, we first randomly generate the routing and departure times of DRT services. If these newly generated DRT services do not exist in $D$, they will be added to $D$.

**Destroy operator 6: Insert new DRT services into the pool of DRT services on the segments with high passenger demand.** This operator sorts all the segments according to the number of passengers on segments in the optimal solution of the previous iteration. Then, new DRT services on the segments with high passenger demand will be inserted into $D$.

**Destroy operator 7: Insert new DRT services into the pool of DRT services on the segments with low passenger demand.** This operator is similar to the Destroy operator 6. The only difference is that we first identify segments with low passenger demand. The reason for designing this operator is that by transporting waiting passengers via DRT services, FLS trips can skip these segments to improve operational efficiency.

As the iterations proceed, the quantity of the decision variable $\pi_{uvtm}(w)$ correspondingly expands. To mitigate the increase in computational difficulty due to large-scale variables, we design a pruning mechanism focused on infrequently operated DRT services. Specifically, if a DRT service in the pool $D$ fails to appear in the optimal solution for certain consecutive iterations, this DRT service will be removed from $D$.

**(3) Repair operators.** Skipping stops, modifying the departure time at the first stop, and altering the dwell time at stops can potentially result in the minimum headway being violated, and thus making the solution infeasible. To ensure the feasibility of solutions, we introduce the following two repair operators to generate a set of feasible neighbor solutions at each iteration.

**Repair operator 1: Adjust the shifting plan of FLS trips at the first stop.** This operator is used to modify the headway that violates constraints (17) by changing the shifting time of FLS trips. To be specific, if two trips $i$ and $i+1$ violate the minimum headway restriction at stops $u$, these two trips will be selected and then the shifting times $\iota_{i1}$ will decrease or $\iota_{(i+1)1}$ increase.

**Repair operator 2: Adjust the dwell time of FLS trips.** This operator is similar to the Repair operator 1. The difference is that we randomly decrease the dwell time of trip $i$ or increase the dwell time of trip $i+1$ at one of the upstream stop of stop $u$.

**(4) Adaptive searching strategy.** In the context of the ALNS algorithm, a feasible solution is generated through the application of a sequence of destroy and repair operators at each iteration. To select the most effective destroy and repair operators at each iteration, we design an adaptive searching strategy. We use $k$ to denote the index of destroy operators or repair operators, and define $\delta_k^d$ and $\delta_k^r$ to denote the weights of the destroy and repair operators, respectively. These weights are dynamically modified based on their performance efficacy in optimizing the objective function. $c_k^d$ and $c_k^r$ to denote the scores of the destroy operators and repair operators. $\phi$ is used to control the rate of change of the weight of each operator with respect to the quality of the solution. Following Yin et al. (2021), the weights of destroy and repair operators are updated according to the following formulas

$$\delta_k^d = (1-\phi)\delta_k^d + \phi c_k^d / \sum_{k=1}^{|\mathcal{K}^d|} c_k^d \qquad \forall k \in \mathcal{K}^d, \tag{56}$$

$$\delta_k^r = (1-\phi)\delta_k^r + \phi c_k^r / \sum_{k=1}^{|\mathcal{K}^r|} c_k^r \qquad \forall k \in \mathcal{K}^r, \tag{57}$$

where $\mathcal{K}^d$ and $\mathcal{K}^r$ are the sets of the destroy and repair operators, respectively.

**(5) Termination criteria.** We establish two termination criteria for the algorithm: (1) The search process terminates if the current number of iterations exceeds a predefined maximum limitation; (2) If the current best objective value remains unchanged for a pre-defined tolerance associated with the number of iterations, the corresponding solution can be output as an approximate optimal solution.

## 6. Numerical experiments

To verify the performance of our proposed approaches, we conduct numerical experiments on a small realistic case study and a real-world case study in this section. The proposed algorithm is coded in Python on a Windows 11 personal computer with 12th Gen Intel(R) Core(TM) i7-12700H and 64G RAM. GUROBI 9.5.1 is used to solve the entire model on the small realistic case study and the second subproblem. The source code can be found at https://github.com/Wodenxdy/TTRDCP.git.

### 6.1. Small-case study

The small realistic case study considers a bus line with six stops named as stops $A$, $B$, $C$, $D$, $E$ and $F$, which is illustrated in Fig. 6. A total of four FLS trips are taken into account to satisfy the passenger demand. We consider a study time horizon of 60 one-minute intervals. The running time on each segment of FLS trips and DRT services is time-varying. The dwell time at each stop ranges between 1 and 2 time intervals if this trip or DRT service docks, while the shifting time of FLS trips falls within $[-2, 2]$ time intervals. The minimum headway is set to one time interval. Each MU has a capacity of 15 passengers, and the maximum number

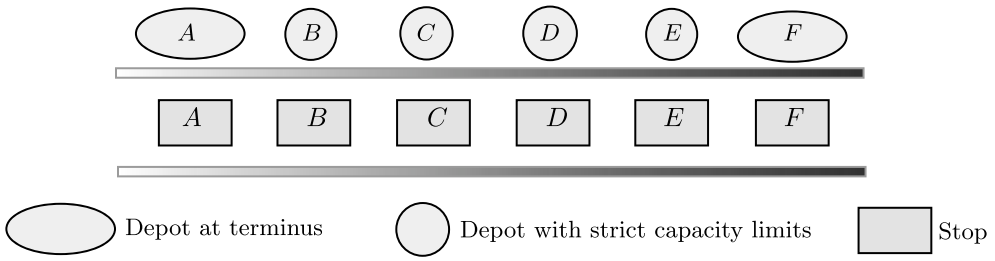
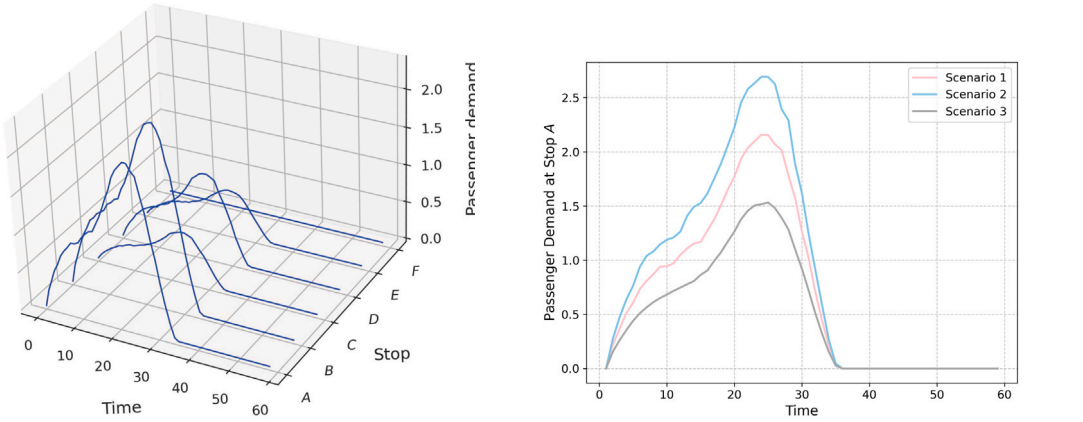



**Fig. 6.** Illustration of the line investigated in the small-case study.




(a) Passenger demand in scenario 1 (b) Passenger demand at stop A among all scenarios

**Fig. 7.** The time-dependent and uncertain passenger demand in the small case study.

of MUs in an MV assigned to FLS trips and DRT services is three. The initial numbers of MUs in depots associated with stops $B$, $C$, $D$ and $E$ are set to be zero. For simplicity, we assume that the limitations on the capacity of these depots are the same and set to three MUs. Besides, we consider three scenarios to capture the stochastic nature of passenger demand where the number of passenger demand are high, medium and low, respectively. The nominal probability distribution is set to $\rho_0 = \{0.4, 0.3, 0.3\}$. The scenario-based, time-dependent passenger groups are generated at random using one-minute intervals, as illustrated in Fig. 7. More specifically, Fig. 7(a) presents the temporal variation in passenger demand at all stops in scenario 1 and Fig. 7(b) shows the variation in passenger demand across all scenarios over time at stop $A$.

After defining the aforementioned parameter settings, we conduct two sets of numerical experiments. The first set of experiments aims to demonstrate the potential performance improvements achieved through the joint optimization of timetabling, routing and dynamic-capacity allocation of FLS trips and DRT services. The second set of experiments is designed to compares the developed algorithm with the benchmark solver GUROBI.

### 6.1.1. Comparison among various operational strategies

In this subsection, we concentrate on evaluating the superiority of the integrated optimization of timetabling and vehicle scheduling in an inter-modal transportation system. The study time horizon is set as $|\mathcal{T}| = 60$, and the number of FLS trips $|\mathcal{I}|$ is four. We explore five different settings for the weighting coefficients associated with passengers' traveling time and operational costs, represented by the ratios $\frac{\zeta_{pa}}{\zeta_{op}} = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{10}$. All instances are solved to optimality with GUROBI.

For comparison, three operational strategies are considered: (1) The first one, serving as our benchmark, represents the conventional operational strategy employed in real-world operations that excludes DRT services. This employs a uniform headway timetable along with fixed capacity allocation of FLS trips and is denoted as *Fixed Timetable and Capacity* (FTC). (2) The second strategy aims to jointly optimize timetables and dynamic capacity allocations of FLS trips while disregarding DRT services. In other words, this strategy corresponds to optimizing timetables and dynamic capacity allocations on a single-mode bus line. This strategy is denoted as *Optimized Timetable and Flexible Capacity* (OT-FC). (3) The third strategy is the integrated optimization of timetabling of FLS trips, routing of DRT services, and vehicle scheduling within the inter-modal transportation network, which is referred to as *Optimized Timetable, Flexible Capacity, and DRT Services* (OT-FC-DRT). It is worth noting that the computational intensity associated with the third strategy is much higher than that of the first two. For instance, using GUROBI, the second strategy is solved within 132 s, whereas the third strategy requires 43,710 s for the same instance.

**Table 2**
Computational results among three operational strategies.

| Ratio of weighting coefficients | Operational strategy | Optimal objective value | Expected value of passengers' costs (unit: $) | Expected value of operating costs (unit: $) | Number of utilized MUs | Average waiting time of passengers (unit: min) | Average in-vehicle time of passengers (unit: min) |
|---|---|---|---|---|---|---|---|
| 1:1 | FTC | 540.31 | 446.13 | 54.65 | 12 | 2.53 | 15.09 |
| | OT-FC | 511.65 | 469.37 | 42.28 | 8 | 3.04 | 13.68 |
| | OT-FC-DRT | 476.96 | 377.32 | 66.32 | 11 | 2.41 | 12.31 |
| 1:2 | FTC | 594.97 | 446.13 | 54.65 | 12 | 2.53 | 15.09 |
| | OT-FC | 554.87 | 431.66 | 42.57 | 7 | 3.06 | 13.68 |
| | OT-FC-DRT | 538.04 | 399.53 | 50.77 | 9 | 2.76 | 12.77 |
| 1:3 | FTC | 649.62 | 446.13 | 54.65 | 12 | 2.53 | 15.09 |
| | OT-FC | 597.98 | 431.66 | 42.57 | 7 | 3.06 | 13.68 |
| | OT-FC-DRT | 589.22 | 403.24 | 49.24 | 9 | 2.87 | 12.78 |
| 1:5 | FTC | 758.93 | 446.13 | 54.65 | 12 | 2.53 | 15.09 |
| | OT-FC | 684.20 | 431.65 | 42.57 | 7 | 3.06 | 13.68 |
| | OT-FC-DRT | 683.27 | 415.16 | 46.53 | 7 | 3.21 | 12.73 |
| 1:10 | FTC | 1032.21 | 446.13 | 54.65 | 12 | 2.53 | 15.09 |
| | OT-FC | 893.16 | 440.67 | 41.42 | 6 | 2.87 | 14.35 |
| | OT-FC-DRT | 893.16 | 440.67 | 41.42 | 6 | 2.87 | 14.35 |

Table 2 provides a detailed performance comparison for the investigated instances. We report the optimal objective value, the passengers' and operator costs in the optimal solution, the number of utilized MUs, as well as the average waiting and in-vehicle times of passengers. To compute the average waiting and in-vehicle times of passengers, we use the following formulas:

$$\text{Average waiting time of passengers} = \sum_{w\in\mathcal{W}}(\rho(w)\sum_{p\in\mathcal{P}_w} n_p \cdot WT_p / \sum_{p\in\mathcal{P}_w} n_p),$$

$$\text{Average in-vehicle time of passengers} = \sum_{w\in\mathcal{W}}(\rho(w)\sum_{p\in\mathcal{P}_w} n_p \cdot TP_p / \sum_{p\in\mathcal{P}_w} n_p).$$

From the experimental results of Table 2, we can derive the following observations. (1) The *Fixed Timetable and Capacity* strategy yields the worst performance among all instances, with respect to the objective values, the passengers' costs and the number of utilized MUs, compared with the results obtained by adopting the other two operational strategies. The main reasons is that this strategy adopts a uniform-headway timetable and fixed capacity which is not optimized according to the time-varying passenger demand. (2) Optimizing the timetable and capacity allocations of FLS trips is beneficial in reducing both the number of utilized MUs and passengers' traveling costs, in comparison to the results derived from adopting the *Fixed Timetable and Capacity* strategy. Although implementing the skip-stop strategy increases average waiting time of passengers, it results in a reduction in the in-vehicle time. (3) The OT-FC-DRT approach leads to higher operating costs than the OT-FC strategy, but it works better for improving service quality (e.g., lower passenger costs and the average in-vehicle time). To be specific, it can be seen that although the OT-FC-DRT strategy decreases the passengers' costs in comparison to the OT-FC strategy, the number of utilized MUs is increased. A second observation is that when the ratio of weighting coefficients changes to emphasize operating expenses (e.g., from 1 to $\frac{1}{10}$), the differences in the objective values and the number of utilized MUs and the optimal objective values between OT-FC and OT-FC-DRT strategies narrow down. When the ratio becomes to $\frac{1}{10}$, the numbers of utilized MUs and other results are the same. The reason is that operating DRT service reduces passengers' costs because it provides direct and non-stop service to passengers, but increases operating costs because the number of FLS trips are fixed and operating DRT service inevitably adds extra costs for using MUs. As the weight on operating costs increases relative to the weight on passenger costs, the objective function becomes more skewed towards minimizing operating costs at the expense of passengers' costs. With a weighting of 1:5 or 1:10, the operational strategy shifts to reducing operating costs, even if this means sacrificing the benefits of the direct, non-stop services provided by DRT. This observation suggests that when operators are extremely concerned about their expenses, operating costs would only be used to run FLS services.

More in general, these computational results let us conclude that adopting the *Optimized Timetable, Flexible Capacity, and DRT Services* is able to reduce the weighed sum of passengers' and operating costs, as well as improve the quality of service by setting the appropriate weighting coefficients in the objective function. As such, our approach gives operators the opportunity to enhance the level of service without increasing the sum of passengers' and operating costs or, conversely, decreasing this value by introducing DRT services.

### 6.1.2. Comparison between different solution methods

To gain more insight into the effectiveness of the proposed algorithm, we perform a total of five instances by varying essential parameters in this subsection. Specifically, we gradually increase the number of FLS trips and time intervals, denoted as $|\mathcal{I}|$ and $|\mathcal{T}|$, from 2 to 6 and from 40 to 80, respectively. All other parameters are set in accordance with Section 6.1.1.

In Table 3, we present the results by stating the objective value, the optimality gap, and the computational time for three solution methodologies: FTC, GUROBI and ALNS+GUROBI. The FTC approach utilizes GUROBI to solve the model based on the

**Table 3**
Computational results for FTC, GUROBI, and the proposed hybird algorithm with a time limit of 3600 s.

| Instance (# of FLS trips - # of time intervals) | Solution method | Objective value | Optimality gap (%) | Computational time (unit: second) |
|---|---|---|---|---|
| 2–40 | FTC | 184.01 | – | – |
| | GUROBI | 169.61 | 0.00 | 82.27 |
| | ALNS+GUROBI | 169.99 | – | 30.20 |
| 3–50 | FTC | 349.51 | – | – |
| | GUROBI | 312.76 | 6.55 | 3600.00 |
| | ALNS+GUROBI | 312.89 | – | 165.67 |
| 4–60 | FTC | 540.31 | – | – |
| | GUROBI | 505.63 | 20.1 | 3600.00 |
| | ALNS+GUROBI | 485.32 | – | 466.28 |
| 5–70 | FTC | 747.27 | – | – |
| | GUROBI | 660.80 | 32.60 | 3600.00 |
| | ALNS+GUROBI | 649.10 | – | 792.74 |
| 6–80 | FTC | 868.58 | – | – |
| | GUROBI | – | – | 3600.00 |
| | ALNS+GUROBI | 735.90 | – | 2241.78 |

**Table 4**
The values of other involved parameters in this case study.

| Parameters | Nations | Values and units |
|---|---|---|
| Number of stops | $\lvert\mathcal{V}\rvert$ | 18 |
| Number of FLS trips | $\lvert\mathcal{I}\rvert$ | 12 |
| The maximum formation of an MV | $\lvert\mathcal{M}\rvert$ | 5 |
| Capacity of one MU | $C$ | 30 passengers |
| Minimum shifting time | $\underline{\iota}$ | −2 min |
| Maximum shifting time | $\bar{\iota}$ | 2 min |
| Minimum dwell time | $\underline{s}$ | 1 min |
| Maximum dwell time | $\bar{s}$ | 2 min |
| Minimum headway | $\underline{h}$ | 3 min |
| Maximum number of MUs stored in the depot | $\bar{N}_u$ | 6 |
| Weighting coefficient of passengers' waiting time | $\xi_1$ | 1.5 |
| Weighting coefficient of passengers' in-vehicle time | $\xi_2$ | 1 |

*Fixed Timetable and Capacity* strategy, serving as a benchmark for comparison. Besides, GUROBI and ALNS+GUROBI are employed to solve model (54). We observe that our developed ALNS+GUROBI algorithm yields superior time-efficiency than GUROBI, even for the relatively small scale instances (e.g., $|\mathcal{I}| = 2$). As the instance sizes expand, the ALNS+GUROBI algorithm continues to outperform GUROBI, achieving noticeable improvements in both solution quality and computational efficiency. For instances with more than 4 trips and 60 time intervals, the optimality gaps returned by GUROBI are larger than 20% after the 1-hour time limit of computation, while better solutions can be obtained by the ALNS+GUROBI algorithm in up to 2242 s. Besides, the convergence tendency of the objective values of the ALNS+GUROBI algorithm is illustrated in Fig. 8. Clearly, our proposed algorithm can converge within a finite number of iterations in all instances.

### 6.2. Real-world case study

In this section, we further consider a realistic case study based on the practical data of the Beijing Bus Line 468, which consists of 18 stops and is depicted in Fig. 9. The time horizon considered in this instance is 7:30 AM–10:00 AM, which includes both morning peak hours and a part of off-peak hours. The time horizon is discretized with a unit of 1 min in order to balance the real-world operating requirements and computational efficiency. The time-dependent and uncertain passenger demand is taken from historically detected Automatic Fare Collection data, and time-varying running times on segments are processed following Xia et al. (2023). To characterize the randomness of passenger demand, we construct five scenarios with the uncertain probability distribution, where the nominal probability distribution is set as $\rho_0 = \{0.2, 0.2, 0.2, 0.2, 0.2\}$. Following Chen et al. (2019) and Dai et al. (2020), we set the other parameter settings, which are listed in Table 4.

For this instance, the benchmark we consider is the practical timetable typically used in the Beijing bus system, which operates with the uniform headway and fixed-capacity vehicles. Using the aforementioned parameter settings, two sets of numerical experiments are conducted to demonstrate the performance of our proposed methodologies and gain more insights into the trade-off between passengers' and operators' costs. The difference between these experiments is the settings of weighting coefficients associated with passengers' and operating costs. Specifically, in the first set, we set the ratio of these coefficients as 1:1, i.e., $\zeta_{pa} : \zeta_{op} = 1 : 1$, and denote its results as the *Passenger-centred solution*. Conversely, this ratio changes to $1 : 5$ in the subsequent set, and its results are denoted as *Trade-off solution*. Because GUROBI performs poorly and is unable to solve the problems after 12 h of computation, we do not use it for these experiments.

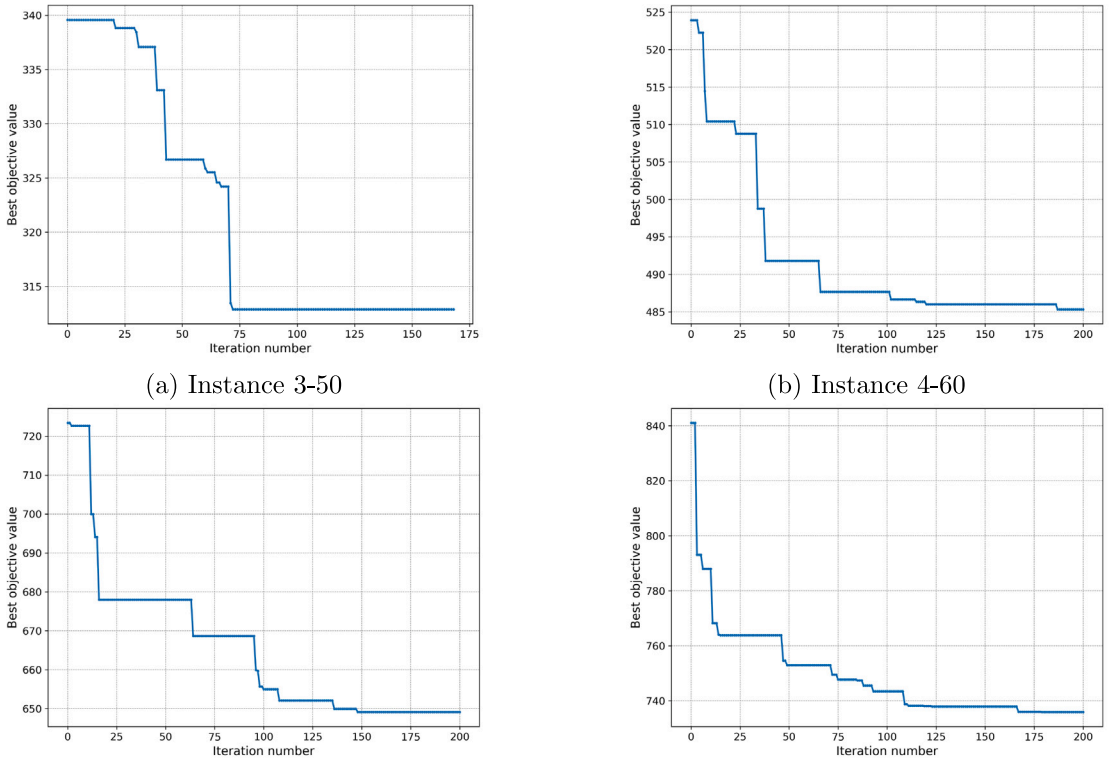



**Fig. 8.** Convergence tendency of the objective values.

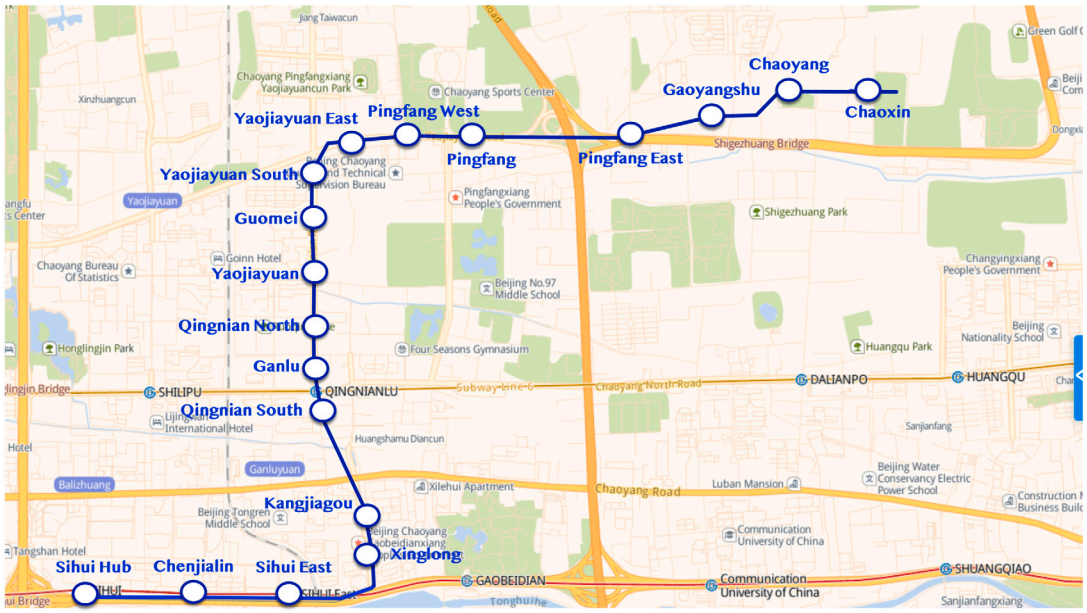



**Fig. 9.** The investigated FSL line of Beijing Bus Line 468.

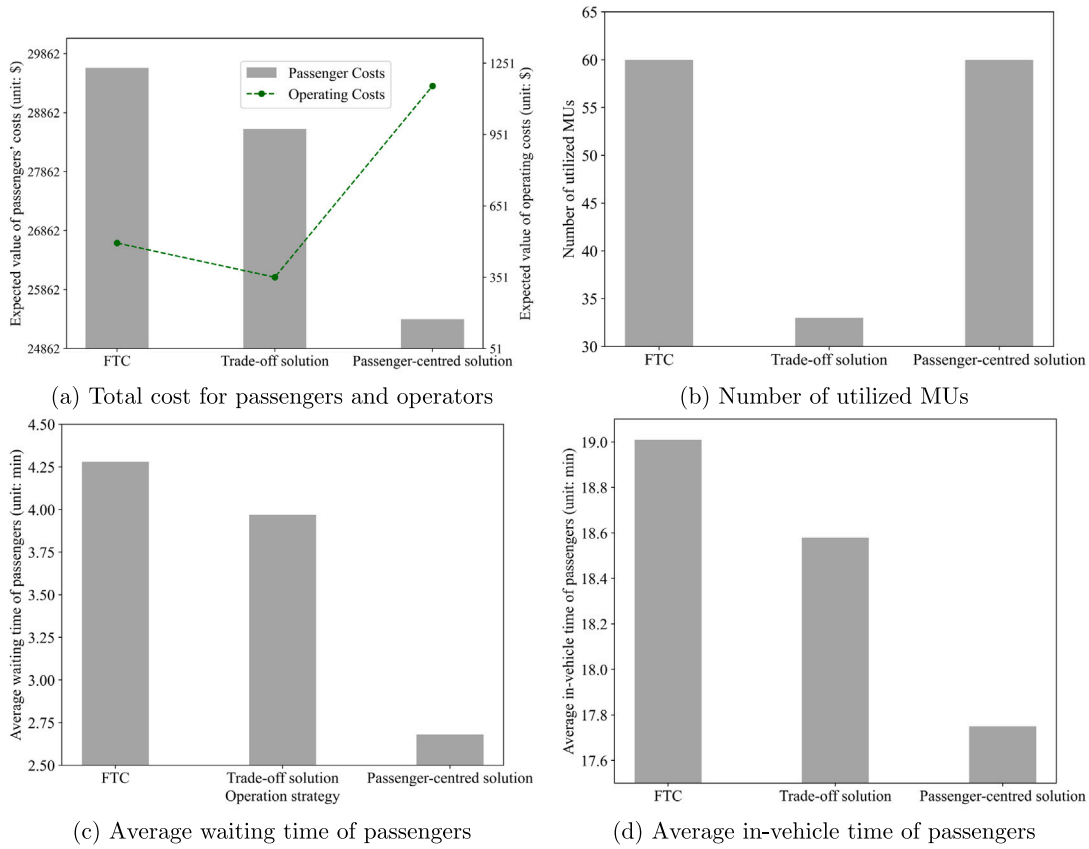



**Fig. 10.** Performance comparison among different solutions.

Fig. 10 shows the computational results of the benchmark (denoted as FTC), the *Trade-off solution*, and the *Passenger-centred solution* on the real-world instance, involving the objective value, the number of utilized MUs, the average waiting time of passengers, and the average in-vehicle time of passengers. From the results in Fig. 10(a), we derive the following two observations: (1) The practical timetable yields the worst performance compared to the proposed integrated timetabling and vehicle scheduling approach independent of weighting coefficients, In terms of expected values of passengers' and operating cost, the proposed integrated timetabling and vehicle scheduling approach performs better than the practical timetable. This is independent of weighting coefficients related to passengers' and operators' costs. Besides, the objective value of the *Trade-off solution* is decreased about 6.0% compared to that of the FTC. (2) Compared to the *Trade-off solution*, the *Passenger-centred solution* results in a reduction in passenger costs from 28587.84 $ to 25362.18 $, while operating costs increase from 350.78 $ to 1155.15 $. In other words, a 229.3% surge in operating costs is exchanged for an 11.3% decrease in passenger costs. This observation If the overarching aim is to promote public transport usage by making it more passenger-friendly (even at higher operational costs), then the Passenger-centred solution might be preferable. On the other hand, if the system has budgetary constraints or aims to be financially self-sustaining, such a steep rise in operating costs might be seen as prohibitive. This finding can provide operators with a management suggestion to set the weights in accordance with the actual requirements. If operators aim to promote public transport usage by making it more passenger-friendly even at higher operational costs, then the *Passenger-centred solution* might be preferable. Otherwise, the *Trade-off solution* is more recommended.

Fig. 10(b) depicts the numbers of utilized MUs among this solutions. The results are in line with those of the previous experiment: the number of required MUs increases in the *Passenger-centred solution*, relative to the *Trade-off solution*. Fig. 10(c) and (d) present the average waiting and in-vehicle times of passengers. The key insight is that the integrated timetabling and vehicle scheduling method results in a considerable reduction in passengers' average waiting and in-vehicle times, compared with adopting the practical timetable and fixed-capacity vehicles. The main reason is that, our proposed methodology is able to dynamically and flexibly adjust capacity allocations at different times, stops and within the whole intermodal urban transit network to improve the match with dynamic and uncertain passenger demand.

For clarify, Fig. 11 and Table 5 present the detailed results in the *Trade-off solution*. To be specific, Fig. 11 depicts the robust timetable for FLS trips and the corresponding dynamic capacity allocation in scenario 1. It can be seen that, the formations of MVs assigned to trips change during operations, illustrating the necessity of enabling adjust capacity dynamically at different times and

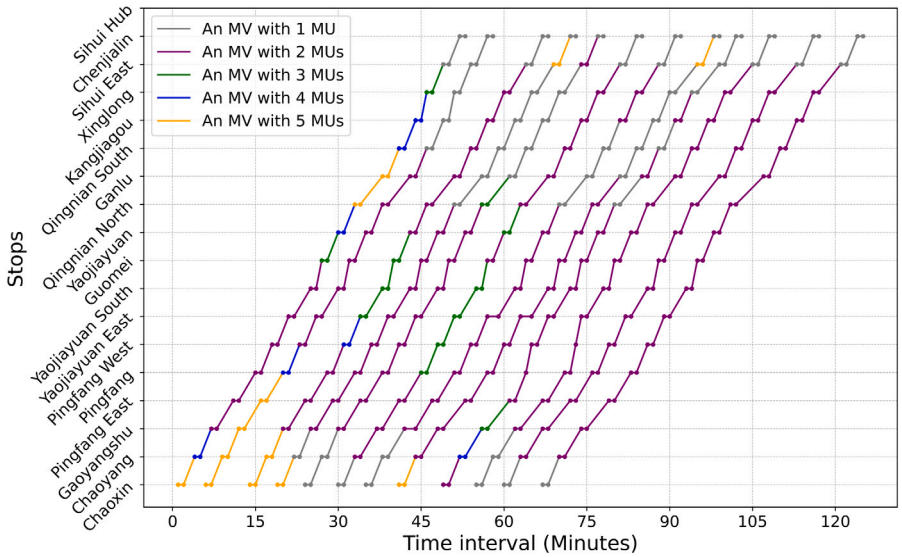



**Fig. 11.** The optimized timetable for FLS trips and the corresponding dynamic capacity allocation in scenario 1 in the *Trade-off solution*.

**Table 5**
The operated DRT services in scenario 1 in the *Trade-off solution*.

| # of DRT service | Origin stop | Destination stop | Departure time | # of DRT service | Origin stop | Destination stop | Departure time |
|---|---|---|---|---|---|---|---|
| 1 | Chaoyang | Pingfang East | 41 | 9 | Yaojiayuan | Qingnian North | 81 |
| 2 | Chaoyang | Pingfang | 50 | 10 | Qingnian North | Ganlu | 66 |
| 3 | Chaoyang | Qingnian North | 6 | 11 | Ganlu | Qingnian South | 54 |
| 4 | Gaoyangshu | Qingnian North | 10 | 12 | Qingnian South | Xinglong | 73 |
| 5 | Gaoyangshu | Sihui East | 10 | 13 | Qingnian South | Sihui East | 88 |
| 6 | Pingfang | Guomei | 79 | 14 | Kangjiagou | Sihui East | 107 |
| 7 | Pingfang West | Yaojiayuan South | 30 | 15 | Sihui East | Chenjialin | 101 |
| 8 | Guomei | Yaojiayuan | 40 | | | | |

stops. Besides, in Table 5, we present the results of the optimized DRT services, stating the origin stop, the destination stop, and the departure time. It can be observed that a total of 15 DRT services are operated among various time intervals. Therefore, we draw the conclusion that DRT services are essential in a public transportation network with strong dynamics to efficiently serve time-dependent passengers.

### 6.3. Out-of-sample performance of DRO-TTVSP and SP-TTVSP

In this section, we follow Xia et al. (2023) and Shehadeh (2023) to conduct the out-of-sample testing procedure to evaluate the out-of-sample performance of the SP-TTVSP and DRO-TTVSP models based on the real-world case study. Since this study considers the uncertainty inherent in the probability distributions of scenarios, we randomly generate multiple sample sets with varying numbers of scenarios, denoted by $N$, specifically $N \in \{10, 50, 100, 500, 1000\}$. We solve the DRO-TTVSP (54) and the SP-TTVSP (51) based on the in-sample scenarios to obtain DRO and SP solutions, respectively. The in-sample scenarios referred to here are those utilized in Section 6.2. To ensure a fair comparison, both models are solved 10 times using the proposed algorithm combining ALNS and GUROBI. Subsequently, the obtained 10 DRO solutions and 10 SP solutions are input in sequence to the randomly generated samples to assess their respective out-of-sample performance. Lastly, we compute and compare the average out-of-sample performance of the 10 solutions.

Table 6 presents the out-of-sample performance of the DRO and SP solutions across various sample sets, stating the number of scenarios in each sample set, the optimization method, the statistical indicators of the objective values including the mean and median values, the standard deviation, and the minimum and maximum values. In terms of mean and median values, we find that the values obtained under the DRO solution are in all instances smaller than those obtained by the SP solution. This suggests that the DRO approach is more effective in generating higher-quality solutions with smaller objective values than the SP approach. The second observation is that the three indicators, standard deviation, maximum and minimum, exhibit the same tendency as in the previous observation. From these results we can conclude that the DRO approach performs less volatility and more robustness than the SP method. In conclusion, the DRO-TTVSP model generates solutions that are not only more effective but also more reliable compared to the SP-TTVSP model.

**Table 6**
Statistical values for out-of-sample performance.

| # of scenarios | Method | Mean value | Median value | Standard deviation | Minimum value | Maximum value |
|---|---|---|---|---|---|---|
| 10 | DRO | 35 322.22 | 35 352.63 | 170.63 | 35 044.86 | 35 527.32 |
| | SP | 35 989.08 | 36 019.91 | 177.53 | 35 700.16 | 36 208.32 |
| 50 | DRO | 35 313.44 | 35 331.46 | 197.43 | 34 835.42 | 35 860.94 |
| | SP | 35 978.05 | 35 996.34 | 202.90 | 35 491.47 | 36 555.82 |
| 100 | DRO | 35 319.26 | 35 333.52 | 195.10 | 34 835.42 | 35 860.94 |
| | SP | 35 982.41 | 36 005.32 | 201.24 | 35 482.17 | 36 555.82 |
| 500 | DRO | 35 305.00 | 35 309.25 | 203.66 | 34 690.41 | 35 906.36 |
| | SP | 35 967.69 | 35 972.4 | 210.74 | 35 333.84 | 36 571.56 |
| 1000 | DRO | 35 299.80 | 35 296.92 | 209.06 | 34 690.41 | 35 916.12 |
| | SP | 35 962.34 | 35 960.12 | 215.78 | 35 333.84 | 36 571.56 |

Furthermore, we explore the value of robustness from the perspective of out-of-sample disappointment. As defined in the related literature (e.g., Wang et al. 2020, Van Parys et al. 2021, Shehadeh 2023), the out-of-sample disappointment quantifies the degree to which the out-of-sample expense exceeds the model's optimal value, which, as defined in Wang et al. (2020), can be computed using the following formula:

$$\max\left\{\frac{AOBJ - EOBJ}{AOBJ}, 0\right\} \times 100,$$

where $EOBJ$ and $AOBJ$ are the objective value of DRO-TTVS (SP-TTVS) model and the out-of-sample objective value of implementing the DRO (SP) solution, respectively. $EOBJ$ and $AOBJ$ can be regarded as the estimated and actual costs of implementing the DRO (SP) solution (Shehadeh, 2023). If the out-of-sample disappointment is 0, the objective value of the DRO-TTVS or SP-TTVS model is equal to or larger than the corresponding out-of-sample objective value, indicating that the model is more conservative and avoids underestimating costs. Otherwise, the model has a higher level of overoptimism.

Fig. 12 presents the normalized histograms of the out-of-sample disappointments of the DRO and SP solutions among various sample sets. It can be seen that, on average, the DRO solution result in considerably smaller out-of-sample disappointments. In addition, the number of the out-of-sample disappointment's values that are 0 under the DRO solution is more than that number under the SP solution. These findings suggest that, in comparison to the SP-TTVSP model, the DRO-TTVSP model can provide a more robust estimate of actual costs that arise in practice. From the aforementioned observations, we can conclude the proposed DRO-TTVSP model allows operators to develop timetables and vehicle schedules that are more realistic to operational needs in an inter-modal transportation network.

## 7. Conclusions

This paper addressed the collaborative optimization of timetabling and vehicle scheduling in an integrated FLS and DRT system with time-dependent and uncertain passenger demand. We have proposed a DRO model to generate robust timetables for FLS trips and scenario-related vehicle schedules within this intermodal urban transit network utilizing MVs, aiming to minimize the expectation of passengers' and operating costs. Our proposed approach allows MVs to be (de)coupled at each stop considering the re-routing of decoupled MUs and the flexible circulations of MUs across different transportation modes. We have introduced the required binary variables to cope with the complexities of this integrated problem under demand uncertainty and finally formulated it as an MILP formulation. To obtain high-quality solutions for real-case problems, we have developed a hybrid solution method that combines a tailored ALNS algorithm with GUROBI, which incorporates the constrained compass search and simulated annealing algorithms.

Computational results based on a virtual line and the Beijing bus line illustrate the value of our approaches. Compared with adopting the timetable with uniform headway and fixed-capacity vehicles, we are able to find timetables and vehicle schedules requiring considerably fewer vehicles while slightly reducing passengers' costs when operators focus more on operating costs than passengers' costs. As such, our approach allows operators to save costs without decreasing the quality of services. Moreover, we are able to find timetables and vehicle schedules with a lower sum of the expectation of passengers' and operating costs compared with optimizing only the timetable and dynamic-capacity allocations of FLS trips. Based on the real-case instances, the computational results show that our proposed method can reduce the expectation of passengers' and operating costs by about 6.0%, compared to results derived from the practical timetable with uniform headway and fixed-capacity vehicles. The results of the real-world instances also indicate that the out-of-sample performance of the DRO approach outperforms the SP method. Additionally, the developed algorithm outperforms GUROBI even in small-scale cases, such as an instance with six stops, four FLS trips, and 60 time intervals.

In further research, it would be interesting to incorporate the individual choices, preferences, and behaviors into our models to obtain user-optimal timetables and vehicle schedules of an intermodal urban transit network. Another interesting direction is to explore the studied problem at the network level, taking multiple FLS lines into account, and develop effective algorithms such as dualizing coupling constraints and using Lagrangian multipliers.

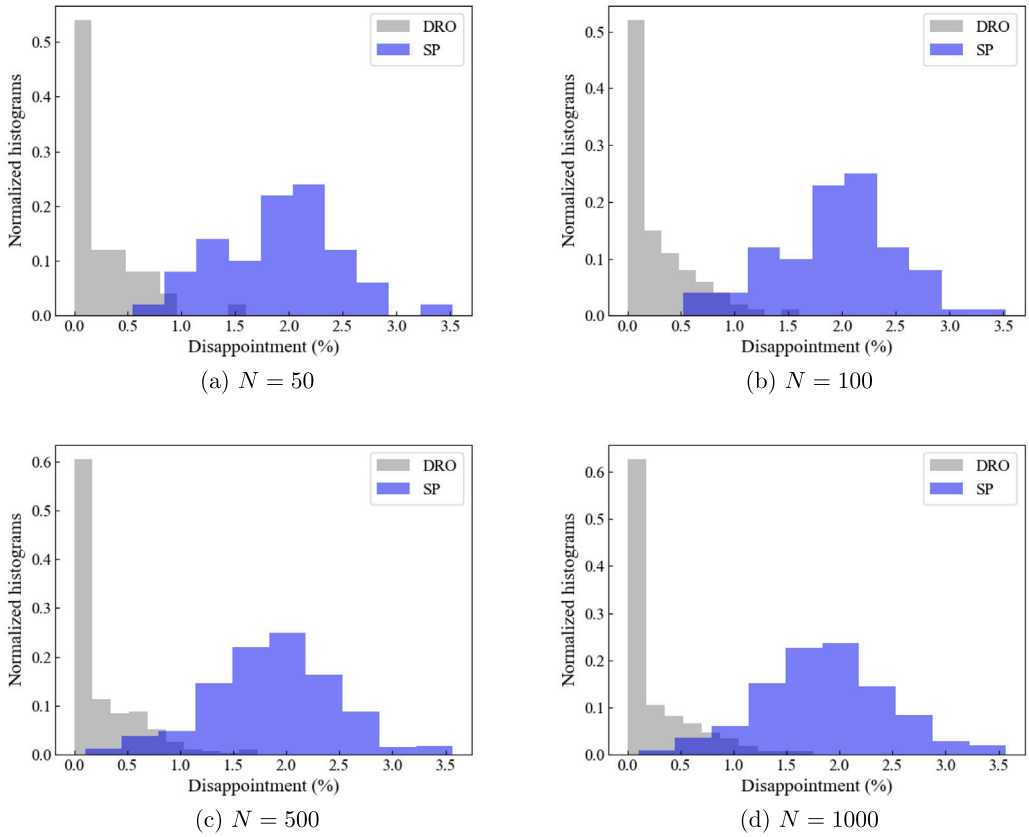



**Fig. 12.** Normalized histograms of out-of-sample disappointments among various sample sets.

## CRediT authorship contribution statement

**Dongyang Xia:** Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Jihui Ma:** Validation, Supervision, Funding acquisition, Data curation. **Sh. Sharif Azadeh:** Writing – review & editing, Validation, Supervision, Investigation, Conceptualization.

## Declaration of competing interest

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

## Acknowledgments

The first author would like to thank the program of China Scholarship Council (No. 202207090003) for the financial support to his visiting Ph.D. research in Delft University of Technology, the Netherlands. The authors are grateful to the Special issue Editor and two anonymous reviewers for their valuable comments that led to a much improved version of this paper.

## Appendix A

Tables 7 and 8 introduce the notations of sets, parameters and dependent variables in our mathematical formulations.

## Appendix B

In this section, we first linear the nonlinear constraints in the proposed model (50).

**Table 7**
Notations of sets and parameters.

| Notation | Description |
|---|---|
| **Sets** | |
| $\mathcal{U}$ | Set of stops, $\mathcal{U} = \{1, 2, \dots, \lvert\mathcal{U}\rvert\}$, indexed by $u$, $v$ |
| $\mathcal{I}$ | Set of FLS trips, $\mathcal{I} = \{1, 2, \dots, \lvert\mathcal{I}\rvert\}$, indexed by $i$ |
| $\mathcal{T}$ | Set of discretized time intervals, $\mathcal{T} = \{1, 2, \dots, \lvert\mathcal{T}\rvert\}$, indexed by $t$ |
| $\mathcal{W}$ | Set of scenarios, $\mathcal{W} = \{1, 2, \dots, \lvert\mathcal{W}\rvert\}$, indexed by $w$ |
| $\mathcal{P}_w$ | Set that containing all passenger groups in scenario $w$, $\mathcal{P} = \{1, 2, \dots, \lvert\mathcal{P}_w\rvert\}$, indexed by $p$ |
| $\mathcal{M}$ | Set of the number of MUs that can be contained in an MV, $\mathcal{M} = \{1, 2, \dots, \lvert\mathcal{M}\rvert\}$, indexed by $m$ |
| **Parameters** | |
| $\Delta$ | Duration of each time interval |
| $\hat{r}_{uvt}$ | Running time from stops $u$ to $v$ of an MV assigned to a FLS trip or a DRT service with a departure time $t$ |
| $r_{iu}^{dec}$ | Deceleration time of trip $i$ when it arrives at stop $u$ |
| $r_{iu}^{acc}$ | Acceleration time of trip $i$ when it leaves stop $u$ |
| $\underline{\iota}$ | Minimum shifting time |
| $\bar{\iota}$ | Maximum shifting time |
| $\underline{s}$ | Minimum dwell time |
| $\bar{s}$ | Maximum dwell time |
| $\underline{h}$ | Minimum headway |
| $C$ | Capacity of one MU |
| $N_u$ | Number of MUs initially stored in the depot attached to stop $u$ |
| $\hat{N}_u$ | Maximum number of MUs stored in the depot attached to stop $u$ |
| $\varphi^T$ | The equivalent monetary value per unit of passengers' traveling time (unit: $) |
| $\varphi_{uvm}$ | Operating costs of utilizing an MV equipped with $m$ MUs on segment between stops $u$ and $v$ |
| $O_p$ | Origin of passenger group $p$ |
| $D_p$ | Destination of passenger group $p$ |
| $AT_p$ | Arrival time of passenger group $p$ |
| $n_p$ | Number of passengers in passenger group $p$ |
| $\rho(w)$ | The probability of scenario $w$ |
| $\xi_1$, $\xi_2$ | Weighting coefficients related to the waiting and in-vehicle times of passengers |
| $\zeta_{pa}$, $\zeta_{op}$ | Weighting coefficients related to the traveling time of passengers and operating costs |
| $M$ | Large positive constant |
| $\epsilon$ | Tiny positive constant |

**Table 8**
Notations of dependent variables.

| Notation | Description |
|---|---|
| $a_{iu}$ | Arrive time of trip $i$ at stop $u$ |
| $d_{iu}$ | Departure time of trip $i$ from stop $u$ |
| $h_{iut}$ | Binary indicator, $h_{iut} = 1$ if the MV assigned to trip $i$ leaves stop $u$ exactly at time interval $t$; $h_{iut} = 0$, otherwise |
| $g_{iut}$ | Binary indicator, $e_{iut} = 1$ if the MV assigned to trip $i$ arrives stop $u$ exactly at time interval $t$; $g_{iut} = 0$, otherwise |
| $o_{iu}^{\mathrm{FLS}}(w)$ | Number of in-vehicle passengers on trip $i$ when it departs from stop $u$ in scenario $w$ |
| $l_{iu}^{\mathrm{FLS}}(w)$ | Number of passengers alighting from trip $i$ when it arrives at stop $u$ in scenario $w$ |
| $b_{iuv}^{\mathrm{FLS}}(w)$ | Number of passengers who board trip $i$ at stop $u$ and head to stop $v$ in scenario $w$ |
| $b_{uvt}^{\mathrm{DRT}}(w)$ | Number of passengers heading to stop $v$ who board the DRT service with a route covering stops $u$ and $v$ as well as a departure time $t$ in scenario $w$ |
| $\vartheta_{pi}^{\mathrm{FLS}}(w)$ | A binary variable indicating whether the MV assigned to trip $i$ can potentially transport passenger group $p$ in scenario $w$ |
| $\vartheta_{puvt}^{\mathrm{DRT}}(w)$ | A binary variable indicating whether the DRT service serving stops $u$ and $v$ with a departure time $t$ can potentially transport passenger group $p$ |

(*continued on next page*)

**Table 8** (*continued*).

| Notation | Description |
|---|---|
| $\varsigma_{pi}^{\mathrm{FLS}}(w)$ | Waiting time of passenger group $p$ if they board trip $i$ |
| $\varsigma_{puvt}^{\mathrm{DRT}}(w)$ | Waiting time of passenger group $p$ if they board the DRT service serving stops $u$ and $v$ with a departure time $t$ |
| $\varphi_{pi}^{\mathrm{FLS}}$ | A binary variable indicating whether passenger group $p$ board trip $i$ or not |
| $\varphi_{puvt}^{\mathrm{DRT}}$ | A binary variable indicating whether passenger group $p$ board the DRT service serving stops $u$ and $v$ with a departure time $t$ can potentially transport passenger group $p$ or not |
| $WT_p$ | Waiting time of passenger group $p$ |
| $TP_p$ | In-vehicle time of passenger group $p$ |
| $N_{ut}(w)$ | Number of available MUs stored in the depot attached to stop $u$ at time interval $t$ in scenario $w$ |
| $AV_{ut}(w)$ | Number of MUs arriving at the depot attached to stop $u$ at time interval $t$ in scenario $w$ |
| $DV_{ut}(w)$ | Number of MUs departing from the depot attached to stop $u$ at time interval $t$ in scenario $w$ |

Specifically, constraints (21), (23)–(26), (41)–(42) are all nonlinear. By introducing auxiliary variables $\chi_{iuv}$ and $\alpha_{iuv}$, constraints (21) is linearized as follows:

$$\begin{cases} \chi_{iuv} - y_{iu} \le 0 & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ \chi_{iuv} - y_{iv} \le 0 & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ y_{iu} + y_{iv} - \chi_{iuv} \le 1 & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ \alpha_{iuv} \le d_{iu} & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ \alpha_{iuv} \le M \cdot \chi_{iuv} & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ \alpha_{iuv} \ge d_{iu} - M(1 - \chi_{iuv}) & \forall i \in \mathcal{I}, u, v \in \mathcal{U}, \\ M(\vartheta_{pi}^{\mathrm{FLS}}(w) - 1) \le \alpha_{iuv} - AT_p \le M \cdot \vartheta_{pi}^{\mathrm{FLS}}(w) - \epsilon_0 & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, v = D_p, \\ \chi_{iuv} \in \{0,1\}, \alpha_{iuv} \in [0, M] & \forall i \in \mathcal{I}, u, v \in \mathcal{U}. \end{cases} \tag{58}$$

Next, to linearize the nonlinear constraints (23) and (24), variables $\beta_{iup}$ and $\dot{\beta}_{puvt}$ are newly introduced here. We have

$$\begin{cases} \beta_{iup}(w) \le d_{iu} & \forall i \in \mathcal{I}, u \in \mathcal{U}, w \in \mathcal{W}, p \in \mathcal{P}_w, \\ \beta_{iup}(w) \le M \cdot \vartheta_{pi}^{\mathrm{FLS}}(w) & \forall i \in \mathcal{I}, u \in \mathcal{U}, w \in \mathcal{W}, p \in \mathcal{P}_w, \\ \beta_{iup}(w) \le d_{iu} - M \cdot (1 - \vartheta_{pi}^{\mathrm{FLS}}(w)) & \forall i \in \mathcal{I}, u \in \mathcal{U}, w \in \mathcal{W}, p \in \mathcal{P}_w, \\ \varsigma_{pi}^{\mathrm{FLS}}(w) = \beta_{iup} - AT_p \cdot \vartheta_{pi}^{\mathrm{FLS}}(w) + M(1 - \vartheta_{pi}^{\mathrm{FLS}}(w)) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, \\ \dot{\beta}_{puvt}(w) - \eta_{uvt}(w) \le 0 & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p, \\ \dot{\beta}_{puvt}(w) - \vartheta_{pi}^{\mathrm{FLS}}(w) \le 0 & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p, \\ \eta_{uvt}(w) + \vartheta_{pi}^{\mathrm{FLS}}(w) - \dot{\beta}_{puvt}(w) \le 1 & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p, \\ \varsigma_{puvt}^{\mathrm{DRT}}(w) = t\dot{\beta}_{puvt}(w) - AT_p \cdot \vartheta_{puvt}^{\mathrm{DRT}}(w) + M(1 - \vartheta_{puvt}^{\mathrm{DRT}}(w)) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, t \in \mathcal{T}, u = O_p, v = D_p, \\ \beta_{iup}(w) \in [0, M] & \forall p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, \\ \dot{\beta}_{puvt}(w) \in \{0,1\} & \forall p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, v = D_p, t \in \mathcal{T}. \end{cases} \tag{59}$$

The nonlinear constraints (25) can be equivalently transformed into the following system of linear inequations by introducing binary variables $\gamma_{pi}$ and $\dot{\gamma}_{puvt}$, i.e.,

$$\begin{cases} WT_p \le \varsigma_{pi}^{\mathrm{FLS}}(w) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{I}, \\ WT_p \le \varsigma_{puvt}^{\mathrm{DRT}}(w) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p, v = D_p, t \in \mathcal{T}, \\ \varsigma_{pi}^{\mathrm{FLS}}(w) \le WT_p - M(1 - \gamma_{pi}(w)) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{I}, \\ \varsigma_{puvt}^{\mathrm{DRT}}(w) \le WT_p - M(1 - \dot{\gamma}_{puvt}(w)) & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p, v = D_p, t \in \mathcal{T}, \\ \sum_{i \in \mathcal{L}} \gamma_{pi}(w) + \sum_{t \in \mathcal{T}} \dot{\gamma}_{puvt}(w) \ge 1 & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, u = O_p, v = D_p, \\ \gamma_{pi}(w) \in \{0,1\} & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{I}, \\ \dot{\gamma}_{puvt}(w) \in \{0,1\} & \forall w \in \mathcal{W}, p \in \mathcal{P}_w, i \in \mathcal{L}, u = O_p, v = D_p, t \in \mathcal{T}. \end{cases} \tag{60}$$

By introducing variables $\phi_{piu}$, $\dot{\phi}_{piu}$ and $\tilde{\phi}_{puvt}$, constraints (26) can be reformulated as the following linear form

$$\begin{cases} TP_p = \sum_{i\in\mathcal{I}}(\dot{\phi}_{piv}(w) - \phi_{piu}(w)) + \sum_{t\in\mathcal{T}} \hat{r}_{uvt}\cdot\varphi^{\mathrm{DRT}}_{puvt}(w) & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u=O_p, v=D_p,\\ \phi_{piu}(w)\le d_{iu} & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \phi_{piu}(w)\le M\cdot\varphi^{\mathrm{FLS}}_{pi}(w) & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \phi_{piu}(w)\le d_{iu}-M(1-\varphi^{\mathrm{FLS}}_{pi}(w)) & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \dot{\phi}_{piu}(w)\le a_{iu} & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \dot{\phi}_{piu}(w)\le M\cdot\varphi^{\mathrm{FLS}}_{pi}(w) & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \dot{\phi}_{piu}(w)\le a_{iu}-M(1-\varphi^{\mathrm{FLS}}_{pi}(w)) & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, u\in\mathcal{U},\\ \phi_{piu}(w),\dot{\phi}_{piu}(w)\in[0,M] & \forall w\in\mathcal{W}, p\in\mathcal{P}_w, i\in\mathcal{I}, u\in\mathcal{U}, \end{cases} \tag{61}$$

Lastly, we introduce variables $\kappa_{iut}$ and $\dot{\kappa}_{iut}$ to linearize constraints (41)–(42). The linear form can be expressed as follows:

$$\begin{cases} AV_{ut}(w) = \sum_{i\in\mathcal{L}}\kappa_{iut} + \sum_{v\in\mathcal{U}}\sum_{m\in\mathcal{M}} m\pi_{vu(t-\hat{r}_{uvt})m}(w) & \forall w\in\mathcal{W}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ DV_{ut}(w) = \sum_{i\in\mathcal{L}}\dot{\kappa}_{iut} + \sum_{v\in\mathcal{U}}\sum_{m\in\mathcal{M}} m\pi_{uvtm}(w) & \forall w\in\mathcal{W}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \kappa_{iut}\le x_{i(u-1)}(w) & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \kappa_{iut}\le M\cdot g_{iut} & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \kappa_{iut}\ge x_{i(u-1)}(w)-(1-M\cdot g_{iut}) & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \dot{\kappa}_{iut}\le x_{iu}(w) & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \dot{\kappa}_{iut}\le M\cdot h_{iut} & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \dot{\kappa}_{iut}\ge x_{iu}(w)-(1-M\cdot h_{iut}) & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T},\\ \kappa_{iut},\dot{\kappa}_{iut}\in[0,M] & \forall w\in\mathcal{W}, i\in\mathcal{L}, 2\le u\le|\mathcal{U}|-1, t\in\mathcal{T}. \end{cases} \tag{62}$$

## Appendix C

In this appendix, we present the Proof of Proposition 1.

**Proof.** Suppose the probability distribution vector $\rho$ belongs to discrepancy-based ambiguity set $\mathscr{P}$, then we have

$$\begin{aligned} &\max_{\rho\in\mathscr{P}} \mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})^T\rho\\ &= \mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})\rho_0 + \max_{\varepsilon}\left\{\mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})^T\rho\varepsilon \,\middle|\, \mathbf{e}^T\rho\varepsilon=0, \rho_0+\rho\varepsilon\ge 0, \|\varepsilon\|_1\le 1\right\}\\ &= \mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})^T\rho_0 + \Gamma^*(\mathbf{OBJ}), \end{aligned}$$

where $\Gamma^*(\mathbf{OBJ})$ is the optimal value of the following convex program

$$\max_{\varepsilon}\left\{\mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})^T\rho\varepsilon \,\middle|\, \mathbf{e}^T\rho\varepsilon=0, \rho_0+\rho\varepsilon\ge 0, \|\varepsilon\|_1\le 1\right\}. \tag{63}$$

Besides, the Lagrange function of the problem (63) is

$$L(\varepsilon,\theta,\zeta,\upsilon) = \mathbf{OBJ}(\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi})^T\rho\varepsilon + \theta^T(\rho_0+\rho\varepsilon) + \zeta\mathbf{e}^T\rho\varepsilon + \upsilon(1-\|\varepsilon\|_1),$$

where $\theta$, $\zeta$, and $\upsilon$ represent the dual variables corresponding to the constraints in program (63), i.e., $\rho_0+\rho\varepsilon\ge 0, \mathbf{e}^T\rho\varepsilon=0, \|\varepsilon\|_1\le 1$, respectively.

Then, we can further obtain the Lagrange dual function of the model (63):

$$\begin{aligned} g(\theta,\zeta,\upsilon) &= \max_{\varepsilon} L(\varepsilon,\theta,\zeta,\upsilon)\\ &= \theta^T\rho_0 + \upsilon + \max_{\varepsilon}\left\{\mathbf{OBJ}^T\rho\varepsilon + \theta^T\rho\varepsilon + \zeta\mathbf{e}^T\rho\varepsilon - \upsilon\|\varepsilon\|_1\right\}\\ &= \theta^T\rho_0 + \upsilon + \max_{\varepsilon}\left\{(\mathbf{OBJ}^T\rho + \theta^T\rho + \zeta\mathbf{e}^T\rho)\varepsilon - \upsilon\|\varepsilon\|_1\right\}\\ &= \theta^T\rho_0 + \upsilon + f^*(\mathbf{OBJ}^T\rho + \theta^T\rho + \zeta\mathbf{e}^T\rho), \end{aligned}$$

where

$$f^*(\mathbf{OBJ}^T\rho + \theta^T\rho + \zeta\mathbf{e}^T\rho) = \begin{cases} 0 & \text{if } \|\mathbf{OBJ}^T\rho + \theta^T\rho + \zeta\mathbf{e}^T\rho\|_\infty \le \upsilon\\ \infty & \text{otherwise.} \end{cases}$$

Therefore, in essence, the dual problem of the model (63) is:

$$\min_{\boldsymbol{\theta},\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi},\upsilon,\zeta} \left\{ \boldsymbol{\theta}^T \boldsymbol{\rho}_0 + \upsilon \middle| \|\mathbf{OBJ}^T \boldsymbol{\varrho} + \boldsymbol{\theta}^T \boldsymbol{\varrho} + \zeta \mathbf{e}^T \boldsymbol{\varrho}\|_\infty \le \upsilon, \boldsymbol{\theta} \ge 0, \zeta \ge 0 \right\}. \tag{64}$$

To sum up, the equivalent form of $\max_{\mathbf{p}\in\mathscr{P}} \left[\mathbf{OBJ}(\boldsymbol{\iota}, \mathbf{x}, \mathbf{y}, \mathbf{z})^T \mathbf{p}\right]$ can be formulated as follows

$$\begin{aligned} \min_{\boldsymbol{\theta},\boldsymbol{\iota},\mathbf{s},\mathbf{y},\mathbf{z},\mathbf{e},\mathbf{x},\mathbf{q},\boldsymbol{\eta},\boldsymbol{\pi},\upsilon,\zeta} \quad & \mathbf{OBJ}^T \boldsymbol{\rho}_0 + \boldsymbol{\theta}^T \boldsymbol{\rho}_0 + \upsilon \\ \text{s.t.} \quad & \|\mathbf{OBJ}^T \boldsymbol{\varrho} + \boldsymbol{\theta}^T \boldsymbol{\varrho} + \zeta \mathbf{e}^T \boldsymbol{\varrho}\|_\infty \le \upsilon, \\ & \boldsymbol{\theta} \ge 0, \zeta \ge 0. \end{aligned} \tag{65}$$

The other constraints are not related to the distributional ambiguity and can be directly added to the computationally tractable reformulation. The proof is thus complete. □

## References


Basciftci, B., Van Hentenryck, P., 2023. Capturing travel mode adoption in designing on-demand multimodal transit systems. Transp. Sci. 57 (2), 351–375.

Chen, Y., An, K., 2021. Integrated optimization of bus bridging routes and timetables for rail disruptions. European J. Oper. Res. 295 (2), 484–498.

Chen, Z., Li, X., Zhou, X., 2019. Operational design for shuttle systems with modular vehicles under oversaturated traffic: Discrete modeling method. Transp. Res. B 122, 1–19.

Chen, Z., Li, X., Zhou, X., 2020. Operational design for shuttle systems with modular vehicles under oversaturated traffic: Continuous modeling method. Transp. Res. B 132, 76–100.

Chen, X., Wang, Y., Ma, X., 2021a. Integrated optimization for commuting customized bus stop planning, routing design, and timetable development with passenger spatial-temporal accessibility. IEEE Trans. Intell. Transp. Syst. 22 (4), 2060–2075.

Chen, X., Wang, Y., Wang, Y., Qu, X., Ma, X., 2021b. Customized bus route design with pickup and delivery and time windows: Model, case study and comparative analysis. Expert Syst. Appl. 168, 114242.

Cortés, C.E., Gil, C., Gschwender, A., Rey, P.A., 2023. The bus synchronization timetabling problem with dwelling times. Transp. Res. B 174, 102773.

Dai, Z., Liu, X.C., Chen, X., Ma, X., 2020. Joint optimization of scheduling and capacity for mixed traffic with autonomous and human-driven buses: A dynamic programming approach. Transp. Res. C 114, 598–619.

Fu, Z., Chow, J., 2023. Dial-a-ride problem with modular platooning and en-route transfers. Transp. Res. C 152, 104191.

Galarza Montenegro, B.D., Sörensen, K., Vansteenwegen, P., 2021. A large neighborhood search algorithm to optimize a demand-responsive feeder service. Transp. Res. C 127, 103102.

Gkiotsalitis, K., Van Berkum, E., 2020. An analytic solution for real-time bus holding subject to vehicle capacity limits. Transp. Res. C 121, 102815.

Gu, Y., Chen, A., 2023. Modeling mode choice of customized bus services with loyalty subscription schemes in multi-modal transportation networks. Transp. Res. C 147, 104012.

Guo, Q.W., Chow, J., Schonfeld, P., 2018. Stochastic dynamic switching in fixed and flexible transit services as market entry-exit real options. Transp. Res. C 94, 288–306.

Ibarra-Rojas, O.J., Giesen, R., Rios-Solis, Y.A., 2014. An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. Transp. Res. B 70, 35–46.

Khan, Z.S., He, W., Menéndez, M., 2023. Application of modular vehicle technology to mitigate bus bunching. Transp. Res. C 146, 103953.

Kim, M.E., Schonfeld, P., 2014. Integration of conventional and flexible bus services with timed transfers. Transp. Res. B 68, 76–97.

Leffler, D., Burghout, W., Jenelius, E., Cats, O., 2021. Simulation of fixed versus on-demand station-based feeder operations. Transp. Res. C 132, 103401.

Lin, J., Nie, Y.M., Kawamura, K., 2023. An autonomous modular mobility paradigm. IEEE Intell. Transp. Syst. Mag. 15 (1), 378–386.

Liu, Z., Homem de Almeida Correia, G., Ma, Z., Li, S., Ma, X., 2023. Integrated optimization of timetable, bus formation, and vehicle scheduling in autonomous modular public transport systems. Transp. Res. C 155, 104306.

Luo, S., Nie, Y.M., 2020. Paired-line hybrid transit design considering spatial heterogeneity. Transp. Res. B 132, 320–339.

Ma, W., Zeng, L., An, K., 2023. Dynamic vehicle routing problem for flexible buses considering stochastic requests. Transp. Res. C 148, 104030.

Narayan, J., Cats, O., Van Oort, N., Hoogendoorn, S., 2020. Integrated route choice and assignment model for fixed and flexible public transport systems. Transp. Res. C 115, 102631.

NEXT Future Transportation Inc., 2019. NEXT is now. https://www.next-future-mobility.com/. (Accessed 12 September 2023).

NEXT Future Transportation Inc., 2023. Dubai experiments the future of transportation, with NEXT. https://www.next-future-mobility.com/post/dubai-with-next-experiments-with-the-future-of-transportation/. (Accessed 12 September 2023).

Niu, H., Zhou, X., 2013. Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. Transp. Res. C (ISSN: 0968-090X) 36, 212–230.

Rahimian, H., Mehrotra, S., 2022. Frameworks and results in distributionally robust optimization. Open J. Math. Optim. 3, 1–85.

Sadrani, M., Tirachini, A., Antoniou, C., 2022. Optimization of service frequency and vehicle size for automated bus systems with crowding externalities and travel time stochasticity. Transp. Res. C 143, 103793.

Sharif Azadeh, S., Atasoy, B., Ben-Akiva, M.E., Bierlaire, M., Maknoon, M., 2022a. Choice-driven dial-a-ride problem for demand responsive mobility service. Transp. Res. B 161, 128–149.

Sharif Azadeh, S., Van der Zee, J., Wagenvoort, M., 2022b. Choice-driven service network design for an integrated fixed line and demand responsive mobility system. Transp. Res. A 166, 557–574.

Shehadeh, K.S., 2023. Distributionally robust optimization approaches for a stochastic mobile facility fleet sizing, routing, and scheduling problem. Transp. Sci. 57 (1), 197–229.

Shi, X., Li, X., 2021. Operations design of modular vehicles on an oversaturated corridor with first-in, first-out passenger queueing. Transp. Sci. 55 (5), 1187–1205.

Steiner, K., Irnich, S., 2020. Strategic planning for integrated mobility-on-demand and urban public bus networks. Transp. Sci. 54 (6), 1616–1639.

Tian, Q., Lin, Y.H., Wang, D.Z., 2023. Joint scheduling and formation design for modular-vehicle transit service with time-dependent demand. Transp. Res. C 147, 103986.

Tian, Q., Lin, Y.H., Wang, D.Z., Liu, Y., 2022. Planning for modular-vehicle transit service system: Model formulation and solution methods. Transp. Res. C 138, 103627.

United Nations, 2019. Shifting demographics: A visual guide. https://www.un.org/en/un75/shifting-demographics. (Accessed 3 September 2023).

Van Parys, B.P., Esfahani, P.M., Kuhn, D., 2021. From data to decisions: Distributionally robust optimization is optimal. Manage. Sci. 67 (6), 3387–3402.

Wang, H., 2019. Routing and scheduling for a last-mile transportation system. Transp. Sci. 53 (1), 131–147.

Wang, S., Chen, Z., Liu, T., 2020. Distributionally robust hub location. Transp. Sci. 54 (5), 1189–1210.

Wang, Y., Lin, X., He, F., Li, M., 2022. Designing transit-oriented multi-modal transportation systems considering travelers' choices. Transp. Res. B 162, 292–327.
Wu, W., Liu, R., Jin, W., Ma, C., 2019. Stochastic bus schedule coordination considering demand assignment and rerouting of passengers. Transp. Res. B 121, 275–303.
Xia, D., Ma, J., Sharif Azadeh, S., Zhang, W., 2023. Data-driven distributionally robust timetabling and dynamic-capacity allocation for automated bus systems with modular vehicles. Transp. Res. C 155, 104314.
Yin, J., D'Ariano, A., Wang, Y., Yang, L., Tang, T., 2021. Timetable coordination in a rail transit network with time-dependent passenger demand. European J. Oper. Res. 295 (1), 183–202.
Zhang, W., Xia, D., Liu, T., Fu, Y., Ma, J., 2021. Optimization of single-line bus timetables considering time-dependent travel times: A case study of Beijing, China. Comput. Ind. Eng. 158, 107444.