https://doi.org/10.1016/j.tre.2021.102580

Optimal pricing and seat allocation schemes in passenger railway systems

Guangming Xu[a], Linhuan Zhong[a], Xinlei Hu[a], Wei Liu[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.: Tel.: +852 3400 2386; Email: wei.w.liu@polyu.edu.hk*

Abstract

This paper examines optimal pricing and seat allocation schemes in passenger railway systems, where ticket pricing and seat allocation (or capacity allocation) are both Origin-Destination specific. We consider that the demand is sensitive to the ticket price, and a non-concave and non-linear mixed integer optimization model is then formulated for the ticket pricing and seat allocation problem to maximize the railway ticket revenue. To find the optimal solution of the ticket revenue maximization problem effectively, the proposed non-concave and non-linear model is reformulated such that the objective function and constraints are linear with respect to the decision variables or the logarithms of the decision variables. The linearized model is then further relaxed as a mixed-integer programing problem (MILP). Based on the above linearization and relaxation techniques, a globally optimal solution can be obtained by iteratively solving the relaxed MILP and adopting the range reduction scheme. Two numerical examples are presented for illustration.

Keywords: Passenger railway system; Pricing; Seat allocation; Global optimization

1. Introduction

Railway is a high-capacity travel mode for passengers with medium-to-long distance journeys in many countries, e.g., in China with a very dense population. It is reported that China has a railway network with over 1,210,000 km track at the end of 2015, and passenger traffic volume is 3,004.7 billion passenger kilometer in 2015.[1] Particularly, China has the world's longest high-speed railway network with over 19,000 km of track in service as of January 2016.[2] These facts highlight that railway systems have been playing a crucial role in passenger transportation nowadays and support social and economic activities (Jiao et al., 2020; Yang et al., 2020; Zhang et al., 2020).

Given the limited resources in railway networks and costly operation (especially high-speed trains), it is of significant interest for a railway operator to maximize the ticket revenue by optimizing the spatiotemporal resource allocation, such as line planning, train scheduling, pricing, and seat allocation.[3] The revenue management (RM) is a long-standing problem for many transportation sectors. For comprehensive reviews of RM in transportation, one may refer to Mcgill and Van Ryzin (1999), Talluri and van Ryzin (2004). Revenue management was initially introduced after the deregulation of the airline industry of the United States in 1970s (Ciancimino et al., 1999). In the past several decades, many studies examined the airline revenue management problem (Belobaba, 1987; Subramanian et al., 1999; Tong and Topaloglu, 2014). For instance, recently, Terciyanlı and Avșar (2019) proposed the alternative risk-averse approaches for seat inventory control in airline networks. Fard et al. (2019) developed a dynamic programming approach for solving the seat overbooking problem of airlines. Much less attention has been paid to RM in passenger railway systems. For the railway system, an itinerary is usually built up by several legs, where each leg is identified by two consecutive stations traversed by a certain train. Thus, the RM of railway system may be regarded as a multi-leg single-fare problem when compared to the airline sector (Ciancimino et al., 1999).

In recent years, there is a growing literature on railway revenue management. Several empirical studies have demonstrated that the RM plays an important role in railway transportation industry (Abe, 2007; Armstrong and Meissner, 2010; Wang et al., 2012). In particular, Armstrong and Meissner (2010) provided an overview and some detailed discussions on railway RM. For railway RM, ticket pricing and seat allocation problems have been studied, while these two aspects are often treated separately (Hetrakul and

[1] Transportation development in China. <http://www.gov.cn/xinwen /2016-12/29/content_5154095.htm>

[2] Chinese high speed network to double in the latest master plan. <http://www.railwaygazette.com/news/infrastructure/single-view/view/chinese-high-speed-network-to-double-in-latest-master-plan.html.>

[3] Ticket prices fluctuate, allowing the railway to dock flexibly with the market. <http://www.peoplerail.com/rail/show-466-418896-1.html.>

Cirillo, 2014; Qiu and Lee, 2019). The seat allocation (or capacity allocation) problem is to determine the number of seats of a train to be allocated to each Origin-Destination (OD) pair, i.e., determine the supplies for different markets, where one OD pair can be regarded as one market. The ticket pricing strategy is to manage or regulate the interaction between demand and supply. The seat allocation and pricing problems are intercorrelated and complementary to each other. This has already been recognized in RM studies (e.g., McGill and van Ryzin, 1999).

However, for the seat allocation problem in passenger railway systems, the ticket prices are often assumed to be fixed and only the capacity allocation is optimized. In this context, Ciancimino et al. (1999) studied a multi-leg seat inventory problem and developed both deterministic linear programming model and probabilistic nonlinear programming model for railway yield management. Following the deterministic linear programming model in Ciancimino et al. (1999), Wang et al. (2012) further proposed a mixed-integer linear optimization model for seat allocation and train dispatching in a single line high-speed railway system. Jiang et al. (2015) developed seat allocation models with dynamic adjustment based on short-term demand forecasting. More recently, Luo et al. (2016) developed an integer linear programming model for multi-train seat allocation problem with different stopping patterns. Yuan et al. (2018) introduced a bid price approach for seat inventory control problem. Yan et al. (2020) further optimized the seat allocation based on the flexible train composition. Following the probabilistic nonlinear programming model in Ciancimino et al. (1999), You (2008) incorporated the pricing discount and developed an efficient heuristic approach to determine the ticket booking limits of the railway seat inventory control system. Wang et al. (2016) formulated the seat allocation problem with single-stage and multi-stage decisions as two stochastic programming models that incorporate the passenger choice behaviors.

The joint optimization of pricing and seat allocation in railway networks has not received sufficient attention. As far as the authors know, Ongprasert (2006) was among the earliest to examine the seat allocation problem in relation to railway RM and discussed the combination of discounted ticket fare and seat allocation. Hetrakul and Cirillo (2014) jointly optimized the pricing and seat allocation for railway systems, using multinomial logit model and latent class to capture the ticket purchasing times of passengers. Hu et al. (2020) established a nonlinear programming model for joint optimization of pricing and seat allocation in high-speed rail system, where it followed the Davidon-Fletcher-Powell method to design a gradient-based algorithm. For the joint optimization problem of ticket pricing and seat allocation, the mathematical models are often non-concave and non-linear. None has proposed a global optimization solution procedure yet. While joint optimization of railway pricing and seat allocation is studied to a very limited extent, there are many studies on railway network passenger assignment (Xu et al., 2018a,b; Xu et al., 2021), and urban transit or rail service network design (e.g., Li et al., 2012; Li et al., 2018; Canca et al., 2019; Zhou et al., 2021), and passenger-freight integrated urban rail service design (e.g., Li et al., 2021).

The joint optimization problem of pricing and seat allocation has received more attention in the airline market. For instance, Kuyumcu and Garcia-Diaz (2000) considered joint pricing and seat allocation problem for airline networks and formulated the 0-1 integer programming models to optimize the decision (accept or reject) for each passenger and the price structure for each origin-destination pair. Bertsimas and de Boer (2002) studied a joint pricing and seat allocation problem in airline revenue management, where the optimization problem is not always concave but can be concave for certain types of demand distributions and the iterative non-linear optimization algorithm adopted does not always guarantee the solution optimality. Cote et al. (2003) built a bi-level programming model to jointly solving the pricing and seat allocation problem with fixed demand, where the upper-level deals with the seat allocation and the lower-level deals with the train fares. Chew et al. (2009) developed a discrete time dynamic programming model to jointly optimizing pricing and seat allocation in order to maximize the expected revenue for a single product with a predetermined lifetime, where the stochastic demand has a mean as a linear function of price and the authors used an enumeration method to find the optimal solution. More recently, Cizaire (2011) proposed both deterministic and stochastic models to solve the joint optimization problems of airline fare and seat allocation for two products and two timeframes.

In airline pricing and seat allocation problems, an airplane usually only serves passengers with one leg, i.e., one origin and one destination (e.g., Cote et al., 2003; Chew et al., 2009). For railway systems, especially high-speed railways, there can be many trains serving a large number of stations (i.e., multiple OD pairs) and different trains can have different stopping patterns (i.e., each train might serve a different set of stations). Therefore, in the seat allocation problem for railway systems, one need to accommodate the train-specific and OD-specific capacity constraints resulting from the seat allocation scheme. Moreover, the pricing control and seat allocation (quantity control) jointly govern the demand and thus the railway revenue, where the seat allocation scheme constrains the demand for each OD pair and the pricing further manages the demand. More critically, in the joint optimization problem of pricing and seat allocation, we have both pricing variables and train-and-OD-specific seat allocation variables that define the capacity constraints. To solve such a problem for railway systems is challenging, given that the problem size is larger and the problem structure is more complicated when compared to aviation system. This indeed motivates the current study to propose an effective iterative algorithm to obtain the optimal solution for pricing and seat allocation, where the joint optimization problem of pricing and seat allocation can be formulated as a non-concave and non-linear mixed integer programming model.

In Table 1, we present a summary of studies regarding the joint optimization of train service pricing and seat allocation that have been reviewed in this section and highlight the contribution of this paper against the existing literature. In particular, this paper develops a joint optimization modelling framework of pricing and seat allocation for railway systems. We maximize the ticket revenue of the railway system considering elastic demand and multiple trains with multiple stopping patterns.[4] The demand is assumed as an exponential function of the train service price. A non-concave and non-linear mixed integer model is developed for the train service pricing and seat allocation optimization problem. In order to find the optimal solution of the pricing and seat allocation problem, the proposed non-concave and non-linear model is reformulated and relaxed as a mixed-integer programing problem (MILP). An optimal solution is then obtained by iteratively solving the relaxed MILP and adopting a range reduction scheme. The method is tested and illustrated with two numerical examples: a toy network example and a real-world network of Ninghang railway. Its advantages are also shown through comparison with the solvers embedded in MATLAB. This study contributes to the literature as follows. (i) This study is the first to propose the linear relaxation technique and interval reduction scheme to solve the joint optimization problem of train service pricing and seat allocation in railway system under elastic demand, which is formulated as a non-concave and non-linear mixed integer optimization model. The effectiveness and applicability of the proposed method is demonstrated with both toy network and real-world network examples. (ii) This study illustrates how to obtain the upper and lower bounds of the railway pricing and seat allocation optimization problem. This adds further examples to the literature on how these techniques can be utilized to solve railway system optimization problems.

[4] Railway systems might have different objectives (e.g., social welfare maximization or revenue maximization). This paper focuses on the case of revenue maximization. Indeed, railway revenue/yield management has been considered by many studies (Ciancimino et al., 1999; Hetrakul et al., 2014; Wang et al., 2016; Canca et al., 2019; Hu et al., 2020). Some studies considered both the railway revenue and social equity (Zhan et al., 2020). In China, the central government has reformed the national railway organization and established the China National Railway Group Co., Ltd to improve its economic efficiency and market competitiveness. In particular, the central government has delegated the pricing power of high-speed railway service to railway enterprises, enabling them to price freely within a certain range given by the government in response to factors such as market supply, demand and competition with other modes. This is the case considered in the current study.

Table 1. A summary of studies on joint optimization problems of pricing and seat allocation in aviation or railway systems

Authors	Travel mode	Demand model	Multiple legs	Multiple trains	Multiple stop patterns	Simultaneous optimization	Model	Solution algorithm	Global optimization algorithm
Weatherford (1997)	Airline	Linear function of price with cross elasticities	×	--	--	√	Non-concave non-linear programming model	Spreadsheet-based nonlinear optimizer/Fletcher-Reeves-Polak-Ribiere algorithm	×
Kuyumcu and Garcia-Diaz (2000)	Airline	Normally distributed demand	√	--	--	×	0-1 integer non-linear programming model	Software: CPLEX	×
Bertsimas and de Boer (2002)	Airline	Function of Price	√	--	--	×	Non-concave non-linear model	An iterative non-linear optimization algorithm	×
Cote et al. (2003)	Airline	Fixed demand	√	--	--	×	Bi-level programming model	Heuristics algorithm	×
Ongprasert (2006)	Railway	Nested logit	√	×	×	×	Linear programming model	Qprog program in GAUSS software	√
Chew et al. (2008)	Airline	Linear function of price	×	--	--	×	Dynamic programming model	An enumeration algorithm/heuristics algorithm	×
Cizaire (2011)	Airline	Function of price	×	--	--	√	Non-concave non-linear model	An interior point algorithm in / Powell's algorithm/ heuristic algorithm	×
Hetrakul and Cirillo (2014)	Railway	Log-linear demand functions	√	√	×	√	Non-concave non-linear model	Software: LINGO	×
Hu et al. (2020)	Railway	Exponential demand function	√	√	√	√	Non-concave non-linear model	Davidon–Fletcher–Powell method	×
This study	Railway	Exponential demand function	√	√	√	√	Non-concave non-linear mixed integer programming model	A linearization-based optimization algorithm	√

The rest of the paper is organized as follows. Section 2 summarizes the notations and the major model assumptions. Section 3 formulates the joint optimization model of train service pricing and seat allocation. Section 4 designs the solution algorithm and proves the optimality of the algorithm. Two numerical examples (on a small toy network and a real-world regional network, respectively) are provided in Section 5. Finally, Section 6 concludes the paper.

2. Basic considerations

In this section, we firstly list the notations and then summarize the major assumptions for the joint optimization problem of railway service pricing and seat allocation.

2.1 Notations

We list the major notations in the following.

Sets and indices	
w	OD pair
W	set of OD pairs (with $w \in W$)
K	total number of trains
K_w	set of trains serving the OD pair w
k	index of a train (where $k = 1,2, \cdots K$)
W_k	set of OD pairs served by train k
L	total number of (rail track) sections
l	a section (where $l = 1,2, \cdots L$), which is the rail track link between stations
Parameters	
$\underline{p}_w$	the lower bound of rail service price for OD pair w
$\bar{p}_w$	the upper bound of rail service price for OD pair w
δ_{wl}^k	a binary variable, which equals one if OD pair w served by train k covers the rail section l and 0 otherwise.
δ_w^k	a binary variable, which equals one if OD pair w is served by train k and 0 otherwise
$\bar{Q}_w$	the potential demand for OD pair w
$Q_w(p_w)$	elastic demand function with respect to ticket price p_w for OD pair w
η_w	a parameter in the demand function
c_k	the capacity for train k
Variables	
p_w	train service price for OD pair w
b_w^k	number of seats assigned to OD pair w in train k
x_w	number of passengers purchasing the tickets between OD pair w

2.2 Assumptions

We now summarize the main assumptions for the joint optimization problem of railway service pricing and seat allocation and briefly discuss them below.

A1. (Elastic demand) The demand for a given OD pair decreases with the rail service price.
A2. (OD specific pricing) The rail service prices are OD pair specific, but not train specific.
A3. Each train has only one seat class.
A4. Ticket overbooking is not considered.

The travel demand is governed by many different factors, such as the service price, service quality (e.g., travel time, service reliability, comfort), which can be modeled as a function of the generalized travel cost (including both monetary and non-monetary costs). This paper considers that other factors such as service quality are given, and the demand is then only a function of the service price and decreases with the price (Assumption A1). This treatment is similar to some existing studies, e.g., Hu et al. (2020) and Yan et al. (2020).

The OD-specific pricing in Assumption A2 reflects the current practice in China, while it is noteworthy that the proposed model can be readily modified to incorporate train-specific pricing.

While Assumption A3 assumes a single seat class for a train, for the case with multiple seat classes in one train, the proposed model is readily applicable where a train with multiple seat-classes can be regarded as multiple trains and each train has one specific seat class.

Assuming no overbooking in Assumption A4 reflects the current railway system practice in China. In railway passenger transportation in China, ticket purchase and seat selection are completed at the same time, and the ticket purchase is train-and-seat-specific. It follows that there will be no overbooking. However, a future study may consider railway system overbooking in order to accommodate demand uncertainty and ticket cancelation issues, which are similar to the aviation market.

3. Problem formulation

In this section, we first illustrate the railway service pricing and seat allocation problem and then present the formulations for jointly optimizing service pricing and seat allocation in the railway system.

3.1 Problem description

This study concerns the joint optimization of pricing and seat allocation in order to maximize railway revenue, where the train scheduling is given. The demand is sensitive to the pricing. The railway operator can adjust the number of seats in a train that is allocated to any OD pair served by this train and the corresponding price (or fare) for each OD pair in order to maximize the revenue of the railway system.

We use a simple example below to illustrate the necessity of jointly optimizing pricing and seat allocation in order to maximize revenue. As shown in Fig. 1, we consider a network with one train traveling from A to B and then to C with one cabin class, where the train has 20 seats in total (train capacity). There are three OD pairs: AB, BC and AC. We can see that AC and AB share the same leg AB, while AC and BC share the same leg BC. The demand for a specific OD pair decreases with the service price for this OD pair. In particular, for OD pairs AB, BC and AC, the demand functions are 25−2×price, 20−1×price, 24−2×price, respectively.

The railway operator can adjust the number of seats assigned to each OD pair and the corresponding service price. For illustration, we consider three solutions, as summarized in Table 2. It is evident that inappropriate seat allocation and pricing will yield inefficiency in revenue and the railway operator should choose the seat allocation scheme and pricing in Case 2 in order to generate the maximal revenue, i.e., 246 CNY, among the three cases.

While joint optimization of pricing and seat allocation can help increase system revenue, to solve such a problem in a network with complicated stopping patterns of multiple trains is challenging, especially when pricing and seat allocation are intercorrelated. This paper aims to develop a global optimization method for pricing and seat allocation schemes in passenger railway systems.

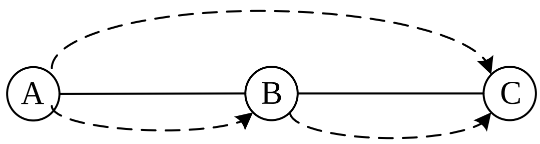


Fig. 1. The network with three stations: an illustrative example

Table 2. Three cases of service pricing and seat allocation schemes

OD	Service pricing and seat allocation schemes								
	Case 1			Case 2			Case 3		
	Ticket fare	Seat allocation	Demand	Ticket fare	Seat allocation	Demand	Ticket fare	Seat allocation	Demand
AB	5	10	15	7	11	11	8	10	9
AC	10	10	10	11	9	9	15	10	5
BC	6	10	12	7	11	10	10	10	4

Profit (CNY)	210	246	187

3.2 Model formulation

Consider a railway network with many lines and trains. We define the rail track link between two stations as the "section" and denote the set of sections by L. Denote the set of trains by K. For each section $l \in L$, there can be multiple trains running on it. Moreover, denote the set of OD pairs by W. For each OD pair $w \in W$, there can be multiple trains serving it.

For each OD pair $w \in W$, denote the total potential demand as $\bar{Q}_w > 0$. The demand for a given OD pair (Shi et al., 2014; Yan et al., 2020; Hu et al., 2020) decreases with the rail service price for this OD pair, which is given as follows:

$$Q_w(p_w) = \bar{Q}_w \cdot \exp(-\eta_w \cdot p_w), w \in W \tag{1}$$

where p_w is the rail service price and η_w is a coefficient for OD pair $w \in W$. Note that the demand function should be appropriately calibrated, where many existing studies provided approaches to solve the demand calibration problem based on real-world data (Sancho, 2009; Flötteröd et al., 2011; Osorio, 2019). In this paper, the coefficient η_w should be calibrated, which is often related to alternative travel modes (e.g., airline or highway) and passengers' socioeconomic attributes. As discussed in Section 2.2, we consider that other factors that affect travel demand such as service quality are given and fixed, and the demand is then only a function of the service price. Moreover, we adopt the exponential function for modeling the demand.[5] Furthermore, we consider that the rail service price p_w for OD pair $w \in W$ is bounded (which is subject to local policies, e.g., the fares in China's high-speed rail system are subject to related regulations[6]), i.e.,

$$\underline{p}_w \leq p_w \leq \bar{p}_w, \qquad \forall w \in W \tag{2}$$

where $\underline{p}_w$ and $\bar{p}_w$ are the lower and upper bounds and $\bar{p}_w \geq \underline{p}_w \geq 0$.

We now further discuss the seat allocation scheme of trains. For train $k \in K$, denote its seat capacity by c_k. We further define b_w^k as the number of seats assigned to OD pair

---

[5] This study is not restricted to the proposed exponential function form for the travel demand. The main feature that should be respected is that the demand should decrease with respect to ticket price. In addition, if the demand function is linear, convex or concave with respect to the ticket price, the proposed bounding techniques and solution approach in Section 4 will still be applicable, while using the logarithmic functions in Section 3.3 might not be necessary and relevant anymore.

[6] National Railway Administration of the People's Republic of China.
< http://www.nra.gov.cn/jgzf/flfg/gfxwj/zt/other/201602/t20160222_21192.shtml >

w in train k. Then the seat capacity constraints can be described as

$$\sum_{w\in W_k} \delta_{wl}^k \cdot b_w^k \le c_k, \qquad \forall l \in L, k \in K \tag{3}$$

$$b_w^k \ge 0, \qquad \forall k \in K, w \in W \tag{4}$$

$$b_w^k \text{ integer}, \qquad \forall k \in K, w \in W \tag{5}$$

where W_k is the set of OD pairs served by train k and δ_{wl}^k is a binary variable, which equals one if OD pair w served by train k covers section l and zero otherwise. It should be noted that W_k is determined based on the stopping patterns of trains. The total number of seats/tickets assigned to OD pair w, i.e., b_w, can be given as follows:

$$b_w = \sum_{k\in K_w} b_w^k, \qquad \forall w \in W \tag{6}$$

where K_w is the set of trains serving the OD pair w. It is noteworthy that the above OD pair specific and train specific seat allocation is an important feature of railway systems when compared to aviation systems.

For each OD pair $w \in W$, the demand $Q_w(p_w)$ given in Eq. (1) is further constrained by b_w in Eq. (6). The realized demand between OD pair w, i.e., x_w, can be given as follows:

$$x_w = \min\{\lfloor Q_w(p_w) \rfloor, b_w\} \tag{7}$$

or

$$\begin{cases} x_w = b_w, & if\ Q_w(p_w) - b_w > 0 \\ x_w = \lfloor Q_w(p_w) \rfloor, & if\ Q_w(p_w) - b_w \le 0 \end{cases} \tag{8}$$

where $\lfloor x \rfloor$ is equal to the largest integer that is no greater than x. The ticket revenue from OD pair w then can be calculated as $(p_w \cdot x_w)$.

We are now ready to formulate the rail service pricing and seat allocation problem. The objective is to maximize the total ticket revenue. The "Mathematical Programming model of Pricing and Seat allocation optimization" (MPPS) can be written as follows:

$$\max Z = \sum_{w\in W} p_w \cdot x_w \tag{9}$$

s.t. $\quad$ Eqs. $(1) - (7)$

where the objective function in Eq. (9) is the ticket revenue from all OD pairs. Similar revenue function has been adopted in many existing studies (Ciancimino et al., 1999; Hetrakul et al., 2014; Wang et al., 2016; Canca et al., 2019; Hu et al., 2020).

Furthermore, the constraint in Eq. (7) can be replaced by the following:

$$-L \cdot \sigma_w \le Q_w(p_w) - b_w \le L \cdot (1 - \sigma_w) \tag{10}$$
$$-L \cdot (1 - \sigma_w) \le x_w - \lfloor Q_w(p_w) \rfloor \le L \cdot (1 - \sigma_w) \tag{11}$$
$$-L \cdot \sigma_w \le x_w - b_w \le L \cdot \sigma_w \tag{12}$$

where L is a large positive constant and σ_w is a binary variable indicating whether $Q_w(p_w)$ is greater than b_w, i.e., if $\sigma_w = 1$, then $Q_w(p_w) \le b_w$ and $x_w = \lfloor Q_w(p_w) \rfloor$; otherwise, $Q_w(p_w) \ge b_w$ and $x_w = b_w$. Alternatively, we can simply add the following three constraints to replace Eq. (7)

$$x_w \le Q_w(p_w) \tag{13}$$
$$x_w \le b_w \tag{14}$$
$$x_w \text{ integer} \tag{15}$$

We can either adopt Eqs. (10)-(12) or Eqs. (13)-(15) to replace the original constraint in Eq. (7). In this paper, we adopt Eqs. (13)-(15) since less variables and inequalities are involved. Therefore, the MPPS can be rewritten as follows:

$$\max Z = \sum_{w \in W} p_w \cdot x_w \tag{16}$$

s.t. Eqs. $(1) - (6)$ and $(13) - (15)$.

As the objective function is quadratic and the constraint in Eq. (1) is non-linear, the above MPPS model is a non-concave and non-linear model, where the non-concavity is further illustrated in Appendix A.

3.3 Model reformulation with logarithmic functions

We now introduce the linearization techniques adopted to facilitate solving the MPPS model. In particular, this subsection will first reformulate the MPPS model such that the objective function and many of the constraints will be linear with respect to the decision variables, and only a few constraints will be nonlinear (but is linear with respect to the logarithms of the decision variables). It can be seen below that, given the exponential demand function, using logarithm is a simple and straightforward way to reduce the nonlinearity involved in the model formulation, especially in the objective function (Section 4 will further discuss how to deal with the remaining nonlinearity in the constraints of the proposed model).

For the non-linear objective function in Eq. (16), we define $y_w = p_w \cdot x_w$, then the objective function in Eq. (16) can be represented by

$$Z = \sum_{w \in W} y_w \tag{17}$$

$$y_w = p_w \cdot x_w, \qquad \forall w \in W \tag{18}$$

By applying logarithm on both sides of Eq. (18), we have

$$\ln(y_w) = \ln(p_w) + \ln(x_w), \qquad \forall w \in W \tag{19}$$

Let $L_{xw} = \ln(x_w)$, $L_{yw} = \ln(y_w)$, and $L_{pw} = \ln(p_w)$, then

$$L_{yw} = L_{pw} + L_{xw}, \qquad \forall w \in W \tag{20}$$

With the demand function in Eq. (1), by applying logarithm on both sides of Eq. (13), we have

$$\ln(x_w) \leq \ln\big(Q_w(p_w)\big) = \ln(\bar{Q}_w) - \eta_w \cdot p_w, \qquad \forall w \in W \tag{21}$$

Eqs. (1) and (13) then can be replaced by Eq. (21), where Eq. (21) can be rewritten as

$$L_{xw} \leq \ln(\bar{Q}_w) - \eta_w \cdot p_w, \qquad \forall w \in W \tag{22}$$

In summary, the original MPPS model can be reformulated (termed as RMPPS model) as follows:

$$\max Z = \sum_{w \in W} y_w \tag{23}$$

s.t.

$$L_{yw} = L_{pw} + L_{xw}, \qquad \forall w \in W \tag{24}$$

$$L_{xw} \leq \ln(\bar{Q}_w) - \eta_w \cdot p_w, \qquad \forall w \in W \tag{25}$$

$$L_{xw} = \ln(x_w), \qquad \forall w \in W \tag{26}$$

$$L_{yw} = \ln(y_w), \qquad \forall w \in W \tag{27}$$

$$L_{pw} = \ln(p_w), \qquad \forall w \in W \tag{28}$$

$$\underline{p}_w \leq p_w \leq \bar{p}_w, \qquad \forall w \in W \tag{29}$$

$$\sum_{w \in W_k} \delta_{wl}^k \cdot b_w^k \leq c_k, \qquad \forall l \in L, k \in K \tag{30}$$

$$b_w^k \geq 0, \qquad \forall k \in K, w \in W \tag{31}$$

$$x_w \leq \sum_{k \in K_w} b_w^k, \qquad \forall w \in W \tag{32}$$

$$b_w^k \text{ integer}, \qquad \forall k \in K, w \in W \tag{33}$$

$$x_w \text{ integer}, \qquad \forall w \in W \tag{34}$$

In the above reformulated model (RMPPS), the objective function in Eq. (23) and the constraints in Eqs. (24), (25), (29)-(34) are linear, and the constraints in Eqs. (26)-(28) are non-linear but the nonlinearity only involves the logarithms of the decision variables. Section 4 will further discuss how to deal with the nonlinear constraints in Eqs. (26)-(28) in order to provide lower and upper bounds for the model.

4. Solution algorithm

To solve the RMPPS model proposed in Section 3, this section first discusses how to further relax the logarithmic function based on the cut scheme of the variable interval in Subsection 4.1. Then in Subsection 4.2, we transform the RMPPS model into a relaxed mix-integer linear programming problem in order to obtain the upper bound of the RMPPS with the relaxation of the logarithmic function (from Subsection 4.1) and construct the feasible solution as the lower bound of the RMPPS. Moreover, in Subsection 4.3 we adopt a range reduction technique, which is coupled with the relaxed mix-integer linear programming problem, to further decrease the computation cost of the algorithm. In Subsection 4.4, we describe the detailed process of the solution algorithm and discuss its convergence to the globally optimal solution.

4.1 Linear relaxation

In Subsection 3.2, the original MPPS is reformulated into the RMPPS model with the objective function and some constraints being linear in terms of the decision variables. The RMPPS is still non-linear considering the logarithms. However, the nonlinearity of RMPPS only relates to the logarithm function, i.e., $\ln(x_w)$, $\ln(y_w)$, and $\ln(p_w)$. To ease the presentation, we define $H = \{x_w, y_w, p_w, \forall\, w \in W\}$ and $h_w \in H$ might be used to indicate x_w, y_w, or p_w.

Similar to existing studies (e.g., Wang and Lo, 2010; Wang et al., 2015; Liu and Wang, 2015; Liu et al., 2019), a piecewise linear relaxation is introduced, as shown in Fig. 2. We take the logarithm function $L_{hw} = \ln(h_w), h_w \in H$ as an example to elaborate the linear relaxation. Denote $\underline{h}_w$ and $\overline{h}_w$ as the predefined lower and upper bounds of h_w. In particular, for $h_w = x_w$, the lower and upper bounds, i.e., $\underline{x}_w$ and $\overline{x}_w$ can be set to be zero and $\bar{Q}_w \cdot \exp\left(-\eta_w \cdot \underline{p}_w\right)$. For $h_w = y_w$, its lower bound $\underline{y}_w$ can be set to be zero. Moreover, with Eqs. (13) and (18), we have

$$y_w = p_w \cdot x_w \le p_w \cdot \bar{Q}_w \cdot \exp(-\eta_w \cdot p_w) \tag{35}$$

Then, the upper bound $\overline{y}_w$ can be set as

$$\overline{y}_w = \begin{cases} \dfrac{\bar{Q}_w \cdot \exp(-1)}{\eta_w}, & if\ \underline{p}_w \le \dfrac{1}{\eta_w} \le \bar{p}_w \\ max\left\{p_w \cdot \bar{Q}_w \cdot \exp(-\eta_w \cdot p_w)\middle| p_w = \underline{p}_w, \bar{p}_w\right\}, & otherwise \end{cases} \quad (36)$$

The interval $[\underline{h}_w, \overline{h}_w]$ can be further divided uniformly into $N-1$ intervals by the set of points h_w^n as given in Eq. (37). As shown in Fig. 2, the tangential support is constructed at each point h_w^n and the curve chords are formed by connecting two adjacent points h_w^n and h_w^{n+1} for $n = 1,2,\cdots N-1$. The linear relaxation of $\ln(h_w)$ is set to be the region below all tangent lines and above all curve chords. Then the relaxation of $\ln(h_w)$ with breakpoints h_w^n, $n = 1,2,\cdots,N$ can be constructed as follows:

$$L_{hw} \le \ln(h_w^n) - 1 + \frac{h_w}{h_w^n}, \forall h_w^n = \underline{h}_w + \frac{\overline{h}_w - \underline{h}_w}{N-1} \cdot (n-1), n = 1,2,\cdots,N \quad (37)$$

$$\sum_{n=1}^{N} \theta_{hw}^n \cdot h_w^n = h_w \quad (38)$$

$$\sum_{n=1}^{N} \theta_{hw}^n \cdot \ln(h_w^n) \le L_{hw} \quad (39)$$

$$\sum_{n=1}^{N} \theta_{hw}^n = 1 \quad (40)$$

$$\theta_{hw}^n \ge 0, \qquad n = 1,2,\cdots,N \quad (41)$$

$$\theta_{hw}^n \le \lambda_{hw}^{n-1} + \lambda_{hw}^n, \quad n = 2,3,\cdots,N-1\,; \theta_{hw}^1 \le \lambda_{hw}^1\,; \theta_{hw}^N \le \lambda_{hw}^{N-1} \quad (42)$$

$$\sum_{n=1}^{N-1} \lambda_{hw}^n = 1 \quad (43)$$

$$\lambda_{hw}^n = \{0,1\}, \quad n = 1,2,\cdots,N-1 \quad (44)$$

In Eq. (37), as the right-hand side denotes all the tangent lines, Eq. (37) represents the upper bound of $\ln(h_w)$, i.e., L_{hw} are below the tangent lines. If h_w is within the interval $[h_w^{n^*}, h_w^{n^*+1}]$, then Eqs. (40)-(44) mean that only the values of $\theta_{hw}^{n^*}$ and $\theta_{hw}^{n^*+1}$ are no less than zero and other values of θ_{hw}^n are all equal to zero. Then the left-hand side of Eq. (39) represents the curve chord from $\left(h_w^{n^*}, \ln\left(h_w^{n^*}\right)\right)$ to $\left(h_w^{n^*+1}, \ln\left(h_w^{n^*+1}\right)\right)$. Therefore, Eqs. (38)-(44) together constrain L_{hw} to be greater than those defined by all curve chords.

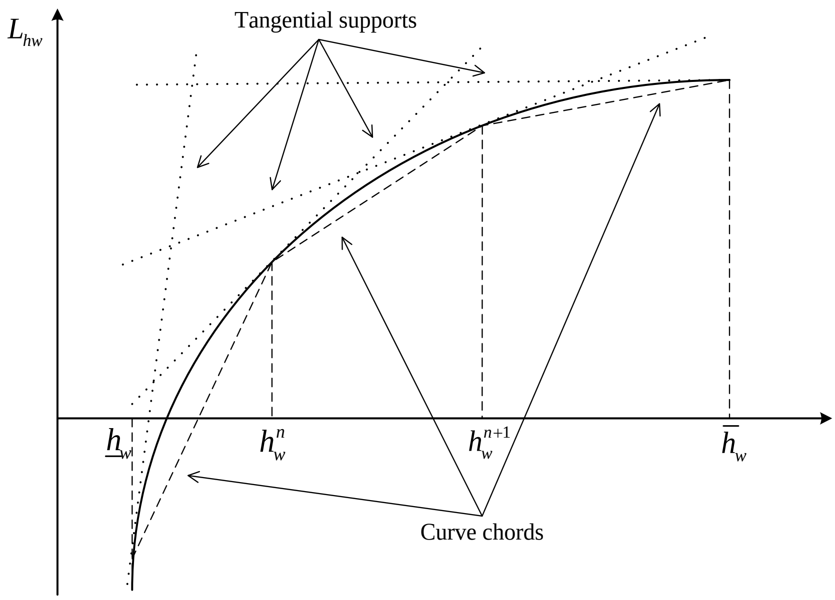


Fig. 2. Linear relaxation of the logarithm function

4.2 Relaxed mixed-integer linear program

Given the predefined breakpoints of the variables x_w, y_w and p_w, the linear relaxation in Section 4.1 transforms the non-linear constraints in Eqs. (26)-(28) into the linear constraints in Eqs. (37)-(44). With the above linear relaxation, the RMPPS is relaxed into the following mixed-integer linear program (RMILP).

$$\max Z = \sum_{w \in W} y_w \tag{45}$$

s.t.

$$L_{yw} = L_{pw} + L_{xw}, \quad \forall w \in W \tag{46}$$

$$L_{xw} \leq \ln(\bar{Q}_w) - \eta_w \cdot p_w, \quad \forall w \in W \tag{47}$$

$$\underline{p}_w \leq p_w \leq \bar{p}_w, \quad \forall w \in W \tag{48}$$

$$\sum_{w \in W_k} \delta_{wl}^k \cdot b_w^k \leq c_k, \quad \forall l \in L, k \in K \tag{49}$$

$$b_w^k \geq 0, \quad \forall k \in K, w \in W \tag{50}$$

$$x_w \leq \sum_{k \in K_w} b_w^k, \quad \forall w \in W \tag{51}$$

$$b_w^k \text{ integer}, \quad \forall k \in K, w \in W \tag{52}$$

$$x_w \text{ integer}, \quad \forall w \in W \tag{53}$$

$$\text{Constraints in Eqs. (37)-(44) for } h_w \in H \tag{54}$$

We now further show that through utilizing the above relaxed mixed-integer linear program (RMILP) we can obtain lower and upper bounds for the original MPPS model.

In the original MPPS, the decision variables are $p_w,\ w \in W$ and $b_w^k, w \in W_k, k \in K$. To ease the presentation, we define the variables p_w and b_w^k in the vector forms by $P = \{p_w, w \in W\}$ and $B = \{b_w^k, w \in W_k, k \in K\}$, respectively. Then the original MPPS model in Eqs. (23)-(34) can be written as (M1) as follows:

$$\text{(M1)}: \max Z_1 = F_1(P, B) \tag{55}$$

s.t.

$$G(P, B) \leq 0 \tag{56}$$

where $G(P, B)$ represents the constraints in Eqs. (24)-(34). Let (P^*, B^*) be the global optimal solution of model M1 and Z_1^* be the corresponding objective function value.

Given the value of P which satisfies Eq. (29), the model M1 becomes an integer linear program for optimizing the seat allocation B, which can be readily solved by existing linear program solvers. We can define the objective function value to be a function of P, i.e., $F_1(P)$ and its solution to be $\tilde{B}$. When the value of B is given, which satisfies Eqs. (30)-(31), the model M1 becomes a problem of optimizing the price P. Similarly, we can define the objective function value to be a function of B, i.e., $F_1(B)$ and its solution to be $\tilde{P}$. In particular, this problem can be transformed into the problems of optimizing the price p_w for each OD pair w, i.e.,

$$(\text{M1}_w): \max y_w = p_w \cdot x_w \tag{57}$$

s.t.

$$x_w \leq Q_w(p_w) \tag{58}$$

$$x_w \leq b_w \tag{59}$$

$$x_w \text{ integer} \tag{60}$$

$$\underline{p}_w \leq p_w \leq \bar{p}_w \tag{61}$$

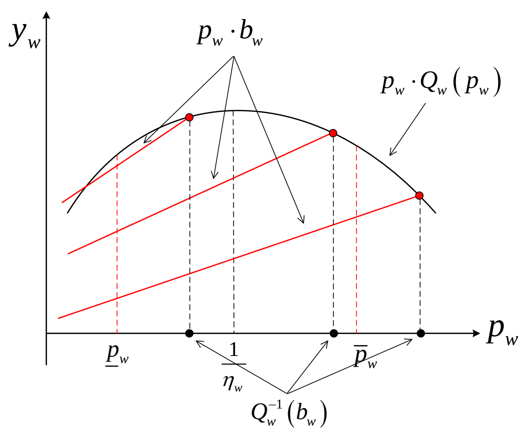
where b_w is given. For solving the model M1_w, we first denote $Q_w^{-1}(\cdot)$ to be the inverse function of $Q_w(p_w)$. If $p_w \leq Q_w^{-1}(b_w)$, then $Q_w(p_w) \geq b_w$ and $y_w = p_w \cdot b_w$, which corresponds to the red solid lines in Fig. 3. The other case with $Q_w(p_w) < b_w$ and $y_w = p_w \cdot Q_w(p_w)$ corresponds to the black solid curves in Fig. 3. One can verify that

- When $\underline{p}_w \leq \frac{1}{\eta_w} \leq \bar{p}_w$, as shown in Fig. (3a): If $Q_w^{-1}(b_w) \leq \frac{1}{\eta_w}$, the optimal price $p_w^* = \frac{1}{\eta_w}$; if $\frac{1}{\eta_w} \leq Q_w^{-1}(b_w) \leq \bar{p}_w$, $p_w^* = Q_w^{-1}(b_w)$; if $Q_w^{-1}(b_w) \geq \bar{p}_w$, $p_w^* = \bar{p}_w$.
- When $\underline{p}_w \leq \bar{p}_w \leq \frac{1}{\eta_w}$, as shown in Fig. (3b): the optimal price $p_w^* = \bar{p}_w$.

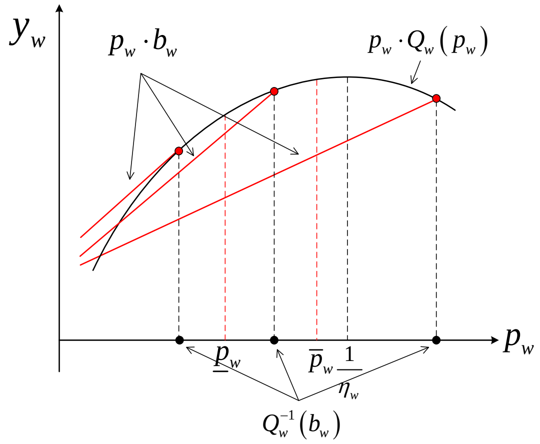
- When $\frac{1}{\eta_w} \le \underline{p}_w \le \bar{p}_w$, as shown in Fig. (3c): If $Q_w^{-1}(b_w) \le \underline{p}_w$, $p_w^* = \underline{p}_w$; if $\underline{p}_w \le Q_w^{-1}(b_w) \le \bar{p}_w$, $p_w^* = Q_w^{-1}(b_w)$; if $Q_w^{-1}(b_w) \ge \bar{p}_w$, $p_w^* = \bar{p}_w$.

Thus, the solution of model $\mathrm{M1}_w$ can be given as

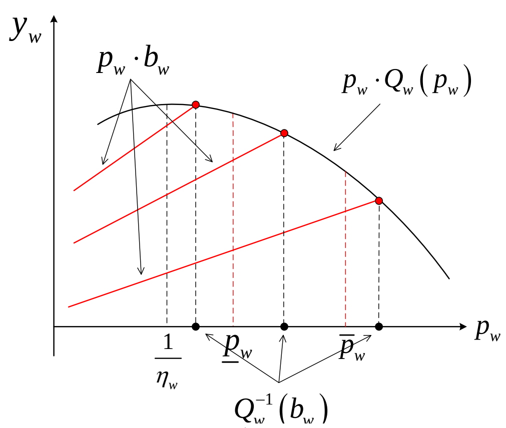
$$p_w^* = \begin{cases} \frac{1}{\eta_w}, & \underline{p}_w \le \frac{1}{\eta_w} \le \bar{p}_w, Q_w^{-1}(b_w) \le \frac{1}{\eta_w} \\ Q_w^{-1}(b_w), & \underline{p}_w \le \frac{1}{\eta_w} \le \bar{p}_w, \frac{1}{\eta_w} \le Q_w^{-1}(b_w) \le \bar{p}_w \\ \bar{p}_w, & \underline{p}_w \le \frac{1}{\eta_w} \le \bar{p}_w, Q_w^{-1}(b_w) \ge \bar{p}_w \\ \bar{p}_w, & \underline{p}_w \le \bar{p}_w \le \frac{1}{\eta_w} \\ \underline{p}_w, & \frac{1}{\eta_w} \le \underline{p}_w \le \bar{p}_w, Q_w^{-1}(b_w) \le \underline{p}_w \\ Q_w^{-1}(b_w), & \frac{1}{\eta_w} \le \underline{p}_w \le \bar{p}_w, \underline{p}_w \le Q_w^{-1}(b_w) \le \bar{p}_w \\ \bar{p}_w, & \frac{1}{\eta_w} \le \underline{p}_w \le \bar{p}_w, Q_w^{-1}(b_w) \ge \bar{p}_w \end{cases} \tag{62}$$



(a) $\underline{p}_w \le \frac{1}{\eta_w} \le \bar{p}_w$



(b) $\underline{p}_w \le \bar{p}_w \le \frac{1}{\eta_w}$



(c) $\frac{1}{\eta_w} \le \underline{p}_w \le \bar{p}_w$

Fig. 3. The illustration of how to solve model $\mathrm{M1}_w$

Let $\widetilde{H}^m$ denote the set of breakpoints for variables x_w, y_w and p_w, $w \in W$ at the iteration number m during the iterative algorithm to be introduced in Section 4.4. In a similar way, the RMILP can also be rewritten as follows (M2):

$$\text{(M2)}: \max Z_2 = F_2\left(P, B, \widetilde{H}^m\right) \tag{63}$$

s.t.

$$\Theta(P, B, \widetilde{H}^m) \leq 0 \tag{64}$$

where $\Theta(P, B, \overline{H}^m)$ represents the linear constrains in Eqs. (46)-(54). We denote (P^m, B^m) as the solution of the above model M2. With P^m and B^m we can obtain $F_1(P^m)$ and $F_1(B^m)$, respectively. It is obvious that

$$\max\{F_1(P^m), F_1(B^m)\} \leq F_1(P^*, B^*) \leq F_2\left(P^m, B^m, \widetilde{H}^m\right) \tag{65}$$

From the above analysis and procedure, it is evident that the upper and lower bounds of the optimal solution can be obtained.

4.3 Range reduction scheme

This subsection further introduces the construction of the set of breakpoints and proposes the range reduction scheme to decrease the computation cost of the algorithm (to be detailed in Section 4.4) for finding the optimal solution.

In model M2 with the set of breakpoints $\widetilde{H}^m$, we denote $\widetilde{H}^m_{hw}$ as the set of breakpoints for the variable $h_w \in H$ at the iteration number m. For the initial step, i.e., $m = 1$, we have

$$\widetilde{H}^1_{hw} = \left\{ h^n_w \middle| h^n_w = \underline{h}_w + \frac{\overline{h}_w - \underline{h}_w}{N-1} \cdot (n-1), n = 1,2,\cdots,N \right\} \tag{66}$$

At the iteration number $m+1$, we can construct the set of breakpoints $\overline{H}^{m+1}$ as follows:

$$\triangle\, \widetilde{H}^m_{hw} = \left\{ \frac{h^n_w + h^{n+1}_w}{2} \middle| h^n_w, h^{n+1}_w \in \widetilde{H}^m_{hw}, n = 1,2,\cdots,N-1 \right\} \tag{67}$$

$$\widetilde{H}^{m+1}_{hw} = \triangle\, \widetilde{H}^m_{hw} \cup \widetilde{H}^m_{hw} \tag{68}$$

To ease the notation burden, we still use N to present the number of breakpoints at iteration number m, and the breakpoints for each variable are sorted in the increasing order of their values. However, the above method involves a growing number of breakpoints.

We now introduce the range reduction scheme to reduce the feasible region and the number of breakpoints. Without loss of generality, we use the variable $h_w \in H$ for illustration. Denote $\underline{h}_w^m$ and $\bar{h}_w^m$ to be the lower and upper bounds of variable h_w at iteration number m. We will solve the following model M3 to obtain the new lower and upper bounds of variable h_w, i.e.,

$$\text{(M3): } \underline{h}_w^{m+1} = \min_{(P,B)} h_w, \quad \bar{h}_w^{m+1} = \max_{(P,B)} h_w \tag{69}$$

s.t.

$$\Theta(P, B, \widetilde{H}^m) \leq 0 \tag{70}$$

$$\underline{h}_w^m \leq h_w \leq \bar{h}_w^m \tag{71}$$

$$F_2\left(P, B, \widetilde{H}^m\right) \geq \max\{F_1(P^m), F_1(B^m)\} \tag{72}$$

Proposition 1. The global solution of the model M1 will not be eliminated by the range reduction scheme governed by Eqs. (69)-(72) (model M3), i.e., the global solution h_w^* is within $\left[\underline{h}_w^{m+1}, \bar{h}_w^{m+1}\right]$.

Proof. The proof of Proposition 1 is provided in Appendix B.

Using the model M3, the feasible region of variable h_w can be reduced, and we can construct the set of breakpoints with the new bounds. Let u and v be positive integers such that

$$h_w^{u-1} \leq \underline{h}_w^{m+1} \leq h_w^u \tag{73}$$

$$h_w^v \leq \bar{h}_w^{m+1} \leq h_w^{v+1} \tag{74}$$

where $h_w^n \in \widetilde{H}_{hw}^m$. We construct the set of breakpoints for variable h_w at the iteration number $m+1$, then

$$\begin{aligned} \widetilde{H}_{hw}^{m+1} = &\left\{h_w^n \in \widetilde{H}_{hw}^m \middle| n = u, u+1, \cdots, v\right\} \\ &\cup \left\{\frac{h_w^n + h_w^{n+1}}{2} \middle| h_w^n, h_w^{n+1} \in \bar{H}_{hw}^m, n = u, 2, \cdots, v-1\right\} \\ &\cup \left\{\underline{h}_w^{m+1}, \bar{h}_w^{m+1}\right\} \end{aligned} \tag{75}$$

and $\widetilde{H}^{m+1} = \bigcup_{h_w \in H} \widetilde{H}_{hw}^{m+1}$. With the above method of constructing the set of breakpoints, we denote Ω_m to be the feasible region of the model M2, and the following Proposition 2 holds.

Proposition 2. With the method for updating the set of breakpoints governed by Eq. (67), we have $\Omega_m \supset \Omega_{m+1}$ and the set of optimal objective function values $\left\{F_2\left(P^m, B^m, \widetilde{H}^m\right)\right\}$ obtained based on the model M2 is a monotonically decreasing series.

Proof. The proof of Proposition 2 is provided in Appendix C.

4.4 Solution algorithm

With the analysis in Sections 4.1-4.3, we can develop a globally optimal solution algorithm for model M1, where the details are summarized below.

Step 0: Initialization.

Use a large enough value as the function upper bound $\bar{Z}_1^0$ and a small enough value as the lower bound $\underline{Z}_1^0$;

Let the iteration number $m = 1$;

Set the initial number of breakpoints $N = 3$, and construct the initial set of breakpoints, $\widetilde{H}^1$ with Eq. (66).

Step 1: Solve the relaxed model.

Solve the model M2 using the MILP algorithm to obtain the optimal solution (P^m, B^m) and the optimal objective function value Z_2^m;

Solve the objective function values $F_1(P^m)$ and $F_1(B^m)$, and get the corresponding solutions $\tilde{B}^m$ and $\tilde{P}^m$, respectively;

Update the objective function bounds: $\bar{Z}_1^m = \min\{\bar{Z}_1^{m-1}, F_2(P^m, B^m, \widetilde{H}^m)\}$ and $\underline{Z}_1^m = \max\{\underline{Z}_1^{m-1}, F_1(P^m), F_1(B^m)\}$;

Step 2: Check the convergence.

If $\frac{|\bar{Z}_1^m - \underline{Z}_1^m|}{\bar{Z}_1^m} \leq \varepsilon$, then stop; otherwise, go to Step 3.

Step 3: Update the breakpoint set

Choose a group of variables for further range reductions and form the set H^m;

For each variable $h_w \in H^m$:

Reduce the range of the variable h_w by solving the model M3 and update the variable range $[\underline{h}_w^{m+1}, \bar{h}_w^{m+1}]$;

Calculate u and v with Eqs. (73)-(74);

Update the set of breakpoints and obtain $\widetilde{H}_{hw}^{m+1}$ with Eq. (75).

For each variable $h_w \in H \backslash H^m$, update the set of breakpoints and obtain $\widetilde{H}_{hw}^{m+1}$ with Eqs. (67) and (68).

Calculate $\widetilde{H}^{m+1} = \bigcup_{h_w \in H} \widetilde{H}_{hw}^{m+1}$;

Set $m = m + 1$ and go to Step 1.

It is noteworthy that if the proposed algorithm executes the range reduction method for all variables in each iteration, it will cost substantial computation time. Thus, in each iteration, the algorithm may choose a subset of variables to apply the range reduction method, which can be based on the variation trend of each variable over the iterations.

Proposition 3. When the iteration number $m \to +\infty$, the proposed algorithm guarantees the convergence to the globally optimal solution of the original model MPPS or model M1.

Proof. The proof of Proposition 3 is provided in Appendix D.

5. Numerical studies

Now we turn to numerical illustrations for the proposed model and algorithm. We firstly present a toy network example for illustration, and then test the developed method on a real-world regional network in China.

All numerical tests are conducted on a personal computer with Intel® Core (TM) 3.00 GHz processor and 16.00 GB RAM and Windows 10 Home Edition operating system (64-bit). The YALMIP-R20190425 together with MATLAB R2018b is used to conduct the numerical tests. The commercial solver GUROBI optimization studio 8.1.1 (IBM ILOG, 2018) is adopted to solve all RMILP problems, whereas the free solver FMINCON from the MATLAB platform is applied to solve all the nonlinear problems.

5.1. A toy network example

We adopt a railway track network shown in Fig. 4. There are six stations S_1, S_2, S_3, S_4, S_5 and S_6 and four trains k_1, k_2, k_3 and k_4 running in the network. For this toy network example, all numerical settings are assumed. There are 15 OD pairs in total, and the total potential demand for all OD pairs is 11250. The stopping pattern of each train is shown in Fig. 4.

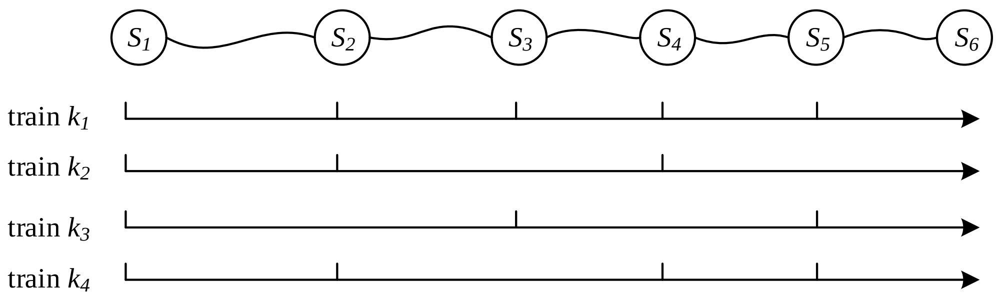


Fig. 4. A railway track network with four running trains in the toy network example

We set the convergence parameter as $\varepsilon = 5 \times 10^{-4}$ (for the solution algorithm in Section 4.4) and the capacity of each train as 400 persons (or seats), i.e., c_k=400 seats per train, $k \in K$. The OD-specific parameters are summarized in Table 3. The potential demands for all OD pairs are summarized in Table 4.

Table 3. A summary of OD-specific parameters in the toy network example

OD pair	$\underline{p}_w$	$\overline{p}_w$	value of η_w
(S_1, S_2)	50	100	0.0109
(S_1, S_3)	80	160	0.0093
(S_1, S_4)	140	280	0.0095
(S_1, S_5)	210	420	0.0132
(S_1, S_6)	260	520	0.0106
(S_2, S_3)	30	60	0.0135
(S_2, S_4)	90	180	0.0109
(S_2, S_5)	160	320	0.0093
(S_2, S_6)	210	420	0.0095
(S_3, S_4)	60	120	0.0132
(S_3, S_5)	130	260	0.0106
(S_3, S_6)	180	360	0.0135
(S_4, S_5)	70	140	0.0093
(S_4, S_6)	120	240	0.0095
(S_5, S_6)	50	100	0.0132

Table 4. Potential demands for all OD pairs in the toy network example

OD pair	S_1	S_2	S_3	S_4	S_5	S_6
S_1	0	600	450	750	1500	750
S_2	0	0	750	600	450	450
S_3	0	0	0	600	750	900
S_4	0	0	0	0	600	900
S_5	0	0	0	0	0	1200
S_6	0	0	0	0	0	0

Given the above setting, we implemented the proposed method for solving this toy network example. The optimal pricing and seat allocation solution is shown in Table 5, and the seat allocation of each section is shown in Table 6. One can verify that the pricing and seat allocation scheme meets all problem constraints. Fig. 5 further shows the convergence process of the lower and upper bounds of model, where the proposed method yields a globally optimal solution (upper and lower bounds converge to $Z = 3.141 \times 10^5$). The total CPU time for solving this toy network example is 28.326s.

Table 5. Optimal price and seat allocation scheme solved by the proposed method in the toy network example

OD pair	p_w	x_w	$Q_w(p_w)$	b_w	b_w^k			
					b_w^1	b_w^2	b_w^3	b_w^4
(S_1, S_2)	93.29	217	217.04	717	380	222	0	115
(S_1, S_3)	104.80	169	169.80	268	0	0	268	0
(S_1, S_4)	140.19	198	198.00	198	20	178	0	0
(S_1, S_5)	210.65	93	93.00	285	0	0	0	285

(S_1,S_6)	261.30	47	47.00	132	0	0	132	0
(S_2,S_3)	60.00	333	333.64	333	333	0	0	0
(S_2,S_4)	91.20	222	222.04	222	0	222	0	0
(S_2,S_5)	160.66	101	101.00	101	0	0	0	101
(S_2,S_6)	210.35	61	61.00	61	47	0	0	14
(S_3,S_4)	76.00	220	220.01	333	333	0	0	0
(S_3,S_5)	130.03	189	189.00	189	0	0	189	0
(S_3,S_6)	180.22	79	79.00	79	0	0	79	0
(S_4,S_5)	107.85	220	220.06	353	353	0	0	0
(S_4,S_6)	120.31	287	287.00	400	0	400	0	0
(S_5,S_6)	76.49	437	437.22	928	353	0	189	386

Table 6. The seat allocation of each section solved by the proposed method in the toy network example

Section no.	Section	b_l^k			
		Train k_1	Train k_2	Train k_3	Train k_4
1	(S_1,S_2)	400	400	400	400
2	(S_2,S_3)	400	400	400	400
3	(S_3,S_4)	67	400	132	400
4	(S_4,S_5)	400	400	400	400
5	(S_5,S_6)	400	400	400	400

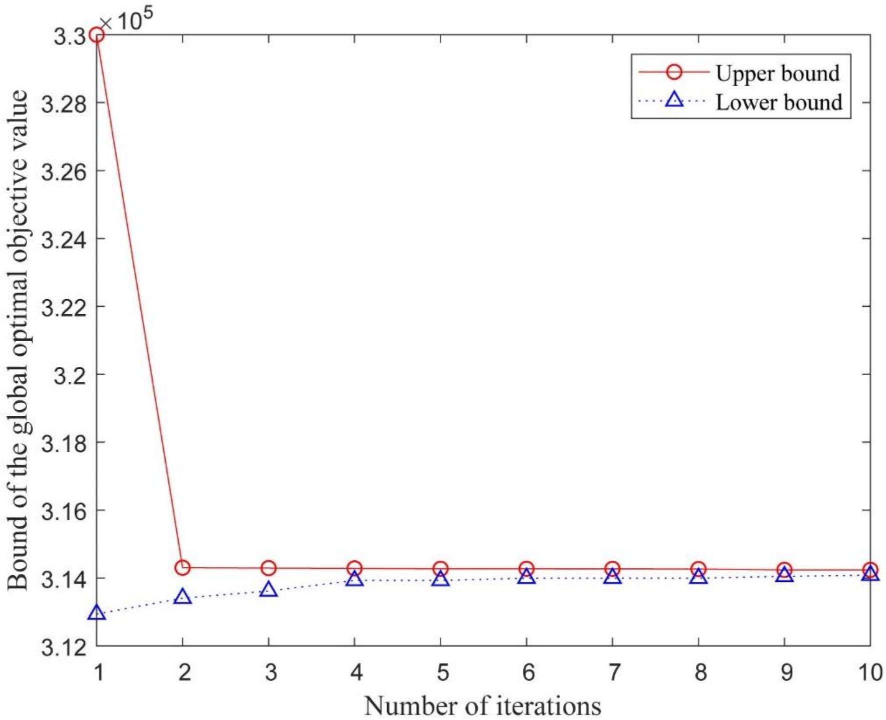


Fig. 5. Updating process of upper and lower bounds in the toy network example

For comparison purpose, we also tested "FMINCON" in MATLAB R2018b, which is used for solving the non-linear constrained optimization problem with four different algorithms: trust region reflective, interior point, active set, and sequential quadratic programming (SQP). As the MPPS problem in this paper involves integers, the above four algorithms embedded in "FMINCON" in MATLAB R2018b cannot produce a feasible solution. We relaxed all the integer variables (i.e., b_w^k and x_w) in the MPPS model as continuous variables and then solved the relaxed model labelled as "MPPS_1" with the four algorithms embedded in "FMINCON". The solutions obtained from "FMINCON" are listed in Appendix E. In particular, trust region reflective algorithm and interior point algorithm produced the same feasible solution (with an objective function value of $Z = 5.209 \times 10^4$, a CPU time of 11.313s for trust region reflective and a CPU time of 11.405s for interior point). The SQP algorithm produced another feasible solution that is more effective than the other three algorithms in FMINCON (with an objective function value of $Z = 3.146 \times 10^5$ and a CPU time of 2.333s), and the activate-set algorithm produces the worst (and very inefficient) solution (omitted in Appendix E), which might be due to that activate-set algorithm mainly focus on quadratic programming.

For comparison purpose, we also implemented the proposed method for solving the MPPS_1 model (with continuous variables) for the toy network example. The pricing and seat allocation solution is shown in Table 7. The seat allocation for each section is shown in Table 8. One can verify that the pricing and seat allocation scheme meets all constraints. The objective function value is $Z = 3.146 \times 10^5$. As can be seen, the objective function value of MPPS_1 model with continuous variables is only slightly larger than the original MPPS model with integer variables (i.e., $3.146 \times 10^5 > 3.141 \times 10^5$). It is also noted that the solution of MPPS_1 model is quite different from the solution of the original MPPS model, and we cannot obtain the MPPS solution by rounding the solution of the MPPS_1. This indeed highlights the importance to explicitly solve the integer models.

Moreover, the solution of the MPPS_1 model produced by our method is much better than those from trust-region-reflective, interior-point and activate-set, and is very close to the SQP algorithm. This is because, MPPS_1 model is non-concave and non-linear, trust-region-reflective and interior-point algorithms may have solved a local optimum rather than a global optimum. The SQP algorithm is very powerful for solving non-linear optimization problems which can handle any degree of non-linearity including non-linearity in the constraints (Nocedal and Wright, 2006). It is noted that the two solutions obtained by SQP (with similar objective function values) and our method are different, which indicate that the optimal solution of this problem might be non-unique. Note that the SQP relies on several derivatives (either analytically or numerically), and it becomes quite cumbersome for large-scale problems with many variables or constraints. As tested in the next subsection for the real-world example, even if we consider the MPPS_1 model with continuous variables, the four algorithms in FMINCON function cannot find an optimal solution.

Table 7. Optimal price and seat allocation scheme for the relaxed model MPPS_1 solved by the proposed method in the toy network example

OD pair	p_w	x_w	$Q_w(p_w)$	b_w	b_w^k			
					b_w^1	b_w^2	b_w^3	b_w^4
(S_1,S_2)	91.34	221.70	221.70	221.70	0	0	0	221.70
(S_1,S_3)	106.14	167.70	167.70	167.70	0	0	167.70	0
(S_1,S_4)	140.00	198.36	198.36	198.36	0	68.34	0	130.01
(S_1,S_5)	210.00	93.81	93.81	93.81	0	0	93.81	0
(S_1,S_6)	260.00	47.66	47.66	47.66	9.76	0	37.70	0
(S_2,S_3)	60.00	333.64	333.64	333.64	333.64	0	0	0
(S_2,S_4)	90.00	224.96	224.96	224.96	0	0	0	224.96
(S_2,S_5)	160.00	101.62	101.62	101.62	56.60	0	0	45.02
(S_2,S_6)	210.00	61.21	61.21	61.21	0	61.21	0	0
(S_3,S_4)	74.99	222.98	222.98	222.98	222.98	0	0	0
(S_3,S_5)	130.00	189.06	189.06	189.06	0	0	189.06	0
(S_3,S_6)	180.00	79.23	79.23	79.23	0	0	79.23	0
(S_4,S_5)	106.33	223.19	223.19	223.19	0	0	0	223.19
(S_4,S_6)	120.00	287.84	287.84	287.84	0	287.84	0	0
(S_5,S_6)	75.95	440.32	440.32	440.32	0	0	40.23	400.00

Table 8. The seat allocation of each section for the relaxed model MPPS_1 solved by the proposed method in the toy network example

Section no.	Section	b_l^k			
		Train k_1	Train k_2	Train k_3	Train k_4
1	(S_1,S_2)	9.76	68.34	299.40	351.71
2	(S_2,S_3)	400.00	129.55	299.40	400.00
3	(S_3,S_4)	289.34	129.55	400.00	400.00
4	(S_4,S_5)	66.36	349.04	400.00	268.21
5	(S_5,S_6)	9.76	61.21	157.45	400.00

We now further illustrate the computation efficiencies of the four algorithms in

FMINCON and our method under varying demand levels for MPPS_1. We here define θ as a scale parameter for potential OD demand, and let:

$$\tilde{Q}_w = \theta \cdot \bar{Q}_w$$

where $\bar{Q}_w$ is the benchmark potential demand level. We solved the MPPS problem for the potential demands $\tilde{Q}_w$, which is different under different values of θ. Specifically, the value of θ increases from 0.1 to 3. We calculate objective values of the methods of FMINCON (denoted by Z) and those of our method (denoted by Z^*) with different values of θ, and evaluate the ticket revenue ratio σ defined as follows:

$$\sigma = \frac{Z}{Z^*}$$

We also record the CPU times by different methods and the results are summarized in Table 9.

Table 9. Optimal objective values and CPU times of the four algorithms in FMINCON and the proposed method in the toy network example

θ	FMINCON for MPPS_1												Our method			
	Trust-region-reflective			Interior-point			Activate-set			SQP			MPPS_1		MPPS	
	Objective value	CPU time(s)	$\sigma(\%)$	Objective value	CPU time(s)	$\sigma(\%)$	Objective value	CPU time(s)	$\sigma(\%)$	Objective value	CPU time(s)	$\sigma(\%)$	Objective value	CPU time(s)	Objective value	CPU time(s)
0.1	3.146E+04	4.184	100.0	3.146E+04	3.761	100.0	1.169E-04	2.359	0.0	3.146E+04	1.344	100.0	3.146E+04	7.422	3.100E+04	76.700
0.2	6.299E+04	5.633	100.0	6.299E+04	5.682	100.0	1.169E-04	2.536	0.0	6.292E+04	3.294	99.9	6.299E+04	8.463	6.246E+04	83.001
0.3	9.437E+04	9.870	100.0	9.437E+04	9.843	100.0	1.169E-04	2.278	0.0	9.437E+04	1.685	100.0	9.437E+04	8.391	9.399E+04	96.158
0.4	1.258E+05	8.736	100.0	1.258E+05	8.694	100.0	1.169E-04	2.328	0.0	1.258E+05	1.913	100.0	1.258E+05	8.499	1.254E+05	84.294
0.5	1.573E+05	5.995	100.0	1.573E+05	5.976	100.0	1.169E-04	2.279	0.0	1.573E+05	1.862	100.0	1.573E+05	8.555	1.566E+05	60.872
0.6	1.887E+05	9.131	100.0	1.887E+05	9.085	100.0	1.169E-04	2.310	0.0	1.887E+05	2.929	100.0	1.887E+05	8.748	1.883E+05	90.908
0.7	2.202E+05	8.451	100.0	2.202E+05	8.457	100.0	1.169E-04	2.307	0.0	2.202E+05	2.194	100.0	2.202E+05	8.896	2.195E+05	101.777
0.8	2.517E+05	9.323	100.0	2.517E+05	8.859	100.0	1.169E-04	2.317	0.0	2.517E+05	1.996	100.0	2.517E+05	8.977	2.512E+05	87.263
0.9	1.733E+05	11.792	61.4	1.733E+05	11.761	61.4	1.169E-04	2.302	0.0	2.822E+05	2.220	100.0	2.822E+05	6.627	2.820E+05	69.527
1	5.209E+04	11.313	0.2	5.209E+04	11.405	0.2	1.169E-04	2.375	0.0	3.146E+05	2.333	100.0	3.146E+05	6.716	3.141E+05	28.326
1.5	3.493E+05	10.228	75.6	3.493E+05	10.125	75.6	1.169E-04	2.337	0.0	4.620E+05	4.296	100.0	4.620E+05	7.930	4.618E+05	29.387
2	4.196E+05	7.065	74.3	4.148E+05	7.032	73.4	1.169E-04	2.375	0.0	5.647E+05	8.508	99.9	5.648E+05	6.786	5.658E+05	54.240
2.5	3.619E+05	11.116	63.3	3.619E+05	11.181	63.3	1.169E-04	2.434	0.0	5.714E+05	2.771	100.0	5.714E+05	4.559	6.438E+05	65.545
3	6.703E+05	11.810	93.9	6.703E+05	11.772	93.9	1.169E-04	2.413	0.0	7.118E+05	3.065	99.7	7.141E+05	4.486	7.092E+05	83.277

We summarize several main observations from Table 9. First, we can see that the activate-set algorithm in FMINCON is always inefficient for solving this problem. Second, for the trust-region-reflective and interior-point, when θ varies from 0.1 to 0.8, the objective values are very close to those obtained by our method, while $\theta \geq 0.9$, our method performs better. Third, for trust-region-reflective and interior-point, the revenue under $\theta = 1$ (a higher potential demand) is smaller than those under $\theta = 0.9$ (a lower potential demand). This implies that these two algorithms may converge into a local optimal solution for the MPPS_1 problem. Fourth, with the increase of θ, the demand levels for all OD pairs increase, and the capacity of trains become less sufficient and the solution space becomes more complicated. The two algorithms, i.e., trust-region-reflective and interior-point, in the function of FMINCON, more likely converge into local optimal solutions, while the proposed method always provides the best solution. Fifth, the SQP algorithm in the function of FMINCON provides solutions of similar quality to our method for this small example under varying demand levels. Sixth, the average CPU time (7.503s) of our method is slightly less than those of the algorithms of region-reflective (8.903s) and interior-point (8.831s), but larger than that of SQP (2.886s). It shows that the built-in function FMINCON of MATLAB with SQP algorithm has its advantage in computing efficiency for the small example. With the increase of θ, the CPU time of the proposed method for MPPS_1 do not have a clear trend, which means that proposed method might not depend on the demand level. We also adopt our method to solve the small example of MPPS under different demand levels. As can be seen from Table 9, with our method, the CPU times for MPPS model are about ten times of those of MPPS_1 under different values of θ. This is because, the MPPS model involves more integer variables when solving the upper and lower bounds than the MPPS_1 model.

5.2. A real-world regional network: Ninghang railway

This section applies the model and algorithm on a real-world regional network, i.e., Ninghang railway. Ninghang Railway includes 11 stations (please refer to Fig. 6). The schedule from 6:00 am to 12:00 am on October 20, 2019 is used in the example. There are 24 high-speed trains running on the network and we set each train's seat capacity to be 1000. The train running diagram is showed in Fig. 7. For this network, there are 55 OD pairs. To ease the presentation, we use S_1, S_2, …, S_{11} to represent Nanjing South station, Jiangning station, …, and Hangzhou East station. We also summarize the price bounds and the potential demands for all OD pairs shown in Table 10 and Table 11 respectively. The upper and lower bounds of the price are usually governed by local policies, and we choose them based on information of the tariffs, mileage and relevant policies in China[7]. The total potential demand for all OD pairs is 189,600. The potential demand values are chosen with reference to the historical ticket booking data from China Academy of Railway Sciences Corporation Limited for the Ninghang railway. The parameters of the elastic demand functions are assumed, which are summarized in Table 12. Note that for the Ninghang railway network example, we also tested the four algorithms in the function of FMINCON to solve the MPPS and MPPS_1 models. For the MPPS model, MATLAB returned that the problems cannot be solved. For the MPPS_1 model, Activate-set and SQP algorithms still cannot solve it, while trust-region-reflective and interior-point algorithms can obtain the same feasible solution for MPPS_1 model and the objective function value is 1.1248×10^5, which is much smaller than the global optimal objective function value 4.2689×10^6 obtained by the proposed algorithm.

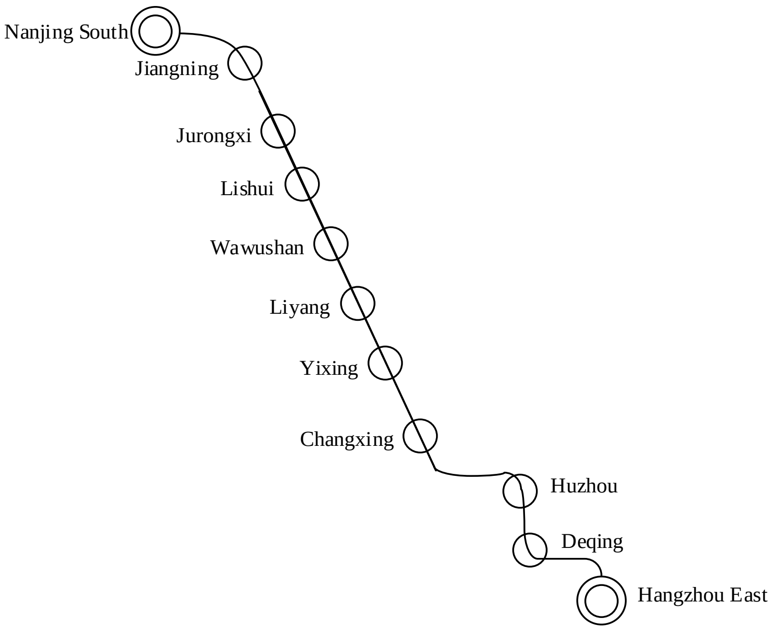


Fig. 6. Ninghang railway track network

[7] National Railway Administration, Notice of the National Development and Reform Commission on the reform and improvement of passenger fare policies for high-speed rail EMUs. Accessed on 10 August 2016. <http://www.nra.gov.cn/jgzf/flfg/gfxwj/zt/other/201602/t20160222_21192.shtml>

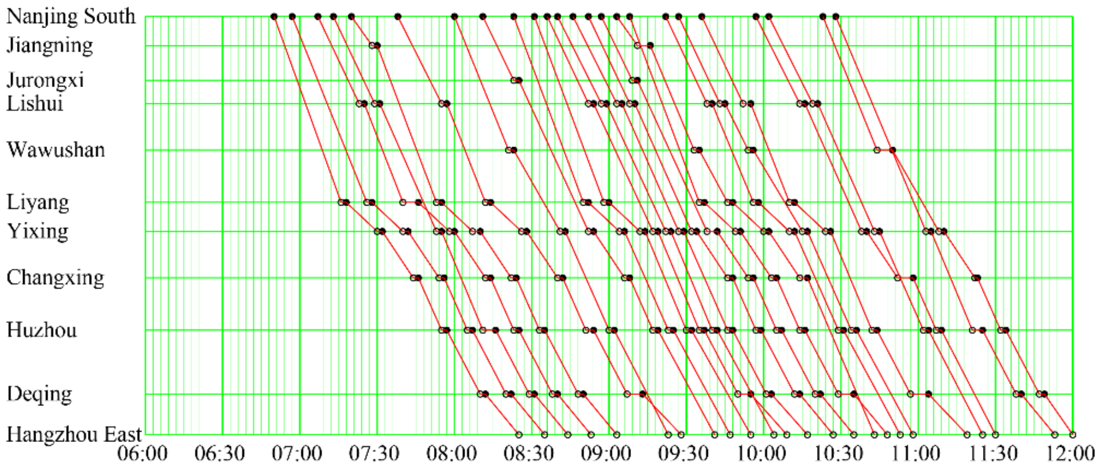


Fig. 7. The train running diagram for Ninghang railway track network under a given schedule

Table 10. The price bounds for all OD pairs of Ninghang railway network (Unit: CNY)

Lower/Upper price	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	12/24	32/64	46/92	68/136	98/196	129/258	165/330	185/370	221/442	256/512
S_2	0	0	20/40	34/68	56/112	86/172	117/234	153/306	173/346	209/418	244/488
S_3	0	0	0	26/52	48/96	78/156	109/218	145/290	165//330	201/402	236/472
S_4	0	0	0	0	22/44	52/104	83/166	119/238	139/278	175/350	210/420
S_5	0	0	0	0	0	30/60	61/122	97/194	117/234	153/306	188/376
S_6	0	0	0	0	0	0	31/62	67//134	87/174	123/246	158/316
S_7	0	0	0	0	0	0	0	36/72	56/112	92/184	127/254
S_8	0	0	0	0	0	0	0	0	20/40	56/112	91/182
S_9	0	0	0	0	0	0	0	0	0	36/72	71/142
S_{10}	0	0	0	0	0	0	0	0	0	0	35/70
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 11. Potential OD demand for the Ninghang railway network

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	2700	3450	8400	4500	7500	15600	9300	15300	9900	6900
S_2	0	0	300	1500	900	3150	1200	900	2400	2100	1800
S_3	0	0	0	1200	1500	900	2100	1800	1200	300	2400
S_4	0	0	0	0	300	2700	900	2400	2400	1800	600
S_5	0	0	0	0	0	300	900	300	2700	300	300
S_6	0	0	0	0	0	0	2700	1800	1500	1200	900
S_7	0	0	0	0	0	0	0	12000	11400	3600	3300
S_8	0	0	0	0	0	0	0	0	3600	1800	2400
S_9	0	0	0	0	0	0	0	0	0	6900	1200
S_{10}	0	0	0	0	0	0	0	0	0	0	9300
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 12. Parameters in the elastic demand function for all OD pairs in the Ninghang railway network example

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	0.0380	0.0170	0.0180	0.0170	0.0160	0.0190	0.0140	0.0185	0.0075	0.0065
S_2	0	0	0.0260	0.0370	0.0392	0.0082	0.0081	0.0075	0.0093	0.0094	0.0095
S_3	0	0	0	0.0265	0.0371	0.0282	0.0058	0.0085	0.0092	0.0088	0.0076
S_4	0	0	0	0	0.0183	0.0385	0.0068	0.0089	0.0087	0.0075	0.0077
S_5	0	0	0	0	0	0.0182	0.0171	0.0171	0.0075	0.0175	0.0082
S_6	0	0	0	0	0	0	0.0188	0.0085	0.0082	0.0075	0.0068
S_7	0	0	0	0	0	0	0	0.0181	0.0195	0.0075	0.0063
S_8	0	0	0	0	0	0	0	0	0.0184	0.0171	0.0075
S_9	0	0	0	0	0	0	0	0	0	0.0182	0.0076
S_{10}	0	0	0	0	0	0	0	0	0	0	0.0188
S_{11}	0	0	0	0	0	0	0	0	0	0	0

We implemented the proposed method to solve the MPPS model for this Ninghang railway example. The optimal pricing and seat allocation solution is presented in Tables 13-15. We take the train departing at 10:02 as an example to illustrate the seat allocation solution (refer to Table 16 and Table 17). From Table 16 and Table 17, one can verify that the seat allocation scheme for this train meets the seat capacity constraints of trains (the solution is feasible). Fig. 8 further shows the convergence process of lower and upper bounds (converged after 29 iterations given the tolerance value $\varepsilon = 5 \times 10^{-4}$ for convergence check) when using the proposed method. The total CPU time is 677.723s under the tolerance value $\varepsilon = 5 \times 10^{-4}$ for convergence check. It is about six times of that for solving the MPPS_1 model with continuous variables (110.220s).

Table 13. Optimal rail service prices for all OD pairs (MPPS model for Ninghang railway example)

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	24.00	58.98	55.36	68.00	98.00	129.00	165.00	185.00	221.00	256.00
S_2	0	0	40.00	68.00	56.00	86.00	117.00	153.00	173.00	209.00	244.00
S_3	0	0	0	52.00	96.00	156.00	173.00	145.00	165.00	201.00	236.00
S_4	0	0	0	0	44.00	52.00	83.00	119.00	139.00	175.00	210.00
S_5	0	0	0	0	0	54.94	61.00	97.00	134.00	153.00	188.00
S_6	0	0	0	0	0	0	53.17	118.03	121.90	133.72	158.00
S_7	0	0	0	0	0	0	0	55.05	56.00	133.18	158.75
S_8	0	0	0	0	0	0	0	0	40.00	58.28	133.85
S_9	0	0	0	0	0	0	0	0	0	55.05	131.68
S_{10}	0	0	0	0	0	0	0	0	0	0	53.11
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 14. The OD-specific optimal seat allocation scheme (MPPS model for Ninghang railway example)

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	1444	1522	6950	2000	3541	1344	923	2696	1887	1693
S_2	0	0	0	0	57	49	4	61	85	294	177
S_3	0	0	0	0	0	0	899	123	50	51	399
S_4	0	0	0	0	134	2000	2554	77	716	484	119
S_5	0	0	0	0	0	110	1928	57	988	20	64
S_6	0	0	0	0	0	0	4065	660	552	723	307
S_7	0	0	0	0	0	0	0	4430	3825	1326	1213
S_8	0	0	0	0	0	0	0	0	3862	664	1805
S_9	0	0	0	0	0	0	0	0	0	6173	6601
S_{10}	0	0	0	0	0	0	0	0	0	0	10741
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 15. The OD demand pattern under the optimal pricing and seat allocation solution (MPPS model for Ninghang railway example)

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	1084	1265	3101	1416	1563	1344	923	499	1887	1306
S_2	0	0	0	0	57	49	4	61	85	294	177
S_3	0	0	0	0	0	0	769	123	50	51	399
S_4	0	0	0	0	134	364	223	77	716	484	119
S_5	0	0	0	0	0	110	317	57	988	20	64
S_6	0	0	0	0	0	0	993	660	552	440	307
S_7	0	0	0	0	0	0	0	4430	3825	1326	1213
S_8	0	0	0	0	0	0	0	0	1724	664	879
S_9	0	0	0	0	0	0	0	0	0	2533	441
S_{10}	0	0	0	0	0	0	0	0	0	0	3426
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 16. The seat allocation scheme of the train departing at 10:02 (MPPS model for the Ninghang railway example)

OD pair	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
S_1	0	0	0	0	0	0	183	0	817	0	0
S_2	0	0	0	0	0	0	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	0	0
S_4	0	0	0	0	0	0	0	0	0	0	0
S_5	0	0	0	0	0	0	0	0	0	0	0
S_6	0	0	0	0	0	0	0	0	0	0	0
S_7	0	0	0	0	0	0	0	0	183	0	0
S_8	0	0	0	0	0	0	0	0	0	0	0
S_9	0	0	0	0	0	0	0	0	0	0	1000

S_{10}	0	0	0	0	0	0	0	0	0	0	0
S_{11}	0	0	0	0	0	0	0	0	0	0	0

Table 17. The seat allocation of each section of the train departing at 10:02 (MPPS model for the Ninghang railway example)

Section no.	Section	Section capacity	Section allocation b_l^k
1	(S_1, S_2)	1000	1000
2	(S_2, S_3)	1000	1000
3	(S_3, S_4)	1000	1000
4	(S_4, S_5)	1000	1000
5	(S_5, S_6)	1000	1000
6	(S_6, S_7)	1000	1000
7	(S_7, S_8)	1000	1000
8	(S_8, S_9)	1000	1000
9	(S_9, S_{10})	1000	1000
10	(S_{10}, S_{11})	1000	1000

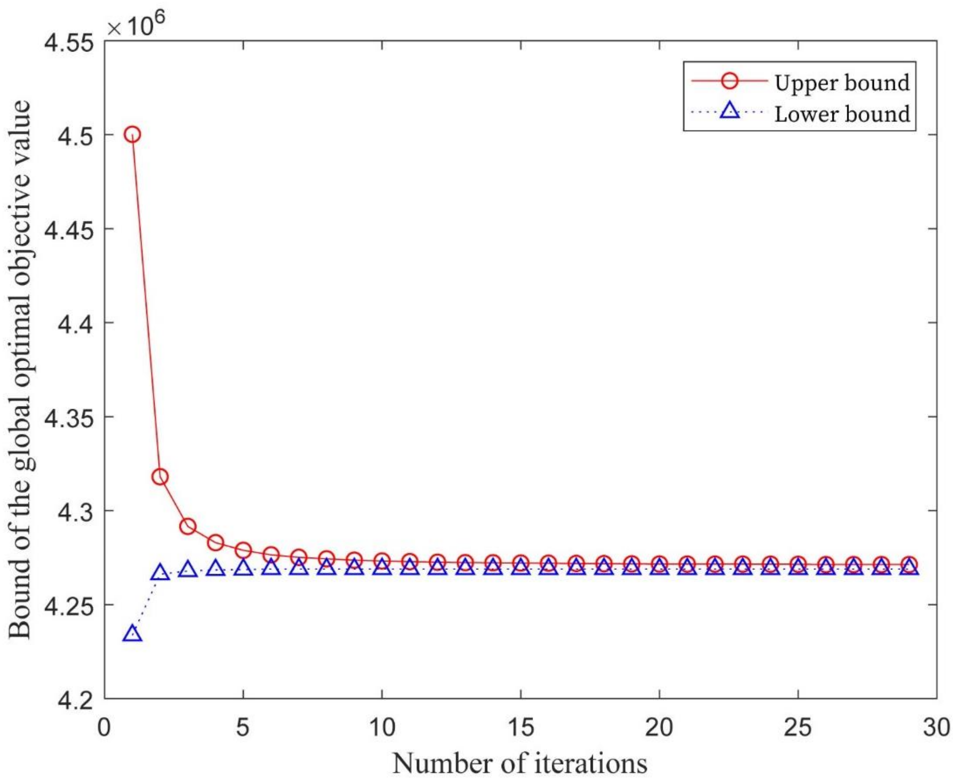


Fig. 8. Update process of upper and lower bounds (MPPS model for the Ninghang railway example)

In the above analysis, we used the train schedules (from 6:00 am to 12:00 am) on October 20, 2019 for the Ninghang railway example. We also tested how the computation time might vary against the number of trains in the network and shown in Fig. 9. When the number of trains is small, we see the CPU time varies and does not necessarily increase with the number of trains. When the number of trains is relatively large (greater than 18), the CPU time increases with the number of trains.

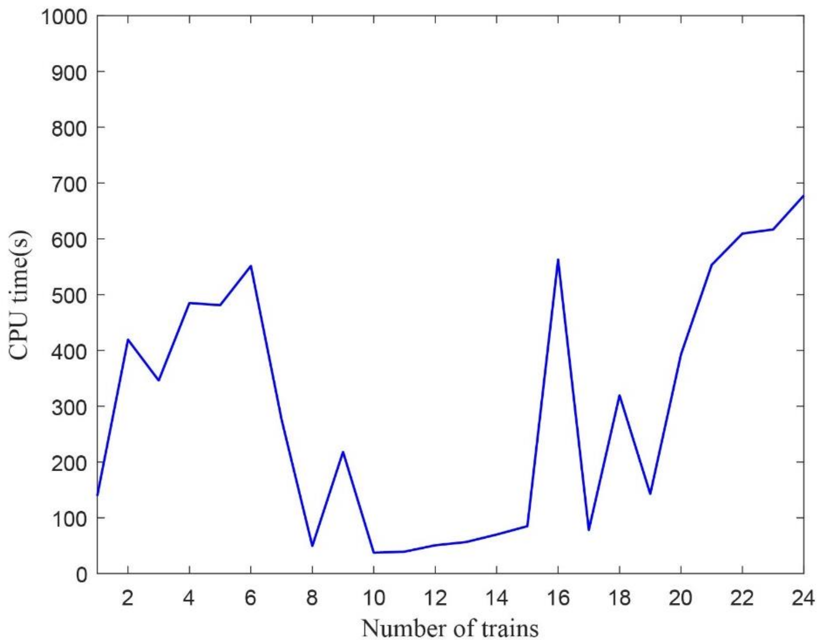


Fig. 9. The variation of the CPU time against the number of trains

6. Conclusion

In this paper, an optimization model for jointly optimizing railway service pricing and seat allocation scheme is introduced and a solution algorithm that produces the globally optimal solution is proposed. The objective is to maximize the ticket revenue of railway network considering elastic demand and multiple trains with multiple stopping patterns, where the demand decreases with respect to the service price. In order to find the globally optimal solution, the objective function and some constraints of the optimization model for railway service pricing and seat allocation are linearized, while only a few constraints involving logarithm functions are still nonlinear. With the relaxation of these logarithm functions, the linearized model is further relaxed as a mixed-integer programing problem (MILP). By coupling the relaxed MILP with a range reduction scheme, a solution algorithm is then designed, and its convergence to the global optimum is illustrated.

This study can be further extended in several ways. Firstly, this study assumes that the demand function is known and deterministic. A future study may further consider the case with stochastic demand (An and Lo, 2015), where we may only know the distribution of the demand given the train service price and level of service. We expect that a robust optimization approach has to be adopted, where similar linear approximation techniques in this paper can still be used in sub-problems of the main robust optimization problem. Moreover, when the demand information is not fully known in advance, a rolling horizon approach might be developed to accommodate real-time inputs. Secondly, this paper assumes that the rail service price for the same OD pair is independent of ticket booking time (this reflects the current practice in China). The proposed method in this paper can be further extended for cases with ticket-booking-time-dependent demand (Niu and Zhou, 2013; Niu et al., 2015). Last but not least, this study jointly considers pricing and seat allocation while a future study can further explore the joint optimization problem of pricing, seat allocation, line planning and scheduling. We expect that the additional interactions among pricing, seat allocation, line planning and scheduling due to the further consideration of line planning and scheduling will add further complexity for both the model formulation and solution approach.

Acknowledgement. We would like to thank the anonymous referees for their valuable comments on an earlier version of this paper, which helped us to improve both the technical quality and exposition of this paper. This study was supported by grants from the National Natural Science Foundation of China (72171236, 71701216, 71871226), the Natural Science Foundation of Hunan Province (2020JJ5783), and the Australian Research Council through the Discovery Early Career Researcher Award (DE200101793).

Appendix A. Illustration of the non-concavity of the model MPPS (or Model M1)

For the model MPPS to be concave, its objective function should be concave, i.e., the Hessian matrix of the objective function should be positive definite (Boyd et.al 2004). We can write down the Hessian matrix of the objective function as follows:

$$A = \begin{bmatrix} \frac{\partial Z^2}{\partial x_1 \partial x_1} & \cdots & \frac{\partial Z^2}{\partial x_1 \partial x_W} & \frac{\partial Z^2}{\partial x_1 \partial p_1} & \cdots & \frac{\partial Z^2}{\partial x_1 \partial p_W} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Z^2}{\partial x_W \partial x_1} & \cdots & \frac{\partial Z^2}{\partial x_W \partial x_W} & \frac{\partial Z^2}{\partial x_W \partial p_1} & \cdots & \frac{\partial Z^2}{\partial x_W \partial p_W} \\ \frac{\partial Z^2}{\partial p_1 \partial x_1} & \cdots & \frac{\partial Z^2}{\partial p_1 \partial x_W} & \frac{\partial Z^2}{\partial p_1 \partial p_1} & \cdots & \frac{\partial Z^2}{\partial p_1 \partial p_W} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Z^2}{\partial p_W \partial x_1} & \cdots & \frac{\partial Z^2}{\partial p_W \partial x_W} & \frac{\partial Z^2}{\partial p_W \partial p_1} & \cdots & \frac{\partial Z^2}{\partial p_W \partial p_W} \end{bmatrix}_{2W \times 2W}$$

$$= \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \\ 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \end{bmatrix}_{2W\times 2W}$$

For the above Hessian matrix, one can verify that its eigenvalue set $\lambda(A) = \{-1, -1, \dots, -1, 1, 1, \dots, 1\}_{2W}$, and thus the matrix is not positive definite. Therefore, the objection function of the model MPPS is not concave, and thus the model is non-concave.

Appendix B. Proof of Proposition 1

We prove Proposition 1 by a contradiction. Without loss of generality, we assume that h_w^* is within $[\underline{h}_w^m, \underline{h}_w^{m+1})$, then

$$F_1(P^*, B^*) > \max\{F_1(P^m), F_1(B^m)\}$$

It is obvious that (P^*, B^*) is also a feasible solution of M2, then

$$F_2(P^*, B^*, \widetilde{H}^m) \geq F_1(P^*, B^*)$$

We then have

$$F_2(P^*, B^*, \widetilde{H}^m) > \max\{F_1(P^m), F_1(B^m)\}$$

As $h_w^* < \underline{h}_w^{m+1}$, it contradicts to the fact that $\underline{h}_w^{m+1}$ is a solution of model M3 with the constraint in Eq. (72). Therefore, Proposition 1 is true. □

Appendix C. Proof of Proposition 2

To provide some intuitions, the updating of chord curves is illustrated in Fig. C1 with the updating of breakpoint set. Fig. C1 shows that the region defined by the red solid curve chords and the curve of $\ln(h_w)$ is a subset of that defined by the black dashed curve chords and the curve of $\ln(h_w)$. The updating of tangential supports is similar, i.e., the region defined by the tangential supports (the short dash lines in Fig.C1) with the updating breakpoint set and the curve of $\ln(h_w)$ is a subset of that defined the tangential supports with the previous breakpoint set and the curve of $\ln(h_w)$. Thus, with the above method of updating the set of breakpoints, $\Omega_m \supset \Omega_{m+1}$. As $\Omega_m \supset \Omega_{m+1}$, then $F_2(P^m, B^m, \widetilde{H}^m) \geq F_2(P^{m+1}, B^{m+1}, \widetilde{H}^{m+1})$. Therefore, Proposition 2 is true. □

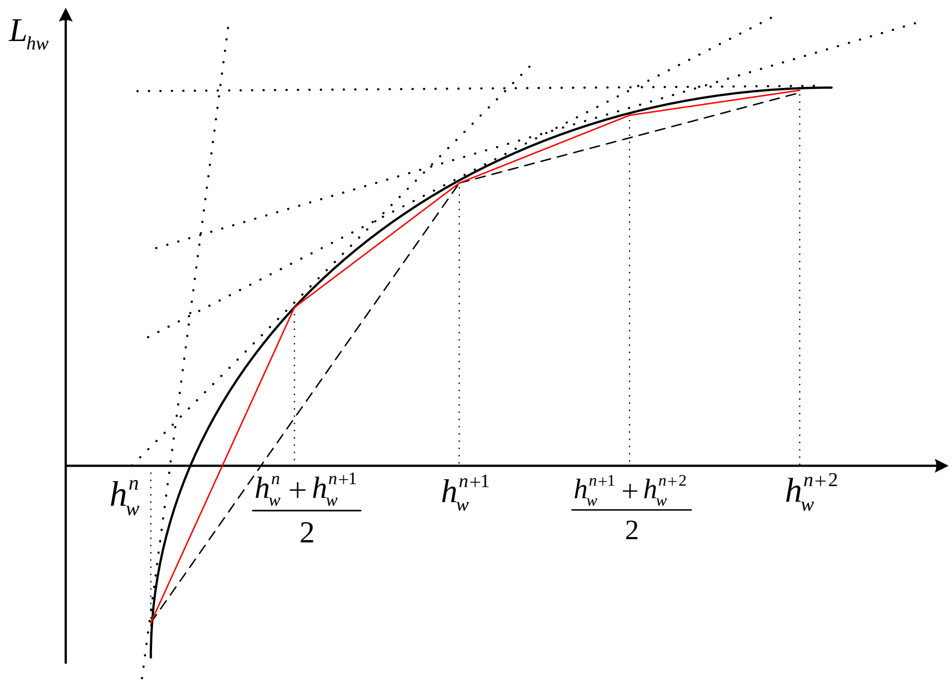


Fig. C1. The illustration of breakpoints updating

Appendix D. Proof of Proposition 3

The proof here follows a similar logic as that in Wang et al. (2015) and Ng (2017). Denote the optimal objective function value of the MPPS model (or model M1) by Z_1^*. As the RMILP model (or model M2) have a larger solution space than that of the MPPS model (or model M1), the objective function value $F_2\left(P^m, B^m, \widetilde{H}^m\right)$ solved based on the RMILP model (or model M2) is always the upper bound of the MPPS model (or model M1), i.e., $F_2\left(P^m, B^m, \widetilde{H}^m\right)$ is no less than Z_1^*, where P^m and B^m are the corresponding rail service pricing and seat allocation solution. With Proposition 2, we know that the set of optimal objective function values $\left\{F_2\left(P^m, B^m, \widetilde{H}^m\right)\right\}$ is a monotonically decreasing series with respect to the iteration number m. From the algorithm in Section 4.4, we have $\bar{Z}_1^m = \min\left\{\bar{Z}_1^{m-1}, F_2\left(P^m, B^m, \widetilde{H}^m\right)\right\} \geq Z_1^*$, so $\{\bar{Z}_1^m\}$ is also a monotonically decreasing series where the following holds

$$\bar{Z}_1^1 \geq \bar{Z}_1^2 \geq \cdots \geq \bar{Z}_1^m \geq \cdots \geq Z_1^*$$

Moreover, with the increasing of the iteration number m, Eqs. (37)-(44) will drive L_{hw} to approach $\ln(h_w)$ and the solution of RMPPS model will approach that of the original MPPS model. If the optimal solution is still not obtained, then the solution of P^m and B^m in RMPPS will be updated with the range reduction technique and the updating of breakpoint sets in Step 3, and the proposed algorithm can update the bounds

$\bar{Z}_1^m$. Thus, when the number of iterations approaches infinity, we have $\lim_{m\to\infty} \bar{Z}_1^m = Z_1^*$, and (P^m, B^m) will approach the optimal solution (P^*, B^*).

We also know the objective function value $F_1(P^m)$ and $F_1(B^m)$ are solved under given P^m and B^m, respectively, so $F_1(P^m)$ and $F_1(B^m)$ are both lower bounds of the MPPS model (or model M1), i.e., they are always no greater than Z_1^* based on Eq. (57). From the algorithm in Section 4.4, we also have $\underline{Z}_1^m = \max\{\underline{Z}_1^{m-1}, F_1(P^m), F_1(B^m)\} \le Z_1^*$, which further implies that $\{\underline{Z}_1^m\}$ is a monotonically increasing series, i.e.,

$$\underline{Z}_1^1 \le \underline{Z}_1^2 \le \cdots \le \underline{Z}_1^m \le \cdots \le Z_1^*$$

Moreover, one can verify that if $P^m = P^*$, then $(P^m, \tilde{B}^m)$ is an optimal solution of the MPPS model (or model M1); and if $B^m = B^*$, then $(\tilde{P}^m, B^m)$ is also an optimal solution of the MPPS (or model M1). Thus, when the number of iterations approaches infinity, we have $\lim_{m\to\infty} \underline{Z}_1^m = Z_1^*$. The above indicates that the proposed method will yield lower bound and upper bound that converge to the exact global optimal solution of the original model MPPS (or model M1). □

Appendix E. Non-integer solutions (for MPPS_1) obtained by "FMINCON" for the toy network example

Table E1. MPPS_1 solution solved by the trust region reflective or interior-point algorithm in FMINCON

OD pair	p_w	x_w	$Q_w(p_w)$	b_w	b_w^k			
					b_w^1	b_w^2	b_w^3	b_w^4
(S_1, S_2)	99.99	36.56	201.75	964.93	310.22	369.46	0	285.25
(S_1, S_3)	159.98	22.42	101.63	464.03	78.84	0	385.19	0
(S_1, S_4)	140.09	25.64	198.19	138.10	0.38	23.42	0	114.29
(S_1, S_5)	385.63	9.19	9.19	15.09	0.30	0	14.70	0.09
(S_1, S_6)	434.82	7.47	7.47	7.74	0.69	6.87	0	0.17
(S_2, S_3)	59.99	61.41	333.69	313.88	313.88	0	0	0
(S_2, S_4)	179.91	19.93	84.42	243.09	4.84	31.56	0	206.69
(S_2, S_5)	317.37	11.30	23.51	18.89	0.26	0	0	18.63
(S_2, S_6)	417.39	8.48	8.48	22.27	0.25	20.59	0	1.43
(S_3, S_4)	119.99	17.43	123.11	321.20	321.19	0	0	0
(S_3, S_5)	258.36	13.87	48.49	100.03	11.34	0	88.70	0
(S_3, S_6)	324.18	10.93	10.93	55.89	0.15	0	55.74	0
(S_4, S_5)	139.99	25.75	163.21	451.19	263.32	0	0	187.87
(S_4, S_6)	239.42	14.99	92.56	570.15	121.77	366.02	0	82.36
(S_5, S_6)	99.99	36.55	320.60	934.14	276.95	0	344.21	312.97

Table E2. MPPS_1 solution solved by the SQP algorithm in FMINCON

OD pair	p_w	x_w	$Q_w(p_w)$	b_w	b_w^k			
					b_w^1	b_w^2	b_w^3	b_w^4
(S_1, S_2)	91.74	220.73	220.73	220.73	141.33	79.21	0.00	0.19
(S_1, S_3)	107.53	165.55	165.55	260.52	35.93	0.00	224.59	0.00
(S_1, S_4)	140.00	198.36	198.36	332.15	0.03	71.19	0.00	260.93
(S_1, S_5)	210.00	93.81	93.81	93.81	0.00	0.00	56.59	37.22
(S_1, S_6)	260.00	47.66	47.66	47.67	0.00	47.67	0.00	0.00
(S_2, S_3)	60.00	333.64	333.64	333.64	333.64	0.00	0.00	0.00
(S_2, S_4)	91.74	220.73	220.73	245.19	19.88	225.31	0.00	0.00
(S_2, S_5)	160.00	101.62	101.62	101.83	0.00	0.00	0.00	101.83
(S_2, S_6)	210.00	61.21	61.21	61.21	10.46	50.75	0.00	0.00
(S_3, S_4)	75.76	220.73	220.73	296.68	296.68	0.00	0.00	0.00
(S_3, S_5)	130.00	189.06	189.06	333.76	0.00	0.00	333.76	0.00
(S_3, S_6)	180.00	79.23	79.23	79.23	69.58	0.00	9.65	0.00
(S_4, S_5)	107.53	220.73	220.73	424.84	319.96	0.00	0.00	104.88
(S_4, S_6)	120.00	287.84	287.84	305.19	0.00	301.56	0.00	3.63
(S_5, S_6)	75.76	441.46	441.46	1077.14	290.42	0.00	390.35	396.37

References

Abe, I., 2007. Revenue management in the railway industry in Japan and Portugal: a stakeholder approach. Technology and Policy Program. MIT, MA.

An, K. and Lo, H.K., 2015. Robust transit network design with stochastic demand considering development density. Transportation Research Part B: Methodological, 81, 737-754.

Armstrong, A. and Meissner, J., 2010. Railway revenue management: overview and models. Tech. Rep. Lancaster University Management School.

Belobaba, P.P., 1987. Airline yield management an overview of seat inventory control. Transportation Science, 21(2), 63-73.

Bertsimas, D. and de Boer, S., 2002. Joint network pricing and resource allocation. Working paper. https://www.mit.edu/~dbertsim/papers.html.

Boyd, S., Boyd, S. P., & Vandenberghe, L, 2004. Convex optimization. Cambridge university press.

Canca, D., De-Los-Santos, A., Laporte, G. and Mesa, J.A., 2019. Integrated railway rapid transit network design and line planning problem with maximum profit. Transportation Research Part E: Logistics and Transportation Review, 127, 1-30.

Chew, E.P., Lee, C. and Liu, R., 2009. Joint inventory allocation and pricing decisions for perishable products. International Journal of Production Economics, 120(1), 139-150.

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

Cizaire, C., 2011. Optimization Models for Joint Airline Pricing and Seat Inventory Control: Multiple Products, Multiple Periods. Ph.D. Dissertation. Department of Aeronautics and Astronautics. Massachusetts Institute of Technology.

Cote, J., Marcotte, P. and Savard, G., 2003. A bilevel modeling approach to pricing and fare optimization in the airline industry. Journal of Revenue and Pricing Management 2(1), 23.

Fard, F.A., Sy, M. and Ivanov, D., 2019. Optimal overbooking strategies in the airlines using dynamic programming approach in continuous time. Transportation Research Part E: Logistics and Transportation Review, 128, 384-399.

Flötteröd, G., Bierlaire, M., and Nagel, K., 2011. Bayesian demand calibration for dynamic traffic simulations. Transportation Science, 45(4), 541-561.

Hetrakul, P. and Cirillo, C., 2014. A latent class choice based model system for railway optimal pricing and seat allocation. Transportation Research Part E: Logistics and Transportation Review, 61, 68-83.

Hu, X., Shi, F., Xu, G. and 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., Zhang, L. and 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.

Jiao, J., Wang, J., Zhang, F., Jin, F. and Liu, W., 2020. Roles of accessibility, connectivity and spatial interdependence in realizing the economic impact of high-speed rail: Evidence from China. Transport Policy, 91, 1-15.

Kuyumcu, A. and Garcia-Diaz, A., 2000. A polyhedral graph theory approach to revenue management in the airline industry. Computers and Industrial Engineering 38 (3), 375–395.

Li, C., Ma, J., Luan, T.H., Zhou, X. and Xiong, L., 2018. An incentive-based optimizing strategy of service frequency for an urban rail transit system. Transportation Research Part E: Logistics and Transportation Review, 118, 106-122.

Li, Z., Shalaby, A., Roorda, M.J. and Mao, B., 2021. Urban rail service design for collaborative passenger and freight transport. Transportation Research Part E: Logistics and Transportation Review, 147, 102205.

Li, Z.C., Lam, W.H., Wong, S.C. and Sumalee, A., 2012. Design of a rail transit line for profit maximization in a linear transportation corridor. Transportation Research Part E: Logistics and Transportation Review, 48(1), 50-70.

Liu, H., Szeto, W.Y. and Long, J., 2019. Bike network design problem with a path-size logit-based equilibrium constraint: Formulation, global optimization, and matheuristic. Transportation Research Part E: Logistics and Transportation Review, 127, 284-307.

Liu, H. and Wang, D.Z., 2015. Global optimization method for network design problem with stochastic user equilibrium. Transportation Research Part B: Methodological, 72, 20-39.

Luo, H., Nie, L. and He, Z., 2016. Modeling of multi-train seat inventory control based on revenue management. In Logistics, Informatics and Service Sciences (LISS), 2016 International Conference on (pp. 1-6). IEEE.

McGill, J.I. and Van Ryzin, G. J., 1999. Revenue management: Research overview and prospects. Transportation Science, 33(2), 233-256.

Niu, H. and Zhou, X., 2013. Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. Transportation Research Part C: Emerging Technologies, 36, 212-230.

Niu, H., Zhou, X. and Gao, R., 2015. Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: Nonlinear integer programming models with linear constraints. Transportation Research Part B: Methodological, 76, 117-135.

Ng, K.F., 2017. Sustainable housing and railway developments over space and time. Ph.D. Dissertation. Civil Engineering. The Hong Kong University of Science and Technology, Hong Kong, China.

Nocedal, J. and Wright, S., 2006. Numerical optimization. Springer Science & Business Media.

Ongprasert, S., 2006. Passenger behavior on revenue management systems of inter-city transportation. Ph.D. Dissertation. Graduate School of Engineering. Kochi University of Technology, Japan.

Osorio, C., 2019. High-dimensional offline origin-destination (OD) demand calibration for stochastic traffic simulators of large-scale road networks. Transportation Research Part B: Methodological, 124, 18-43.

Qiu, X, and Lee, C.Y., 2019. Quantity discount pricing for rail transport in a dry port system. Transportation Research Part E: Logistics and Transportation Review, 122, 563-580.

Sancho, F., 2009. Calibration of CES functions for real-world multisectoral modeling. Economic Systems Research, 21(1), 45-58.

Shi, F., Xu, G.M., Liu, B. and Huang, H., 2014. Optimization method of alternate traffic restriction scheme based on elastic demand and mode choice behavior. Transportation Research Part C: Emerging Technologies, 39, 36-52.

Subramanian, J., Stidham, S. and Lautenbacher, C.J., 1999. Airline yield management with overbooking, cancellations, and no-shows. Transportation Science, 33(2), 147-167.

Talluri, K.T. and van Ryzin, G.J., 2004. The Theory and Practice of Revenue Management. Springer, New York.

Terciyanlı, E. and Avşar, Z. M., 2019. Alternative risk-averse approaches for airline network revenue management. Transportation Research Part E: Logistics and Transportation Review, 125, 27-46.

Tong, C. and Topaloglu, H., 2014. On the approximate linear programming approach for network revenue management problems. INFORMS Journal on Computing, 26(1), 121-134.

Wang, D.Z. and Lo, H.K., 2010. Global optimum of the linearized network design problem with equilibrium flows. Transportation Research Part B: Methodological, 44(4), 482-492.

Wang, D.Z., Liu, H. and Szeto, W.Y., 2015. A novel discrete network design problem formulation and its global optimization solution algorithm. Transportation Research Part E: logistics and Transportation Review, 79, 213-230.

Wang, X., Wang, H. and 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.

Wang, Y., Lan, B.X. and Zhang, L., 2012. A revenue management model for high-speed railway. In: Ni, Y.-Q., Ye, X.-W. (Eds.), Proceedings of the 1st International Workshop on High-Speed and Intercity Railways, Lecture Notes in Electrical Engineering, vol. 147. Springer, Berlin, Heidelberg, pp. 95–103 (Chapter 9).

Weatherford, L.R., 1997. Using prices more realistically as decision variables in perishable-asset revenue management problems. Journal of Combinatorial Optimization, 1(3), 277-304.

Xu, G., Liu, W., Wu, R. and Yang, H., 2021. A double time-scale passenger assignment model for high-speed railway networks with continuum capacity approximation. Transportation Research Part E: Logistics and Transportation Review, 150, 102305.

Xu, G., Liu, W. and Yang, H., 2018a. A reliability-based assignment method for railway networks with heterogeneous passengers. Transportation Research Part C: Emerging Technologies, 93, 501-524.

Xu, G., Yang, H., Liu, W., and Shi, F., 2018b. Itinerary choice and advance ticket booking for high-speed-railway network services. Transportation Research Part C: Emerging Technologies, 95, 82-104.

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

Yang, Z., Li, C., Jiao, J., Liu, W. and Zhang, F., 2020. On the joint impact of high-speed rail and megalopolis policy on regional economic growth in China. Transport Policy, 99, 20-30.

You, P. C., 2008. An efficient computational approach for railway booking problems. European Journal of Operational Research, 185, 811-824.

Yuan, W., Nie, L., Wu, X. and Fu, H., 2018. A dynamic bid price approach for the seat inventory control problem in railway networks with consideration of passenger transfer. PLoS One, 13(8), e0201718.

Zhan, S., Wong, S. C., and 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. and 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.

Zhou, Y., Yang, H., Wang, Y. and Yan, X., 2021. Integrated line configuration and frequency determination with passenger path assignment in urban rail transit networks. Transportation Research Part B: Methodological, 145, 134-151.