Xu, G., Zhou, P., Liu, W., & Gao, Y. (2025). Seat allocation optimization for railway systems with equity considerations. Transportmetrica B: Transport Dynamics, 13(1), 2448707. https://doi.org/10.1080/21680566.2024.2448707

Optimization of seat allocation scheme for railway systems with equity considerations

Guangming Xu[a], Peilan Zhou[a], Wei Liu[b,*], Yihan Gao[b]

[a] *School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China*

[b] *Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong, China*

[*] *Corresponding author. Email: wei.w.liu@polyu.edu.hk (W. Liu)*

Abstract

This paper investigates the railway seat allocation problem with a focus on equity considerations. We aim to distribute railway capacity more fairly among passengers from different Origin-Destination (OD) pairs while ensuring operational profitability. We first develop a Mixed Integer Linear Programming (MILP) model for scenarios with deterministic demand. We then further extend our study by formulating Stochastic Programming (SP) and Distributionally Robust Optimization (DRO) models for scenarios with demand uncertainty. Additionally, we derive the deterministic equivalent of the DRO model using a box ambiguity set. Furthermore, we explore the relationships between the proposed DRO and SP models, both of which can be efficiently solved using widely available MILP solvers like GUROBI. To validate our approach, we perform numerical studies on a small-scale toy example and the Zhengzhou-Xi'an high-speed railway corridor. The results demonstrate that the proposed optimization methods improve passenger equity across OD pairs, where the DRO model yields high-quality and stable solutions.

Keywords: Railway system; Seat allocation; Equity; Demand uncertainty; MILP

Acknowledgments. We would like to thank the referees for their helpful comments. This study is partially supported by the National Natural Science Foundation of China (72171236, 71871226), the MTR Research Funding Scheme (PTU-24016), and The Hong Kong Polytechnic University (P0039246, P0040900, P0041316).

1. Introduction

Railway transportation plays an important role in the rapid development of countries around the world (Chai et al., 2018). For instance, by the end of 2023, China's railway network had expanded to a total track length of 159,000 km, including an impressive 45,000 km of high-speed railways.[1] This extensive rail network offers efficient transportation for passengers and accommodates the increasing travel demand, especially for those with medium-to-long trip distances. However, the significant fixed costs of constructing railway infrastructure often lead to financial deficits for many railway lines. As a result, it is both important and necessary for railway operators to develop planning and operational strategies that can improve system efficiency and profitability.

Revenue management or RM (Çetiner, 2013) has been extensively studied in many industries. Previous research (Abe, 2008; Armstrong & Meissner, 2010) has provided comprehensive overviews specific to railway RM. In the context of RM, seat allocation is a critical component, determining numbers of seats assigned to all origin-destination (OD) pairs on specific trains. Unlike airline services, where flights typically serve a single OD pair, railway systems often feature complex supply and demand structures. For instance, a train can serve multiple OD pairs, and different trains on the same line may have different stopping plans and schedules. Consequently, each OD pair can be viewed as a distinct market that is governed by the seat allocation (Xu et al., 2022a). An effective seat allocation scheme can mitigate inefficient competition for limited train capacity, thereby enhancing both capacity utilization efficiency and profitability of railway operations.

Ciancimino et al. (1999) addressed the railway seat allocation problem by employing deterministic linear and probabilistic nonlinear programming models, approaching it as a multi-leg, single-fare revenue management problem. You (2008) further expanded the analysis to incorporate two fare classes—full fare and discounted fare—building on the framework proposed by Ciancimino et al. (1999). These earlier studies mainly

[1] Railway development in China to 2023.
< http://www.china-railway.com.cn/xwzx/zhxw/202403/t20240315_134819.html>

concentrated on the seat allocation problem for a single train or a single line. In recent years, however, the rapid development of high-speed railways has prompted a growing number of studies to examine seat allocation from a network perspective (Zhao & Zhao, 2019). Some researchers have introduced more flexible strategies to accommodate system variations (Jiang et al., 2015; Yan et al., 2020; Zhao et al., 2023). Additionally, recognizing the interaction between pricing and seat allocation (Hu et al., 2020), some studies have jointly optimized these aspects (Hu et al., 2020; Yan et al., 2020; Xu et al., 2022a). Other research has simultaneously optimized seat allocation and train stopping plans (Han & Ren, 2020; Xu et al., 2023). However, existing studies on railway seat allocation have seldom addressed equity issues across heterogeneous passengers from different OD pairs, despite the importance of equity in transportation service provision for social good (Litman, 2002).

Although equity considerations in railway seat allocation schemes have not been widely explored, the integration of equity principles into transportation systems has garnered increasing interest in recent decades (Bertolaccini, 2013). For example, Chen & Subprasom (2007) highlighted the unequal distribution of benefits among road users from different destinations. Some studies have investigated resource-sharing within urban rail transit networks (Shang et al., 2018; Yin et al., 2022). The accessibility and social impacts of high-speed railways have been evaluated in many studies (Kim & Sultana, 2015; Chen & Haynes, 2017; Zhou et al., 2018; Cavallaro et al., 2020; Ren et al., 2020; Zhang et al., 2020). However, few studies have developed optimization tools to enhance equity in railway systems. Notably, Zhan et al. (2020) examined social equity-based timetabling and ticket pricing for high-speed railways, while Xia et al. (2023) explored train stop planning and ticket pricing optimization with equity considerations. These studies primarily focused on the cost of rail travel, whereas others (Shao et al., 2022) concentrated on equity in passenger travel time. Unlike these studies, this research emphasizes fair resource distribution within the rail system to reduce inequality in travel opportunities among passengers. We contribute to the existing literature by developing seat allocation methods that enhance operator profit while ensuring that a certain level of equity is maintained.

Railway seat allocation optimization depends on demand information, which is inherently uncertain and varies daily, making accurate prediction challenging. This uncertainty is influenced by various factors, including seasonal changes, weather conditions, special events (such as holidays or emergencies), and the time-varying preferences of travelers. Some studies have addressed this uncertainty in railway seat allocation. For example, Wang et al. (2016) assumed a non-uniform Poisson process for demand between different OD pairs and developed a stochastic programming model. More recently, Xu et al. (2022b) optimized train capacity allocation for railways with mixed passenger and freight flows, considering uncertain passenger and freight demand to maximize expected system revenue. However, due to the ex-ante nature of transport system decisions, obtaining the true probability distribution of uncertain parameters remains challenging, leading to less robust results with Stochastic Programming (SP) approaches. To address this issue, Distributionally Robust Optimization (DRO) methods have gained attention for handling uncertain parameters (Sun et al., 2014; Rahimian & Mehrotra, 2022). Initially introduced by Delage & Ye (2010) for uncertainty in moments (mean and covariance matrix), the efficacy of DRO methods has been reported by many. Consequently, researchers in various fields have increasingly adopted DRO models (Agra & Rodrigues, 2022; Zheng et al., 2022). However, to the best of our knowledge, DRO has rarely been applied to railway revenue management problems.

In summary, this paper advances the literature by integrating equity considerations into railway seat allocation schemes and developing robust optimization approaches to better manage demand uncertainty. This is especially pertinent because, in China and many other countries, railways are frequently funded wholly or partially by central and local governments (i.e., public funding). Therefore, service profitability and social welfare, including equity issues, are critical aspects that need to be addressed (Camporeale et al., 2019; Zhan et al., 2020). However, in transportation practice, social equity is often overlooked (Manaugh et al., 2015) and relevant tools are not available.

In this paper, we propose an innovative approach to optimizing railway seat allocation that balances revenue maximization and equity considerations. Our contributions are as follows: (i) Equitable seat allocation. We formulate a seat allocation model for a multi-train, multi-station railway line[2] with multiple OD pairs. This model incorporates

[2] Our model is primarily designed for trains, focusing on the equitable distribution of seat resources on

service level constraints to more equitably distribute railway resources among passengers from different OD pairs, aiming to maximize expected revenue while maintaining equity. (ii) Max-Min equity principle. This principle is applied in scenarios with limited resources that require fair allocation. Equity is evaluated by enhancing the minimum passenger accessibility to mitigate extreme disparities. (iii) Handling demand uncertainty. An SP model is introduced to address the complete probability distribution of passenger demand. Additionally, when only partial demand distribution information is available, we propose a tailored DRO model. This DRO model is then converted into an equivalent deterministic formulation that can be efficiently solved using existing linear solvers. (iv) Practical case evaluation. The proposed models are tested on both a toy railway line and the Zhengzhou-Xi'an high-speed rail corridor.

The remaining sections of this paper are organized as follows. Section 2 introduces the concept of equity in seat allocation. Section 3 outlines the proposed models, beginning with a description of the deterministic demand model and the Stochastic Programming (SP) model used for comparison. It then presents the Distributionally Robust Optimization (DRO) model to address uncertainty. In Section 4, we conduct numerical studies using both a small illustrative example and a larger real-world case. Section 5 discusses the seat allocation model under time-varying demand. Finally, Section 6 concludes the paper.

2. Equity in railway seat allocation

The distribution of passenger demand in a railway service network varies across different OD pairs, while the system capacity remains limited. This heterogeneity, coupled with capacity constraints, can result in inefficient competition among passengers from different OD pairs sharing the same train capacity. The seat allocation scheme in the railway is crucial for managing this competition and enhancing efficiency by distributing capacities across the network. The allocation of tickets or capacities to each OD pair directly impacts the accessibility of passengers from different OD pairs,

---

the trains rather than the lines themselves. Optimization methods effective for a single line can also be extended to larger networks, necessitating the management of more OD pairs and trains. Consequently, our model is equally applicable to rail networks.

making both efficiency and fairness essential. Optimizing seat allocation should aim not only to manage inefficient and unnecessary competition but also to ensure fairness across the network for the greater social good. To underscore the importance of considering both efficiency and social equity in seat allocation optimization, we provide a simple example below.

In Fig. 1, a single train with a seating capacity of 10 travels sequentially from station A to B, and then to C. We consider three OD pairs: AB, BC, and AC, with ticket fares and demands as follows: AB (fare: 2, demand: 10), BC (fare: 2, demand: 20), and AC (fare: 3, demand: 30). To maximize revenue, the optimal seat allocation would be 10 seats each to AB and BC, with no seats allocated to AC. This results in a total revenue of 40 and serves 20 passengers. However, this allocation neglects the demand for OD pair AC, which can be considered "unfair" and severely limits accessibility for passengers traveling between A and C. Alternatively, if we aim to "maximize equity," we could allocate 4 seats each to AB and BC, and 6 seats to AC. This results in a total revenue of 34, serving 14 passengers, with the service distribution being 40% for AB, 20% for BC, and 20% for AC. This example underscores the importance of equitable seat allocation across different OD pairs while fully utilizing the train's capacity. Clearly, there is a trade-off between efficiency/profitability and equity.

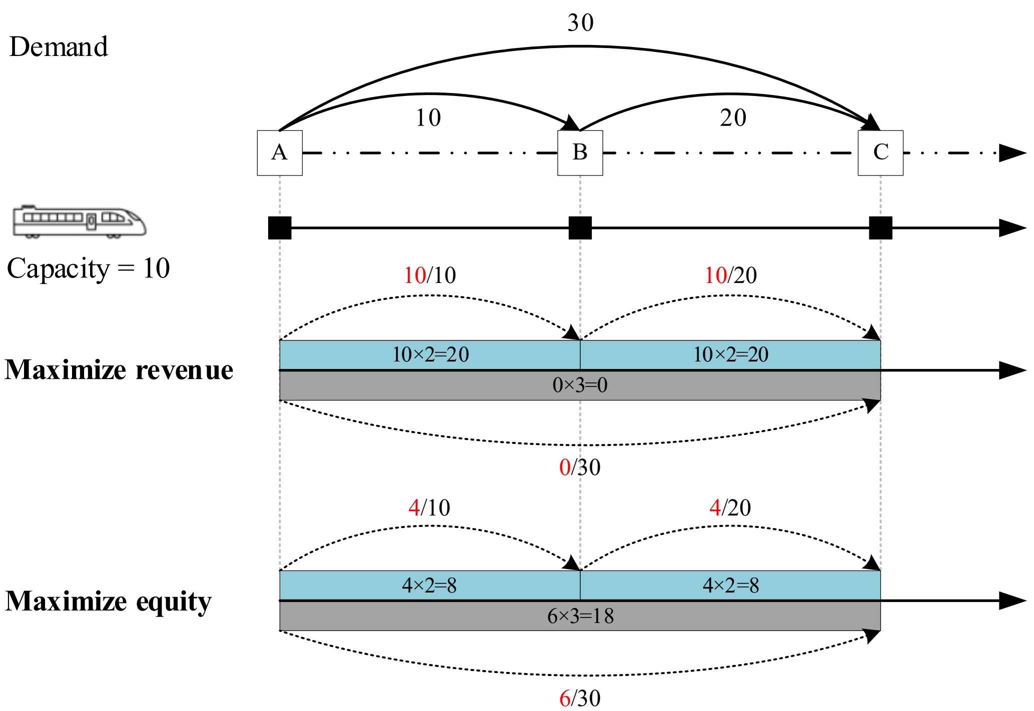


Fig. 1. An example: maximize revenue vs "maximize equity" in seat allocation.

The Gini coefficient (Gini, 1912) is a widely used metric for quantifying disparities in income or wealth distribution within a population. Delbosc & Currie (2011) applied this metric to evaluate equity implications in transportation systems. Other equity-related metrics, such as the relative mean deviation (Ogryczak, 2009), the gap between the best and worst values (Martens et al., 2012), and a lexicographic framework (Liu & Papageorgiou, 2013), have also been utilized. However, these traditional fairness metrics often struggle to accommodate the complexities of rail systems and are particularly inflexible in resource allocation problems. This study employed the min-max approach, which has proven effective. For instance, Li et al. (2019) used min-max fairness to reduce maximum passenger waiting times, thereby enhancing overall system equity. Similarly, Zhao et al. (2021) employed maximum waiting time as a min-max indicator to assess equity implications when optimizing train schedules.

We present the notations for sets, parameters, and decision variables used in the subsequent modeling in Table 1.

Table 1. Notations used in the model.

Notation	Description
Sets and indices	
S	Set of stations
K	Set of trains
M	Set of sections
RS	Set of OD pairs
(i,j)	OD pair: from origin station i to destination station j, where $i,j \in S$
m	Index of sections
k	Index of trains
Parameters	
d_{ij}	Potential passenger demand for OD pair (i,j)
c_k	Seat capacity of train k
δ_{ij}^{km}	Binary variable that equals 1 if the OD pair (i,j) served by train k covers section m and 0 otherwise
h_{ij}	Seat price for OD pair (i,j)

λ	The social equity weighting factor for seat allocation
Decision variables	
x_{ij}	Total number of seats assigned to OD pair (i,j)
x_{ij}^{k}	Number of seats assigned to OD pair (i,j) of train k
Auxiliary variables	
e_{ij}	Proportion of passenger demand satisfied by seats allocated to OD pair (i,j)
θ	Minimum value among e_{ij}
y_{ij}	Actual demand on OD pair (i,j)

This study incorporates equity considerations in the following manner. Let the variable x_{ij} and parameter d_{ij} represent the seat allocation and demand for any OD pair (i,j), respectively, where $(i,j) \in RS$. We can introduce the variable e_{ij} to indicate the proportion of seat supply meeting the demand.

$$e_{ij} = \frac{x_{ij}}{d_{ij}}, \forall (i,j) \in RS \tag{1}$$

The metric e_{ij} can be regarded as the level of accessibility to the railway for passengers from OD pair (i,j). We try to improve equity by improving the minimum e_{ij} among all OD pairs, i.e., we try to improve the θ defined below

$$\theta = \min_{(i,j)\in RS} e_{ij} \tag{2}$$

In the model formulations to be discussed in Section 3, we will try to maximize θ and then constraint (2) is replaced by the constraint below, i.e.,

$$e_{ij} \geq \theta, \forall (i,j) \in RS \tag{3}$$

Alternatively speaking, we try to improve the minimum level of accessibility for all OD pairs.

3. Model formulations

In this section, we introduce the model formulations for seat allocation in rail systems, incorporating considerations of social equity.

3.1 Problem description

Consider a railway line comprising $|S|$ stations and $|K|$ trains, with station 1 and station S representing the first and terminal stations along the line, respectively, as shown in Fig. 2. Passengers of each OD pair can be served by multiple trains. A section, denoted as m, is defined as the link between two consecutive stations, i.e., we have $|S|-1$ sections along the line. We list some major assumptions in this study below.

A1. The railway service price for each OD pair is predetermined. Our focus on equity primarily involves balancing railway supply and demand, thus we concentrate on the seat allocation problem.

A2. Consider homogeneous seats within a single train, consistent across the entire fleet (Xu et al., 2022a). While trains may have multiple seating categories in reality, our proposed model can treat each seating category as a separate train, each containing only one specific type of seat. This approach allows our model to effectively address the seat allocation problem.

A3. Passenger ticket cancellations and transfers are not considered (Yan et al., 2020). Equity is only addressed during the ticket sales phase, without accounting for changes that may occur afterward. Ignoring these factors does not impact passengers' perception of equity.

We will consider both deterministic demand and stochastic demand cases.

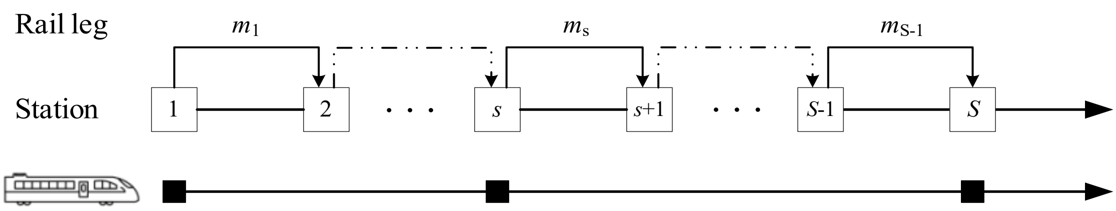


Fig. 2. An example railway line and a given train stopping plan.

3.2 Model formulation for the deterministic demand

We begin by addressing the deterministic demand scenario and develop a seat allocation optimization model, incorporating both problem constraints and the objective function.

Model Constraints. The seat allocation model has the following constraints: train capacity constraints, seat allocation conservation constraints, social equity constraints, and actual demand constraints.

(Train capacity constraints) In seat allocation, a single seat cannot be simultaneously assigned to two or more overlapping OD pairs. For example, a seat cannot be allocated to both the first station to the second station and the first station to the last station at the same time. However, it can be assigned to both the first station to the second station and the second station to the last station, as there is no overlap between these OD pairs. Correspondingly, in the model, for any rail section m (between two adjacent stations), the total number of seats on train k allocated to all OD pairs covering the section m should not exceed the train capacity c_k, i.e.,

$$\sum_{(i,j)\in RS} x_{ij}^k \delta_{ij}^{km} \leq c_k, k \in K, m \in M \tag{4}$$

$$x_{ij}^k \in N, \forall (i,j) \in RS, k \in K \tag{5}$$

where the binary variable δ_{ij}^{km} is given based on train stopping planning, i.e., subject to the incidence relation among section m, OD pair (i,j) and train k. Specifically, δ_{ij}^{km} equals one if the OD pair (i,j) served by train k encompasses section m and zero otherwise.

(Seat allocation conservation) For each OD pair $(i,j) \in RS$, there might be multiple trains serving it. Consequently, the total number of seats allocated to each OD pair is equal to the sum of that on all trains serving this OD pair, i.e.,

$$\sum_{k\in K} x_{ij}^k = x_{ij}, \forall (i,j) \in RS \tag{6}$$

where x_{ij} is the total number of seats allocated to the OD pair $(i,j) \in RS$.

(Actual demand constraints) For OD pair (i,j), we denote y_{ij} as the actual demand (i.e., the total number of tickets sold given the ticket price). We then have

$$y_{ij} = \min\{x_{ij}, d_{ij}\}, \forall (i,j) \in RS \tag{7}$$

i.e., the actual demand is the minimum between the total number of seats allocated x_{ij} and the potential demand d_{ij}. Given an objective of revenue maximization (to be discussed below), the nonlinear constraint (7) can be converted to the following two linear constraints (8) and (9)

$$y_{ij} \leq x_{ij}, \forall (i,j) \in RS \tag{8}$$

$$y_{ij} \leq d_{ij}, \forall (i,j) \in RS \tag{9}$$

(Equity constraints) Constraints related to equity have been discussed in Section 2, i.e., constraints (1) and (3).

Objective function. This study addresses both railway profitability (for financial sustainability) and social equity. Our objective is to optimize the train seat allocation strategy to balance revenue generation with equitable service distribution. Previous studies (Li et al., 2019; Ning et al., 2021) and the simple example in Section 2 have already demonstrated the trade-off between efficiency/profitability and service equity. We simultaneously consider social equity through the proposed equity indicator θ and ticket revenue, aiming to achieve a balance between the two, i.e.,

$$\max Z_1 = \theta \tag{10}$$

$$\max Z_2 = \sum_{(i,j)\in RS} h_{ij} y_{ij} \tag{11}$$

To simplify the bi-objective problem, we consider a linear combination of the above two objectives, where we add a weight λ for the equity metric. Then, we can define the following seat allocation model:

$$\begin{gathered} \max Z = \lambda\theta + \sum_{(i,j)\in RS} h_{ij} y_{ij} \\ s.t. \quad \text{Constraints} \quad (1), (3) - (6), (8) - (9) \end{gathered} \tag{12}$$

When λ is set as zero, the above model reduces to revenue maximization. Note that the operator may choose a different value for λ that reflects its preference.

3.3 Stochastic programming model

The deterministic model discussed earlier assumes a scenario with fixed and fully known demand. However, demand is often uncertain in practice. To address this, we introduce a stochastic programming model that accounts for demand variability. Specifically, we adopt a scenario-based approach, where each scenario represents a possible realization of demand distributions across all OD pairs. We use w to denote a scenario and $d_{ij}(w)$ to represent the demand from station i to station j under the scenario w. Note that $d_{ij}(w)$ will replace the deterministic potential demand d_{ij}. Let $p_w \geq 0$ denotes the probability of the occurrence of scenario w, where $\sum_{w\in\Omega} p_w = 1$, and Ω contains all possible scenarios. The probability vector is $\boldsymbol{p} =$

$(p_1, p_2, \cdots, p_w, \cdots, p_{|\Omega|})^T$, where $w \in \Omega$.

Under the aforementioned stochastic demand setting, our objective is to maximize the expected value across all demand scenarios $w \in \Omega$, i.e.,

$$\max E_{\boldsymbol{p}} \left[\lambda\theta(\boldsymbol{w}) + \sum_{(i,j)\in RS} h_{ij} y_{ij}(\boldsymbol{w}) \right] \tag{13}$$

where $E_{\boldsymbol{p}}[\cdot]$ is the expected value operator. $\theta(\boldsymbol{w})$ and $y_{ij}(\boldsymbol{w})$ replace θ and y_{ij} from the deterministic model, making them scenario-dependent. The other problem constraints should be adjusted accordingly, i.e.,

$$e_{ij}(w) = \frac{x_{ij}}{d_{ij}(w)}, \forall (i,j) \in RS, w \in \Omega \tag{14}$$

$$e_{ij}(w) \geq \theta(w), \forall (i,j) \in RS, w \in \Omega \tag{15}$$

$$y_{ij}(w) \leq x_{ij}, \forall (i,j) \in RS, w \in \Omega \tag{16}$$

$$y_{ij}(w) \leq d_{ij}(w), \forall (i,j) \in RS, w \in \Omega \tag{17}$$

It is noteworthy that we may have full and accurate information regarding $\boldsymbol{p} = (p_1, p_2, \cdots, p_w, \cdots, p_{|\Omega|})^T$, i.e., the probability distributions for stochastic demand. This motivates us to further develop a distributionally robust optimization modeling framework (to be discussed in Section 3.4).

3.4 Distributionally robust optimization model

The Distributionally Robust Optimization (DRO) method enhances the generality of optimization results by assuming that uncertain parameters follow distributions within a set of probability distributions, known as the ambiguity set. This approach has been widely adopted in various studies (Goh & Sim, 2010; Sun et al., 2014; Wang et al., 2021; Rahimian & Mehrotra, 2022). Building on these studies, we discard the assumption of a perfect probability distribution for uncertain demand. Instead, we design a DRO framework for the seat allocation problem, aiming to identify the "worst-case" passenger distribution within the ambiguity set.

We consider that only partial probabilistic information regarding uncertain demand is available. The passenger demand $d_{ij}, (i,j) \in RS$, is a random variable following a probability distribution $\boldsymbol{p}$ with finite support $\mathbb{p}$. Note that $\mathbb{p}$ is the given ambiguity set, where $\boldsymbol{p} \in \mathbb{p}$. Unlike stochastic programming, which optimizes the expected value of a given function under a specific probability distribution, Distributionally Robust Optimization (DRO) focuses on the worst-case distribution within an ambiguity set. In DRO, the distribution of random factors is treated as a decision variable. The objective of the DRO model is to optimize the expected value against the most adverse probability distribution within the defined ambiguity set. Therefore, the DRO formulation for this problem can be expressed as follows:

$$\max\left\{\min_{\boldsymbol{p}\in\mathbb{p}} E_{\boldsymbol{p}}\left[\lambda\theta(\boldsymbol{w}) + \sum_{(i,j)\in RS} h_{ij}y_{ij}(\boldsymbol{w})\right]\right\}$$
$$s.t.\quad \text{Constraints } (4)-(6),(14)-(17) \tag{18}$$

where the expression $\min_{\boldsymbol{p}\in\mathbb{p}} E_{\boldsymbol{p}}$ depends on the characteristics of ambiguity sets associated with discrete probability distributions. Our objective is to maximize the minimum expected value within the specified ambiguity set.

3.4.1 Ambiguity set

Ambiguity sets should possess two essential attributes: (i) they should leverage existing knowledge of the parameter distribution, including historical data viewed as samples from the true, albeit unknown, distribution, and (ii) they should facilitate the development of computationally feasible formulations solvable by tools like CPLEX. The DRO model (18), by its nature, remains computationally challenging for generic ambiguity sets. To mitigate this complexity, we introduce a box ambiguity set that delineates potential distributions of uncertain demand. This set satisfies the aforementioned criteria and features a clear, concise, and flexible structure. Several studies have successfully applied the box ambiguity set to describe uncertainty and have demonstrated its effectiveness. For instance, Ma et al. (2020) utilized the box ambiguity set to manage demand uncertainty in supply chain management. Lu et al. (2022) employed it to characterize urban rail transit demand and constructed a computationally tractable DRO model. Wang et al. (2021) addressed uncertain demand issues in disaster management. These studies have shown that the DRO method based on the box ambiguity set outperforms traditional methods in terms of solution quality and stability.

The box ambiguity set is characterized by partial distributional information about uncertain demand, denoted as $\mathbb{p} = \{\boldsymbol{p} = \boldsymbol{p_0} + \boldsymbol{\pi} | \boldsymbol{e^T\pi} = 0, \|\boldsymbol{\pi}\|_\infty \leq \varphi\}$. The probability distribution $\boldsymbol{p}$ is discrete and can be expressed as $\boldsymbol{p} = \{\Omega, \boldsymbol{\pi_p}\}$. We assume that different distributions in the box ambiguity set share a common set of scenarios Ω, with $\boldsymbol{\pi_p} = \left(\pi_p^1, \cdots, \pi_p^w, \cdots, \pi_p^{|\Omega|}\right)^T$ and π_p^w indicating the probability of scenario $w \in \Omega$ under probability distribution $\boldsymbol{p}$ - lacks precise knowledge but is presumed to be encapsulated within the ambiguity set $\mathbb{p}$. Additionally, the nominal distribution $\boldsymbol{p_0}$ is the distribution with the highest probability; $\boldsymbol{e}$ stands for the unit vector; $\boldsymbol{\pi} \in R^S$ is a perturbation vector which constitutes different probability distributions within the ambiguity set when combined with $\boldsymbol{p_0}$; $\|\boldsymbol{\pi}\|_\infty$ signifies the infinite norm, reflecting the peak absolute value among the elements of $\boldsymbol{\pi}$; $\varphi \in [0,1]$ is the known upper threshold of fluctuation. In addition, $\boldsymbol{e}^T\boldsymbol{\pi} = 0$ is equivalent to $\sum_{w=1}^{|\Omega|} \pi_p^w = 1$, which ensures that the sum of the probabilities of each scenario w under the probability distribution $\boldsymbol{p}$ equals one.

3.4.2 Deterministic equivalent formulations

The DRO model with the box ambiguity set is characterized as a max-min problem. To address this model effectively, it is helpful to obtain an equivalent deterministic reformulation of the DRO model (18). We introduce the notation $Q(x, d(\boldsymbol{w})) = \lambda\theta(\boldsymbol{w}) + \sum_{(i,j)\in RS} h_{ij} y_{ij}(\boldsymbol{w})$ to ease the presentation, where w represents distinct scenarios under probability distribution $\boldsymbol{p}$, and $w \in \Omega$. Consequently, we can derive the following:

$$\min_{\boldsymbol{p}\in\mathbb{p}} E_{\boldsymbol{p}}\left[\lambda\theta(\boldsymbol{w}) + \sum_{(i,j)\in RS} h_{ij} y_{ij}(\boldsymbol{w})\right] = \min_{\boldsymbol{p}\in\mathbb{p}} E_{\boldsymbol{p}}\left[Q\left(x, d(\boldsymbol{w})\right)\right] \tag{19}$$

Under the box ambiguity set, formula (19) constitutes a linear programming problem. Given the equation $\boldsymbol{p} = \boldsymbol{p_0} + \boldsymbol{\pi}$, where $\boldsymbol{p_0}$ represents the nominal distribution of the discrete probability, the selection of probability distribution $\boldsymbol{p}$ is equivalent to the decision on the perturbation vector $\boldsymbol{\pi}$. Thus, formula (19) can be given by

$$\begin{aligned}
\min_{\boldsymbol{p}\in\mathbb{p}} E_{\boldsymbol{p}}\left[Q\left(x, d(\boldsymbol{w})\right)\right] &= \min_{\boldsymbol{p}\in\mathbb{p}} Q\left(x, d(\boldsymbol{w})\right)^T \boldsymbol{p} \\
&= \min_{\boldsymbol{p}\in\mathbb{p}} Q\left(x, d(\boldsymbol{w})\right)^T (\boldsymbol{p_0} + \boldsymbol{\pi}) \\
&= Q\left(x, d(\boldsymbol{w})\right)^T \boldsymbol{p_0} + \min_{\boldsymbol{\pi}} \left\{ Q\left(x, d(\boldsymbol{w})\right)^T \boldsymbol{\pi} \middle| \boldsymbol{e}^T\boldsymbol{\pi} = 0, \|\boldsymbol{\pi}\|_\infty \leq \varphi \right\}
\end{aligned} \tag{20}$$

where $\|\boldsymbol{\pi}\|_\infty = \max_{w\in\Omega} |\pi_w|$.

Furthermore, $\min_{\boldsymbol{\pi}} \left\{ Q\big(x, d(\boldsymbol{w})\big)^T \boldsymbol{\pi} | \boldsymbol{e}^T \boldsymbol{\pi} = 0, \|\boldsymbol{\pi}\|_\infty \leq \varphi \right\}$ can be rewritten as

$$\begin{aligned} \min_{\boldsymbol{\pi}} \; & Q\big(x, d(\boldsymbol{w})\big)^T \boldsymbol{\pi} \\ s.t. \;\; & \boldsymbol{e}^T \boldsymbol{\pi} = 0 \\ & -\boldsymbol{\pi} \leq \boldsymbol{\Psi} \\ & \boldsymbol{\pi} \leq \boldsymbol{\Psi} \end{aligned} \tag{21}$$

where $\boldsymbol{\Psi} = \varphi \boldsymbol{e}$, which is the upper bound vector of the perturbation vector $\boldsymbol{\pi}$.

Applying the strong duality theory of linear programming, we introduce the pairwise variable μ and the pairwise vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, corresponding to the three constraints in Eq. (21), respectively. We can obtain the pairwise problem of the above as follows:

$$\begin{aligned} \max_{\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}} \; & \boldsymbol{\Psi}^T \boldsymbol{\alpha} + \boldsymbol{\Psi}^T \boldsymbol{\beta} \\ s.t. \;\; & \boldsymbol{e}\mu - \boldsymbol{\alpha} + \boldsymbol{\beta} = Q\big(x, d(\boldsymbol{w})\big) \\ & \boldsymbol{\alpha} \leq 0 \\ & \boldsymbol{\beta} \leq 0 \end{aligned} \tag{22}$$

Overall, the objective function of the original model (18) can be equivalently represented as the following linear programming:

$$\begin{aligned} \max_{\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}} \; & Q\big(x, d(\boldsymbol{w})\big)^T \boldsymbol{p_0} + \boldsymbol{\Psi}^T \boldsymbol{\alpha} + \boldsymbol{\Psi}^T \boldsymbol{\beta} \\ s.t. \;\; & \boldsymbol{e}\mu - \boldsymbol{\alpha} + \boldsymbol{\beta} = Q\big(x, d(\boldsymbol{w})\big) \\ & \boldsymbol{\alpha} \leq 0 \\ & \boldsymbol{\beta} \leq 0 \end{aligned} \tag{23}$$

Following collation, we have successfully derived a deterministic equivalent representation of the model (18) under the box ambiguity set. The variable-based representation is summarized as follows:

$$\max \sum_{w \in \Omega} \left[\lambda\theta(w) + \sum_{(i,j) \in RS} h_{ij} y_{ij}(w) \right] p_0(w) + \sum_{w \in \Omega} \Psi(w)\alpha(w) + \sum_{w \in \Omega} \Psi(w)\beta(w) \tag{24}$$

$$s.t. \;\; \mu - \alpha(w) + \beta(w) = \lambda\theta(w) + \sum_{(i,j) \in RS} h_{ij} y_{ij}(w), w \in \Omega \tag{25}$$

$$\alpha(w) \leq 0, w \in \Omega \tag{26}$$

$$\beta(w) \leq 0, w \in \Omega \tag{27}$$

$$\text{Constraints } (4) - (6), (14) - (17)$$

From the derivations presented, we have formulated a mixed-integer linear programming (MILP) equivalent of the DRO model. Existing solvers like CPLEX and GUROBI, known for their efficiency in handling large-scale linear programs, can be used to solve this MILP problem. Note that when the upper limit of fluctuation, denoted by φ, within the box ambiguity set of the DRO model, is set to zero, the DRO model reduces to the SP model. This indicates that the DRO model is essentially a generalization of the SP model.

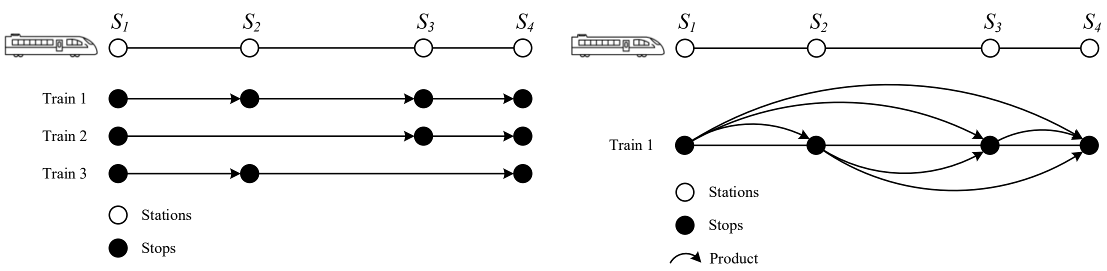
4. Numerical studies

In this section, we test and evaluate the proposed models using both a small-scale example and a large-scale case study. All numerical analyses were conducted on a computer equipped with an Intel® Core™ i7-11700 CPU @2.50GHz, 16.00 GB RAM, and Windows 11 Home Edition OS (64-bit). The MILP problems were solved using the GUROBI 10.0.0 solver on the MATLAB 2021b platform.

4.1 A small toy example

4.1.1 Data description

A toy line example, with a few trains and stations, is presented to illustrate the seat allocation strategy for all train services and highlight key attributes of the proposed approach. As depicted in Fig. 3, a unidirectional railway consists of 4 stations. Three trains depart from the origin station S_1, and arrive at the terminal station S_4, with each train having a capacity of 650 passengers. The train stopping plan for each train is specified in Fig. 3(a), while Fig. 3(b) provides train service information for the specific OD pairs served by train 1. Table 2 lists the fares for all OD pairs.



(a) Stopping plans for multiple trains. (b) A train serving multiple ODs.

Fig. 3. Train service network in the small case.

Table 2. Ticket prices (unit: CNY) for all OD pairs in the toy example.

OD pair	S_1	S_2	S_3	S_4
S_1		12	27	35
S_2			17	24
S_3				9
S_4				

Seat allocation plans are intended for relatively long-term use and must accommodate demand uncertainty over extended periods. To simulate various real-life situations—such as increased demand during holidays or peak travel seasons, decreased demand due to bad weather or emergencies, and regular daily demands—we consider three discrete demand scenarios (i.e., scenarios 1, 2, and 3). This study introduces a standardized set of demand metrics for each OD pair listed in Table 3 as passenger demand under scenario 2. In contrast, the demands for scenarios 1 and 3 are derived by decreasing or increasing the standardized demand with a scaling factor ξ, where ξ ranges between zero and one.

Table 3. Potential passenger demand for all OD pairs under the standard scenario.

OD pair	S_1	S_2	S_3	S_4
S_1		460	1100	1740
S_2			660	840
S_3				320
S_4				

4.1.2 Comparative analysis of different models

In the deterministic demand case (DP model), demand is defined as the expected value derived from multiple uncertain scenarios. The DRO model incorporates three demand scenarios as previously mentioned. Moreover, when the fluctuation upper bound φ within the box ambiguity set is set as zero, the DRO model converges to the traditional SP model. In particular, p_0 is set as $(0.3,0.5,0.2)^T$, the value of the equity conversion coefficient λ in the objective function is set as 10,000, and the fluctuation upper threshold φ in the DRO model is set to 0.2.

The detailed results are presented in Table 4. For all models, the MIPGap configuration parameter in GUROBI is set to 0.00%. Specifically, the computational time required to achieve the optimal solution for deterministic demand is less than 0.2 seconds. The third and fourth columns of Table 4 display the performance metrics under the SP and DRO models across the three demand scenarios. Specifically, the third column indicates the minimum level of service (expressed as a supply/demand ratio) identified by the model for all OD pairs. For comparison, we apply the optimal solution of the DP model to various scenarios with demand ambiguity and then assess the minimum service level and the expected revenue. The results are presented in Table 5.

Table 4. Results of different models.

Model	Demand scenario w	Minimum ratio θ	Total revenue (unit: CNY)	Objective value	CPU time (unit: second)
DP	-	0.25	69,185	71,650.99	0.19
	$w = 1$	1.00	41,030		
SP	$w = 2$	0.33	63,388	62,962.24	0.29
	$w = 3$	0.31	68,488		
	$w = 1$	1.17	41,030		
DRO	$w = 2$	0.40	60,971	58,849.31	0.26
	$w = 3$	0.32	61,829		

Table 5. Results of the DP model for different demand scenarios.

Demand scenario w	Minimum ratio θ	Total revenue (unit: CNY)
$w = 1$	0.36	35,081
$w = 2$	0.19	68,852
$w = 3$	0.18	69,185

Our objective is to maximize the minimum service level among OD pairs to enhance equity, meaning that a higher minimum service level corresponds to a more equitable seat allocation. As demonstrated, incorporating demand uncertainties results in improved performance, achieving both greater equity (a higher minimum service level θ) and increased revenue. Moreover, the DRO model delivers the most robust performance, exhibiting minimal variation across different scenarios. In contrast, the DP model exhibits more fluctuations in performance across scenarios.

Additionally, we created a box plot of e_{ij}, as shown in Fig. 4 to provide a comprehensive understanding of the satisfaction levels across all OD pairs, illustrating the range from the lowest to the highest values. In Fig. 4, the DP_DRO model represents the application of the optimal solution of the DP model to three different demand scenarios. It can be observed that the DRO model, while improving the worst-case scenario, allocates more seats to passengers, thereby fully utilizing the train capacity.

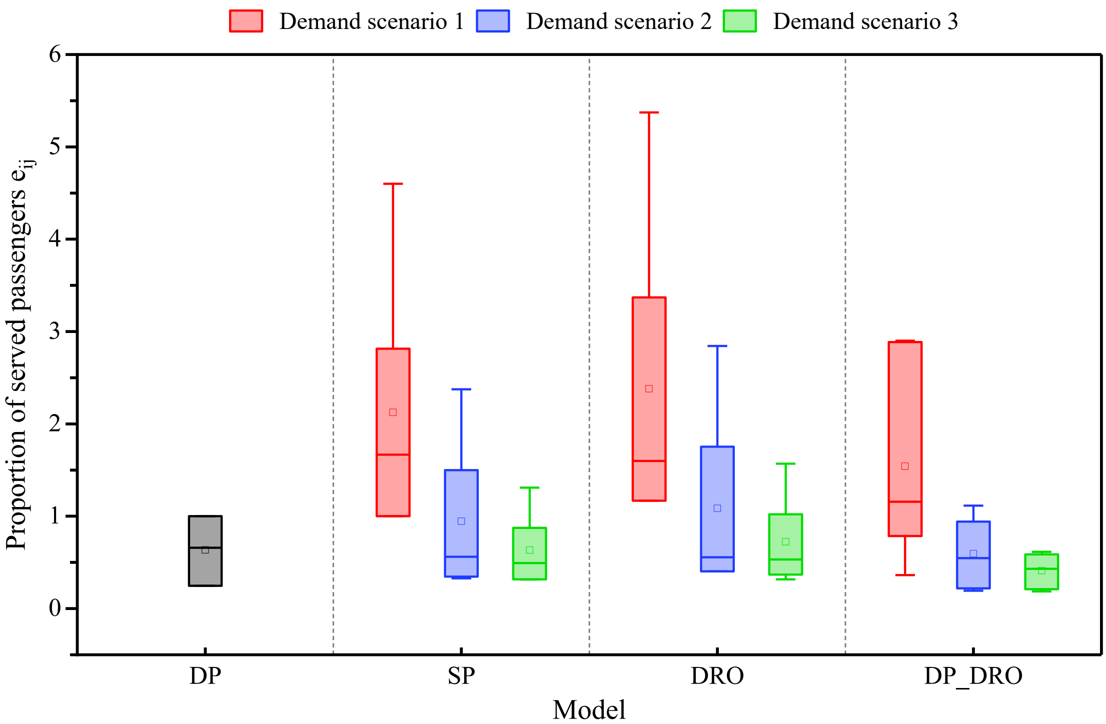


Fig. 4. A box plot of the proportion of served passengers e_{ij} under different models.

4.1.3 Sensitivity analysis of the DRO model against uncertainty

To investigate the influence of φ (an upper bound regulating fluctuations) on the results derived from the DRO model, we vary parameter values as follows: $\varphi = \{0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1\}$, with λ fixed at 30,000. Additionally, we introduce the concept of distributional robustness (metric PDR) to gauge the difference in objective values between the DRO and SP models, i.e.,

$$\mathrm{PDR} = \frac{\mathrm{SP}^* - \mathrm{DRO}^*}{\mathrm{SP}^*} \times 100\% \tag{28}$$

where DRO^* represents the optimal objective value for the DRO model, and SP^* denotes the optimal objective value for the SP model. The results presented in Table 6 show that the objective function value of the DRO model decreases with the upper limit φ of fluctuation. This indicates that increasing ambiguity leads to worse performance. Notably, the increase in φ does not cause a significant decrease in the objective value, meaning that the proposed DRO model produces quite robust results.

Table 6. Comparison of SP and DRO models under different parameters φ.

λ	φ	Objective value	PDR (%)
30,000	0	77,441.14	-
	0.01	77,392.06	0.06
	0.02	77,342.97	0.13
	0.03	77,293.89	0.19
	0.04	77,244.80	0.25
	0.05	77,195.87	0.32
	0.06	77,158.29	0.37
	0.07	77,120.72	0.41
	0.08	77,083.15	0.46
	0.09	77,052.42	0.50
	0.10	77,020.86	0.54

We calculate the PDR associated with the DRO model within the box ambiguity set. The results in Table 6 indicate that the PDR does not exceed 1%, meaning the DRO model requires only a minimal trade-off to effectively address the inherent uncertainties related to probability distributions. These findings highlight the robust performance of the proposed DRO model, particularly in scenarios with limited information on demand distributions.

(Out-of-sample analysis) To assess the validity of the solutions derived from the DRO and SP models, we perform an out-of-sample analysis. In this analysis, the fluctuation limit φ in the DRO model is set to 0.06, while other parameters remain unchanged. The experiments involve the initial three scenarios and randomly generated N sets of scenario probabilities, where we have $N \in \{10, 50, 100, 1000, 10000\}$. The models are optimized using the known dataset, establishing optimal decisions for each model. Then the optimal decisions are tested with distinct sets of test scenarios. The results, presented in Table 7, include out-of-sample objective function values. The last column illustrates the performance gap (PDR) of the DRO model compared to the SP model in terms of average objective value.

Table 7. Comparison of out-of-sample results.

N	Model	Out-of-sample objective value					PDR (%)
		Average value	25th percentile	75th percentile	Minimum value	Range	
10	SP	77,957.23	77,698.46	78,360.47	76,975.03	1,739.43	0.32
	DRO	77,708.47	77,416.28	78,163.75	77,092.10	1,595.76	
50	SP	77,958.08	77,523.75	78,492.72	76,158.71	3,170.22	0.38
	DRO	77,659.16	77,434.71	78,024.08	76,257.96	2,143.49	
100	SP	78,104.97	77,666.40	78,579.66	76,097.63	3,336.62	0.59
	DRO	77,641.16	77,319.37	78,030.34	76,167.59	2,538.38	
1,000	SP	78,102.07	77,639.18	78,596.08	76,038.17	3,484.32	0.50
	DRO	77,713.92	77,351.36	78,093.12	76,051.02	2,826.78	
10,000	SP	78,104.43	77,649.02	78,608.95	75,789.43	3,891.95	0.51
	DRO	77,706.08	77,345.56	78,105.83	75,857.95	3,000.66	

The DRO model distinguishes itself from the SP model by incorporating ambiguity sets, enabling it to effectively manage uncertainties in probability distributions. In contrast, the SP model is only applicable when the probability distribution is fully known. Consequently, the DRO model consistently exhibits smaller extreme deviations compared to the SP model in scenarios where partial information about the probability distribution is unavailable. This suggests that the DRO model can mitigate fluctuations in the optimal objective value, thereby enhancing its robustness. Additionally, the DRO model consistently outperforms the SP model in worst-case scenarios, demonstrating its ability to deliver better results under unfavorable conditions. Notably, these performance improvements come with a reduction of no more than 0.6% in the average objective value.

4.2 A real-world regional railway: Zhengzhou-Xi'an high-speed railway

We now test the proposed models on the Zhengzhou-Xi'an high-speed railway corridor in China. Fig. 5 illustrates the Zhengzhou-Xi'an high-speed railway, which begins at the eastern Zhengzhou hub and passes through key municipalities such as Luoyang, Sanmenxia, and Weinan, before terminating at the western Xi'an hub, covering a total distance of 523 kilometers. The dataset for this example includes train operation and passenger demand records from Zhengzhou East Station to Xi'an North Station. Network-related data is sourced from *China Railway Zhengzhou Bureau Group Co Ltd*. This study involves a fleet of 18 trains operating along this corridor, each with a capacity of 650 passengers. The train stopping plans are shown in Fig. 6.

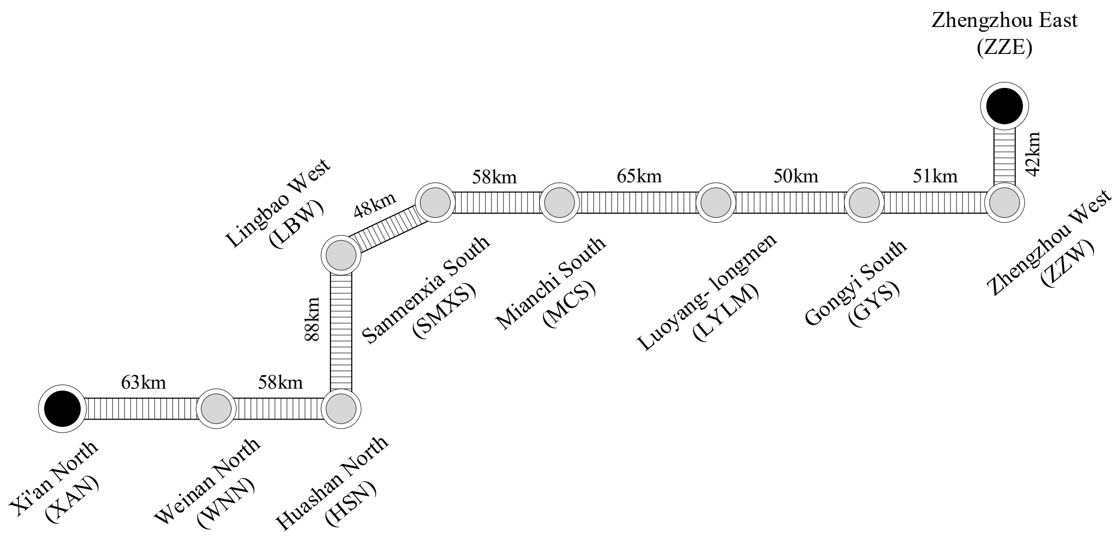


Fig. 5. Zhengzhou–Xi'an high-speed railway corridor (station abbreviations will be used later to ease the presentation).

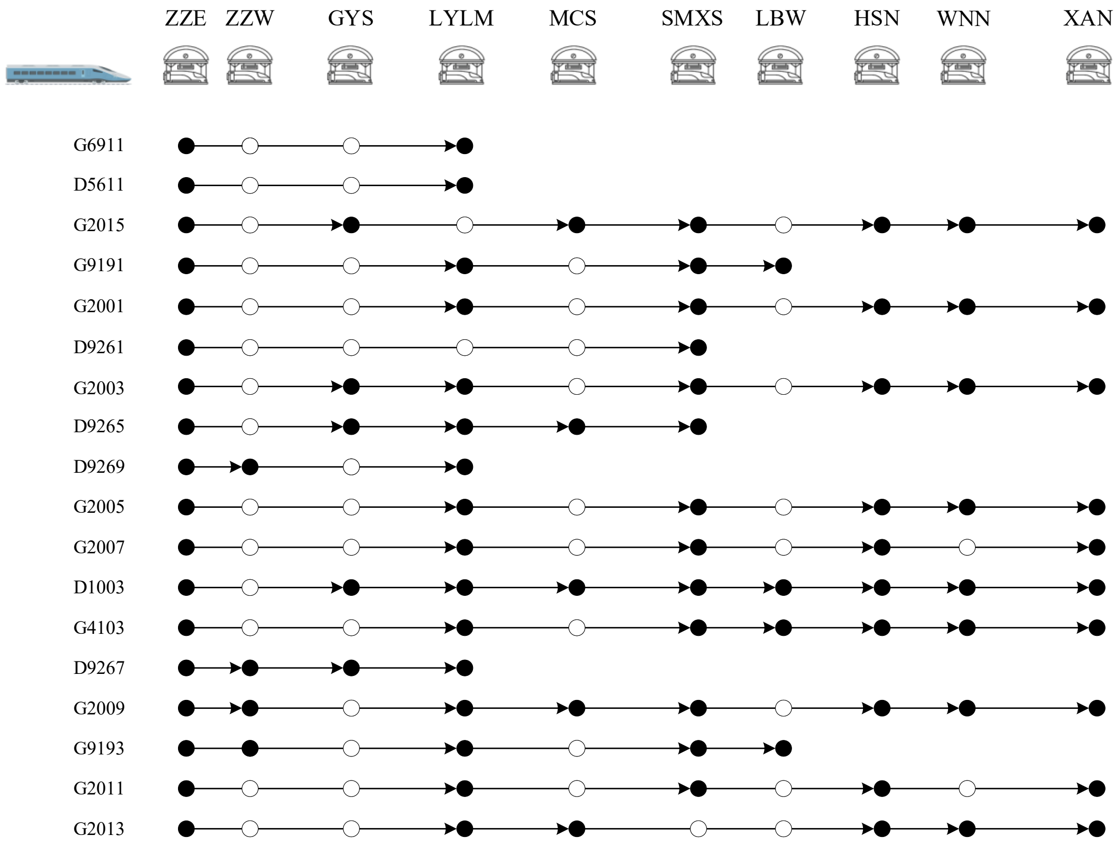


Fig. 6. Train stopping plan of 18 trains running on the Zhengzhou–Xi'an high-speed railway corridor (solid circle: a train will stop at this station; hollow circle: a train will not stop at this station).

4.2.1 Numerical setting

Based on historical ticket purchase data provided by the *China Railway Zhengzhou Bureau Group Co Ltd*, we summarize fare and passenger demand data for the Zhengzhou-Xi'an high-speed railway corridor on May 14, 2016, in Tables 8 and 9, respectively. To assess the robustness of the proposed DRO model, we outline 20 distinct passenger demand scenarios. Notably, we reference the experimental framework established by Qi et al. (2018) to evaluate the efficacy of the DRO model across a range of passenger demand scenarios, considering diverse conservation value settings. Specifically, the passenger demand data within each scenario, denoted as $w \in \Omega$, is generated by randomly perturbing the passenger demand between OD pairs in the reference standard scenario (i.e., Table 9). This perturbation is quantified by an integer parameter Δ_{ij}^{w}, within the range of [4%,16%]. Additionally, the parameter φ governing the DRO model is set as 0.05.

Table 8. Rail service prices (unit: CNY) for each OD pair.

Station	ZZE	ZZW	GYS	LYLM	MCS	SMXS	LBW	HSN	WNN	XAN
ZZE		19.5	42.5	65.5	95.5	121.5	144.5	184.5	211	239
ZZW			20.5	46.5	75.5	102.5	124.5	165.5	192	221
GYS				24.5	54.5	79.5	99.5	139.5	169.5	199
LYLM					29.5	54.5	79.5	119.5	144.5	174.5
MCS						24.5	49.5	89.5	114.5	144.5
SMXS							19.5	64.5	89.5	119.5
LBW								39.5	64.5	94.5
HSN									24.5	54.5
WNN										29.5
XAN										

Table 9. Potential passenger demand (unit: per day) for each OD pair.

Station	ZZE	ZZW	GYS	LYLM	MCS	SMXS	LBW	HSN	WNN	XAN
ZZE		18	42	2,634	236	508	30	62	86	1,324
ZZW			2	34	4	20	8	14	24	20
GYS				28	12	86	54	64	78	104
LYLM					104	744	72	238	126	2,000
MCS						10	6	36	52	86
SMXS							90	46	44	780
LBW								16	6	160
HSN									260	1,934
WNN										4,912
XAN										

4.2.2 Results and analysis

The proposed methods were tested using the GUROBI solver. Fig. 7 visually illustrates the objective function values against different combinations of λ and φ values. It is worth noting that a clear trend emerges that the objective function value decreases as λ and φ increase, consistent with observations from earlier experiments in the small toy example.

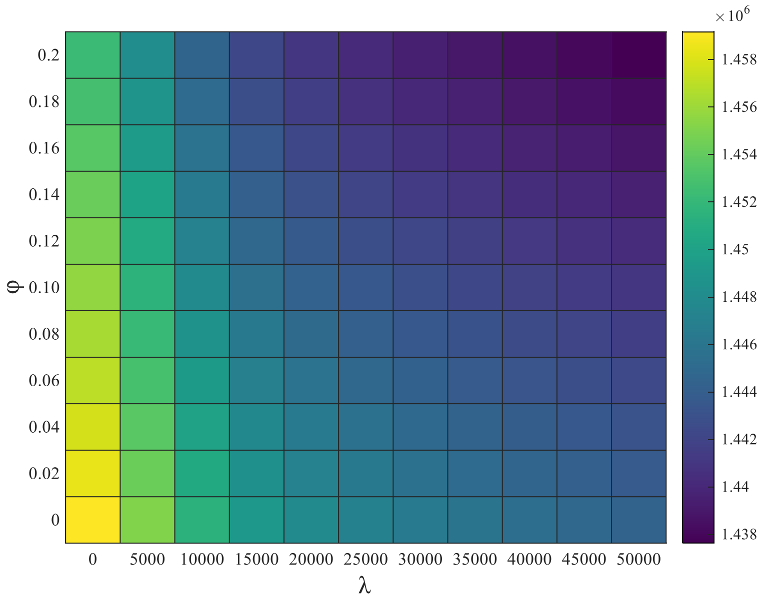


Fig. 7. Variation of objective function values against λ and φ values.

We first investigate the impact of different λ values (with 0, infinity, and 20,000) on the optimal seat allocation scheme, resulting in distinct configurations shown in Figs. 8, 9, and 10. The average computation time for these analyses is 19.16 seconds. Evidently, the DRO model demonstrated highly competitive performance when applied to a real-world railway network. Figs. 8 and 9 illustrate the seat allocation schemes when the objectives are primarily revenue maximization and social equity, respectively, with objective values of 1,449,818.90 and 5,619,669.82). When the model focuses solely on revenue maximization, seat allocation tends to be more uneven. For example, the OD pair ZZE-LYLM, which has high demand, receives an excessive number of seats, while the OD pair HSN-WNN, which also has substantial demand but lower fares, receives no seats. This raises equity concerns and potential passenger dissatisfaction. Conversely, when the focus shifts to maximizing social equity, a more equitable distribution of seats is achieved, accommodating demand across all OD pairs and minimizing disparities in service accessibility. However, this more equitable allocation results in reduced revenue.

The seat allocation scheme that balances both objectives is illustrated in Fig. 10, with an objective value of 1,453,261.98. This scheme considers the financial interests of the railway company while addressing varying demand levels across OD pairs, with a focus on maximizing social equity. Fig. 11 provides a detailed breakdown of the seat allocation for each train on the Zhengzhou-Xi'an high-speed railway. It is evident that the capacities of multiple trains have been distributed to specific OD pairs with substantial demand, rather than concentrating most of the capacity on one or two trains. This approach enhances flexibility, enabling a fair and efficient distribution of capacity across the network.

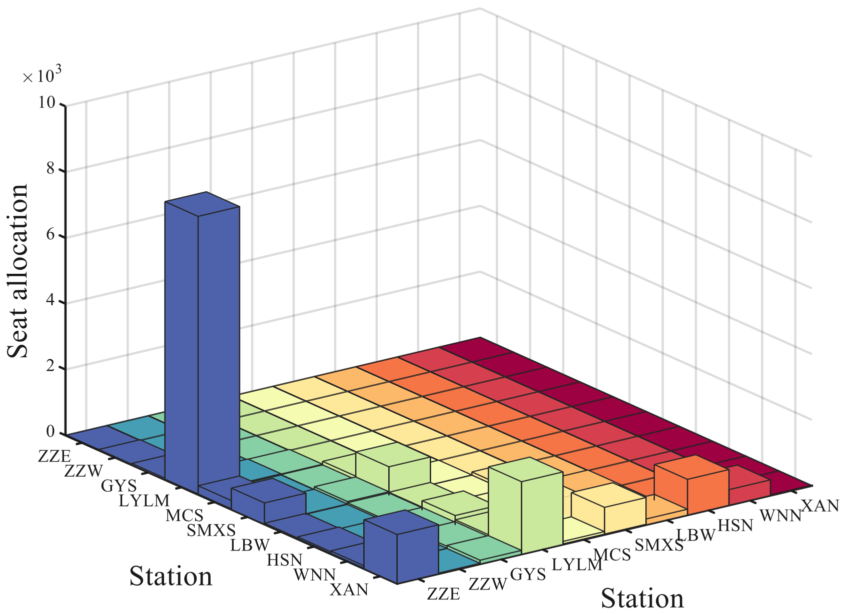


Fig. 8. Seat allocation scheme when λ =0.

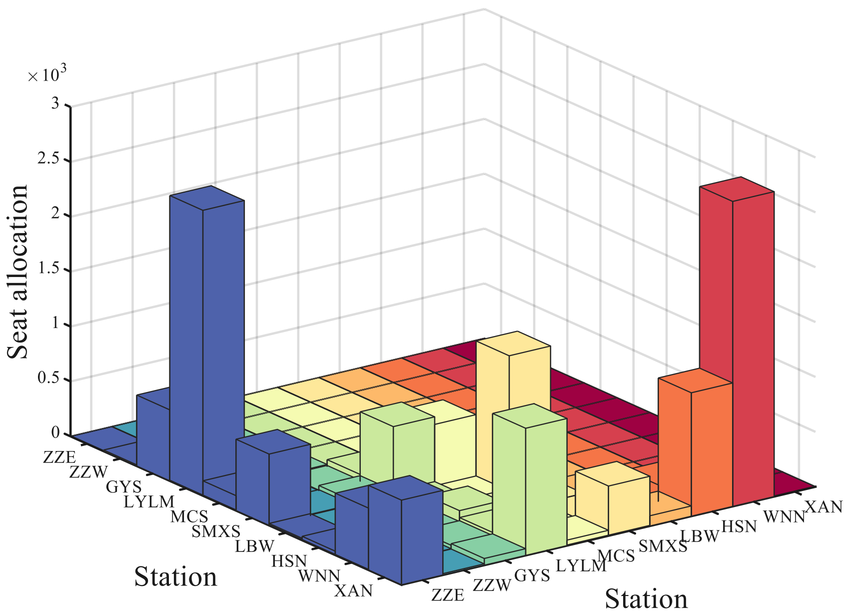


Fig. 9. Seat allocation scheme when λ approaches infinity.

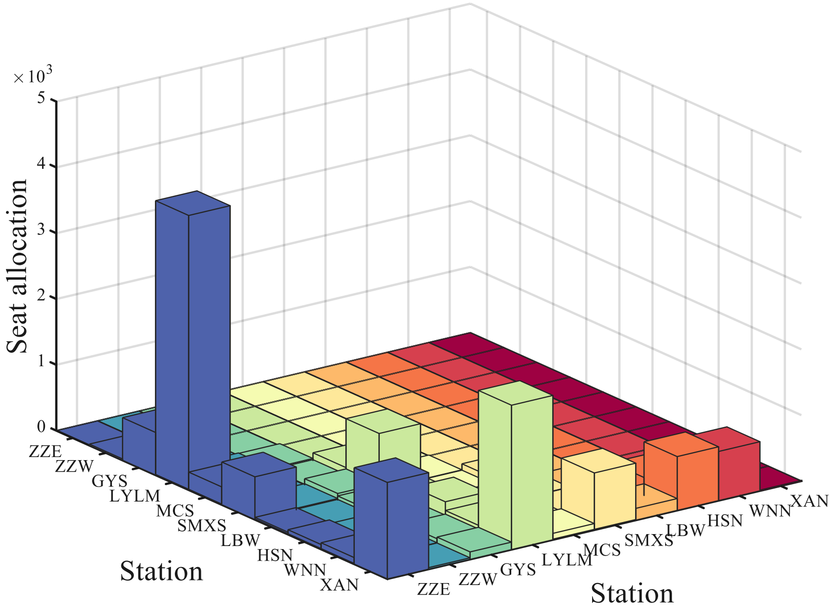


Fig. 10. Seat allocation scheme when λ =20,000.

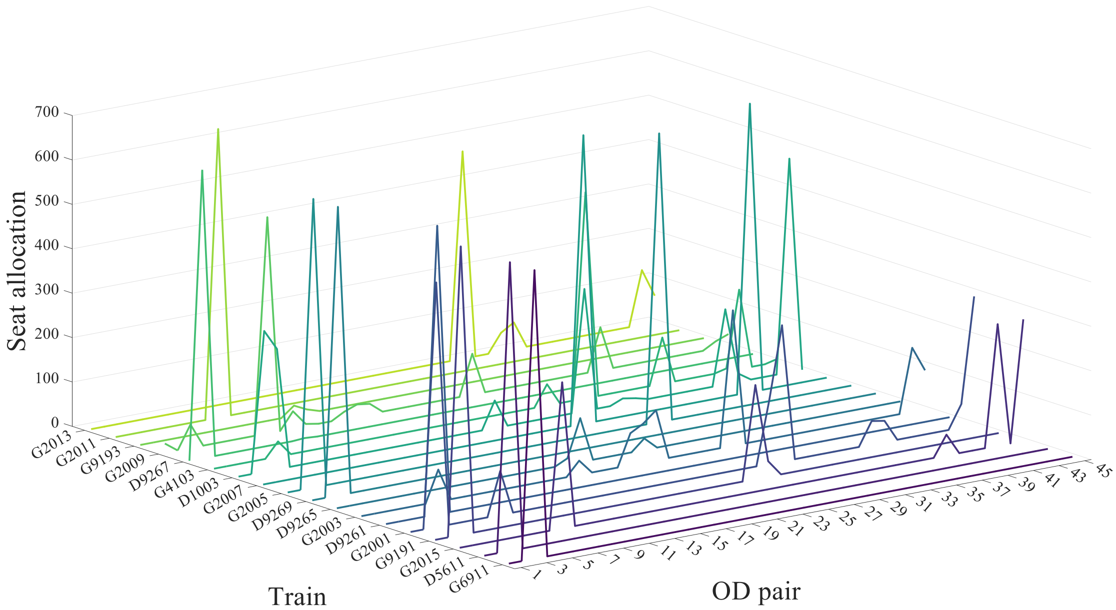


Fig. 11. Seat allocation for 18 trains on the Zhengzhou-Xi'an high-speed railway when λ =20,000 (45 pairs of ODs in order from origin to destination).

Our analysis includes exploring 20 different demand scenarios, where the DRO model consistently achieves the optimal solution in a short computation time. To further test the computational efficiency of the DRO model, we assessed its performance under an increasing number of scenarios. We consider the scenario set Ω with $|\Omega| \in \{20, 200, 2000\}$. In Table 10, we recorded performance metrics associated with the optimal solution under these different scenarios. Even as the number of uncertain demand scenarios expanded from 20 to 2,000, the computation time for obtaining the optimal solution did not exceed 40 minutes. Notably, as the number of scenarios increased, there was a concurrent rise in the uncertainty of the probability distribution, leading to a slight decrease in the objective values for both the SP and DRO models. This observation underscores the capability of the proposed approaches to handle demand uncertainty and limited information effectively.

Table 10. Comparison of the results of the DRO model under different scenario sets.

Scenario	Model	Objective value	CPU time (s)	Gap
20	SP	1,463,148.80	18.63	0.0000%
	DRO	1,453,608.27	19.25	0.0003%
200	SP	1,460,631.00	62.44	0.0000%
	DRO	1,449,488.18	65.02	0.0001%
2,000	SP	1,460,407.69	2,116.71	0.0000%
	DRO	1,449,025.02	2,348.29	0.0000%

5. Extension: Seat allocation model under time-varying passenger demand

In this study, we developed seat allocation models that consider demand uncertainty and resource allocation equity among OD pairs. In addition, passengers' ideal departure times (related to their activities) vary across OD pairs. And the transport service capacity of trains also differs across time periods. For railway travel, passengers typically prefer train tickets that align with their ideal departure times, making it particularly important to incorporate desired departure times into demand modeling. Consequently, we constructed seat allocation models under time-varying demand based on passengers' ideal departure times.

Specifically, following the work of Qi et al. (2021) and Zhang et al. (2021), we divided the entire planning horizon into $|T| \in N$ time intervals based on passengers' ideal or desired departure times. Each time interval is denoted as t, where $t \in T$, and T represents the set of all passengers' desired departure time intervals. The potential passenger demand for OD pair (i,j) in time interval t is represented as d_{ij}^t. The binary parameter τ_i^{kt} indicates whether train k departs from station i during interval t, where $\tau_i^{kt}=1$ if train k departs from station i during t, and $\tau_i^{kt}=0$ otherwise. The decision variable x_{ij}^{kt} represents the number of seats allocated by train k to the OD pair (i,j) in t, and y_{ij}^t denotes the actual demand during each time interval t. Based on these definitions, the mathematical model for seat allocation under time-varying demand can be formulated as follows:

(T_DP):

$$\max Z=\lambda\theta+\sum_{(i,j)\in RS}\sum_{t\in T}h_{ij}y_{ij}^t \tag{29}$$

$$s.t.\ \ x_{ij}^t=\sum_{k\in K_{ij}^t}x_{ij}^{kt}, \forall(i,j)\in RS, t\in T \tag{30}$$

$$e_{ij}^t=\frac{x_{ij}^t}{d_{ij}^t}, \forall(i,j)\in RS, t\in T \tag{31}$$

$$e_{ij}^t\geq\theta, \forall(i,j)\in RS, t\in T \tag{32}$$

$$x_{ij}^{kt}\left(1-\tau_i^{kt}\right)=0, \forall(i,j)\in RS, i\in S, k\in K, t\in T \tag{33}$$

$$\sum_{(i,j)\in RS}\sum_{t\in T}x_{ij}^{kt}\delta_{ij}^{km}\leq c_k, k\in K, m\in M \tag{34}$$

$$y_{ij}^t\leq x_{ij}^t, \forall(i,j)\in RS, t\in T \tag{35}$$

$$y_{ij}^t\leq d_{ij}^t, \forall(i,j)\in RS, t\in T \tag{36}$$

$$x_{ij}^{kt}\in N, \forall(i,j)\in RS, k\in K, t\in T \tag{37}$$

In this model, the objective function aims to jointly maximize passenger equity and ticket revenue. Constraints (30) and (31) evaluate the satisfaction of demand based on seat allocation decisions for each OD pair during different time intervals. Additionally, K_{ij}^t represents the set of all trains departing from station i to j during the time interval t, and x_{ij}^t is the total number of seats allocated during this interval. The parameter θ in constraint (32) denotes the minimum service level across both temporal and spatial dimensions. Constraint (33) ensures that no seats are allocated to trains outside their operating times. Constraint (34) addresses train capacity limitations, while constraints (35) and (36) evaluate the actual passenger demand for each time interval t. Finally, constraint (37) imposes an integer restriction on the decision variables.

Furthermore, we can extend the stochastic programming model discussed in Section 3.3 by incorporating the uncertainty of time-varying demand. We introduce the following SP model, where the demand parameter for each OD pair in different time intervals follows a probability distribution and is denoted as $d_{ij}^{t}(w)$ in scenario w.

(T_SP):

$$\max E_p\left[\lambda\theta(\boldsymbol{w}) + \sum_{(i,j)\in RS}\sum_{t\in T} h_{ij}y_{ij}^{t}(\boldsymbol{w})\right] \tag{38}$$

$$s.t.\ \ x_{ij}^{t} = \sum_{k\in K_{ij}^{t}} x_{ij}^{kt}, \forall(i,j)\in RS, t\in T \tag{39}$$

$$e_{ij}^{t}(w) = \frac{x_{ij}^{t}}{d_{ij}^{t}(w)}, \forall(i,j)\in RS, t\in T, w\in\Omega \tag{40}$$

$$e_{ij}^{t}(w) \geq \theta(w), \forall(i,j)\in RS, t\in T, w\in\Omega \tag{41}$$

$$x_{ij}^{kt}\left(1-\tau_{i}^{kt}\right) = 0, \forall(i,j)\in RS, i\in S, k\in K, t\in T \tag{42}$$

$$\sum_{(i,j)\in RS}\sum_{t\in T} x_{ij}^{kt}\delta_{ij}^{km} \leq c_k, k\in K, m\in M \tag{43}$$

$$y_{ij}^{t}(w) \leq x_{ij}^{t}, \forall(i,j)\in RS, t\in T, w\in\Omega \tag{44}$$

$$y_{ij}^{t}(w) \leq d_{ij}^{t}(w), \forall(i,j)\in RS, t\in T, w\in\Omega \tag{45}$$

$$x_{ij}^{kt} \in N, \forall(i,j)\in RS, k\in K, t\in T \tag{46}$$

Similarly, as discussed in Section 3.4, we address the uncertainty of time-varying demand by introducing the following DRO model, using a box ambiguity set.

(T_ DRO):

$$\max \sum_{w\in\Omega}\left[\lambda\theta(w) + \sum_{(i,j)\in RS}\sum_{t\in T} h_{ij}y_{ij}^{t}(w)\right]p_0(w) + \sum_{w\in\Omega}\Psi(w)\alpha(w) + \sum_{w\in\Omega}\Psi(w)\beta(w) \tag{47}$$

$$s.t.\ \mu - \alpha(w) + \beta(w) = \lambda\theta(w) + \sum_{(i,j)\in RS}\sum_{t\in T} h_{ij}y_{ij}^{t}(w), w\in\Omega \tag{48}$$

$$\alpha(w) \leq 0, w\in\Omega \tag{49}$$

$$\beta(w) \leq 0, w\in\Omega \tag{50}$$

$$\text{Constraints } (39)-(46)$$

In summary, to account for passengers' desired departure times, transport capacity has been converted into train service time dependent capacities. Passenger demand during different time intervals is matched with the corresponding trains before allocating seats.

6. Conclusion

In this paper, we optimized the seat allocation problem in railway systems under demand uncertainty. We developed a mathematical model for seat allocation to improve equity among passengers while ensuring ticketing revenue. Specifically, by optimizing the passenger group with the lowest service level in the system, we enhanced fair travel opportunities for passengers across different OD pairs. To address demand uncertainty, we first established a stochastic programming (SP) model. Subsequently, for cases where only partial information on passenger demand probability distributions is available, we developed a scenario-based box ambiguity set method and constructed a distributionally robust optimization (DRO) model. Additionally, we transformed the DRO model into an equivalent mixed-integer linear programming (MILP), enabling the use of commercial solvers like GUROBI to obtain optimal solutions. The proposed method was tested on a small toy example and the real Zhengzhou-to-Xi'an high-speed railway. The results demonstrate that the seat allocation method proposed in this paper significantly improves passenger travel equity while ensuring the economic benefits of the rail company. Compared to optimization strategies that simply maximize revenue, this method effectively avoids severe resource allocation imbalances, while the fairest allocation potentially leads to lower revenue. Moreover, the DRO model under demand uncertainty minimizes risk and achieves better performance in both equity and revenue. Notably, the PDR upper bound associated with the DRO model does not exceed 1%. Compared to the SP model, these performance improvements reduce the average objective value by no more than 0.6%. Finally, we also considered the time-varying nature of passenger demand and provided model extensions that incorporate desired departure times into the seat allocation model.

This study can be further extended in the following ways. Firstly, future research could integrate train stop planning and scheduling with seat allocation optimization to develop more systematic railway operation management schemes. Secondly, when optimizing equity, aspects in addition to OD pairs might be considered, such as equity among passengers with different desired departure times. Also, alternative metrics for measuring equity can be introduced. Furthermore, since passengers' ideal or desired departure times vary across different OD pairs, future research could incorporate passenger choice behavior to further explore seat allocation optimization under time-varying demand. Finally, to address demand uncertainty, more precise ambiguity sets, such as polyhedral or Wasserstein sets, could be employed based on specific problem contexts to enhance the model's robustness.

References

Abe, I., S. M. Massachusetts Institute of Technology. (2008). *Revenue Management in the railway industry in Japan and Portugal: A stakeholder approach*.

Agra, A., & Rodrigues, F. (2022). Distributionally robust optimization for the berth allocation problem under uncertainty. *Transportation Research Part B: Methodological*, *164*.

Armstrong, A., & Meissner, J. (2010). *Railway Revenue Management: Overview and Models*.

Bertolaccini, K. L. (2013). *Assessing the Equity of Transit Supply Distribution in Metropolitan Areas Using Lorenz Curves and Gini Coefficients*.

Camporeale, R., Caggiani, L., & Ottomanelli, M. (2019). Modeling horizontal and vertical equity in the public transport design problem: A case study. *Transportation Research Part A: Policy and Practice*, *125*, 184–206.

Cavallaro, F., Bruzzone, F., & Nocera, S. (2020). Spatial and social equity implications for High-Speed Railway lines in Northern Italy. *Transportation Research Part A: Policy and Practice*, *135*, 327–340.

Çetiner, D. (2013). Selected Topics in Revenue Management. In D. Çetiner, *Fair Revenue Sharing Mechanisms for Strategic Passenger Airline Alliances* (Vol. 668, pp. 3–30). Springer Berlin Heidelberg.

Chai, J., Zhou, Y., Zhou, X., Wang, S., Zhang, Z. G., & Liu, Z. (2018). Analysis on shock effect of China's high-speed railway on aviation transport. *Transportation Research Part A: Policy and Practice*, *108*, 35–44.

Chen, A., & Subprasom, K. (2007). Analysis of regulation and policy of private toll roads in a build-operate-transfer scheme under demand uncertainty. *Transportation Research Part A: Policy and Practice*, *41*(6), 537–558.

Chen, Z., & Haynes, K. E. (2017). Impact of high-speed rail on regional economic disparity in China. *Journal of Transport Geography*, *65*, 80–91.

Ciancimino, A., Inzerillo, G., Lucidi, S., & Palagi, L. (1999). A Mathematical Programming Approach for the Solution of the Railway Yield Management Problem. *Transportation Science*, *33*(2), 168–181.

Delage, E., & Ye, Y. (2010). Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems. *Operations Research*, *58*(3), 595–612.

Delbosc, A., & Currie, G. (2011). Using Lorenz curves to assess public transport equity. *Journal of Transport Geography*, *19*(6), 1252–1259.

Gini, C. W. (1912). Variability and mutability, contribution to the study of statistical distributions and relations. *Studi Economico-Giuridici Della R. Universita de Cagliari*.

Goh, J., & Sim, M. (2010). Distributionally Robust Optimization and Its Tractable Approximations. *Operations Research*, *58*(4), 902–917.

Han, B., & Ren, S. (2020). Optimizing stop plan and tickets allocation for high-speed railway based on uncertainty theory. *Soft Computing*, *24*(9).

Hu, X., Shi, F., Xu, G., & Qin, J. (2020). Joint optimization of pricing and seat allocation with multistage and discriminatory strategies in high-speed rail networks. *Computers & Industrial Engineering*, *148*, 106690.

Jiang, X., Chen, X., Lei, Z., & Zhang, R. (2015). Dynamic Demand Forecasting and Ticket Assignment for High-Speed Rail Revenue Management in China. *Transportation Research Record: Journal of the Transportation Research Board*, *2475*, 37–45.

Kim, H., & Sultana, S. (2015). The impacts of high-speed rail extensions on accessibility and spatial equity changes in South Korea from 2004 to 2018. *Journal of Transport Geography*, *45*, 48–61.

Li, D., Zhang, T., Dong, X., Yin, Y., & Cao, J. (2019). Trade-off between efficiency and fairness in timetabling on a single urban rail transit line under time-dependent demand condition. *Transportmetrica B Transport Dynamics*, 1–29.

Litman, T. A. (2002). Evaluating transportation equity. *World Transport Policy & Practice*, *8*(2), Article 2.

Liu, S., & Papageorgiou, L.G. (2013). Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. *Omega*, *41*(2), 369–382.

Lu, Y., Yang, L., Yang, K., Gao, Z., Zhou, H., Meng, F., & Qi, J. (2022). A distributionally robust optimization method for passenger flow control strategy and train scheduling on an urban rail transit line. Engineering, 12, 202-220.

Ma, L., Liu, Y., & Liu, Y. (2020). Distributionally robust design for bicycle-sharing closed-loop supply chain network under risk-averse criterion. Journal of Cleaner Production, 246, 118967.

Manaugh, K., Badami, M.G., & El-Geneidy, A. M. (2015). Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in North America. *Transport Policy*, *37*, 167–176.

Martens, K., Golub, A., & Robinson, G. (2012). A justice-theoretic approach to the distribution of transportation benefits: Implications for transportation planning practice in the United States. *Transportation Research Part A: Policy and Practice*, *46*(4), 684–695.

Ning, H., Yao, E., & Zhang, J. (2021). Demand-responsive passenger flow control strategies for metro networks considering service fairness and passengers' behavioural responses. *Transportation Research Part C: Emerging Technologies*, *131*, 103335.

Ogryczak, W. (2009). Inequality measures and equitable locations. *Annals of Operations Research*, *167*(1), 61–86.

Qi, J., Cacchiani, V., & Yang, L. (2018). Robust Train Timetabling and Stop Planning with Uncertain Passenger Demand. *Electronic Notes in Discrete Mathematics*, *69*, 213–220.

Qi, J., Cacchiani, V., Yang, L., Zhang, C., & Di, Z. (2021). An Integer Linear Programming model for integrated train stop planning and timetabling with time-dependent passenger demand. Computers & Operations Research, 136, 105484.

Rahimian, H., & Mehrotra, S. (2022). Distributionally Robust Optimization: A Review. *Open Journal of Mathematical Optimization*, *3*, 1–85.

Ren, X., Chen, Z., Wang, F., Dan, T., Wang, W., Guo, X., & Liu, C. (2020). Impact of high-speed rail on social equity in China: Evidence from a mode choice survey. *Transportation Research Part A: Policy and Practice*, *138*, 422–441.

Shang, P., Li, R., Liu, Z., Yang, L., & Wang, Y. (2018). Equity-oriented skip-stopping schedule optimization in an oversaturated urban rail transit network. *Transportation Research Part C: Emerging Technologies*, *89*, 321–343.

Shao, J., Xu, Y., Sun, L., Kong, D., & Lu, H. (2022). Equity-oriented integrated optimization of train timetable and stop plans for suburban railways system.

Computers & Industrial Engineering, 173, 108721.

Sun, H., Gao, Z., Szeto, W. Y., Long, J., & Zhao, F. (2014). A Distributionally Robust Joint Chance Constrained Optimization Model for the Dynamic Network Design Problem under Demand Uncertainty. *Networks and Spatial Economics*, *14*(3–4), 409–433.

Wang, W., Yang, K., Yang, L., & Gao, Z. (2021). Two-stage distributionally robust programming based on worst-case mean-CVaR criterion and application to disaster relief management. *Transportation Research Part E: Logistics and Transportation Review*, *149*, 102332.

Wang, X., Wang, H., & Zhang, X. (2016). Stochastic seat allocation models for passenger rail transportation under customer choice. *Transportation Research Part E: Logistics and Transportation Review*, *96*, 95–112.

Xia, Y., Wang, H., Xu, Y., Zhou, Y., & Wei, Y. (2023). Integrated optimization of train stop planning and ticket pricing with equity concerns of passengers for high-speed railways. *Measurement and Control*, *56*(3–4), 792–812.

Xu, G., Liu, Y., Gao, Y., & Liu, W. (2023). Integrated optimization of train stopping plan and seat allocation scheme for railway systems under equilibrium travel choice and elastic demand. *Transportation Research Part E: Logistics and Transportation Review*, *177*, 103231.

Xu, G., Zhong, L., Hu, X., & Liu, W. (2022a). Optimal pricing and seat allocation schemes in passenger railway systems. *Transportation Research Part E: Logistics and Transportation Review*, *157*, 102580.

Xu, G., Zhong, L., Wu, R., Hu, X., & Guo, J. (2022b). Optimize train capacity allocation for the high-speed railway mixed transportation of passenger and freight. *Computers & Industrial Engineering*, *174*, 108788.

Yan, Z., Li, X., Zhang, Q., & Han, B. (2020). Seat allocation model for high-speed railway passenger transportation based on flexible train composition. *Computers & Industrial Engineering*, *142*, 106383.

Yin, Y., Li, D., Han, Z., Dong, X., & Liu, H. (2022). Maximizing network utility while considering proportional fairness for rail transit systems: Jointly optimizing passenger allocation and vehicle schedules. *Transportation Research Part C: Emerging Technologies*, *143*, 103812.

You, P. S. (2008). An efficient computational approach for railway booking problems. *European Journal of Operational Research*, *185*(2), 811–824.

Zhan, S., Wong, S. C., & Lo, S. M. (2020). Social equity-based timetabling and ticket pricing for high-speed railways. *Transportation Research Part A: Policy and Practice*, *137*, 165–186.

Zhang, F., Yang, Z., Jiao, J., Liu, W., & Wu, W. (2020). The effects of high-speed rail development on regional equity in China. *Transportation Research Part A: Policy and Practice*, 141, 180-202.

Zhang, C., Qi, J., Gao, Y., Yang, L., Gao, Z., & Meng, F. (2021). Integrated optimization of line planning and train timetabling in railway corridors with passengers' expected departure time interval. Computers & Industrial Engineering, 162, 107680.

Zhao, S., Yang, H., & Wu, Y. (2021). An integrated approach of train scheduling and rolling stock circulation with skip-stopping pattern for urban rail transit lines. *Transportation Research Part C: Emerging Technologies*, *128*, 103170.

Zhao, X., Shan, X., & Wu, J. (2023). The Impact of Seat Resource Fragmentation on Railway Network Revenue Management. *Networks and Spatial Economics*, *23*(1), 135–177.

Zhao, X., & Zhao, P. (2019). A seat assignment model for high-speed railway ticket booking system with customer preference consideration. *Transportmetrica A: Transport Science*, *15*(2), 776–806.

Zheng, L., Liu, P., Huang, H., Ran, B., & He, Z. (2022). Time-of-day pricing for toll roads under traffic demand uncertainties: A distributionally robust simulation-based optimization method. *Transportation Research Part C: Emerging Technologies*, *144*, 103894.

Zhou, J., Yang, L., & Li, L. (2018). The implications of high-speed rail for Chinese cities: Connectivity and accessibility. *Transportation Research Part A: Policy and Practice*, 116, 308-326.