



Research
Engineering Management—Article

A Distributionally Robust Optimization Method for Passenger Flow Control Strategy and Train Scheduling on an Urban Rail Transit Line



Yahan Lu, Lixing Yang*, Kai Yang, Ziyou Gao*, Housheng Zhou, Fanting Meng, Jianguo Qi

State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Article history:

Received 8 June 2021

Revised 8 September 2021

Accepted 22 September 2021

Available online 24 December 2021

Keywords:

Passenger flow control

Train scheduling

Distributionally robust optimization

Stochastic and dynamic passenger demand

Ambiguity set

ABSTRACT

Regular coronavirus disease 2019 (COVID-19) epidemic prevention and control have raised new requirements that necessitate operation-strategy innovation in urban rail transit. To alleviate increasingly serious congestion and further reduce the risk of cross-infection, a novel two-stage distributionally robust optimization (DRO) model is explicitly constructed, in which the probability distribution of stochastic scenarios is only partially known in advance. In the proposed model, the mean-conditional value-at-risk (CVaR) criterion is employed to obtain a tradeoff between the expected number of waiting passengers and the risk of congestion on an urban rail transit line. The relationship between the proposed DRO model and the traditional two-stage stochastic programming (SP) model is also depicted. Furthermore, to overcome the obstacle of model solvability resulting from imprecise probability distributions, a discrepancy-based ambiguity set is used to transform the robust counterpart into its computationally tractable form. A hybrid algorithm that combines a local search algorithm with a mixed-integer linear programming (MILP) solver is developed to improve the computational efficiency of large-scale instances. Finally, a series of numerical examples with real-world operation data are executed to validate the proposed approaches.

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The coronavirus disease 2019 (COVID-19) has been greatly affecting people's daily travel patterns, resulting in significant uncertainty regarding passenger demand in urban public transit. As the main artery of public transit, the urban rail transit system transports a large number of passengers shuttling between different lines and stations every day. In some megacities, the passenger demand in peak hours far exceeds the transport capacity at busy stations (especially at some transfer stations), and a great many passengers must wait on the platforms and are even stranded, resulting in a serious crowdedness issue [1]. According to the World Health Organization, people who leave their home during an epidemic period should refrain from crowding to avoid cross-infection. In other words, passenger congestion in the confined environment of stations necessarily increases the chance of cross-infection and decreases travel comfort. Thus, practical and effective measures—such as limiting the vehicle loading capacity,

keeping a safe distance between passengers, and so forth—are needed to ensure safe operation of rail transit with the gradual restoration of social order and the increase in traffic volume brought by the resumption of work and schools.

At present, two types of measures have been applied in China's urban rail transit systems, which do indeed mitigate the risk of congestion at stations with large passenger demand: ① The first measure involves passenger flow control. To cope with large passenger demand, the best choice is to impose passenger flow control; that is, to limit the speed of passenger flow entering the platform and make passengers queue in an orderly manner in station halls, so as to avoid passengers gathering on the platforms and thus reduce the risk of cross-infection while improving operation safety. For example, so far, up to 96 stations in the Beijing rail transit system have imposed routine passenger control strategies. ② Another practical method is demand-oriented train scheduling; this is a traditional approach used to deal with heavily congested passenger flow in peak hours, which has been studied deeply in the past years [2–4]. More precisely, based on detailed information on dynamic passenger demand, this method aims to shorten the train headway to enhance the capacity of the urban rail transit line/system and optimize train schedules so as to minimize the

* Corresponding authors.

E-mail addresses: lxyang@bjtu.edu.cn (L. Yang), zygao@bjtu.edu.cn (Z. Gao).

passenger waiting time at stations. However, the phenomenon of huge passenger flow during peak hours has become normative in some megacities, such that even the maximum departure frequency is insufficient to significantly ease the heavy congestion [5]. In addition, as far as most transfer stations are concerned, the number of transfer passengers in practical operations has far exceeded the volume of outside arrival passengers. Xizhimen Station in Beijing can be taken as an illustration. As shown in Fig. 1, the transfer passenger demand at this station is always greater than the outside arrival passenger demand, which is particularly obvious during peak periods. Although some studies [5–9] have paid a great deal of attention to this problem, most still neglect to control transfer passengers simultaneously with outside arrival passengers. To relieve crowding and enhance efficiency as much as possible on an overall line, there is an urgent need to study passenger flow control strategies combined with train scheduling under the influence of transfer passengers—a need that is not addressed in the existing literature.

In reality, some important information on the urban rail transit system, such as passenger demand, is usually uncertain. To capture this uncertainty, studies [10,11] have used stochastic scenarios with deterministic probability distributions to characterize the randomness of passenger demand, and the scenario-based time-dependent passenger demand can be recorded by the automatic fare collection (AFC) system. However, it is difficult to obtain the accurate probability distribution of the involved scenarios in advance, especially in real-world applications [12]. Based on partial known probability information, the modeling methodology corresponds to the distributionally robust optimization (DRO) approach, which has recently been applied to the field of transportation planning and management. Enlightened by this method, the present study aims to find a robust passenger flow control strategy that works in all passenger demand scenarios with partially known probability. For clarity, Fig. 2 shows the passenger control process, in which the time-dependent passenger demand of two scenarios is represented by blue and red curves, respectively, and a robust passenger flow control strategy (represented by the purple line) is generated to control the passenger entering speed and thereby reduce platform congestion. This decision-making process is associated with the following potential risks incurred by the two-fold uncertainty: ① Dynamic outside arrival passengers and transfer passengers are both highly uncertain, and are described by stochastic scenario sets; and ② in practice, operators might only know the support of the probability distribution information of the scenarios, while precise distribution

information is out of reach. Under these uncertain conditions, risk management [13] has not been sufficiently studied in the existing literature related to the passenger flow control problem, to the best of our knowledge, so it will also be explicitly addressed in our study.

1.1. Literature review

Imbalance of supply and demand is the main cause of congestion on an urban rail transit line. Based on this fact, the existing literature has focused on regulating capacity and demand from the supply and demand sides, respectively, in order to alleviate congestion. On the supply side, the literature mainly optimizes the train schedule to meet the dynamic passenger demand with uneven temporal and spatial distribution, so as to avoid a deficiency or insufficiency of transport capacity; on the demand side, a passenger flow strategy, differential pricing, and other measures are mainly used to guide and limit the passenger flow. In addition, the existing studies that focus on dealing with uncertainty in the field of transportation tend to adopt different methodologies, including stochastic programming (SP), robust optimization (RO), and DRO. The following discussion provides a detailed overview of the state of art in the literature.

1.1.1. Passenger flow control

Due to the increasing gap between the limited transport capacity and the continuous growth of passenger demand, many urban rail transit lines in China are suffering from overcrowding. The question of how to deal with this problem is thus the focus of operation management departments and many researchers. As a substantial increase in transport capacity is almost impossible in the near future, implementing passenger flow control at congested stations or lines is the best choice to reduce crowdedness. Many researchers have devoted time and effort to this problem, and the related studies can be summarized into the following two categories: The first category is based on the mathematical programming approach and the other one is based on the computer simulation method.

- **Passenger flow control based on the mathematical programming approach.** In recent years, an increasing number of researchers have employed the mathematical programming approach to study the passenger flow control problem, which includes three stratification levels: the station level, line level, and network level. Considering the fixed train schedules and the time-dependent passenger demand,

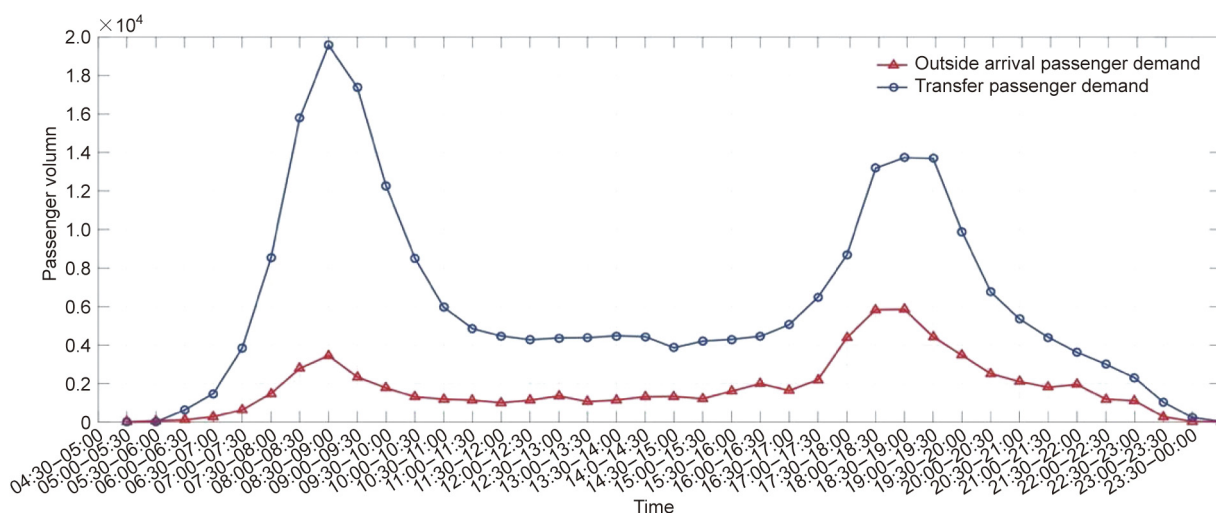


Fig. 1. Illustration of inbound and transfer passengers at Xizhimen Station.

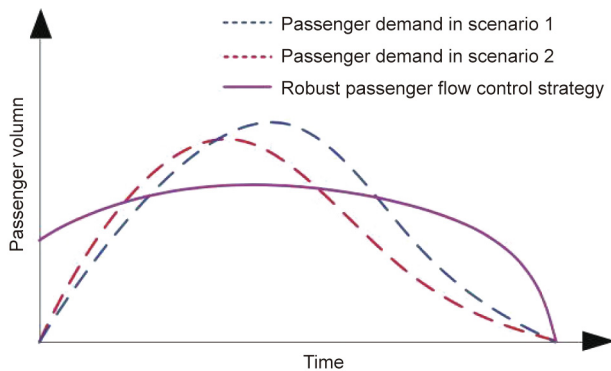


Fig. 2. Illustration of the robust passenger flow control process.

existing studies generally aim to minimize the delay time or the waiting time of passengers. For example, Xu et al. [6] proposed a multi-objective optimization model for an urban rail transit station under uncertain demand in order to calculate the number of passengers that should be controlled through the genetic algorithm. In addition, Xu et al. [8] addressed the problem of passenger flow control in an urban rail transit network during peak hours. Their problem was formulated as a mixed-integer programming model to simultaneously adjust the number of inbound and transfer passengers, which could be solved by the genetic algorithm. Shi et al. [9] developed an integer linear programming (ILP) model to generate an optimal passenger flow control strategy for an urban rail transit system. Based on analyses of the alighting and boarding processes, Wang et al. [14] formulated an integer programming model to achieve the optimal state for the entire urban rail transit line, with a focus on minimizing the average passenger delay. Yuan et al. [15] established a mixed-ILP (MILP) model in order to minimize the total passenger waiting time in an urban rail transit network and solved it by CPLEX software. By introducing a timetable-oriented space-time network representation, Meng et al. [16] proposed an ILP model to produce high-quality passenger flow control strategies for an urban rail transit line.

- Passenger flow control based on the computer simulation method.** In view of the complex dynamic characteristics of the operation state in an urban rail transit system, computer simulation modeling can reveal the occurrence and evolution of passenger flow congestion, and thus give full play to the advantages of simulation modeling in this problem. Numerous published studies use the simulation method to investigate the passenger flow control at stations. Hoogendoorn and Daamen [17] established microscopic pedestrian dynamics models to assess the design safety in large passenger flow scenarios in Lisbon metro stations. To simulate passenger flow at the Xuanwumen Station (a transfer station) in the Beijing rail transit system during peak hours, Zhang et al. [18] used Vissim software, which achieved good performance. Seriani and Fernandez [19] studied the effects of pedestrian traffic management on passengers' boarding and alighting time, and used experimental means to obtain the standards for pedestrian traffic management at the platform and doors of metro vehicles. Fei and Liu [20] used Anylogic software to formulate a simulation model of passenger flow organization and accurately identified the congestion points according to the simulation results. Jiang et al. [21] proposed a coordination model of passenger flow control and train jumping based on the Q -learning approach, which could be efficiently solved by simulation methods.

1.1.2. Train scheduling in urban rail transit

Over the past several decades, a large and growing body of literature has proposed different mathematical models and solution algorithms for train scheduling problem in urban rail transit. For example, to minimize passenger waiting time, Niu and Zhou [4] developed a nonlinear programming model to optimize the train schedule for the whole line, which could be optimized by the genetic algorithm. They then took oversaturated situations into consideration and formulated a mixed-integer nonlinear programming model [22]. Li et al. [23] studied the fairness issue in the train scheduling process. To optimize train schedules by considering min-max fairness and α fairness, they established a mixed-integer programming model that could be solved by a simulated annealing-based adaptive large neighborhood search algorithm. Yin et al. [24] proposed an MILP model to fix the coordinated train scheduling problem with unlimited train boarding capacity. Yang et al. [25] developed a 0–1 ILP model for the last train timetable in an urban rail transit network, which could be effectively solved by a heuristic algorithm based on Lagrangean relaxation. Liu et al. [26] formulated a mathematical optimization model to determine a robust train schedule for maximizing robustness and minimizing total energy consumption, so as to avoid delay propagation as much as possible. Huang et al. [27] considered the rescheduling problem in urban rail transit systems. Two mixed-integer nonlinear programming models with different recovery strategies were formulated to reschedule trains during track disruption, which could be linearized by a hybrid approach combining big- M and time-indexed formulations. Tian and Niu [28] addressed the demand-oriented train scheduling problem under overtaking operations. Their problem was formulated into an ILP model, and a corresponding novel surrogate-dual-variable column generation algorithm was proposed to generate approximate optimal solutions. Mo et al. [29] proposed a mixed-integer nonlinear programming model to generate an optimized train timetable with suitable train speed profiles, as well as an efficient rolling stock plan, aiming to minimize both passenger waiting and travel times. Qi et al. [30] formulated an ILP model for the integrated stop planning and timetabling problem, with the aim of minimizing the total travel time of trains.

Overall, the above studies address the passenger flow control problem and the train scheduling problem separately, thus failing to achieve an optimal system effect or reduce the pressure on urban rail transit operation to the greatest extent. It should be noted that integrated optimization of the above two aspects is much more challenging than dealing with them separately because of the complex coupling relationship between trains and passenger flow. To date, only three papers can be found in the literature that use a mathematical programming approach to carry out an integrated study of these two aspects: Xu [31] formulated a quadratic programming model to generate a multi-station collaborative passenger flow control strategy and a corresponding train schedule; in this model, the decision variables include the number of inbound passengers and the dwell time of trains at certain key stations. Shi et al. [5] established an ILP model to obtain an optimal passenger flow control strategy for the whole line and a corresponding train schedule to improve the service quality, in which the passenger flow control strategy was implemented on each timestamp. Gong et al. [32] implemented the passenger flow control based on the train status. However, none of these studies consider the influence of transfer passengers.

1.1.3. Approaches for dealing with uncertainty

It is well known that transportation activities usually faces a lot of dynamics and uncertainties due to system complexity [1,11,33,34]. With the development of optimization theory and the improvement of computing technologies, an increasing

amount of attention has been paid to how to deal with uncertain parameters. Thus far, three major kinds of approaches—namely, SP, RO, and DRO—have been developed to handle uncertainty. More specifically, the SP method assumes that the probability distribution of uncertain parameters is known in advance. For example, Gong et al. [35] considered the stochastic environment in an urban rail transit system and developed an integer nonlinear programming model to simultaneously optimize the number of train services, the headway settings, and the speed profile selection decisions. Errico et al. [36] constructed a two-stage SP model for the vehicle routing problem with hard time windows, which could be solved by exact branch-cut-and-price algorithms. It is widely recognized that the SP method is based on a given probability distribution. However, the real value of probability distribution is difficult to obtain in practice and usually needs to be estimated through historical data. Due to the existence of estimated errors, major deviations may occur in practical applications, which is a great weakness of this method. The RO approach considers that uncertain parameters are all based on uncertainty sets without any distribution assumption. The aim of the RO approach is to overcome the uncertainty and find the best possible solutions that satisfy all constraints. For example, to handle the uncertainty of passenger demand, Yang et al. [10] proposed an MILP model with the max–min reliability criterion, aiming to increase the number of last-train successful transfer passengers and reduce the total running time. Qi et al. [37] studied the train timetabling and stop-planning problem with uncertain passenger demand. An ILP model was formulated and solved by CPLEX software. Based on the technique of Light Robustness, Cacchiani et al. [38] proposed three MILP models to derive robust train-stop plans and timetables, in order to reduce passengers' inconvenience. Here, we note that a serious drawback of the RO approach is that it is too rigid and conservative; however, it does not require the distribution information to be known in advance, unlike the SP method.

The rapid development of the DRO approach has overcome the aforementioned shortcomings of the SP and RO methods. The DRO approach combines the statistical learning technique with optimization theory; it can obtain a good enough solution by assuming that the parameters obey some possible distributions. The theories underpinning the DRO approach have been well studied over the last decades [39,40]. For example, Delage and Ye [39] produced one of the first studies on this approach, in which the DRO models were formulated under the first-order and second-order moment inequalities, respectively. Esfahani and Kuhn [40] proved that the dual form of a Wasserstein-metric DRO model could be decomposed into finitely convex optimization problems with tractable computational complexity. Recently, this approach has also attracted much attention from certain engineering fields [41–47]. For example, focusing on the planning and scheduling problem under uncertain demand, Shang and You [41] formulated a two-stage DRO model to produce less conservative solutions. For finite and infinite horizon optimal control problems, Van et al. [42] proposed distributionally robust constrained formulations to provide frameworks that could generate robust control policies. Considering the integration of urban and rural public transport systems, Shang et al. [43] developed a DRO model with ambiguous chance constraints regarding the travel time requirement, with the aim of minimizing the total construction cost. For a disaster relief management problem under demand uncertainty, Wang et al. [44] proposed a DRO model to simultaneously optimize the integrated facility location, inventory pre-positioning, and delivery decisions. Yang et al. [45] presented a distributionally robust chance-constrained program to model the last-train coordination planning problem in an urban rail transit system, in which uncertain passenger demand was captured by an ambiguous set. Wang et al. [46] focused on the hub location problem with multiple commodities

under uncertain demand and cost. This problem was modeled as a DRO model, which was further reformulated as a moderate-sized MILP for ease of calculation. Zhang et al. [47] established a DRO model with the Wasserstein distance-based ambiguity set for the vehicle routing problem with time windows, aiming to minimize the risk of late customer arrivals.

Thus far, situations with uncertain passenger demand have received less attention in the area of passenger flow control. As far as we know, there is only one relevant reference in this field that addresses uncertain demand: Meng et al. [11] proposed an ILP model on an urban rail transit line by applying the SP method; in this model, uncertain passenger demand was described using a stochastic scenario set with completely pre-given distribution information. Moreover, there is no relevant theoretical achievement employing the DRO approach in the field of passenger flow control combined with train scheduling.

In summary, the passenger flow control problem has been a hot research topic in recent years, and some mathematical models and simulation methods to address this problem have been well studied. Nevertheless, most of these approaches still separately consider or partially integrate the three aspects of passenger flow control, train scheduling, and SP method. Moreover, no existing study simultaneously deals with the two-fold uncertainty with respect to the passenger demand and scenario probability, even though it is a significant and challenging issue for real-world applications, as it is difficult to predict passenger demand accurately. With this concern, this paper is particularly interested in developing a DRO method for the passenger flow control strategy and train scheduling on an urban rail transit line, as well as a powerful algorithm with fast computing speed.

1.2. The contributions of this study

To fill the abovementioned gaps, this paper aims to make the following contributions to the study of passenger flow control and train scheduling problem on an urban rail transit line:

(1) A DRO model, which focuses not only on both outside arrival passengers and transfer passengers, but also on train schedules, is rigorously formulated. Compared with the existing literature, this model has two significant improvements: ① The proposed model considers the inherent uncertainty of passenger demand. Since passenger demand in the real world is usually uncertain and dynamic, a series of stochastic scenarios are developed to depict this uncertain parameter in this paper. ② The precise probability distribution information of each stochastic scenario does not need to be known completely in advance—that is, only the partial distribution information is needed, which is the main reason why this approach is employed in this paper. Since very few existing studies consider the inherent uncertainty of passenger demand and the characteristics of scenario probability cannot be predicted in advance, this is a novel idea.

(2) Our distributionally robust formulation is a semidefinite programming model, which is a notoriously difficult problem to solve. To convert this model into a computationally tractable form, a discrepancy-based ambiguity set is employed. Furthermore, the equivalent MILP model can be obtained through the strict theoretical proofs. Based on an analysis of the complexity of our formulation, we propose an efficient hybrid algorithm by combining a local search algorithm (LS) with a suitable MILP solver to generate high-quality solutions in an acceptable computation time, especially for large-scale instances.

(3) Finally, to demonstrate the effectiveness and practicability of our proposed approaches, a series of numerical experiments are implemented based on operating data from Line 15 in the Beijing rail transit system. We compare the experimental results of our DRO model with those of the traditional SP model and the

classic RO model. These experiments demonstrate the performance and advantages of the proposed optimization model and solution approaches.

The remainder of this paper is organized as follows. In Section 2, we give a detailed description of the passenger flow control and train scheduling problem on an urban rail transit line. In Section 3, we formulate the problem into a two-stage SP model and a DRO method, respectively. Then, Section 4 proposes a hybrid algorithm based on the LS and a suitable MILP solver. In Section 5, we implement a case study, that is, a large-scale case based on the operating data of Line 15 in the Beijing rail transit system, to show the superiority of our proposed approaches. Finally, Section 6 gives some conclusions and future research.

2. Problem description

Consider an oversaturated urban rail transit line, which consists of a series of common and transfer stations (S^{or} and S^{tr} are used to denote the sets of common and transfer stations, respectively), and a set of physical segments between stations (the k th segment is defined as the connection from stations k to $(k + 1)$). Denote the set of stations on this line as S , then $S = S^{or} \cup S^{tr} = \{1, 2, \dots, |S|\}$. In operations, the transfer between different lines is a common passenger behavior at each transfer station. In order to ease congestion, we aim to collaboratively optimize the train schedule and passenger flow control strategy on this oversaturated line by taking the transfer passenger flow into account. For descriptive clarity, an illustrative urban rail transit system is depicted in Fig. 3, where Line 1 is the focus line, and Lines 2 and 3 are the adjacent lines. Typically, Stations 2, 3, 5 $\in S^{or}$, and Stations 1, 4 $\in S^{tr}$. Passengers on Lines 2 and 3 are allowed to transfer to Line 1 at Stations 1 and 4, respectively. At Stations 2, 3, and 5, the inbound passengers are mainly from the outside of the transit system. At Stations 1 and 4, both outside arrival passengers and transfer passengers from Lines 2 and 3 should be taken into consideration, in the process of optimizing the passenger flow control strategies.

It is widely recognized that passenger demand in an urban rail transit system is time-dependent and varies at different time periods, during which outside arrival passengers continuously come, while transfer passengers are intermittently centralized due to the arrival of trains from adjacent lines at certain intervals. In addition, passenger demand is often uncertain in practical operations due to the influence of various factors. Thus, to describe the dynamic and stochastic passenger demand more realistically, the time horizon is discretized into a series of timestamps with

discretized time granularity t_{unit} in this study, denoted by $T = \{1, 2, \dots, |T|\}$, and a stochastic scenario set—that is, $W = \{\omega_1, \omega_2, \dots, \omega_{|W|}\}$ —is employed to characterize the uncertainty. Note that, since it is difficult to know the probability of each scenario precisely in advance, in order to be more realistic, we characterize this parameter by imprecise discrete probability distributions. Similar to the existing literature [11], the dynamic and stochastic passenger demand is depicted by $P_k^{come,out}(t, \omega)$ and $P_k^{come,tr}(t, \omega)$, which represent the number of outside arrival passengers and the number of transfer passengers from adjacent lines at station k and timestamp t in scenario ω , respectively. It should be noted that the aforementioned passenger demand can be obtained from historical AFC data through travel demand assignment approaches [48].

For an urban rail transit line with uncertain passenger demand, congestion is directly related to the following four factors: outside arrival passengers, transfer passengers from adjacent lines, the schedules of involved trains, and the train capacity. Due to the limited improvement of train capacity in real-world operations, the most effective method to alleviate congestion is to consider the effects of the other three factors from the system perspective, and to carry out collaborative optimization of the train schedules and passenger flow control strategies. In the process of passenger flow control, outside arrival passengers are first required to queue in station halls, and they are then released to the platform to board train i under the robust passenger flow control strategy at station k in scenario ω (denoted by $P_{i,k}^{inc,out}(\omega)$). Nevertheless, the transfer passengers from adjacent lines (i.e., $P_k^{come,tr}(t, \omega)$) are allowed to go to the platform freely and wait for the next train arrivals, as shown in Fig. 4. Once the number of outside arrival and transfer passengers exceed the capacity of the trains, outside arrival passengers who come later than this time must be detained in the station hall; they then gradually enter the platform according to the robust passenger flow control strategy. In this paper, the control strategy is only imposed on outside arrival passengers, while transfer passengers are released to the platform and can take trains freely. In addition, all the passengers on the platform are required to take the train that arrives earliest, without waiting for the next train, in order to guarantee the operation safety.

In summary, the focus of this paper is to develop a multi-scenario mathematical formulation for the robust passenger flow control problem combined with train scheduling on an urban rail transit line, with the goal of minimizing the sum of operating time, the expected number of waiting passengers, and the conditional value-at-risk (CVaR) of congestion. Our formulation also explicitly

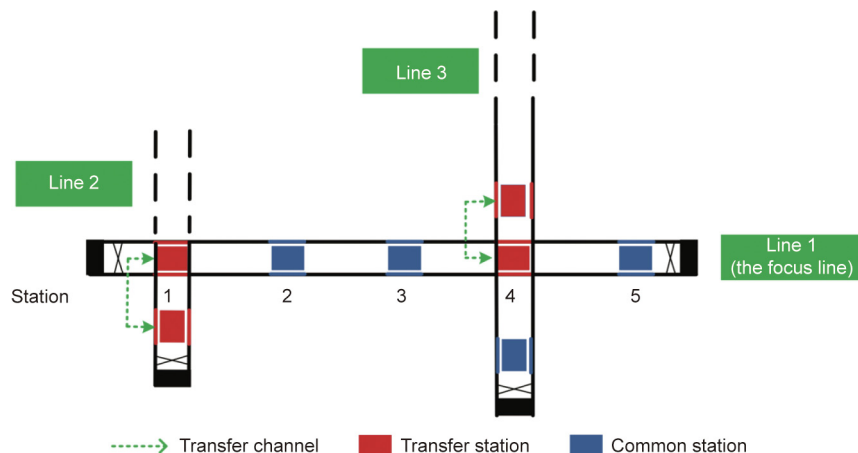


Fig. 3. An illustrative urban rail transit system.

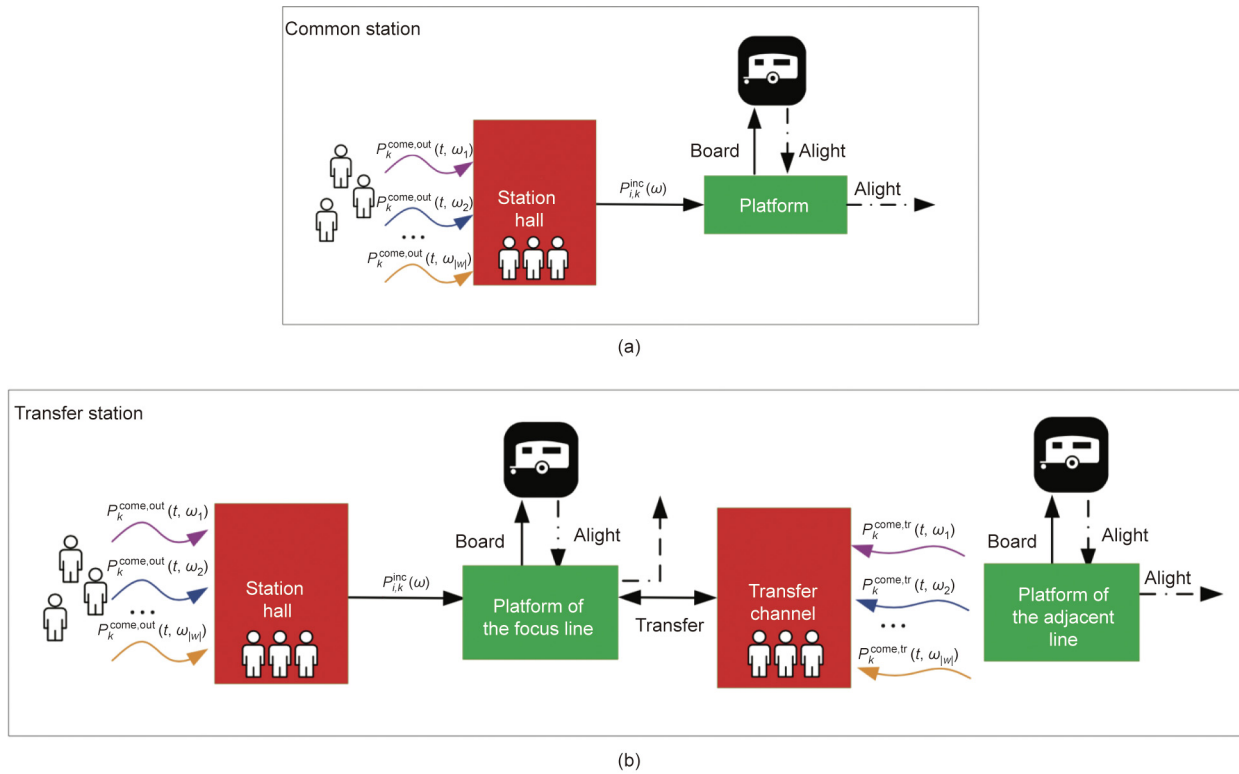


Fig. 4. Robust passenger flow control at common and transfer stations. (a) Robust passenger flow control at a common station; (b) robust passenger flow control at a transfer station.

captures the time-dependent and stochastic passenger demand—that is, the outside arrival passengers and the transfer passengers—and the train capacity, aiming to achieve the optimization of system performance. The proposed mathematical formulation is based on the following three assumptions.

Assumption 1: The outside arrival and transfer passenger demand is known in advance by processing the historical AFC data. A similar assumption has been widely adopted in previous studies [1,8,11].

Assumption 2: The section running time and station dwell time of trains are pre-given by rail managers according to the actual operating data. That is, these parameters are not affected by the numbers of boarding or alighting passengers.

Assumption 3: We only consider the direction from station 1 to station $|S|$ as the operation direction, while the other direction can be operated according to the same optimization method. In addition, the inbound and transfer passengers, who are released to platforms under the robust passenger flow control strategy, can all be taken away by the first arrival train. This assumption is common in passenger flow control and train timetabling studies [4,32].

3. Mathematical formulations

In this section, we first introduce a two-stage SP model for passenger flow control and train scheduling problem with completely known probability information in Section 3.1. We then formulate a novel DRO model with uncertain probability in Section 3.2.

3.1. Two-stage SP model with completely known probability information

To characterize the above problem in a stochastic version, the modeling process will be explicitly discussed in the following

sections. In Section 3.1.1, we first list all the related notations and decision variables in the formulation. The events that describe the dynamic processes of passenger loading and train operations are introduced in Section 3.1.2. In Section 3.1.3, we present a two-stage SP model based on the mean-CVaR criterion.

3.1.1. Notations and decision variables

In this paper, we use boldface letters to represent vectors and V is employed to represent the set of random variables. Uncertainty is modeled via a set of stochastic scenarios, the probability distribution of which is denoted as $\mathbf{p}[\cdot]$. We use $\mathbb{E}_{\mathbf{p}}[x]$ to represent the expectation of x with probability distribution $\mathbf{p}[x]$. Based on these definitions, we next develop a two-stage SP model for the problem of interest, in which the first stage determines the train schedules and the second stage is associated with the movement of passengers. To formulate the model rigorously, five sets of decision variables are defined as follows.

Decision variables:

$t_{i,k}^a$: arrival time of train i at station k (integer variable).

$t_{i,k}^d$: departure time of train i at station k (integer variable).

$d_{i,k}(t)$: departure indicator of in-service train i at station k and timestamp t (binary variable).

$P_{i,k}^{ina,out}(\omega)$: number of outside arrival passengers who actually board train i at station k in scenario ω (integer variable).

$P_{i,k}^{inc,out}(\omega)$: number of outside arrival passengers who are allowed to board train i at station k in scenario ω under the robust passenger flow control strategy (integer variable).

In these decision variables, $t_{i,k}^a$ and $t_{i,k}^d$ determine the arrival and departure times of train i at station k , respectively; $d_{i,k}(t)$ is also a scheduling variable, which corresponds to a train timetable (i.e., the headway between adjacent trains, the arrival and departure times of each train at each station); $P_{i,k}^{ina,out}(\omega)$ and $P_{i,k}^{inc,out}(\omega)$ are

related to the dynamic evolution process of passengers. The detailed sets, parameters, and notations are listed below in order to describe the problem more conveniently.

Sets:

- W : set of involved scenarios, $W = \{\omega_1, \omega_2, \dots, \omega_{|W|}\}$.
- S : set of involved stations on the focus line,
- $S = S^{or} \cup S^{tr} = \{1, 2, \dots, |S|\}$.
- S^{or} : set of involved common stations.
- S^{tr} : set of involved transfer stations.
- I : set of in-service trains, $I = \{1, 2, \dots, |I|\}$.
- T : set of discrete timestamps, $T = \{1, 2, \dots, |T|\}$.

Parameters and notations in the first-stage model:

- k, v : index of stations, $k, v \in S$.
- ω : index of scenarios, $\omega \in W$.
- p_ω : probability of scenario ω .
- i, j : index of trains, $i, j \in I$.
- t : index of time, $t \in T$.
- t_{unit} : discretized time granularity.
- $t_{i,k}^r$: running time of train i from station k to station $(k + 1)$.
- $t_{i,k}^{dw}$: dwell time of train i at station k .
- $h_{i,k}$: departure headway between trains i and $(i + 1)$ at station k .
- $\underline{h}_{i,k}$: minimum headway for two consecutive in-service trains at station k .
- $\bar{h}_{i,k}$: maximum headway for two consecutive in-service trains at station k .
- T_0 : total operating time of the focus line.

Parameters and notations in the second-stage model:

- C : train capacity.
- $P_k^{come.out}(t, \omega)$: number of outside arrival passengers at station k and timestamp t in scenario ω , $P_k^{come.out}(t, \omega) \in V$.
- $P_k^{come.tr}(t, \omega)$: number of passengers who get off trains at station k and make transfer at timestamp t in scenario ω , $P_k^{come.tr}(t, \omega) \in V$.
- $P_{i,k}^{on}(\omega)$: number of in-vehicle passengers in train i when departing from station k in scenario ω , $P_{i,k}^{on}(\omega) \in V$.
- $P_{i,k}^{al}(\omega)$: number of alighting passengers of train i when arriving at station k in scenario ω , $P_{i,k}^{al}(\omega) \in V$.
- $P_{i,k}^{waita.out}(\omega)$: number of passengers who actually wait for train i in the hall of station k in scenario ω , $P_{i,k}^{waita.out}(\omega) \in V$.
- $P_{i,k}^{wait.tr}(\omega)$: number of passengers who wait for making transfers to train i at station k in scenario ω , $P_{i,k}^{wait.tr}(\omega) \in V$.
- $P_{i,k}^{inc}(\omega)$: total number of passengers boarding train i at station k under robust passenger flow control strategy in scenario ω .
- $P_{i,k}^{ina}(\omega)$: total number of passengers actually boarding train i at station k in scenario ω .
- $\delta_{v,k}(\omega)$: ratio of outside arrival passengers who arrive at station v heading to station k in scenario ω , $\delta_{v,k}(\omega) \in V$.

$\beta_{v,k}(\omega)$: ratio of transfer passengers who arrive at station v heading to station k in scenario ω , $\beta_{v,k}(\omega) \in V$.

3.1.2. Passenger loading and train operations

For the sake of clarity, this section introduces a non-increasing binary variable $d_{i,k}(t)$ to denote the departure characteristics of train i at station k and timestamp t . Specifically, $d_{i,k}(t) = 0$ indicates that train i has departed from station k at timestamp t , and $d_{i,k}(t) = 1$ indicates otherwise. As shown in Fig. 5, train i departs from station k at timestamp 2. It is clear that a series of “1” for $d_{i+1,k}(t) - d_{i,k}(t)$ indicates the departure headway between trains i and $(i + 1)$ at station k , during which passengers can board train $(i + 1)$ according to the robust passenger flow control strategy. By using this binary variable, it is possible to formulate the problem of interest as an optimization model with linear forms, as described in the following discussion.

3.1.3. A two-stage SP model

A two-stage SP model based on the mean-CVaR criterion is formulated in this section, in which the distribution of the uncertain passenger demand is assumed to be completely known in advance and is captured by multiple scenarios with deterministic probability distribution. In the first stage, we propose a formulation that is only relevant to train scheduling, shown in the equations below:

$$\min \zeta_1 \cdot T_0 + \zeta_2 \cdot [(1 - \lambda) \mathbb{E}[Q(\mathbf{d}, \xi(\omega))] + \lambda \cdot \text{CVaR}[Q(\mathbf{d}, \xi(\omega))]] \quad (1)$$

$$\text{s.t. } t_{i,k}^a = t_{i,k-1}^d + t_{i,k-1}^r, \quad \forall i \in I, k \in S \setminus \{1\} \quad (2)$$

$$t_{i,k}^d = t_{i,k}^a + t_{i,k}^{dw}, \quad \forall i \in I, k \in S \quad (3)$$

$$h_{i,k} = t_{i+1,k}^d - t_{i,k}^d, \quad \forall i \in I \setminus \{|I|\}, k \in S \quad (4)$$

$$\underline{h}_{i,k} \leq h_{i,k} \leq \bar{h}_{i,k}, \quad \forall i \in I \setminus \{|I|\}, k \in S \quad (5)$$

$$h_{i,k} = \sum_{t \in T} (d_{i+1,k}(t) - d_{i,k}(t)) \cdot t_{unit}, \quad \forall i \in I \setminus \{|I|\}, k \in S \quad (6)$$

$$d_{i,k}(t) \geq d_{i,k}(t + 1), \quad \forall i \in I, k \in S, t, t + 1 \in T \quad (7)$$

$$T_0 = \sum_{i \in I \setminus \{|I|\}} \sum_{k \in S} h_{i,k} + \sum_{i \in I} \sum_{k \in S} t_{i,k}^{dw} + \sum_{i \in I} \sum_{k \in S \setminus \{|S|\}} t_{i,k}^r \quad (8)$$

$$d_{i,k}(t) \in \{0, 1\}, \quad \forall i \in I, k \in S, t \in T \quad (9)$$

where $Q(\mathbf{d}, \xi(\omega))$ is the random optimal value of the second-stage model, \mathbf{d} denotes the first-stage decision vector of train schedules, $\xi(\omega)$ represents the random input data of passenger flows, \mathbb{E} is the expected number of waiting passengers, λ is the tradeoff coefficient that reflects the decision-makers’ risk preference, ζ_1 and ζ_2 are two

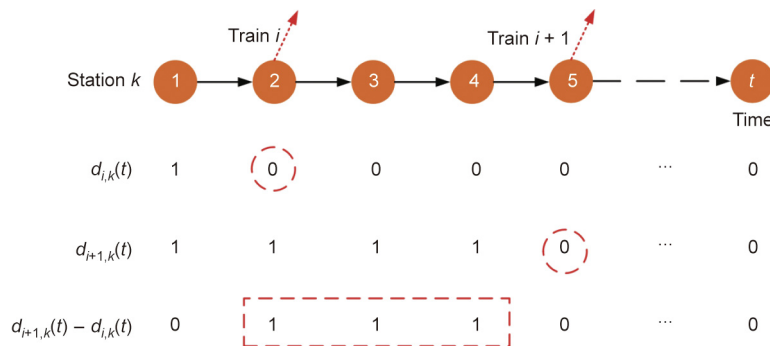


Fig. 5. Illustration of scheduling variables at station k .

cost coefficients, and CVaR denotes the conditional value-at-risk (VaR) criterion, which is a popular risk measure to model the decision-maker’s risk preference. Compared with the VaR criterion, CVaR highlights the average level of excess losses that occur beyond the VaR cutoff point in the distribution. Therefore, CVaR is superior to VaR in reducing uncertainty risk, since more risk-averse decisions can be obtained by adopting the CVaR criterion. This could be particularly helpful in preventing serious situations in which a large number of passengers congregate at an urban rail transit station.

The objective function in Eq. (1) aims to minimize the sum of the operating time (i.e., T_0), the expected number of waiting passengers, and the CVaR for making decisions. In particular, if $\lambda = 0$, the obtained solution will be the result of not taking the CVaR for making decisions into consideration. It is clear that this objective function will minimize the operating time if ζ_2 is set to zero. In this case, this problem turns out to be a supply-side-oriented optimization. However, when considering multi-scenario coupling, only minimizing the operating time will inevitably increase the risk of service unfulfillment. Thus, the CVaR criterion is employed to handle the decision risk caused by uncertainty, and the number of waiting passengers is the most common decision criterion in the passenger flow control problem.

Given the section running time $t_{i,k}^r$ and the station dwell time $t_{i,k}^{dw}$, Eqs. (2) and (3) formulate the arrival and departure times for each train. In particular, $t_{i,1}^a$ is an input parameter that is fixed in advance. Eq. (4) defines the departure headway between two consecutive trains i and $(i + 1)$ arriving at station k . Eq. (5) guarantees the minimum and maximum safe departure headway. By imposing Eq. (6), we can obtain the actual time with respect to the departure headway. Eq. (7) ensures that the 0–1 decision variable $d_{i,k}(t)$ is non-increasing for train i at station k . Eq. (8) is employed to calculate the total operating time for the focus line by summing up the headway, station dwell time, and section running time of each train at all stations. Finally, the domain of the decision variable, $d_{i,k}(t)$, is defined by Eq. (9).

In the second stage, the corresponding robust passenger flow control strategy is optimized, which is formulated as follows:

$$Q(\mathbf{d}, \xi(\omega)) = \min \sum_{i \in I} \sum_{k \in S} P_{i,k}^{waita,out}(\omega) \quad (10)$$

s.t.

$$\sum_{t \in T} P_k^{come,out}(t, \omega) + \sum_{t \in T} P_k^{come,tr}(t, \omega) = \sum_{i \in I} P_{i,k}^{ina}(\omega), \forall k \in S, \omega \in W \quad (11)$$

$$P_{i,k}^{waita,out}(\omega) = \begin{cases} \sum_{t \in T} d_{i,k}(t) \cdot P_k^{come,out}(t, \omega), & \text{if } i = 1 \\ \sum_{t \in T} d_{i,k}(t) \cdot P_k^{come,out}(t, \omega) - \sum_{j=1}^{i-1} P_{j,k}^{ina,out}(\omega), & \text{if } i \in I \setminus \{1\} \end{cases}, \forall k \in S, \omega \in W \quad (12)$$

$$P_{i,k}^{wait,tr}(\omega) = \begin{cases} \sum_{t \in T} d_{i,k}(t) \cdot P_k^{come,tr}(t, \omega), & \text{if } i = 1 \\ \sum_{t \in T} (d_{i,k}(t) - d_{i-1,k}(t)) \cdot P_k^{come,tr}(t, \omega), & \text{if } i \in I \setminus \{1\} \end{cases}, \forall k \in S, \omega \in W \quad (13)$$

$$P_{i,k}^{waita,out}(\omega) \geq 0, \forall i \in I, k \in S, \omega \in W \quad (14)$$

$$P_{i,k}^{wait,tr}(\omega) \geq 0, \forall i \in I, k \in S, \omega \in W \quad (15)$$

$$P_{i,k}^{ina,out}(\omega) \leq P_{i,k}^{waita,out}(\omega), \forall i \in I, k \in S, \omega \in W \quad (16)$$

$$P_{i,k}^{ina}(\omega) = P_{i,k}^{ina,out}(\omega) + P_{i,k}^{wait,tr}(\omega), \forall i \in I, k \in S, \omega \in W \quad (17)$$

$$0 \leq P_{i,k}^{ina,out}(\omega) \leq P_{i,k}^{inc,out}(\omega), \forall i \in I, k \in S, \omega \in W \quad (18)$$

$$P_{i,k}^{inc,out}(\omega) = P_{i,k}^{inc,out}(\omega'), \forall i \in I, k \in S, \omega, \omega' \in W \quad (19)$$

$$P_{i,k}^{inc}(\omega) = P_{i,k}^{inc,out}(\omega) + P_{i,k}^{wait,tr}(\omega), \forall i \in I, k \in S, \omega \in W \quad (20)$$

$$P_{i,k}^{al}(\omega) = \sum_{v=1}^{k-1} (P_{i,v}^{inc,out}(\omega) \cdot \delta_{v,k}(\omega) + P_{i,v}^{wait,tr}(\omega) \cdot \beta_{v,k}(\omega)), \quad (21)$$

$$\forall i \in I, k \in S, \omega \in W$$

$$P_{i,k}^{on}(\omega) = \begin{cases} P_{i,k}^{inc}(\omega), & \text{if } k = 1 \\ 0, & \text{if } k = |S|, \\ P_{i,k-1}^{on}(\omega) + P_{i,k}^{inc}(\omega) - P_{i,k}^{al}(\omega), & \text{if } k \in S \setminus \{1, |S|\} \end{cases}, \forall i \in I, \omega \in W \quad (22)$$

$$P_{i,k}^{on}(\omega) \leq C, \forall i \in I, k \in S, \omega \in W \quad (23)$$

In this stage, we need to formulate the relationship between the train schedule and the robust passenger flow control strategy, by establishing their coupling constraints. The objective function Eq. (10) aims to minimize the total number of waiting passengers for different scenarios. As we consider the direction from station 1 to station $|S|$ as the operation direction, and thus station $|S|$ is the terminal station with no passenger demand, the number of waiting people at this station is zero. By imposing Eq. (11), it is ensured that all passenger demand should be satisfied over the considered time horizon T . The number of passengers waiting for train i in the station hall can be calculated by Eq. (12). Since all outside arrival passengers need to queue in the station hall and wait for boarding permission, the number of waiting passengers is the difference between the cumulative number of outside arrival passengers and the cumulative number of passengers being served. Eq. (13) tracks the number of transfer passengers boarding train i at station k . As the transfer passengers can directly reach the platform, these passengers must have a high priority to get on trains. By using Eqs. (12) and (13), passenger demand and train schedules can be tightly connected in a series of unified formulations. Eqs. (14) and (15) require the number of waiting passengers to be non-negative because the number of passengers who are allowed to get on any train i cannot be greater than the number of outside arrival and transfer passengers, respectively.

Next, we analyze the dynamic passenger loading process, which consists of two types of activities: boarding and alighting. Since multiple stochastic scenarios are considered in this study, a robust passenger flow control strategy, $P_{i,k}^{inc}(\omega)$, which is suitable for all involved scenarios, would be generated. However, due to the diversity of passenger demand, an actual control strategy, $P_{i,k}^{ina}(\omega)$, also exists in each respective scenario. Eq. (16) is used to restrict the number of outside arrival passengers actually boarding train i , which should not exceed the number of waiting passengers. Through Eq. (17), the total number of passengers actually boarding train i at station k can be calculated. Eq. (18) is imposed to ensure that the actual number of boarding passengers is less than or equal to that value under the robust passenger flow control strategy. Eq. (19) ensures that the number of outside arrival passengers allowed to board train i under the robust passenger flow control strategy is equal for each scenario, which greatly improves the chance of practical application in the real world. Eq. (20) makes it possible to obtain the total number of passengers that are allowed to board train i at station k under the robust passenger flow control strategy. Specifically, at common stations, the boarding passengers under the robust passenger flow control strategy only include outside arrival passengers, while the transfer passenger demand is also required to be considered at transfer stations. Eq. (21) makes sure that the number of alighting passengers under the robust passenger

flow control strategy from train i at station k is equal to the number of passengers heading to this station, who are in train i in front of this station. Eq. (22) considers the dynamic variation of in-vehicle passengers under the robust passenger flow control strategy when train i leaves station k . More specifically, when train i arrives at its start terminal (station 1) with empty vehicles, the number of in-vehicle passengers will be equal to the number of boarding passengers under the robust passenger flow control strategy. This value is equal to 0 when train i leaves the terminal station because all passengers are required to get off the train. In comparison, at each intermediate station, the number of in-vehicle passengers is closely related to the total number of in-vehicle passengers at the previous station, the number of alighting passengers at the current station, and the number of boarding passengers at the current station. Moreover, the number of in-vehicle passengers cannot exceed the capacity of the trains, as expressed in Eq. (23).

If the discrete probability distribution of different scenarios can be specified, we can formulate the expected number of waiting passengers as a computationally tractable form, which can be expressed as follows:

$$\mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] = \mathbf{p}^T \cdot \mathbf{Q}(\mathbf{d}, \xi(\omega)) \quad (24)$$

where $\mathbf{p}^T = (p_{\omega_1}, p_{\omega_2}, \dots, p_{\omega_{|W|}})$, $p_{\omega} \geq 0$ is the probability of scenario ω such that $\sum_{\omega \in W} p_{\omega} = 1$, and $\mathbf{Q}(\mathbf{d}, \xi(\omega)) = (Q(\mathbf{d}, \xi(\omega_1)), Q(\mathbf{d}, \xi(\omega_2)), \dots, Q(\mathbf{d}, \xi(\omega_{|W|}))$. Similarly, the third part of the objective function, $\text{CVaR}_{\alpha, \mathbf{p}}[Q(\mathbf{d}, \xi(\omega))]$, can be transformed into the following form:

$$\text{CVaR}_{\alpha, \mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] = \min_{\phi \in \mathbb{R}^+} \left\{ \phi + \frac{1}{1-\alpha} \mathbb{E}_{\mathbf{p}}[\max\{Q(\mathbf{d}, \xi(\omega)) - \phi, 0\}] \right\} \quad (25)$$

in which α denotes decision-makers' risk preferences, and the value of ϕ represents the maximum loss of the function $Q(\mathbf{d}, \xi(\omega))$ at a given value of risk preference α .

For clarity, an auxiliary variable $t(\omega)$ is introduced here, and Eq. (25) can be reformulated as follows:

$$\begin{aligned} \min_{\phi \in \mathbb{R}^+} \quad & \phi + \frac{1}{1-\alpha} \mathbf{p}^T \mathbf{t} \\ \text{s.t.} \quad & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e}\phi \leq \mathbf{t} \\ & \mathbf{t} \geq 0 \end{aligned} \quad (26)$$

where $\mathbf{t} = (t(\omega_1), t(\omega_2), \dots, t(\omega_{|W|}))$, \mathbf{e} is the unit vector.

According to the above analyses, the two-stage SP model can be formally expressed as follows:

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \left[(1-\lambda) \cdot \mathbf{p}^T \cdot \mathbf{Q}(\mathbf{d}, \xi(\omega)) + \lambda \cdot \left(\phi + \frac{1}{1-\alpha} \mathbf{p}^T \mathbf{t} \right) \right] \\ \text{s.t.} \quad & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e}\phi \leq \mathbf{t}, \forall \omega \in W \\ & \mathbf{t} \geq 0 \\ & \text{Eqs. (2)–(23)} \end{aligned} \quad (27)$$

In practice, it is rather difficult to precisely obtain the probability distribution of stochastic scenarios. In such a circumstance, the probability distribution is assumed to be only partially known; in this study, it belongs to an ambiguity set, which will be analyzed in the following sections.

3.2. DRO model with uncertain probability

Based on the notations in the previous sections, we formally propose the following DRO model in Section 3.2.1; next, in Section 3.2.2, we specifically design a discrepancy-based ambiguity set for our developed model. In Section 3.2.3, we further reformulate the DRO model as an MILP model. Finally, we analyze the com-

plexity in terms of the number of decision variables and key constraints in Section 3.2.4.

3.2.1. DRO model

Given the uncertain probability of each stochastic scenario, a DRO model based on the worst-case mean-CVaR criterion is formulated in this section, as shown below:

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \sup_{\mathbf{p} \in P} \left\{ (1-\lambda) \cdot \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \lambda \cdot \text{CVaR}_{\alpha, \mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] \right\} \\ \text{s.t.} \quad & \text{Eqs. (2)–(26)} \end{aligned} \quad (28)$$

in which we use $\sup_{\mathbf{p} \in P} \{\cdot\}$ to denote the worst-case evaluation over all possible distributions, and P represents an ambiguity set characterized by the available partial distributional information.

Below, we discuss the equivalent formulations for the DRO model. First, the objective function in model (Eq. (28)) can be transformed into

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \left(\sup_{\mathbf{p} \in P} \min_{\phi \in \mathbb{R}^+} \left\{ (1-\lambda) \cdot \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \lambda \cdot \left(\phi + \frac{1}{1-\alpha} \mathbb{E}_{\mathbf{p}}[\mathbf{t}] \right) \right\} \right) \end{aligned} \quad (29)$$

According to the strong dual property of $Q(\mathbf{d}, \xi(\omega))$, we can change the order of the operators $\sup_{\mathbf{p} \in P} \min_{\phi \in \mathbb{R}^+} \{\cdot\}$, leading to the following results:

$$\begin{aligned} \sup_{\mathbf{p} \in P} \min_{\phi \in \mathbb{R}^+} \left\{ (1-\lambda) \cdot \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \lambda \cdot \left(\phi + \frac{1}{1-\alpha} \mathbb{E}_{\mathbf{p}}[\mathbf{t}] \right) \right\} \\ = \min_{\phi \in \mathbb{R}^+} \sup_{\mathbf{p} \in P} \left\{ (1-\lambda) \cdot \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \lambda \cdot \left(\phi + \frac{1}{1-\alpha} \mathbb{E}_{\mathbf{p}}[\mathbf{t}] \right) \right\} \\ = \min_{\phi \in \mathbb{R}^+} \lambda \phi + (1-\lambda) \cdot \max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \frac{\lambda}{1-\alpha} \max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[\mathbf{t}] \end{aligned} \quad (30)$$

Then, the proposed two-stage DRO model (Eq. (28)) (i.e., the DRO model) can be equivalently formulated as follows:

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \left\{ \min_{\phi \in \mathbb{R}^+} \lambda \phi + (1-\lambda) \cdot \max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] + \frac{\lambda}{1-\alpha} \max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[\mathbf{t}] \right\} \\ \text{s.t.} \quad & \text{Eqs. (2)–(23)} \end{aligned} \quad (31)$$

Here, the computation processes of $\max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))]$ and $\max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[\mathbf{t}]$ are dependent on the property of the ambiguity sets.

Moreover, the maximum expected number of waiting passengers, that is, $\max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))]$, can be equivalently represented by

$$\max_{\mathbf{p} \in P} \mathbb{E}_{\mathbf{p}}[Q(\mathbf{d}, \xi(\omega))] = \max_{\mathbf{p} \in P} \mathbf{p}^T \cdot \mathbf{Q}(\mathbf{d}, \xi(\omega)) \quad (32)$$

Likewise, the CVaR in the objective function of model (Eq. (28)), that is, $\max_{\mathbf{p} \in P} \text{CVaR}_{\alpha, \mathbf{p}}[Q(\mathbf{d}, \xi(\omega))]$, is equivalent to

$$\begin{aligned} \min_{\phi \in \mathbb{R}^+} \quad & \phi + \frac{1}{1-\alpha} \max_{\mathbf{p} \in P} \mathbf{p}^T \mathbf{t} \\ \text{s.t.} \quad & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e}\phi \leq \mathbf{t} \\ & \mathbf{t} \geq 0 \end{aligned} \quad (33)$$

Note that the DRO model (Eq. (28)) is a semidefinite programming, which is computationally intractable for general ambiguity sets. Hence, to solve it directly using a commercial solver (e.g., CPLEX) within an acceptable computation time, we can transform the DRO model (Eq. (28)) into its computationally tractable form. In the next section, a special ambiguity set is designed to characterize the distribution uncertainty, and the corresponding computationally tractable forms of the DRO model (Eq. (28)) are deduced.

3.2.2. Ambiguity set

In this study, uncertainty is modeled via a series of stochastic scenarios, and the corresponding probability distributions are characterized by a discrepancy-based ambiguity set [49], that is, P , which is defined as follows:

$$P = \{ \mathbf{p} = \mathbf{p}_0 + \boldsymbol{\zeta} \mathbf{e}^T \boldsymbol{\zeta} = 0, \|\boldsymbol{\zeta}\|_\infty \leq \Psi \} \quad (34)$$

To be specific, \mathbf{p}_0 is a vector representing the nominal distribution of the discrete probability, $\boldsymbol{\zeta}$ denotes the fluctuation vector, and Ψ is a real value between 0 and 1. Note that the condition $\mathbf{e}^T \boldsymbol{\zeta} = 0$ is formulated to ensure that the vector \mathbf{p} meets the requirements $\mathbf{e}^T(\mathbf{p}_0 + \boldsymbol{\zeta}) = 1$. That is, the discrepancy-based ambiguity set describes the deviation from a reference (often empirical) distribution. The reasons why we use this ambiguity set are threefold: ① This set is simple and practical. The former part, \mathbf{p}_0 , can be obtained from the data of different dates in practical operations, which is easy to realize in engineering. ② This set is convenient for transformation into a computationally tractable form by applying the related theoretical analysis. ③ This set often possesses asymptotic properties [50].

3.2.3. Deterministic equivalent formulation

Under the designed ambiguity set, the deterministic equivalent formulation can be obtained by invoking the theoretical results of Ref. [49], which are shown as follows.

Theorem 1. Under the discrepancy-based ambiguity set P , the equivalent form of the DRO model (Eq. (28)) can be expressed as follows:

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \left\{ \lambda \phi + (1 - \lambda) \cdot \left[\mathbf{p}_0^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) + \boldsymbol{\Psi}^T \boldsymbol{\eta} + \boldsymbol{\Psi}^T \boldsymbol{\gamma} \right] + \frac{\lambda}{1 - \alpha} \left(\mathbf{p}_0^T \mathbf{t} + \boldsymbol{\Psi}^T \boldsymbol{\eta}' + \boldsymbol{\Psi}^T \boldsymbol{\gamma}' \right) \right\} \\ \text{s.t.} \quad & \mathbf{e} \mu - \boldsymbol{\eta} + \boldsymbol{\gamma} = \mathbf{Q}(\mathbf{d}, \xi(\omega)) \\ & \mathbf{e} \mu' - \boldsymbol{\eta}' + \boldsymbol{\gamma}' = \mathbf{t} \\ & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e} \phi \leq \mathbf{t} \\ & \mathbf{t} \geq 0 \\ & \boldsymbol{\eta} \geq 0, \boldsymbol{\gamma} \geq 0, \boldsymbol{\eta}' \geq 0, \boldsymbol{\gamma}' \geq 0 \\ & \text{Eqs. (2) – (23)} \end{aligned} \quad (35)$$

where $\mu, \boldsymbol{\eta}, \boldsymbol{\gamma}, \mu', \boldsymbol{\eta}', \boldsymbol{\gamma}' \in R \times R^{|\mathcal{W}|} \times R^{|\mathcal{W}|} \times R \times R^{|\mathcal{W}|} \times R^{|\mathcal{W}|}$ are auxiliary variables or vectors, and $\boldsymbol{\Psi} = \mathbf{e} \Psi$.

Proof. Under the discrepancy-based ambiguity set P , Eq. (32) can be transformed into the following equivalent form:

$$\begin{aligned} \max_{\mathbf{p} \in P} \quad & \mathbf{p}^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) = \mathbf{p}_0^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) + \\ \max_{\boldsymbol{\zeta}} \quad & \{ \boldsymbol{\zeta}^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) | \boldsymbol{\zeta}^T \mathbf{e} = 0, \|\boldsymbol{\zeta}\|_\infty \leq \Psi \} \end{aligned} \quad (36)$$

where $\|\boldsymbol{\zeta}\|_\infty = \max_{\omega \in \mathcal{W}} |\zeta_\omega|$.

With the strong duality theory of linear programming, the dual form of Eq. (32) can be formulated as the following MILP model:

$$\begin{aligned} \min_{\mu, \boldsymbol{\eta}, \boldsymbol{\gamma}} \quad & \mathbf{p}_0^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) + \boldsymbol{\Psi}^T \boldsymbol{\eta} + \boldsymbol{\Psi}^T \boldsymbol{\gamma} \\ \text{s.t.} \quad & \mathbf{e} \mu - \boldsymbol{\eta} + \boldsymbol{\gamma} = \mathbf{Q}(\mathbf{d}, \xi(\omega)) \\ & \boldsymbol{\eta} \geq 0, \boldsymbol{\gamma} \geq 0 \end{aligned} \quad (37)$$

Likewise, the dual programming of Eq. (33) is developed as the following linear programming model:

$$\begin{aligned} \min_{\mu', \boldsymbol{\eta}', \boldsymbol{\gamma}'} \quad & \phi + \frac{1}{1 - \alpha} \left(\mathbf{p}_0^T \mathbf{t} + \boldsymbol{\Psi}^T \boldsymbol{\eta}' + \boldsymbol{\Psi}^T \boldsymbol{\gamma}' \right) \\ \text{s.t.} \quad & \mathbf{e} \mu' - \boldsymbol{\eta}' + \boldsymbol{\gamma}' = \mathbf{t} \\ & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e} \phi \leq \mathbf{t} \\ & \mathbf{t} \geq 0 \\ & \boldsymbol{\eta}' \geq 0, \boldsymbol{\gamma}' \geq 0 \end{aligned} \quad (38)$$

Combining models (Eqs. (37) and (38)), under the discrepancy-based ambiguity set $P = \{ \mathbf{p} = \mathbf{p}_0 + \boldsymbol{\zeta} \mathbf{e}^T \boldsymbol{\zeta} = 0, \|\boldsymbol{\zeta}\|_\infty \leq \Psi \}$, the DRO model (Eq. (28)) is equivalent to the following MILP model:

$$\begin{aligned} \min \quad & \zeta_1 \cdot T_0 + \zeta_2 \cdot \left\{ (1 - \lambda) \cdot \left(\mathbf{p}_0^T \mathbf{Q}(\mathbf{d}, \xi(\omega)) + \boldsymbol{\Psi}^T \boldsymbol{\eta} + \boldsymbol{\Psi}^T \boldsymbol{\gamma} \right) + \lambda \cdot \left[\phi + \frac{1}{1 - \alpha} \left(\mathbf{p}_0^T \mathbf{t} + \boldsymbol{\Psi}^T \boldsymbol{\eta}' + \boldsymbol{\Psi}^T \boldsymbol{\gamma}' \right) \right] \right\} \\ \text{s.t.} \quad & \mathbf{e} \mu - \boldsymbol{\eta} + \boldsymbol{\gamma} = \mathbf{Q}(\mathbf{d}, \xi(\omega)) \\ & \mathbf{e} \mu' - \boldsymbol{\eta}' + \boldsymbol{\gamma}' = \mathbf{t} \\ & \mathbf{Q}(\mathbf{d}, \xi(\omega)) - \mathbf{e} \phi \leq \mathbf{t} \\ & \boldsymbol{\eta}' \geq 0, \boldsymbol{\gamma}' \geq 0 \\ & \text{Eqs. (2) – (23)} \end{aligned} \quad (39)$$

By regrouping the terms, model (Eq. (39)) can be represented as model (Eq. (35)). The proof of Theorem 1 is thus complete.

Remark 2: In the case of $\Psi = 0$, the DRO model (Eq. (35)) is equivalent to the two-stage SP model (Eq. (27)).

3.2.4. Complexity analyses

The decision variables in the proposed DRO model (Eq. (35)) can be categorized into three categories: The first category comprises the variables associated with the train schedule—that is, $t_{i,k}^a, t_{i,k}^d, h_{i,k}, d_{i,k}(t)$ —and the second is related to the robust passenger flow control strategy—that is, $P_{i,k}^{\text{ina,out}}(\omega)$ and $P_{i,k}^{\text{inc,out}}(\omega)$. The last category comprises the auxiliary variables, such as $\mu, \boldsymbol{\eta}(\omega), \boldsymbol{\gamma}(\omega), \mu', \boldsymbol{\eta}'(\omega)$, and $\boldsymbol{\gamma}'(\omega)$. Obviously, the scales of the abovementioned decision variables are fully dependent on the number of stations, trains, timestamps, and scenarios. We are particularly interested in analyzing the complexity of the proposed model (Eq. (35)) in detail, as shown in Table 1. To illustrate this problem more clearly, an instance with 40 in-service trains, 20 stations, 180 timestamps, and three scenarios is provided in order to identify the total number of decision variables and constraints. In this case, there are over 150 000 decision variables and 319 980 constraints, which indicates that the proposed model (Eq. (35)) is a fairly complex problem.

4. Solution approaches

Based on the analyses above, if the scale of the problem is small, the proposed model (Eq. (35)) can be solved by means of MILP solvers (e.g., CPLEX) in a short time, due to the linearity of the model. However, it is difficult to solve real-world large-scale instances. Thus, we need to design heuristic algorithms to search for approximate optimal solutions within an acceptable computation time, in order to meet the needs of real-world applications. Therefore, an effective hybrid algorithm combining an LS with an MILP solver

Table 1
Numbers of decision variables and constraints in the DRO model.

Variables or constraints	Total number
Decision variables $t_{i,k}^a, t_{i,k}^d$	$2 \cdot (I \cdot S)$
Decision variables $d_{i,k}(t)$	$ I \cdot S \cdot T $
Decision variables $P_{i,k}^{\text{ina,out}}(\omega), P_{i,k}^{\text{inc,out}}(\omega)$	$2 \cdot (I \cdot S \cdot W)$
Decision variables μ, μ'	2
Decision variables $\boldsymbol{\eta}(\omega), \boldsymbol{\gamma}(\omega), \boldsymbol{\eta}'(\omega), \boldsymbol{\gamma}'(\omega)$	$4 \cdot W $
Train timetable Eq. (2)	$ I \cdot (S - 1)$
Train timetable Eq. (3)	$ I \cdot S $
Train timetable Eqs. (4)–(6)	$3 \cdot (I - 1) \cdot S $
Train timetable Eq. (7)	$ I \cdot S \cdot (T - 1)$
Train timetable Eq. (8)	1
Train timetable Eq. (9)	$ I \cdot S \cdot T $
Robust passenger flow control Eq. (11)	$ S \cdot W $
Robust passenger flow control Eqs. (12)–(23)	$12 \cdot I \cdot S \cdot W $
Remaining constraints in the model Eq. (35)	$8 \cdot W $

will be developed in this section to find high-quality solutions. The framework and the detailed techniques of our hybrid algorithm are introduced in the following discussion.

4.1. Algorithm framework

An algorithm framework is developed to solve the proposed DRO model, following the procedure shown in Fig. 6. In brief, the original problem (Eq. (35)) is decomposed into two subproblems. The first subproblem (denoted by MP) determines the decision variables related to the train schedule for the focus line, while the second subproblem (denoted by SCP) optimizes the control variables for uncertain passenger demand over all the involved scenarios. An LS is designed to search for feasible train schedules. Then, each feasible schedule is taken as the input for subproblem SCP to produce the corresponding robust passenger flow control strategy, which is used to evaluate the generated schedule in subproblem MP. We observe that subproblem SCP is an MILP model that can theoretically be solved to optimality. However, similar to the problem in Ref. [11], the scenarios in subproblem SCP are tightly coupled across different scenarios, making it difficult for the problem to be solved directly by MILP solvers. Therefore, we decompose it here into a series of single-scenario-based subproblems SSP in order to speed up the computation, in which the train schedule is the input for subproblem SCP and SSP. As each subproblem SSP is an MILP problem and does not involve the coupling relationship among different scenarios, it can be efficiently solved using a suitable MILP solver.

Next, the minimum value of the optimal solutions to subproblems SSP, denoted by $P_{i,k}^{inc.out,SSP}$, is treated as the input to subproblem SCP. That is, we first let

$$P_{i,k}^{inc.out,SSP} \leftarrow \min\{P_{i,k}^{ina.out}(\omega_1), P_{i,k}^{ina.out}(\omega_2), \dots, P_{i,k}^{ina.out}(\omega_{|W|})\} \quad (40)$$

and then add the following equation to subproblem SCP.

$$P_{i,k}^{inc.out}(\omega) \geq P_{i,k}^{inc.out,SSP}, \quad \forall i \in I, k \in S, \omega \in W \quad (41)$$

The solution space of the decision variable $P_{i,k}^{inc.out}(\omega)$, which is the key variable that couples the various scenarios together, is greatly reduced by Eq. (41); thus, Eq. (41) acts as a so-called “valid inequality” to the problem of interest. The proposed valid inequality is then taken as the input for subproblem SCP.

4.2. Local search algorithm

The LS is an iterative searching procedure that starts from an initial solution and attempts to find a better solution. It should be noted that, due to the existence of time-related binary variable $d_{i,k}(t)$, the number of decision variables in subproblem MP is huge. For simplicity, this paper assumes that the running time over each section and the dwell time at each station are identical for different in-service trains. To guarantee the headway in operations, it is therefore sufficient to ensure that the departure headway can be kept at the start terminal station. In addition, since the arrival time for the first train is known in advance, the decision variable mentioned above can be converted into the headway between

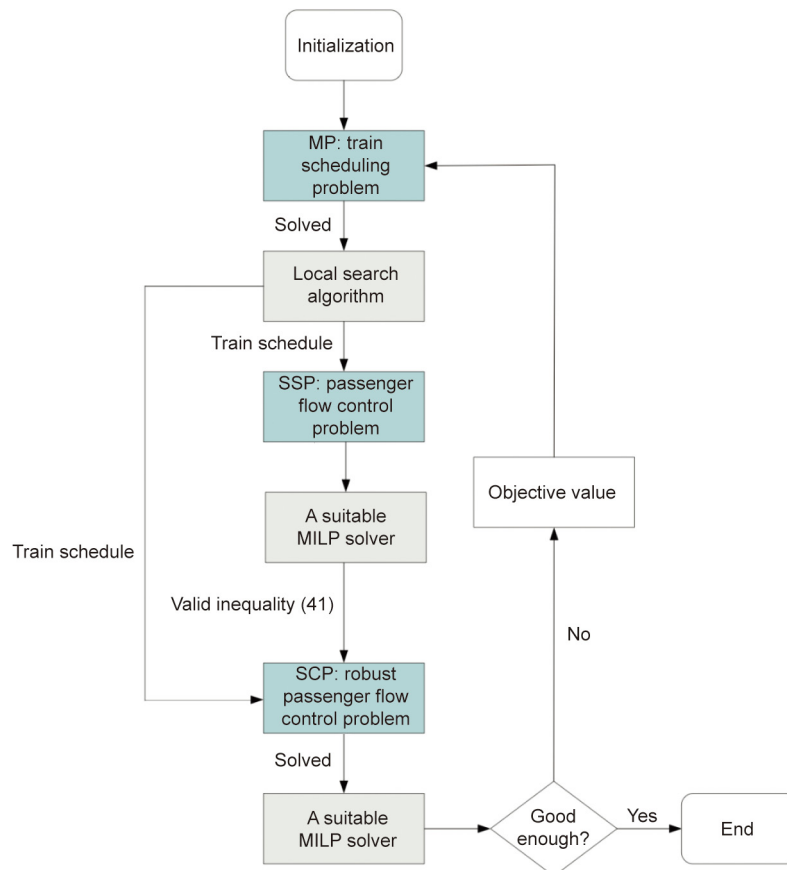


Fig. 6. The procedure for solving the DRO model.

two consecutive trains arriving at the start terminal station. Thus, this paper takes a $(|I| - 1)$ -dimensional headway vector to represent a solution for the train scheduling problem MP. To be specific, we adopt a problem-based solution representation, which consists of only one type of decision variable—that is, the headway vector $H = \{h_{1,1}, h_{2,1}, \dots, h_{i,1}, \dots, h_{|I|-1,1}\}$ —in which $h_{i,1}$ denotes the headway between trains i and $(i + 1)$ at the start terminal station. In this way, the number of core decision variables decreases from $|I| \cdot |S| \cdot |T|$ to $(|I| - 1)$, which significantly reduces the complexity of subproblem MP.

To solve subproblem MP, an initial solution should first be generated. Next, the approximate optimal solution is searched in a greedy way. During the searching process, only the best solution in the neighborhood space can be accepted as a seed solution. Finally, the algorithm is terminated when the current best solution cannot be updated within a given maximum number of iterations, or when the number of search iterations reaches a predetermined threshold. Then, the best solution that has been encountered is outputted as a near-optimal solution. For the sake of description, we introduce some important notations in the algorithm: n denotes the index of iteration, n_{\max} denotes the threshold of the number of iterations, M denotes the maximum number of iterations for which the encountered best solution is not updated, X^n denotes the best solution at iteration n , X^0 denotes the initial solution, m denotes the number of candidate solutions at each iteration, $N(X^{n-1})$ denotes the neighborhood at iteration n , X^* denotes the current best solution, f^* denotes the current best objective value, and f^n denotes the best objective value at iteration n .

4.2.1. Initial solution generation

In this study, the initial solution is generated randomly as the seed solution for the first iteration. That is, we first generate a total of $(|I| - 1)$ headway variables randomly within a given range $[\underline{h}, \bar{h}]$, and thus an initial solution X^0 can be produced, whose feasibility can be tested by Algorithm 1, shown below. If this solution is feasible, it will be set as the initial solution; otherwise, we need to regenerate the headway vector until a feasible solution is found. In the searching process, the initial solution will be treated as the input data for subproblem SCP, and the corresponding objective value is used as the evaluation of this train schedule.

4.2.2. Neighborhood generation

In order to generate the neighborhood for each iteration, a total number m of candidate solutions must be generated. For example,

Algorithm 1. The checking procedure for solution feasibility.

Input: A solution that needs to be checked

Output: The feasibility of the solution

Step 1: According to Eqs. (2)–(5) and the input solution, calculate the arrival and departure times for all in-service trains at each station, then go to Step 2;

Step 2: If $t_{|I|,|S|}^d \in T$, go to Step 3; otherwise, the solution is infeasible, stop;

Step 3: The solution is feasible, stop.

the set of candidate solutions at iteration n can be denoted by $N(X^{n-1}) = \{X^{n,1}, X^{n,2}, \dots, X^{n,c}, \dots, X^{n,m}\}$, where $X^{n,c}$ represents one of the candidate solutions and c denotes the index of candidate solutions. More specifically, we can generate the candidate solutions by adding or subtracting a tiny integer vector on X^{n-1} , according to the following calculation: $X^{n,c} = X^{n-1} + \Delta^n$ where the value of Δ^n is randomly generated.

It is worth mentioning that we need to test the feasibility of each newly generated neighbor solution by means of Algorithm 1. If the solution is feasible, we then insert it into the neighborhood of the current seed solution; otherwise, a new solution must be regenerated until its feasibility is guaranteed.

4.2.3. Algorithm procedure and termination criteria

The procedure of the LS + MILP solver algorithm is reported below as Algorithm 2, which is designed based on the aforementioned specific techniques. Two criteria are employed to terminate the searching process: ① If the current iteration n is greater than a pre-given threshold n_{\max} , the search process should be terminated. ② If the objective value of the currently encountered best solution is not updated within M iterations, the LS is terminated, and the current best solution can be treated as the near-optimal solution. We use y to denote the number of iterations that the current best solution is not updated.

5. Numerical experiments

In this section, a series of numerical experiments based on real-world operation data from the Beijing rail transit system are conducted to demonstrate the effectiveness and performance of our

Algorithm 2. The procedure of the LS + MILP solver algorithm.

Step 1: Initialize the iteration index $n = 1$ and generate a feasible initial solution X^0 . Calculate the objective value f^0 of the initial solution; let $X^{n-1} = X^0, f^* = f^0$ and set $y = 0$, then go to Step 2;

Step 2: Generate neighborhood $N(X^{n-1})$ by the method used in neighborhood generation, then go to Step 3;

Step 3: For each candidate solution $X^{n,c}$, calculate the decision variables associated with train scheduling problem MP, i.e., $t_{i,k}^a$ and $t_{i,k}^d$; input the train schedule to subproblems SSP of the different scenarios, respectively, which can be solved by a suitable MILP solver; then the passenger flow control strategy $P_{i,k}^{\text{ina,out}}(\omega)$ for each scenario $\omega \in W$, can be obtained, so go to Step 4;

Step 4: Let $P_{i,k}^{\text{inc,out,SSP}} \leftarrow \min\{P_{i,k}^{\text{ina,out}}(\omega_1), P_{i,k}^{\text{ina,out}}(\omega_2), \dots, P_{i,k}^{\text{ina,out}}(\omega_{|W|})\}, \forall i \in I, k \in S$; input the train schedule and Eq. (41) to the subproblem SCP; solve it via a suitable MILP solver. Then go to Step 5;

Step 5: Get the current best solution X^n and its corresponding objective value f^n ;

Step 6: Update the best solution. If $f^n < f^*$, let $f^* = f^n, X^* = X^n$, then set $y = 0$; otherwise, set $y = y + 1$, and go to Step 7;

Step 7: Check the termination condition. If $n < n_{\max}$ and $y < M$, then set $n = n + 1$, and go to Step 2; otherwise, output the best solution X^* and the best objective value f^* , then stop.

proposed models and algorithm. More specifically, the stochastic and dynamic passenger demand is obtained by processing historical AFC data, and the running time of each section is deduced from the practical timetable. To solve this large-scale problem within an acceptable computation time, we use the proposed algorithm framework to generate near-optimal solutions for the models. The algorithm is coded by MATLAB 2014a with CPLEX 12.6 on a Windows 10 personal computer with Intel Core i5-10500 central processing unit (CPU) and 16G random-access memory (RAM).

5.1. Data description

As shown in Fig. 7, this set of experiments consider Line 15 of the Beijing rail transit system (i.e., the blue line), which has a total length of 41.4 km and 20 stations. This line connects a residential zone in a suburban town with the downtown working area, and numerous commuters need to ride this line on weekdays, resulting in over-saturation in the morning peak hours in the downtown direction. We only consider the direction from station Fengbo (FB) in the suburbs to station Qinghuadongluxikou (QH) downtown as the experiment environment, for which the station dwell time and the section running time are taken from real operation data, as shown in Table S1 in Appendix A. We select the morning peak hours from 7:00 to 10:00 as the considered time horizon. To characterize the dynamics of the passenger demand, the discretized time granularity is set as $t_{unit} = 60$ s in this series of experiments, and a total of 180 timestamps are involved (i.e., $T = \{1, 2, \dots, 180\}$), so as to balance the real-world operation conditions and the computation efficiency. To depict the uncertainty of passenger demand, three typical stochastic scenarios with unknown probabilities are taken into account, for which the nominal probability distribution is predetermined to be $\mathbf{p}_0^T = (0.20, 0.30, 0.50)$. To determine the ambiguity set, the value of the adjustable parameter Ψ is set as 0.02, 0.06, and 0.10, respectively; the risk-aversion parameter α is chosen from the set $\{0.95, 0.50, 0.05\}$; and the tradeoff coefficient λ is chosen from the set $\{0.10, 0.50, 0.90\}$. For illustrative clarity, the dynamic outside arrival passenger demand in scenario 1 is displayed in Fig. 8. In addition, it should be noted that there are four transfer stations on this focus line: station Wangjing (WJ) connects Lines 15 and 14, station Wangjing West (WJX) connects Lines 15 and 13, station Datunlu East (DTLD) connects Lines 15 and 5, and station Olympic Green (ALPK) connects Lines 15 and 8. These four stations attract many transfer passengers in addition to the outside arrival

demand. In the data preparation stage, the passenger ratios for the outside arrival and transfer passengers at different destinations are given in Tables S2 and S3 in Appendix A, respectively. In the algorithm, the threshold of the number of iterations n_{max} is set as 100. The other relevant parameters are listed in Table S4 in Appendix A, where C is set as the capacity using the 120% full-loading rate of each in-service train.

Using the above parameter settings, a series of numerical experiments are conducted to demonstrate the performance of our proposed approaches. The first set of experiments are carried out to evaluate the validity of the proposed DRO model in Section 5.2. The experiments in Section 5.3 aim to compare the performance of the SP model (Eq. (27)) with that of the DRO model (Eq. (35)). In addition, the experiments in Section 5.4 are used to evaluate the computation results of the classic RO model (Eq. (S2) in Appendix A) and the DRO model (Eq. (35)). Finally, Section 5.5 reports the impact of some key coefficients in the final set of experiments.

5.2. Computation results of the DRO model

To test the proposed DRO model (Eq. (35)), we consider the computation results by using different combinations of parameters α , Ψ , and λ . It should be noted that the cost coefficients ζ_1 and ζ_2 in the objective function are used to make tradeoffs between operations and services. Here, to make the comparison of optimization results more obvious and to consider the difference in the order of magnitude for two objectives, we set ζ_1 as 100 000 and ζ_2 as 1 in this set of experiments, and the detailed sensitivity analyses related to these two coefficients are shown in Section 5.5.

The computation results are displayed in Table 2. For clarity, Fig. 9 presents the optimal objectives obtained by the DRO model with different combinations of parameters. Typically, with the increase of parameter Ψ , the optimal objective of the model (Eq. (35)) gradually worsens. Furthermore, the tradeoff coefficient λ and the optimal objectives have the same monotonicity, and the optimal objective increases when the risk-aversion parameter α increases. These phenomena are consistent with the theoretical results. More specifically, the parameter α represents the level of risk aversion; that is, $\alpha = 1$ represents extreme risk aversion. Therefore, a larger parameter α leads to a more conservative decision plan, which corresponds to a larger objective value. In the ambiguity set, Ψ represents the disturbance range of the uncertain probability. As the parameter Ψ increases, the distribution of

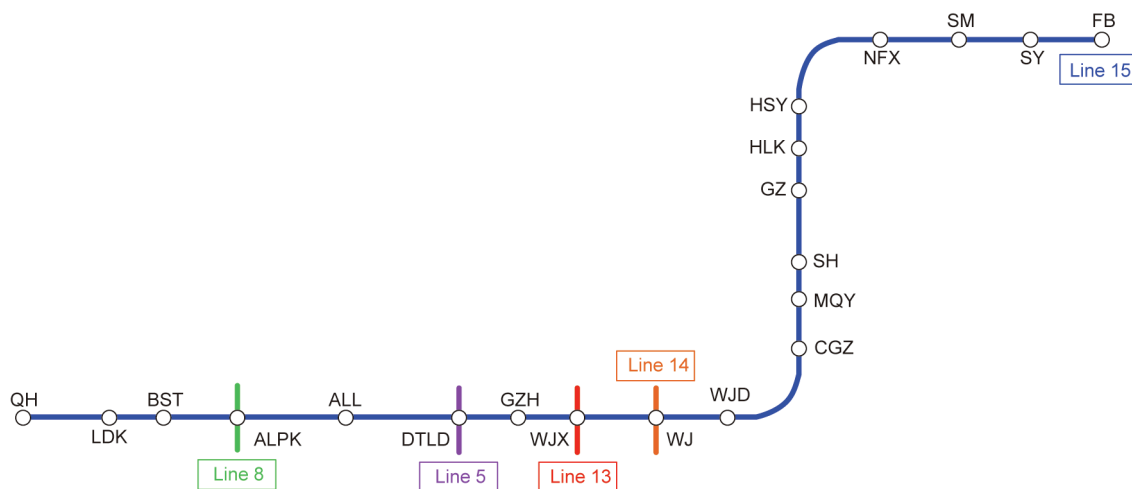


Fig. 7. Illustration of Line 15 of the Beijing rail transit system. QH: Qinghuadongluxikou; LDK: Liudaokou; BST: Beishatan; ALPK: Olympic Green; ALL: Anlilu; DTLD: Datunlu East; GZH: Guanzhuang; WJX: Wangjing West; WJ: Wangjing; WJD: Wangjing East; CGZ: Cuigezhuang; MQY: Maquanying; SH: Sunhe; GZ: Guozhan; HLK: Hualikan; HSY: Houshayu; NFX: Nanfaxin; SM: Shimen; SY: Shunyi; FB: Fengbo.

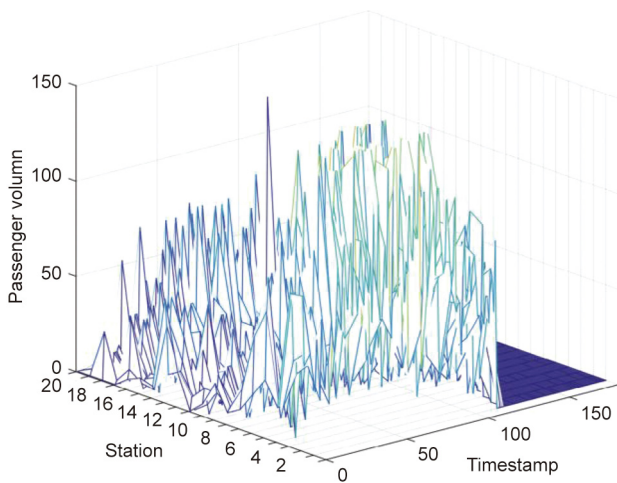


Fig. 8. Illustration of the outside arrival passenger demand of Line 15.

probability information to the decision-maker becomes increasingly coarse, resulting in much worse outcomes for the decision.

For clarity, Fig. 10 displays an optimized train schedule and the number of in-vehicle passengers in each train over each section, where the red lines represent the full-load sections. Clearly, most departure headway during the 7:00–7:42 time period is relatively small, and the full-load situations are all associated with the trains that depart from the start terminal station during this time period, which demonstrates that the train schedule and passenger demand are closely coupled. Next, we are particularly interested in investigating how the objective value converges to the optimal value in the searching process of our proposed algorithm. Since the convergence tendency of the objective values is similar for different instances, we only present the results of the instance with $\alpha = 0.95$, $\lambda = 0.10$, and $\Psi = 0.02$ for illustrative convenience here, as

shown in Fig. 11. It follows from this figure that the best objective value can be iteratively reduced by the proposed LS + MILP solver algorithm, where the solution quality is gradually improved with a fast convergent speed. More specifically, the best objective value decreases drastically in the first several iterations, but is not updated any longer in the last few iterations. Thus, we conclude that the LS + MILP solver algorithm exhibits a good performance in solving the developed DRO model (Eq. (35)) in terms of solution quality, while requiring a reasonable computation time. Furthermore, to clarify the impact of the tradeoff coefficient, λ , in the following context, we will discuss the variations of the expected number of waiting passengers and CVaR when parameter λ changes. We summarize this section with the following observation.

Observation 1. (The necessity of considering different combinations of α , λ , and Ψ .) It follows from the computation results that the combination of parameters α , λ , and Ψ has significant effects on the DRO model. If the decision-makers prefer a lower risk decision, it is better to set a bigger α and λ , and to know the probability distribution as precisely as possible (in which case, Ψ should be small). That is, the level of risk aversion and the disturbance range of uncertain probability are significant factors in the process of optimizing the train schedules and passenger flow control strategies.

Next, we further consider the variation of the expected number of waiting passengers and CVaR with respect to different values of parameter λ . For comparison, we set parameter α as 0.50 and set Ψ as 0.02 and 0.06, respectively. Fig. 12 depicts the variation trend of the expected number of waiting passengers and CVaR with different values of λ . Clearly, for two instances in this figure, the CVaR yields a downward linear increasing tendency in the optimal objective if we increase the coefficient λ , while the expected value shows the opposite trend when α and Ψ are fixed. In addition, when λ is set as 0.5, the risk preference of the decision-makers is neutral. Thus, the expected value and CVaR are very close in this

Table 2
Computation results of the DRO model.

Risk-aversion parameter α	Tradeoff coefficient λ	Adjustable parameter Ψ					
		0.02		0.06		0.10	
		Objective value	CPU time (s)	Objective value	CPU time (s)	Objective value	CPU time (s)
0.95	0.1	17767833.3464	1336.645	17769921.0128	1181.031	17771819.6438	1657.806
	0.5	17771866.8497	1482.181	17772870.7818	1234.462	17773845.5056	1615.507
	0.9	17775309.9180	1352.341	17775484.8424	1251.802	17775659.4888	1646.706
0.05	0.1	17766970.0804	1471.173	17769336.2210	1177.698	17771531.4178	1472.486
	0.5	17767989.3749	1580.820	17770350.0112	1574.838	17772497.5873	1338.949
	0.9	17768991.0045	1476.097	17771316.9863	1169.674	17773422.7690	1461.012

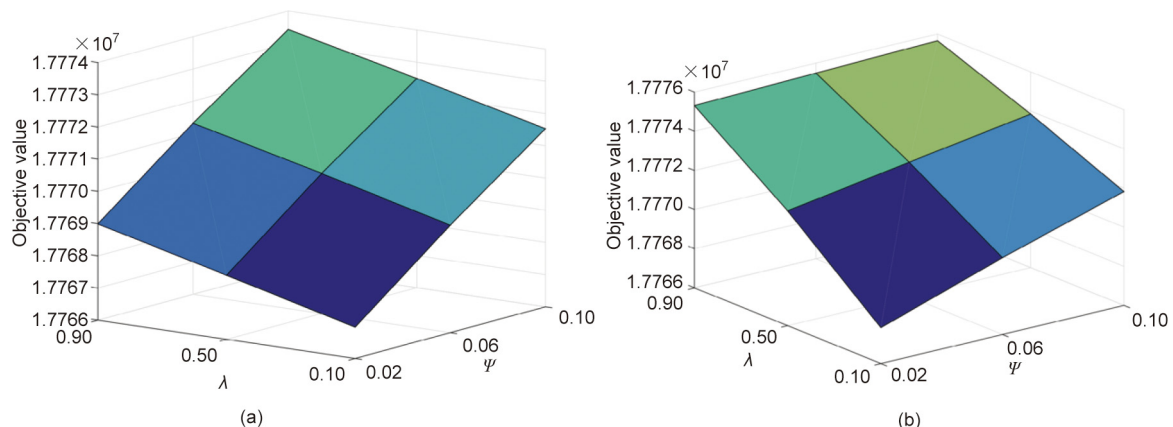


Fig. 9. Computation results of the DRO model. (a) $\alpha = 0.95$; (b) $\alpha = 0.05$.

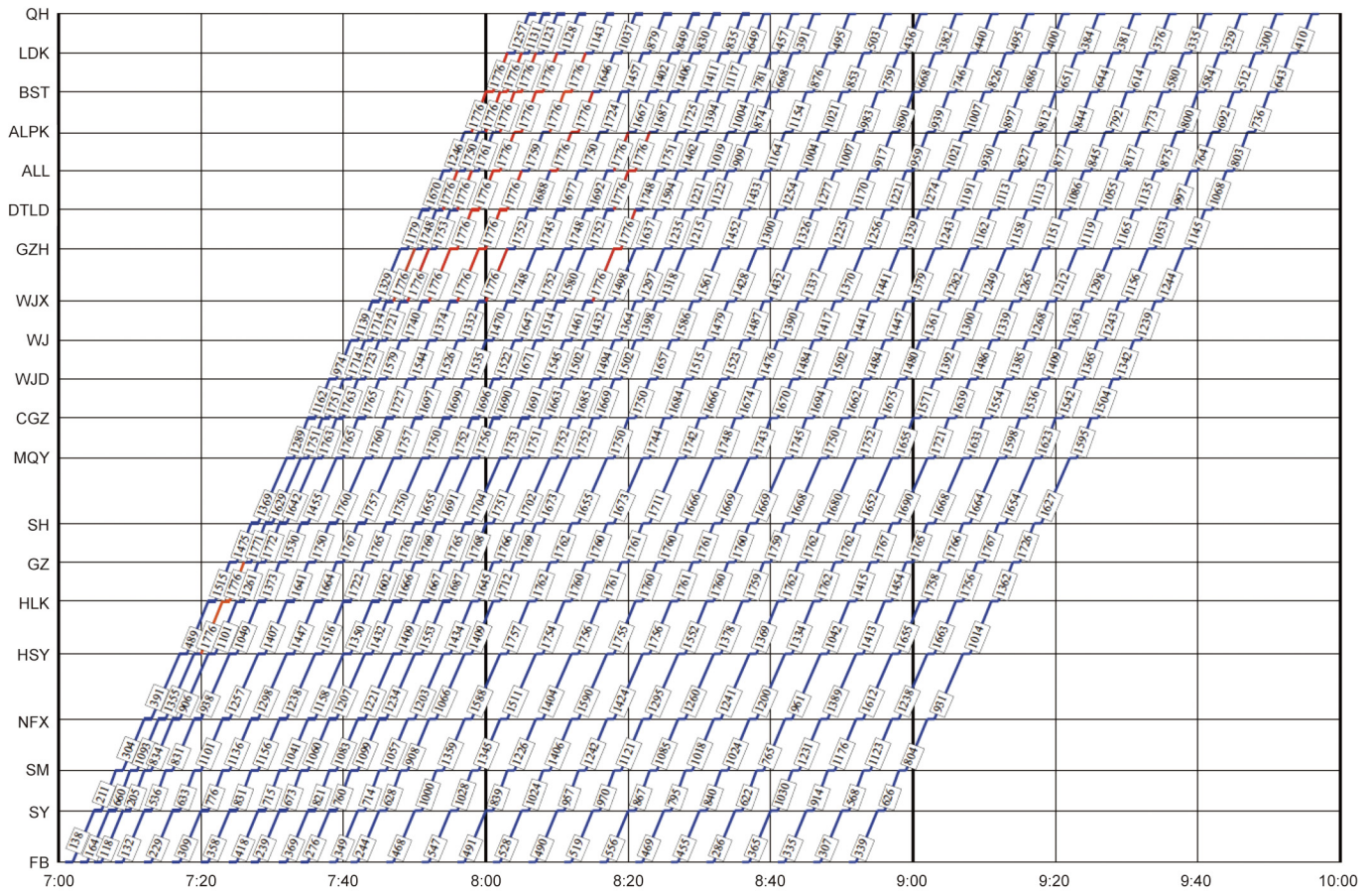


Fig. 10. Train schedule and in-vehicle passengers for the instance when $\alpha = 0.95$, $\lambda = 0.10$, and $\Psi = 0.02$.

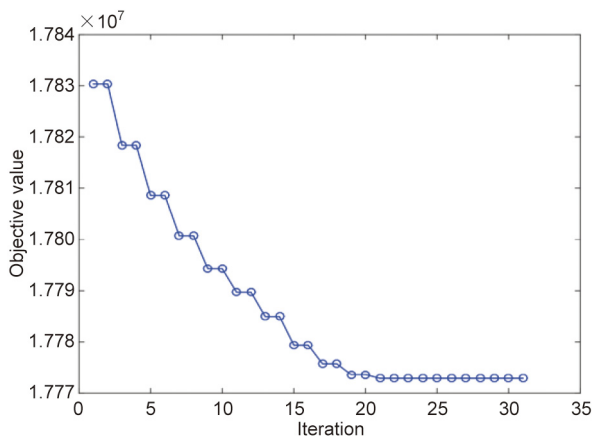


Fig. 11. Convergence tendency of the best objective in the LS + MILP solver algorithm.

case, as shown in Fig. 12. These findings are summarized as the following observation.

Observation 2. (The necessity of considering λ depends on the risk attitude of decision-makers.) The numerical experiments in this section indicate the importance of the tradeoff coefficient λ in the proposed model, as it provides a support for generating a balanced solution that is customized to the actual needs of the decision-makers. The tradeoff coefficient λ provides decision-makers with the opportunity to obtain a personalized and reliable plan. Decision-makers can choose the appropriate values of λ based on their subjective preferences and risk attitudes.

5.3. Comparison between the SP model and the DRO model

In this section, numerical experiments are conducted by solving the two-stage SP model (Eq. (27)). Since the probability distribution is known exactly in this model, the disturbance range of the uncertain probabilities—that is, Ψ —is equal to 0. Therefore, we display the computation results with different combinations of α and λ in Fig. 13. As shown, the optimal objective shows an increasing tendency with an increase in the risk-aversion parameter α if the tradeoff coefficient λ is fixed. In addition, when the level of risk aversion is fixed, the optimal objective follows the same trend with respect to the tradeoff coefficient λ .

It should be noted that, since parameter Ψ is set as 0 in the SP model (Eq. (27)), the objective values should theoretically be less than those of the DRO model (Eq. (35)), with the same parameter settings. Below, we further analyze their differences between the SP model and the DRO model, by conducting a series of comparative experiments. For clarity, the following formula is used to quantify the differences (denoted as Price) in the optimal objective values between the SP model (termed SP^*) and the DRO model (denoted as DRO^*):

$$\text{Price} = \frac{(DRO^* - SP^*)}{SP^*} \cdot 100\%$$

Table 3 provides a detailed performance comparison for various combinations of α , λ , and Ψ .

From the computation results in Table 3, it can be observed that the DRO model always yields a slightly larger objective value than the SP model, since the values of Price are all positive. This demon-

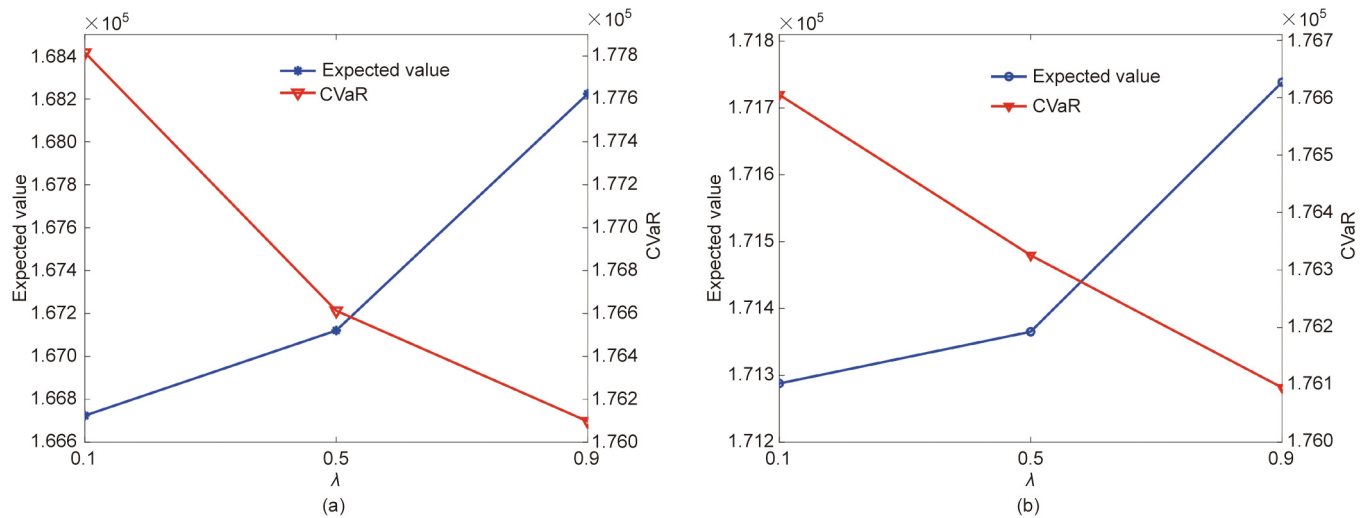


Fig. 12. Expected number of waiting passengers and CVaR with respect to λ . (a) $\alpha = 0.50, \Psi = 0.02$; (b) $\alpha = 0.50, \Psi = 0.06$.

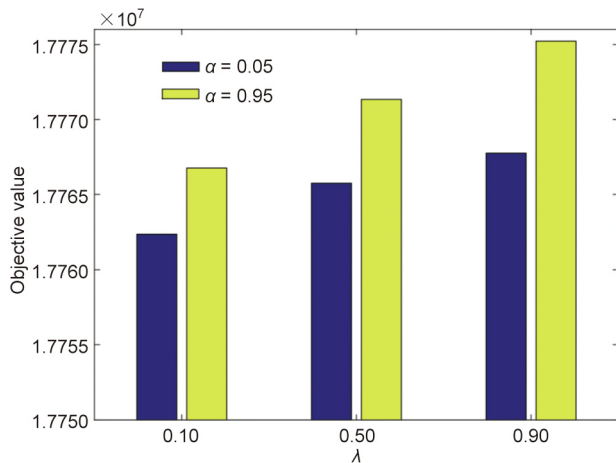


Fig. 13. Computation results of the SP model.

strates that, in an urban rail transit system, the more precise the information of probability distribution we know (i.e., the smaller Ψ is), the better the passenger flow control strategy and train schedule are. In this set of experiments, it can be seen that the maximum Price calculated is not greater than 0.051640%, which corresponds to a very small extra cost in the process of handling distribution uncertainty. Hence, in terms of solution stability, our proposed DRO model is superior to the SP model.

Observation 3. (The superiority of the DRO model in comparison with the SP model.) To summarize, we conclude that

the DRO model (Eq. (35)) utilizes limited distribution information to hedge against the uncertainty more robustly than the SP model (Eq. (27)), while requiring a relatively small Price.

5.4. Comparison between the classic RO model and the DRO model

In this section, we compare the performance of the classic RO model in Eq. (S2) and the DRO model (Eq. (35)). Since the risk measure, CVaR, is not involved in the classic robust method, we set $\lambda = 0$ in the following experiments. For concise descriptions, the classic RO model (Eq. (S2)) is referred to as the RO model. Fig. 14 provides a comparison of the results when adopting different Ψ and α in the models. As shown in Fig. 14, regardless of what Ψ is, the approximate optimal objective of the DRO model is smaller than that of the RO model. The reason for the poor optimization result of the RO model might be that the RO model is designed to optimize the worst-case scenario rather than the expected value, which leads to over-conservative solutions. It should be noted that Ψ represents the disturbance range of the uncertain probability. We thus conclude that the uncertainty can be handled well by our proposed DRO model, regardless of whether it is strong or weak. In addition, we set different values of α to carry out experiments, and the results show that the objective values remain the same when parameter α varies. This finding is consistent with the following analyses: Since α is set as zero in these experiments, the coefficient $\lambda/(1 - \alpha)$ in the objective function is 0 for any value of α , so the corresponding part vanishes.

Observation 4. (The superiority of the DRO model in comparison with the RO model.) From the computational results, we can derive the following observations: ① Our proposed DRO model outperforms the RO model in terms of solution quality

Table 3
The obtained Price in the numerical experiments.

Risk-aversion parameter α	Tradeoff coefficient λ	Adjustable parameter Ψ		
		0.02	0.06	0.10
0.95	0.1	0.006055%	0.017805%	0.028491%
	0.5	0.002938%	0.008588%	0.014072%
	0.9	0.000492%	0.001476%	0.002459%
0.05	0.1	0.025960%	0.039281%	0.051640%
	0.5	0.012523%	0.025810%	0.037899%
	0.9	0.006915%	0.020006%	0.031858%

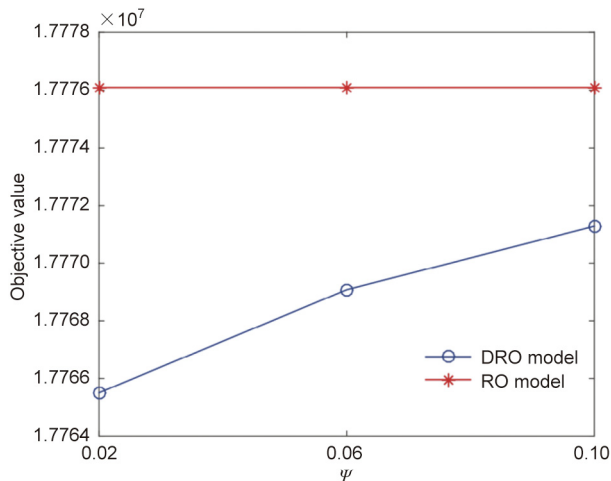


Fig. 14. Comparison of results from the RO model and the DRO model.

and stability. ② When a decision-maker is extremely risk averse, he/she can employ the classic RO method to optimize the worst-case scenario, so as to obtain a conservative solution; otherwise, we recommend the DRO approach with an appropriate combination of parameters, which can be selected flexibly according to the decision-makers' preference and actual requirements.

5.5. Sensitivity analyses with respect to ζ_1 and ζ_2

In this section, we perform sensitivity analyses with respect to the variation of two parameters, ζ_1 and ζ_2 . These two parameters are used in the objective function of the DRO model (Eq. (35)) to balance the operating time (denoted as Obj. 1) related to the interests of the operators and the target relating to passengers (denoted as Obj. 2). To be specific, if the value of ζ_2/ζ_1 goes to ∞ , then Obj. 2 takes precedence over Obj. 1, which implies that more weight is given to passenger interests; otherwise, the proposed model mainly focuses on the optimization of the train schedule for operators. In particular, we aim to discuss how different values of ζ_2/ζ_1 , leading to different train schedules, can affect the dynamic passenger control process. Table 4 reports the results obtained by varying ζ_1 , ζ_2 , and ζ_2/ζ_1 . We list the objective values in the fifth column, and list the values of Obj. 1 and Obj. 2 in the sixth and seventh columns, respectively. As shown, when the value of ζ_2/ζ_1 is between 0.01 and ∞ , the value of Obj. 2 remains constant until ζ_2/ζ_1 drops to 0.00001. The experiment with $\zeta_1 = 1$ and $\zeta_2 = 0$ does not consider the interests of the passengers, but only optimizes the train schedules. We note that the value of Obj. 2 increases drastically in this case. This computation result indicates the importance of collaboratively optimizing the robust passenger flow control strategies and the train schedule in order to ease congestion. In addition, in order to analyze the additional costs caused by the ambiguous distribu-

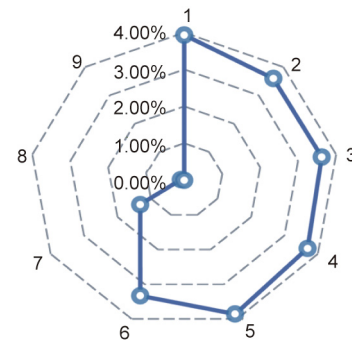


Fig. 15. The prices of distributional robustness under different values of ζ_2/ζ_1 .

tion, we present the Price of each instance in Fig. 15. Clearly, although the value of Price varies considerably with respect to different ζ_1 and ζ_2 , it is still acceptable, since it does not exceed 4.00% at most. Moreover, for instances $\zeta_1 = 1$ and $\zeta_2 = 0$, the Price turns out to be 0, since the objective function does not take Obj. 2 into account.

Observation 5. (Sensitivity to parameters ζ_1 and ζ_2 .) Considering both Obj. 1 and Obj. 2, especially regarding the results of the last experiment, it can be seen that only optimizing the train schedule is insufficient to achieve the system optimum on the considered oversaturated line. To improve the service quality as much as possible, the decision-makers of the urban rail transit system must pay more attention to the collaborative optimization of robust passenger flow control and train scheduling approaches, and moreover, they must carefully determine the value of ζ_2/ζ_1 based on the practical operation needs.

6. Conclusions

To reduce over-saturation of the urban rail transit line, this paper investigated the integrated problem of collaboratively optimizing passenger flow control strategies and train schedules, for a situation in which the passenger demand is uncertain and is characterized by a series of stochastic scenarios with partially known probability distribution. A multi-scenario two-stage DRO model was formulated to generate near-optimal strategies for the passenger flow control and train scheduling under distribution ambiguity. The mean-CVaR criterion was adopted in the objective function, and the goal was to minimize the sum of the operating time, the expected number of waiting passengers, and the CVaR of congestion, the coefficients of which could be adjusted flexibly according to the practical situation in order to achieve the most suitable balance. Since the proposed model is usually computationally intractable, a discrepancy-based ambiguity set was constructed by using partial information on the uncertain probability distribution of the stochastic scenarios. Computationally tractable forms of the proposed model were also deduced. To

Table 4 Sensitivity analyses of ζ_1 and ζ_2 with $\alpha = 0.50$, $\lambda = 0.10$, and $\Psi = 0.02$.

Instance	ζ_1	ζ_2	ζ_2/ζ_1	Objective value	Obj. 1	Obj. 2	CPU time (s)
Case 1	0	1.00000	∞	166970.080	176.000	166970.080	1563.642
Case 2	0.00001	1.00000	100 000	166970.082	176.000	166970.080	1169.072
Case 3	0.00010	1.00000	10 000	166970.098	176.000	166970.080	1355.917
Case 4	0.00100	1.00000	1000	166970.256	176.000	166970.080	1376.100
Case 5	0.01000	1.00000	100	166971.840	176.000	166970.080	1172.356
Case 6	1.00000	1.00000	1	167146.080	176.000	166970.080	1401.422
Case 7	1.00000	0.01000	0.01	1845.701	176.000	166970.080	1181.674
Case 8	1.00000	0.00001	0.00001	177.677	176.000	167697.539	1721.266
Case 9	1.00000	0	0	176.000	176.000	352150.961	1099.210

obtain high-quality solutions for large-scale problems within an acceptable computing time, we developed a hybrid algorithm by combining an LS with a suitable MILP solver. To verify the performance of the proposed approaches, a series of numerical experiments were conducted using actual operation data from a Beijing rail transit line. The computation results show that our proposed DRO model is superior to the SP model in terms of solution stability. In addition, our model is superior in terms of solution quality and solution stability, compared with the classic RO model. These findings validate the proposed DRO method and the LS + MILP solver algorithm.

Future research can be dedicated to the following aspects: ① In the process of characterizing the uncertain probability distribution, a discrepancy-based ambiguity set is used in this study. Thus, a promising extension would be to construct multi-type ambiguity sets for capturing the probability distribution, such as a Gaussian ambiguity set or a polyhedral ambiguity set. ② The study of integrated and interconnected multimodal transport systems remains a hot topic for the future development of urban mobile ecosystems. Hence, the collaborative optimization of the passenger flow control in a multimodal transportation system may be an interesting research direction for comprehensively easing congestion. ③ In addition, it should be noted that this paper only considers the operation environment of an urban rail transit line. An interesting direction for future research would be to investigate efficient algorithms for more complex lines or networks in Beijing and in other representative cities.

Acknowledgments

This work was supported the National Natural Science Foundation of China (71621001, 71825004, and 72001019), the Fundamental Research Funds for Central Universities (2020JBM031 and 2021YJS203), and the Research Foundation of State Key Laboratory of Rail Traffic Control and Safety (RCS2020ZT001).

Compliance with ethics guidelines

Yahan Lu, Lixing Yang, Kai Yang, Ziyou Gao, Housheng Zhou, Fanting Meng, and Jianguo Qi declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2021.09.016>.

References

- [1] Yin J, D'Ariano A, Wang Y, Yang L, Tang T. Timetable coordination in a rail transit network with time-dependent passenger demand. *Eur J Oper Res* 2021;295(1):183–202.
- [2] Barrena E, Canca D, Coelho LC, Laporte G. Exact formulations and algorithm for the train timetabling problem with dynamic demand. *Comput Oper Res* 2014;44:66–74.
- [3] Mesa JA, Ortega FA, Pozo MA. Locating optimal timetables and vehicle schedules in a transit line. *Ann Oper Res* 2014;222(1):439–55.
- [4] Niu HM, Zhou XS. Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. *Transp Res Part C Emerg Technol* 2013;36(16):212–30.
- [5] Shi J, Yang L, Yang J, Gao Z. Service-oriented train timetabling with collaborative passenger flow control on an oversaturated metro line: an integer linear optimization approach. *Transport Res B-Meth* 2018;110:26–59.
- [6] Xu XY, Liu J, Li HY, Jiang M. Capacity-oriented passenger flow control under uncertain demand: algorithm development and real-world case study. *Transport Res E-Log* 2016;87:130–48.
- [7] Jiang M, Li HY, Xu XY, Xu S, Miao JR. Metro passenger flow control with station-to-station cooperation based on stop-skipping and boarding limiting. *J Cent South Univ* 2017;24(1):236–44.
- [8] Xu XY, Li HY, Liu J, Ran B, Qin LQ. Passenger flow control with multi-station coordination in subway networks: algorithm development and real-world case study. *Transportmetrica B* 2018;7(1):446–72.
- [9] Shi J, Yang L, Yang J, Zhou F, Gao Z. Cooperative passenger flow control in an oversaturated metro network with operational risk thresholds. *Transp Res Part C Emerg Technol* 2019;107:301–36.
- [10] Yang S, Yang K, Yang L, Gao Z. MILP formulations and a TS algorithm for reliable last train timetabling with uncertain transfer flows. *J Oper Res Soc* 2018;69(8):1318–34.
- [11] Meng FT, Yang LX, Shi JG, Jiang ZZ, Gao ZY, et al. Collaborative passenger flow control for oversaturated metro lines: a stochastic optimization method. *Transportmetrica A*. In press.
- [12] Ben-Tal A, Ghaoui LE, Nemirovski A. *Robust optimization*. Princeton: Princeton University Press; 2009.
- [13] Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. *J Risk* 1999;2(3):21–42.
- [14] Wang L, Yan X, Wang Y. Modeling and optimization of collaborative passenger control in urban rail stations under mass passenger flow. *Math Probl Eng* 2015;2015:1–8.
- [15] Yuan F, Sun H, Kang L, Wu J. Passenger flow control strategies for urban rail transit networks. *Appl Math Model* 2020;82:168–88.
- [16] Meng F, Yang L, Wei Y, Li S, Gao Z, Shi J. Collaborative passenger flow control on an oversaturated metro line: a path choice approach. *Transportmetrica B* 2020;8(1):376–404.
- [17] Hoogendoorn SP, Daamen W. Design assessment of Lisbon transfer stations using microscopic pedestrian simulation. *Computers in Railways* 2004;74:135–47.
- [18] Zhang JX, Han BM, Li DW. Application of VISSIM in pedestrian simulation of MTR stations. *Comput Simul* 2007;6:239–42. Chinese.
- [19] Seriani S, Fernandez R. Pedestrian traffic management of boarding and alighting in metro stations. *Transp Res Part C Emerg Technol* 2015;53:76–92.
- [20] Fei S, Liu ZL. Simulation analysis of metro congestion points and optimization method. *Urban Mass Transit* 2018;21(7):100–5.
- [21] Jiang ZB, Gu JJ, Fan W, Liu W, Zhu BQ. Q-learning approach to coordinated optimization of passenger inflow control with train skip-stopping on an urban rail transit line. *Comput Ind Eng* 2019;127:1131–42.
- [22] Niu H, Zhou X, Gao R. Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: nonlinear integer programming models with linear constraints. *Transport Res B-Meth* 2015;76:117–35.
- [23] Li DW, Zhang TY, Dong XL, Yin Y, Cao J. Trade-off between efficiency and fairness in timetabling on a single urban rail transit line under time-dependent demand condition. *Transportmetrica B* 2019;7(1):1–29.
- [24] Yin JT, D'Ariano A, Wang YH, Xun J, Su S, Tang T. Mixed-integer linear programming models for coordinated train timetabling with dynamic demand. In: Bogenberger K, Tang T, Knoll A, Dietmayer K, Hu J, Zöllner JM, editors. *IEEE Intelligent Transportation Systems Conference; 2019 Oct 27–30; Auckland, New Zealand*. New York City: Curran Associates; 2019. p. 863–8.
- [25] Yang LX, Di Z, Dessouky MM, Gao ZY, Shi JG. Collaborative optimization of last-train timetable with passenger accessibility: a space-time network design based approach. *Transp Res Part C Emerg Technol* 2020;114:572–97.
- [26] Liu P, Schmidt M, Kong Q, Wagenaar JC, Yang L, Gao Z, et al. A robust and energy-efficient train timetable for the subway system. *Transp Res Part C Emerg Technol* 2020;121:102822.
- [27] Huang Y, Mannino C, Yang L, Tang T. Coupling time-indexed and big-M formulations for real-time train scheduling during metro service disruptions. *Transport Res B-Meth* 2020;133:38–61.
- [28] Tian XP, Niu HM. Optimization of demand-oriented train timetables under overtaking operations: a surrogate-dual-variable column generation for eliminating indivisibility. *Transport Res B-Meth* 2020;142:143–73.
- [29] Mo P, D'Ariano A, Yang L, Veelenturf LP, Gao Z. An exact method for the integrated optimization of subway lines operation strategies with asymmetric passenger demand and operating costs. *Transport Res B-Meth* 2021;149:283–321.
- [30] Qi J, Cacchiani V, Yang L, Zhang C, Di Z. An integer linear programming model for integrated train stop planning and timetabling with time-dependent passenger demand. *Comput Oper Res* 2021;136:105484.
- [31] Xu XY. *Calculation of station service capacity and adaptability in urban rail transit [dissertation]*. Beijing: Beijing Jiaotong University; 2015. Chinese.
- [32] Gong C, Mao B, Wang M, Zhang T. Equity-oriented train timetabling with collaborative passenger flow control: a spatial rebalance of service on an oversaturated urban rail transit line. *J Adv Transp* 2020;2020. 8867404.
- [33] Yang W, Su Q, Huang SH, Wang Q, Zhu Y, Zhou M. Simulation modeling and optimization for ambulance allocation considering spatiotemporal stochastic demand. *J Manage Sci Eng* 2019;4(4):252–65.
- [34] Zhang Z, Liu Y, Tong Q, Guo S, Li D. Evacuation based on spatio-temporal resilience with variable traffic demand. *J Manage Sci Eng* 2021;6(1):86–98.
- [35] Gong C, Shi J, Wang Y, Zhou H, Yang L, Chen D, et al. Train timetabling with dynamic and random passenger demand: a stochastic optimization method. *Transp Res Part C Emerg Technol* 2021;123:102963.
- [36] Errico F, Desaulniers G, Gendreau M, Rei W, Rousseau LM. A priori optimization with recourse for the vehicle routing problem with hard time windows and stochastic service times. *Eur J Oper Res* 2016;249(1):55–66.
- [37] Qi J, Cacchiani V, Yang L. Robust train timetabling and stop planning with uncertain passenger demand. *Electron Notes Discret Math* 2018;69:213–20.

- [38] Cacchiani V, Qi J, Yang L. Robust optimization models for integrated train stop planning and timetabling with passenger demand uncertainty. *Transport Res B-Meth* 2020;136:1–29.
- [39] Delage E, Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper Res* 2010;58(3):595–612.
- [40] Esfahani PM, Kuhn D. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Math Program* 2018;171(1):115–66.
- [41] Shang C, You F. Distributionally robust optimization for planning and scheduling under uncertainty. *Comput Chem Eng* 2018;110:53–68.
- [42] Van BP, Kuhn D, Goulart PJ, Morari M. Distributionally robust control of constrained stochastic systems. *IEEE Trans Automat Contr* 2015;61(2):430–42.
- [43] Shang XT, Yang K, Jia B, Gao Z. Distributionally robust cluster-based hierarchical hub location problem for integration of urban and rural public transport system. *Comput Ind Eng* 2021;155:107181.
- [44] Wang WQ, Yang K, Yang LX, Gao Z. Tractable approximations for the distributionally robust conditional vertex p -center problem: application to the location of high-speed railway emergency rescue stations. *J Oper Res Soc* 2020.
- [45] Yang K, Lu Y, Yang L, Gao Z. Distributionally robust last-train coordination planning problem with dwell time adjustment strategy. *Appl Math Model* 2021;91:1154–74.
- [46] Wang S, Chen Z, Liu T. Distributionally robust hub location. *Transport Sci* 2020;54(5):1189–210.
- [47] Zhang Y, Zhang Z, Lim A, Sim M. Robust data-driven vehicle routing with time windows. *Oper Res* 2021;69(2):469–85.
- [48] Domencich TA, McFadden D. *Urban travel demand: a behavioral analysis*. Oxford: North-Holland Publishing Company Limited; 1975.
- [49] Ma L, Liu Y, Liu Y. Distributionally robust design for bicycle-sharing closed-loop supply chain network under risk-averse criterion. *J Clean Prod* 2020;246:118967.
- [50] Mevissen M, Ragnoli E, Yu JY. Data-driven distributionally robust polynomial optimization. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, editors. *Proceedings of the 26th International Conference on Neural Information Processing Systems*; 2013 Dec 5–10; Lake Tahoe, USA. New York City: Curran Associates; 2013. p. 37–45.