Application of a Seat-based Booking Control Mechanism in Rail Transport with Customer Diversion

Ahmet Z. Bulum
Department of Industrial Engineering
Karabuk University
Karabük, Turkey
ahmetbulum@karabuk.edu.tr

Muharrem Dugenci
Department of Industrial Engineering
Karabuk University
Karabük, Turkey
mdugenci@karabuk.edu.tr

Mumtaz Ipek
Department of Industrial Engineering
Sakarya University
Sakarya, Turkey
ipek@sakarya.edu.tr

Received: 1 July 2022 | Accepted: 10 July 2022

Abstract-The ticket booking control mechanism is a part of the Revenue Management (RM), commonly used in the airline industry. This study aims to optimize seat allocation in the railway industry and compare the performance of three booking control techniques by considering customer behavior. The preferences of customers who cannot find their desired ticket are considered as a customer diversion matrix, which also includes waiting and no-purchase probability. A Ticket Booking System (TBS) with buckets, which assigns seats to buckets, was adapted and implemented on the Turkish railway for the first time. A genetic algorithm that is specifically written to apply the TBS, including customer diversion, is used in simulations to obtain approximate solutions. It is seen that TBS gave successful results with a revenue increase of around 5.8%. We can also suggest, considering customer behavior, that the revenue can be raised by sales in periods.

Keywords-revenue management; railway transportation; customer behavior; genetic algorithm; seat inventory control

I. INTRODUCTION

Revenue Management (RM) can improve the profitability of firms in the service industry. RM has been used for pricing, forecasting, seat inventory control, and overbooking in airlines [1]. As a part of RM, seat inventory control or seat allocation deals with the decision to allocate a finite seat capacity optimal to demands [2]. Authors in [5] compared a deterministic linear and a probabilistic nonlinear programming model for the network problem with non-nested seat allocation. Authors in [6] suggested a multi-train seat inventory control model to control the seat inventory capacity among different trains. Authors in [7] proposed a stochastic dynamic model with some properties related to the profit-to-go functions. Authors in [8] used three seat inventory control methods and modelled them under a single-fare class. With the spread of High-Speed Railway (HSR), RM's applicability in the railway industry has also increased [5, 10]. In the railway industry, RM is utilized as a management technique that aims to maximize the total revenue by finding an optimal strategy controlling the availability and/or the price of train tickets without any changes in the resources (e.g. adding trains or carriages) [11]. Articles regarding capacity allocation and prices in railway transport are mostly originated from countries where HSR is widely used [6, 8, 12–22]. While train tickets were sold formerly in railway stations or by authorized agents, online ticket reservation has become preferable nowadays. According to the classification in [23], there are different types of ticket booking mechanisms, such as Virtual Nesting (VN) and Partitioned Booking Limit Control (PBLC). VN, widely used in the airline industry, uses single-source nested controls to allocate each resource in the network. In practice, both the indexing process of VN and the control logic of mapping products to virtual classes is complex [23]. PBLC is a booking mechanism that allocates a fixed amount of capacity on each resource for every product offered, i.e. the seats are assigned to the tickets at the beginning of the booking horizon. Studies in which PBLC is used mainly aimed to find the optimal allocation of seats [5, 24, 25]. However, because of the fixed use of seats in PBLC, it is considered inefficient when the demand is stochastic and is seldom used in practice [23].

Customer behavior has been acknowledged as an important element for increasing revenue and has been included in modeling RM problems in some previous studies [26, 27]. Authors in [11] proposed a seat-based control mechanism and considered customers' purchase preferences as customer

Corresponding author: Ahmet Z. Bulum

behavior while optimizing seat allocation in China railway. Authors in [28, 29] proved that customer behavior is an essential factor affecting the ticket sale profit: the railway companies can collect more information with online reservation systems, such as the ticket booking time and purchase preferences of customers.

This research concentrates on evaluating the seat allocation performance of booking control mechanisms under diverse customer conditions. Unlike the previous studies, the diverted customer has more than two options: buying another ticket, waiting for an available seat, or leaving without buying. The customer diversion is utilized as a percentage matrix in this study. The ticket reservation process is utilized by the simulation method and a Genetic Algorithm (GA) is designed to cope with the complexity of the process. The methods were evaluated and compared through an experimental case and a real-world case using data from the Turkish railway. One contribution of this study is the application of customer diversion as customer behavior while optimizing seat allocation in the railway industry. Another contribution is the application of a ticket booking mechanism with buckets to the Turkish railway for the first time.

II. TICKET BOOKING MODELING WITH BUCKETS

Figure 1 shows an example of a railway line with three stops. In RM, resources are the seats in a train, and Resources 1 and 2 (Res1, Res2) consist of three seats. Products 1, 2 and 3 (Pr1, Pr2, Pr3) use the same seats but different resources. If a customer books a ticket for Pr3, for example Seat 1 (S1) will be occupied in Res1 and Res2. Then Res1 and Res2 have only two empty seats for booking, S2 and S3. But if the customer books a ticket for Pr1, then S1 will be removed only from Res1.

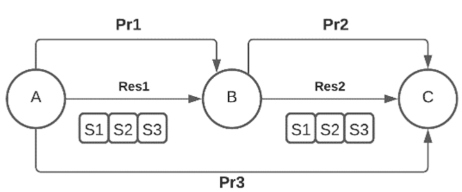

Fig. 1. A representative railway line with three stops.

A product is referred to the used resources, i.e. occupied seats in a train, between an Origin-Destination (O-D) pair. In the reservation process, customers send requests for products. After an accepted request, the seats related to that product will be assigned and occupied. Then different ticketing policies are involved: in PBLC a ticket will be printed immediately, but a ticket is generated before printing in the proposed ticket booking system with buckets (TBS). Modeling of the ticket booking process starts by allocating products and resources to the buckets. A bucket is a combination of tickets, and it contains a seat set and an offered product set. The bucket describes which products can be sold from the available resources [11]. The ticket pool is designed to collect the tickets generated for reusing. Some resources can remain unused after selling a ticket from a bucket. The remaining seats are assigned to appropriate products, which are collected in the ticket pool to sell additional tickets. Here, the exception is the ticket for travel from the first station to the last station.

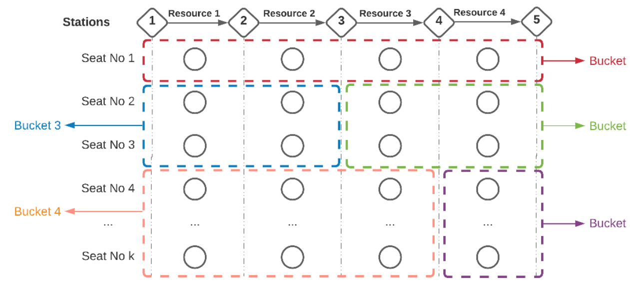

Fig. 2. The working mechanism of TBS.

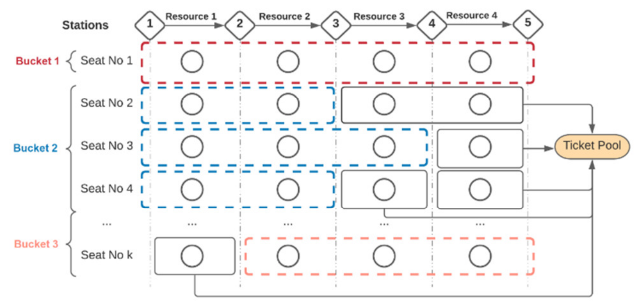

Fig. 3. The working mechanism of TBS.

Figures 2 and 3 show the different working mechanisms of PBLC and the proposed TBS. As seen in Figure 2, in PBLC all seats are assigned to some partitions and all tickets are generated and await their owner. However, in TBS, most seats are assigned to the buckets, and the remaining seats are sold by creating tickets according to the arriving passengers (Figure 3). The general assumptions of seat allocation problems are that the demand for cheap tickets is higher than for the expensive ones, and price-sensitive customers and long-distance travelers arrive earlier than the time-sensitive customers and short-distance travelers. The rules and assumptions made for TBS modeling are [11]: (1) The number of buckets is fixed during the ticket booking period, (2) a seat can be put in only a single bucket, (3) a product can relate to at most one bucket, (4) at the beginning of the reservation period, all the seats must be assigned to a bucket, (5) the ticket pool should be set to be empty before the ticket booking process starts, (6) the ticket pool has priority over the buckets for selling tickets, (7) demand for each product is assumed to be stochastic and representable by a probability distribution (mostly normal), (8) the number of seats is assumed to be fixed during the ticket booking period, and (9) no overbooking is considered. Tickets with lower prices to occupy the seats instead of more expensive ones are undesirable because airline or railway companies always prefer to sell seats at a higher price. Therefore, the allocation of products and seats in the buckets is the most critical factor affecting the profit.

III. FORMULATION AND SIMULATION STEPS

Two parameters must be set at the beginning of the booking horizon to start the ticket booking process: The number of seats reserved for each bucket rc_b and the product set for each bucket, which can be represented by the bucket-product relationship δ_{kb}. The value of δ_{kb} is assumed to be equal to 1 if product k is in bucket b, and 0 otherwise. All necessary notations are given in Table I.

TABLE I. NOTATIONS

C	Total capacity of the train
n_{Buc}	Total number of buckets
n_{Prod}	Total number of products
b	Bucket number (b=1, 2, ..., n_{Buc})
k	Product number (k=1, 2, ..., n_{Prod})
c_b	Seat capacity for bucket b
$\boldsymbol{\delta}$	Bucket-product relation vector ($\delta \in \{0,1\}$)
Pr	Product vector
Buc	Bucket vector
Rev	Total revenue or sale income
t	A small time interval of the ticket booking time horizon T
R_t	Revenue of a time interval t
dec_{ij}	Diversion matrix for products
div_k	Number of the diverted customers of product k
rc_b	Remaining capacity of bucket b
p	Price vector for products/tickets
u	Product/ticket availability vector ($u \in \{0,1\}$)
$dmnd$	Demand vector for products/tickets
m	Mean demand vector for products/tickets
s	Standard deviation vector of demand for products/tickets
$rest_k$	Number of customers who did not buy a ticket for product k yet
rs	Random number representing customer requests a ticket for prod rs
$sale$	Income/revenue vector from sales of each product
TK	Ticket pool that contains the unoccupied seats or tickets
qtk_k	Quantity of seats in ticket pool for product k
$sold$	Vector indicating the number of tickets sold from each product

The objective of the seat allocation problem is to find an optimal combination of parameters (rc, δ) that maximizes total revenue $Rev(rc, \delta)$. It can be written as a mathematical formulation as follows:

$$\max_{rc,\delta} Rev(rc, \delta) \quad (1)$$

$$\sum_{b}^{n_{Buc}} rc_b = C \quad (2)$$

$$\sum_{k}^{n_{Prod}} \delta_{kb} \le 1 \qquad \forall b \quad (3)$$

where $rc_b \in Z^+, \delta_{kb} = \{0,1\}, b = 1, \dots, n_{Buc.}$ The first constraint in (2) is for the capacity to ensure that each seat is placed in only one bucket. The second constraint (3) states that a product can be found in at most one bucket.

The ticket booking process begins when customers arrive to purchase tickets. In an appropriate sampling time, the customer arrival sequence can be approximated as a Bernoulli process. The ticket booking horizon is divided into t time intervals in the simulation process. If we assume that at most one customer can book a ticket in each time interval t, all the customers will arrive in the total ticket booking period T. Thus, $t = 1,2, \dots, T$ and T are equal to the total number of customers.

If R_t denotes the revenue from ticket sale in time interval t, then the total revenue $Rev(rc, \delta)$ is expected to be the sum of each revenue in all time intervals. $Rev(rc, \delta)$ can be written as:

$$Rev(rc, \delta) = Exp\{\textstyle\sum_{t=0}^{T} R_t | rc_{b,0} = c_b, qtk_{k,0} = 0\} \quad (4)$$

At the beginning of the booking process, the initial number of seats is determined and fixed, and the ticket pool is empty ($qtk_{k,0} = 0$ and $rc_{b,0} = c_b$ with $b = 1, \dots, n_{Buc}$ and $k = 1, \dots, n_{Prod}$). c_b represents the seat capacity of each bucket, $rc_{b,t}$ denotes the number of remaining seats in bucket b in time interval t and $qtk_{k,t}$ denotes the remaining number of product k in the ticket pool in time interval t.

Revenue in time interval t can be represented as:

$$R_t = \begin{cases} p_k \cdot u_{k,t}, & \text{if product } k \text{ is sold in period } t \\ 0, & \text{if no product is sold in t or } rest_{k,t} = 0 \end{cases} \quad (5)$$

Revenue R_t depends on the ticket price (p) and the available product (u). The price of product k (p_k) is fixed during the ticket booking process, and $u_{k,t}$ denotes the availability of product k for booking in time interval t. The value of $u_{k,t}$ is 1 if the product k is available for booking in time interval t, and 0 otherwise (8). $rest_{k,t}$ represents the remaining demand, that has not been processed yet, from product k in time interval t and its value can be determined as:

$$rest_{k,t} = \begin{cases} rest_{k,t} - 1 & \text{, if product } k \text{ } is \text{ sold in time period } t \\ rest_{k,t} + 1 & \text{, if } div_{k,t} \ge 1 \\ rest_{k,t} & \text{, otherwise} \end{cases} \quad (6)$$

$div_{k,t}$ denotes the number of diverted customers who could not find tickets for product k in time interval t. $div_{k,t}$ may increase by the customer diversion matrix dec_{ij} (will be discussed below) depending on the availability of product k in time interval t or decrease by 1 if the diverted customers find the ticket they want. Thus, the value of $div_{k,t}$ can be calculated by:

$$div_{k,t} = \begin{cases} div_{k,t} + dec_{kj} \mid j = 1, \dots, n_{\mathrm{Prod}}, \text{ if } u_{k,t} = 0 \text{ and} \\ \qquad \text{no product is sold in time period } t \\ div_{k,t} - 1, \text{ if product } k \text{ is sold in time period } t \end{cases} \quad (7)$$

The availability of a product for booking depends on whether it is offered by any bucket ($\sum rc_{b,t}.\delta_{kb,t} > 0$) or if it is in the ticket pool ($qtk_{k,t} > 0$).

$$u_{k,t} = \begin{cases} 1 & , if \ \sum rc_{b,t}.\delta_{kb,t} > 0 \text{ or } qtk_{k,t} > 0 \\ 0 & \text{, otherwise} \end{cases} \quad (8)$$

The remaining seat capacities (rc_b) and the remaining products in the ticket pool (qtk_k) in time interval t+1 depend on the sold product in time interval t. $rc_{b,t+1}$ and $qtk_{k,t+1}$ are determined as:

$$rc_{b,t+1} = \begin{cases} rc_{b,t} - 1 & \text{, if product k sold in time period t} \\ & \text{and } qtk_{k,t} = 0 \text{ and } \delta_{kb,t} = 1 \\ rc_{b,t} & \text{, otherwise} \end{cases} \quad (9)$$

$$qtk_{k,t+1} = \begin{cases} qtk_{k,t} - 1 & \text{, if product } k \text{ is sold in time} \\ & \text{period } t \text{ and } qtk_{k,t} > 0 \\ qtk_{k,t} + \varphi_{k,t} \ |\forall k & \text{, if product } x \text{ is sold in time} \\ & \text{period } t \text{ and } qtk_{x,t} = 0 \\ qtk_{k,t} & \text{, otherwise} \end{cases} \quad (10)$$

$\varphi_{k,t} \in \{0,1\}$ (for $k = 1, \dots, n_{Prod}$) shows whether there are any available products left for booking when product x is sold. After product x is sold, more than one products can be transferred to the ticket pool. In each time interval t, the maximum number of products to be transferred to the ticket pool is one less than the total number of products.

There are some restrictions in rail transport that are different from those in airlines. There can be as many passengers as the train capacity simultaneously during the whole trip, including the passengers who keep sitting, get on and off the train. These restrictions make the problem very difficult, especially as the number of stops and trains increases. So, revenues cannot be evaluated by simply calculating the difference between demand and capacity. Calculating the objective function by the booking process model can be difficult due to the curse of dimensionality. In addition, the fact that the number of products to be sold can be huge and it is challenging to express the parameter $\varphi_{k,t}$ analytically makes it very hard to find an absolute solution to the problem. Therefore, a GA was applied to solve the problem approximately, its steps will be described below.

A. Customer Diversion Matrix (CDM)

The author in [30] considered the buy-up possibility of customers, i.e. diversion. He defined diversion as the willingness of customers to purchase tickets in a different price class than they initially requested. Other studies with diversion are [31-33]. There are very few studies with customer diversion, and they are all related to the airline industry [34-37]. This study uses a CDM based on the customers' decisions. Table II shows an example of a CDM, which contains the diversion percentage of customers who decide to make a buy-up, wait, or buy nothing. CDM has as many rows as the number of products but has two more columns than the number of products. One of these two extra columns is for customers who prefer to wait and the other for customers who do not want to buy anything. It is not easy to create the dec_{ij} matrix. Whereas it was challenging to create an exact profile in the past, as most sales were made at the box offices in stations, nowadays, a lot of data about the booking process can be obtained thanks to online sales. dec_{ij} is the percentage of customers who want to buy a ticket from Product j because they could not buy a ticket from Product i. For instance, $dec_{23} = 0.4$ means that 40% of the customers who could not find any ticket from Product 2 decide to buy a ticket from Product 3. The total percentage of customers who decide to wait, buy-up, or buy nothing should always be 100%. In other words, the summation of each row in the diversion matrix never exceeds 100%. It is important to note here that in PBLC and First-Come First-Served (FCFS), no waiting list is used, as all tickets are generated before the reservation period begins.

TABLE II. EXAMPLE MATRIX FOR $\mathrm{DEC_{IJ}}$

	1	**2**	**3**	**4**	**5**	**6**	**7(w)**	**8(nb)**
1	0	0	0	0	0	0	0	1
2	0.4	0	0	0	0	0	0	0.6
3	0	0.2	0	0	0	0	0.3	0.5
4	0.1	0.2*	0	0	0	0.2*	0	0.7
5	0	0.2	0	0.2	0	0	0.3	0.3
6	0	0	0	0.3	0	0	0.4	0.3

** Customers who cannot find Product (Ticket) 4 can maketheir trip by purchasing Ticket 2 and Ticket 6 together*

B. Genetic Algorithm (GA)

The GA is an optimization method and one of the adaptive heuristic search algorithms. At first, the solution parameters are encoded in chromosomes. Each chromosome consists of a certain number of gene blocks representing the buckets, and is a solution candidate. The sum of the generated chromosomes constitutes the starting population. If we have a railway line with n stations and seat capacity C on a train, there will be n-1 resources and $n_{Prod} = n(n-1)/2$ products. If the example from Section II is considered, Table III illustrates the way the products and buckets are encoded. There are 3 products and 2 buckets. Pr1 uses Res1 and Res2 and it is placed alone in Bucket 1 (Buc1). Zero stands for the absence of any candidates.

TABLE III. ENCODED BUCKETS AND PRODUCTS FOR A THREE-STOP LINE EXAMPLE

n=3	**Stations** from..to..	**Encoded products**	**Encoded buckets**
Pr1	1->3	Pr1=[1 2]	Buc1=[Pr1 0]
Pr2	1->2	Pr2=[1 0]	Buc2=[Pr2 Pr3]
Pr3	2->3	Pr3=[0 2]	

Buckets have a predetermined number of cells, symbolized as l_{Buc}. The number of seats, i.e. capacities of related buckets c_b, allocated for each bucket is placed in the last cell of a gene, so that the size of a gene should be $l_{Buc} + 1$. As seen in Table IV, a chromosome is encoded as an integer array with the genes arranged in a row. Therefore, the chromosome size depends on both the number of buckets used (n_{Buc}) and the size of a gene.

TABLE IV. ENCODED GENES AND CHROMOSOMES FOR A THREE-STOP LINE EXAMPLE

Parameters:	l_{Buc}=2, n_{Buc}=2, C=100		
Encoded buckets	Buc1=[Pr1 0] Buc2=[Pr2 Pr3]	Buc1=[Pr1 0] Buc2=[Pr2 Pr3]	Buc1=[0 Pr1] Buc2=[Pr2 0]
Encoded genes	Gene1=[Buc1 50] Gene2=[Buc2 50]	Gene1=[Buc1 60] Gene2=[Buc2 40]	Gene1=[Buc1 60] Gene2=[Buc2 40]
Encoded chromosomes	Cr1= [Gene1 Gene2]	Cr2= [Gene1 Gene2]	Cr3= [Gene1 Gene2]

Fitness: The performance of a chromosome is evaluated by fitness, which can be calculated using the objective function. Genetic operators expand the solution space by producing generations with better characteristics. Three basic genetic operators are selection, crossover, and mutation.

Selection: The selection process copies the existing chromosome to the next generation without making any changes in its genetic structure. We used the roulette method for selection in this research.

Crossover: Crossover is the formation of new individuals by randomly combining the structures of two individuals, namely chromosomes. We use the point-to-point uniform crossover rule, in which a coin is flipped for each gene to decide whether or not it will be included in the offspring. Since the chromosomes consist of the bucket and the number of seats allocated to it, the crossover was also performed in two steps according to the crossover rate.

Mutation: A mutation is performed by replacing one or more of the genes of an existing individual with entirely new genes. We adopt the uniform mutation rule. The mutation also takes place in two steps. First step: The products in the genes of the randomly selected chromosome are replaced according to the mutation rate with new products that were never selected. Buckets with empty cells are filled and the filled ones can be empty. Second step: The number of seats in the randomly selected gene of the randomly selected chromosome is increased by 5 or decreased by 5. While performing this, it is essential to check that the number of seats does not fall below zero and does not exceed the capacity.

C. Simulations

In this research, ticket booking is treated as a time-dependent process. Only one customer arrives in a time interval t, which is a small part of the time horizon. So, the time horizon is divided into T time periods. The arrival of customers is simulated according to Poison distribution. A random demand vector is generated with the average demand and standard deviation of demand provided. The customer behavior is simulated by using the CDM. The sales from any product depend on the sales from all other products. Therefore, in our model, the demand for a product consists of partially the customers who initially request that product and partially the customers diverting from other products. Several algorithms have been written in Matlab to simulate the booking process with the procedure shown in Table V. Various random demands are generated as over, under, and close or equal to the average demand to analyze the different scenarios.

In the beginning, the buckets used in the simulation have random initial values that come from the GA. After many simulation rounds, buckets with higher revenue can be achieved in GA via feedback of the simulation results. The values of the GA parameters, i.e. population size, mutation rate and crossover rate, were decided using the experimental design technique proposed by Taguchi [38]. Ticket requests are primarily responded to by sales from TK, i.e. ticket pool. If there is no appropriate ticket in TK, all the buckets are scanned for the product. It is not enough to find the product in a bucket, it is also necessary to have an empty seat for that bucket. After the sale from buckets, some seats can remain unoccupied. So these are added to TK for reuse next time. If there is no available seat for a product, customers who request a ticket for that product can either choose another product, wait for a seat for that ticket, or leave without purchasing (see Figure 4). The simulation algorithm runs until there is no demand and all ticket requests, positive or negative, are met. Evaluation and comparison of the techniques are made over a fictional railway line and a real-life example with different demand scenarios. We aim through simulations to observe how TBS differs from the other two ticket booking mechanisms and the effects of customer behavior on the booking process.

Apart from different random demand scenarios and random arrival sequences, special demand situations were also examined. We specifically set low, high, and equal-to-mean demand scenarios with appropriate random demands. Low and high demand represent the demand that is lower and higher than the capacity respectively. Equal-to-mean demand is equal or random but too close to the average historical demand. Each simulation has been run 100 times with various scenarios. The exhibited results are the average values obtained at the end of the simulation trials.

TABLE V. SIMULATION STEPS

	Start
1.	Set initial parameters: C, Buc, $dmnd$, dec_{ij}, TK
2.	If there is no request for Pr(rs) → Go to end
3.	Else
4.	If TK contains available resource(s) for Pr(rs) → Sale from TK, go to line 2
5.	Else
6.	If Buc contains available resource(s) for Pr(rs)
7.	Sale from Buc
8.	If there is remaining resource(s)/seat(s) → Add the remaining resource(s)/seat(s) to TK, go to line 2
9.	Else
10.	If customer wants to buy nothing (with dec_{ij}) → Add Pr(rs) to no-buy list, go to line 2
11.	Else if customer wants to wait (only in TBS !!!) → Add Pr(rs) to wait list, go to line 2
12.	Else if the customer wants to buy another ticket → Distribute demand to re-buy list by dec_{ij}, go to line 2
13.	Else
14.	Add wait list and rebuy list to $dmnd$, go to line 2
	End

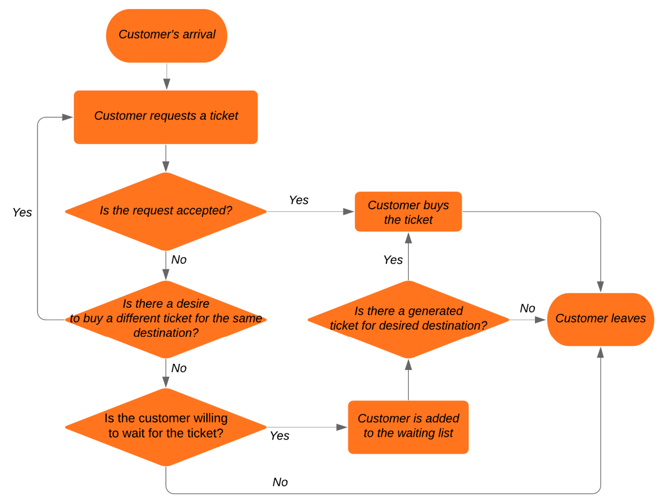

Fig. 4. Flowchart of the booking process steps with customer diversion.

IV. COMPUTATIONAL EXPERIMENTS

A. Fictive Railway Line Example

In this test example, there is a 5-stop railway line with a single train, whose seat capacity is 100. For the stations from A to E, the ticket price in US$ and the demand information of the products is given in Table VI.

TABLE VI. TICKET FARES AND DEMAND INFORMATION

Products	1	2	3	4	5	6	7	8	9	10
	A-E	A-D	A-C	A-B	B-E	B-D	B-C	C-E	C-D	D-E
Price ($)	80	50	30	20	70	40	20	60	30	40
Mean demand	75	50	13	5	38	25	5	37	12	63
St. dev. of demand	11.3	6.7	3.1	2.9	4.2	1.7	2.6	4.7	2.8	7.1

The diversion matrix of this example is generated from random values. As described above, the values in columns 11 and 12 of the 10×12-matrix are the waiting and non-purchase rates respectively. The results for different demand scenarios (under, over, and equal to the mean of the fictional demand) can be seen in Figures 5 and 6. The Figures show the expected revenue and the number of rejected requests for each scenario. In the Figures, TBS_axb stands for TBS with a buckets, each b cells long. The algorithm suggests 3 buckets with a length of 4 cells in this example, but we put some alternative bucket solutions in the Figures. As seen in Figure 5, in all demand levels, TBS gives a better performance than PBLC and FCFS. At every demand level, even if the arrival sequence of customers changes, the performance ranking by revenue is TBS > FCFS > PBLC. The way the integration of the customer's behavior into the booking process affects the income can be seen in Figure 5. While the effects of customer behavior are minimal on TBS, it appears to have the most significant impact on PBLC. Especially at low demand level, sales revenue is significantly increased in PBLC compared to the case without customer behavior. Figure 6 shows that considering customer behavior has not only resulted in improvement in revenue. The number of rejected customers decreased at all demand levels and for all techniques. The amount and rate of change vary according to the customers' demand and arrival sequence.

(a)
(b)
(c)

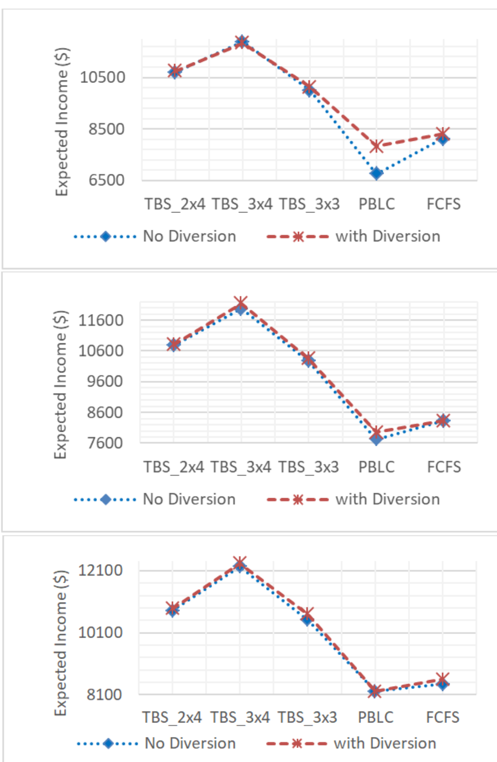

Fig. 5. Comparison of expected income with/without customer diversion: (a) Low demand, (b) mean-equal demand, and (c) high demand.

B. Real-life Railway Example

As a real-life example, the railway line between Ankara and Istanbul in Turkey was examined. The data were obtained from the statistical and annual reports published on the Turkish national railway company website. The incompatible and irrelevant data were extracted or neglected. The company has 7 trains, each with 419 seats, daily departing from Ankara to Istanbul. There are 9 stations on the line, the first station is in Ankara and the last station is in Halkali in Istanbul. There are 36 products (see Table VII) and 8 resources in this example. Price and product numbers (in square brackets) of each O-D pair are listed in Table VII. Average demand and standard deviations of demand (shown in parentheses) of each product can be seen in Table VIII.

(a)
(b)
(c)

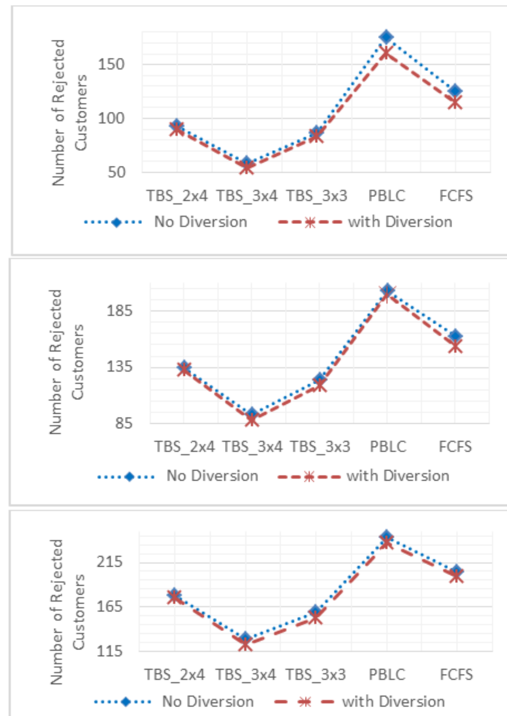

Fig. 6. Comparison graphs showing the number of rejected customers with/without customer diversion: (a) Low demand, (b) mean-equal demand, and (c) high demand.

TABLE VII. TICKET PRICE INFORMATION AND PRODUCT NUMBERS (IN SQUARE BRACKETS) FOR THE REAL-LIFE EXAMPLE

Price (TL)	Polatli	Eskisehir	Bozuyuk	Bilecik	Arifiye	Izmit	Istanbul (S.cesme)	Istanbul (Halkali)
Ankara	19.5 [8]	37.5 [7]	49.5 [6]	61.5 [5]	73.5 [4]	79.5 [3]	85.5 [2]	96 [1]
Polatli		28 [15]	43.5 [14]	55.5 [13]	61.5 [12]	67.5 [11]	79.5 [10]	90 [9]
Eskisehir			19.5 [21]	31.5 [20]	52.5 [19]	52.5 [18]	55.5 [17]	64 [16]
Bozuyuk				19.5 [26]	37.5 [25]	52.5 [24]	55.5 [23]	64 [22]
Bilecik					31.5 [30]	49.5 [29]	55.5 [28]	64 [27]
Arifiye						19.5 [33]	31.5 [32]	38.5 [31]
Izmit							19.5 [35]	26.5 [34]
Istanbul (S.cesme)								24 [36]

TABLE VIII. MEAN DEMAND AND STANDARD DEVIATION OF THE DEMAND (IN PARATHESIS) FOR GIVEN O-D PAIRS

Demand (Std.Dev.)	Polatli	Eskisehir	Bozuyuk	Bilecik	Arifiye	Izmit	Istanbul (S.cesme)	Istanbul (Halkali)
Ankara	8 (4.1)	121 (6.5)	4 (1.7)	6 (2.4)	7 (7.4)	15 (5.0)	219 (9.1)	32 (4.9)
Polatli		4 (4.0)	1 (1.4)	1 (2.8)	1 (1.7)	2 (2.2)	7 (6,6)	1 (7.0)
Eskisehir			1 (1.1)	2 (1.7)	7 (3.2)	17 (5.4)	112 (8.4)	13 (3.2)
Bozuyuk				1 (2.8)	1 (1.3)	1 (1.6)	3 (6.5)	3 (1.7)
Bilecik					2 (2.9)	2 (2.6)	6 (4.5)	4 (3.6)
Arifiye						1 (2.4)	6 (5.7)	4 (5.2)
Izmit							23 (7.4)	20 (10.9)
Istanbul (S.cesme)								1 (3.8)

The diversion matrix, for this example was created based on the opinions of expert. As mentioned above, the diversion data can be obtained online using new technologies and can be replaced easily in TBS. The matrix has a size of 36×38 and contains zero values for some products. Column 37 represents the rate of customers who desire to wait. Similarly, column 38 is the rate of customers who prefer to leave without purchasing. The simulation results are summarized in the tables according to the demand levels and whether customer behavior is considered. The obtained results are shown in Table IX and Figure 7 comparatively.

TABLE IX. CALCULATED EXPECTED REVENUES (IN TL), SOLD TICKETS AND REJECTED CUSTOMERS ACCORDING TO DEMAND SCENARIOS

Revenue (TL)				
Demand	**Diversion**	**TBS_7x7**	**PBLC**	**FCFS**
Low	No	36976	36976	36976
	Yes	36976	36976	36976
Mean-equal	No	39079	38084	37787
	Yes	39154	38159	37811
High	No	42567	40793	38663
	Yes	42634	41020	38742

Sold Tickets (pcs)				
Demand	**Diversion**	**TBS_7x7**	**PBLC**	**FCFS**
Low	No	589	589	589
	Yes	589	589	589
Mean-equal	No	624	618	617
	Yes	629	621	618
High	No	709	659	642
	Yes	719	664	644

Rejected customers (persons)				
Demand	**Diversion**	**TBS_7x7**	**PBLC**	**FCFS**
Low	No	0	0	0
	Yes	0	0	0
Mean-equal	No	35	41	42
	Yes	30	38	41
High	No	95	145	162
	Yes	85	140	160

The algorithm for TBS suggests 7 buckets with a length of 7 cells after trials of many bucket combinations. In low demand, the expected revenues of all booking techniques are the same or very close to each other. One reason for this may be that all customers can get the tickets they want due to the low demand and sufficient capacity. In high and mean-equal demand scenarios, the differences between the booking mechanisms are evident (Figure 7(b)-(c)). TBS shows clearly better performance than PBLC and FCFS. If we take the customer behavior into account in simulations, the changes in revenue can be seen in Figure 8. Again, there are no differences in revenue with or without customer diversion in low-demand scenarios and the reason explained above also applies here. In other demand levels, customer behavior shows its effect. Its most significant impact is on TBS, then on PBLC, and FCFS. In the simulation, experiments were carried out by diversifying customer arrivals. It has been observed that the effect of the changes in the customers' arrival is shallow, and there is no change in the performance order even though there are minor changes in the revenues. Figures 8 and 9 show that incorporating customer behavior into the system increased the expected revenues and decreased the number of rejected customers. In particular, the decrease in the number of rejected customers is quite evident, except for the low demand level (Figure 9(a)).

(a)

(b)

(c)

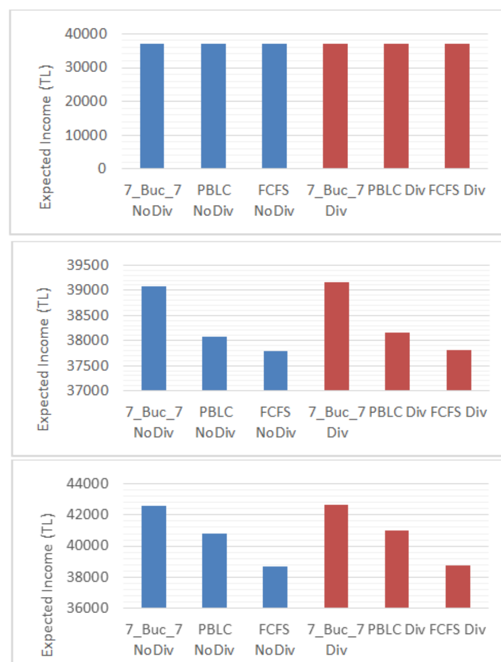

Fig. 7. Expected income with/without customer diversion according to the demand scenarios for the real-life example: (a) Low demand, (b) mean-equal demand, and (c) high demand.

C. *Impact of Customer Behavior*

Simulations were performed with different valued matrices of dec_{ij} to evaluate the customer behavior impact. The tests were carried out in two stages. First, simulations were run with different demand levels while keeping CDM constant. Changes were made in CDM in the second stage while keeping the demand constant.

(a) (b) (c)

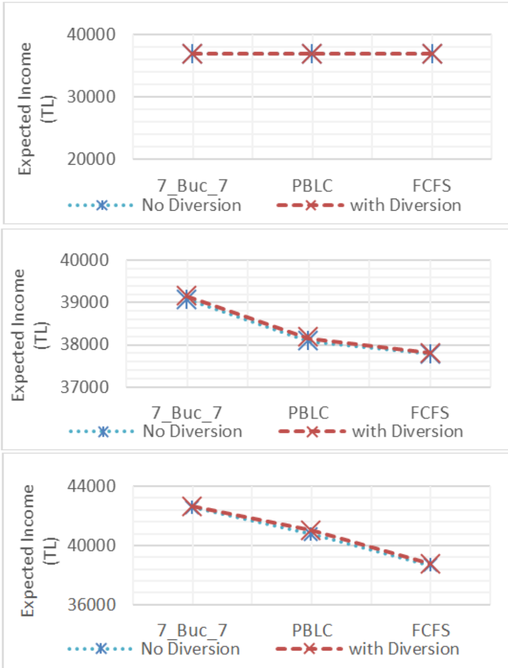

Fig. 8. Comparison of expected revenues with/without customer diversion according to the demand scenarios for the real-life example: (a) Low demand, (b) mean-equal demand, and (c) high demand.

(a) (b) (c)

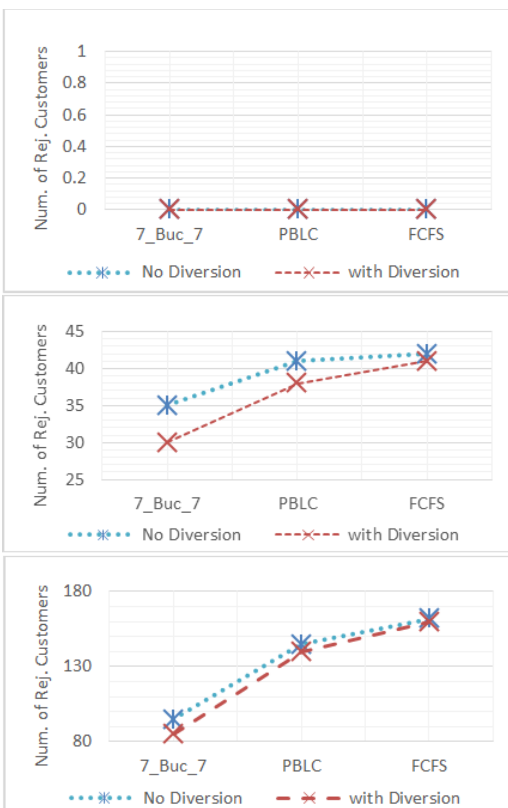

Fig. 9. Changes in the number of rejected customers with/without considering customer diversion according to the demand scenarios: (a) Low demand, (b) mean-equal demand, and (c) high demand.

As can be seen in Figure 10, in low demand, the revenue and the number of tickets sold are not affected by the customer behavior. As the demand increases, the change in income becomes more pronounced and is most significant in TBS. The impact of customer behavior on revenue and ticket sales was primarily monitored in above- and near-average demands. In the tests conducted with high demand, it has been observed that the effect of customer behavior decreases gradually due to the exhaustion of capacity if the demand is too high. In the second phase, changes were made in dec_{ij} by keeping the demand high above average and keeping it constant. It is important to note that the waiting behavior in the PBLC and FCFS methods does not affect revenue and ticket sales, because in these methods, after the capacity of a product is reached, the same product does not open for sale again. In TBS, the re-opening possibility is always present according to the order of the customer's arrival. Except for the waiting or no-buy columns, the changes are limited to cells with zero value. It would be easier to reduce the no-buy percentage by encouraging those who leave without buying anything to buy any product. The changes made in columns 1-36 (columns 37 and 38 are adjusted accordingly) had a more significant impact on revenue and ticket sales in PBLC and FCFS than in TBS.

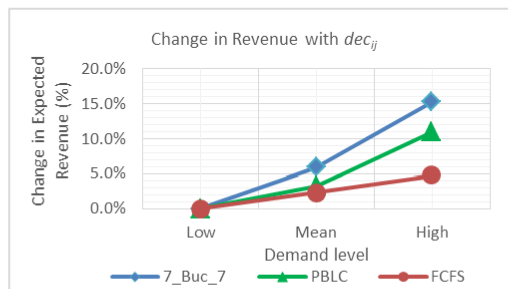

Fig. 10. Changes in the expected revenue with considering customer diversion according to the demand scenarios: (a) Low demand, (b) mean-equal demand, and (c) high demand.

The effect of changes made in the waiting and no-buy columns on revenue and sales can be analyzed by assigning a value of zero to the other columns. So, in Figures 11 and 12, W0/NB0 refers the results without CDM. In the expression W01/NB09, W01 represents that all cells in the waiting column get the value 0.1 and NB09 means that all cells in the no-buy column have the value 0.9. Similarly, numbers in other expressions (W0x/NB0y) represent ratios for columns 37 and 38. It should be noted that since the ratios in the cells represent a probability, their sum should not exceed 1, so the sum of each row cannot exceed 1. It has been observed that changes in the waiting column significantly impact the revenue and sales with TBS. This impact increases positively as the waiting rate increases (see Figures 11-12). In the tests with PBLC and FCFS, the slight changes in revenue and sales seen in the Figures can be interpreted as a result of the change in the no-buy percentage. The modification of CDM has a remarkable impact on revenue and ticket sales in TBS (up to 3.49% and 2.12% respectively) when the waiting rate (column 37) is changed. This will be instrumental for the development of a different sales strategy or policy to be based on the examination of the effects of customer behavior.

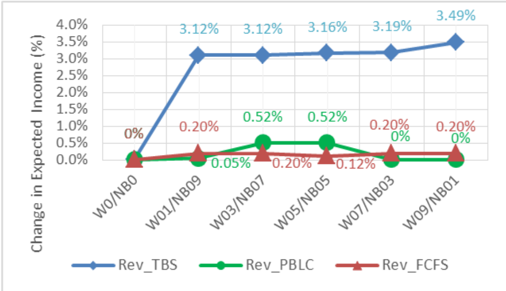

Fig. 11. Changes in expected revenue with various values of the dec_{ij} matrix (in percentage).

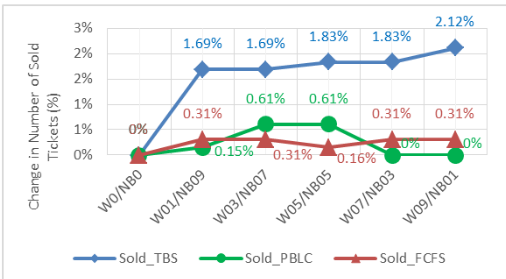

Fig. 12. Changes in the number of sold tickets with various values of the dec_{ij} matrix (in percentage).

V. CONCLUSION

The Turkish railway company sells 593 tickets and has an income of 36.866 TL on average in a day per train (2019). With the proposed algorithm (TBS), approximately 630 tickets can be sold in the case of average demand. An income raise of around 2100 TL per train can be expected with revenue up to 39000 TL, which is a good result compared to the current level. TBS gave the best results among the three examined methods. An annual revenue increase of 5.3M TL can be expected if our proposed model is implemented. Unlike previous research in which only the customers' choice of a train before departure has been taken into account, this study regards customers' preferences when they cannot find the ticket they want. This research concentrates on evaluating the performance of TBS under customer diversion conditions, which is utilized in our research as a percentage matrix. According to the experts' opinion, customers have 3 possibilities: Buying another product, waiting for a seat of the same product or not purchase anything when the desired ticket is unavailable. Adding customer behavior to the simulation process helps us find a better revenue-generating bucket.

This study is the first application of an alternative booking system for Turkish railways and implements a booking system with buckets in Turkish railways. This paper also contributes in terms of considering customer behavior in a booking system with buckets. To the best of our knowledge, customer diversion has never been applied before in the railway industry. The proposed TBS performs better than the PBLC and the FCFS. When the demand is high, all methods have yielded their best results, i.e. the highest revenue. At low demand, the success of all techniques is very close to each other. The positive aspects of the TBS are the more profitable use of the empty seats with the ticket pool and the effective distribution of the capacity to the products through the buckets. On the negative side, finding the optimum buckets takes time and processing power.

If we ignore the misleading results due to the missing data, we can say that the gain will improve with TBS with the available data, even in its pure form. If capacity cannot be increased, it may be wise to encourage customers to wait. Thus, empty seats can be sold to customers who prefer to wait for a second sale-period. Considering the effects of customer behavior, we can argue that selling tickets in two or more periods will increase revenue. In addition to the increase in revenue, the decrease in the number of rejected customers is obviously important in terms of customer satisfaction.

The seat allocation and ticketing processes of high-speed trains are similar to those on airplanes. For this reason, it would be appropriate for railway transportation companies to use technological advances in software to capture customer behavior more precisely. In the railway sector, hybrid algorithms can be tried with artificial intelligence on seat allocation and booking optimization.

ACKNOWLEDGEMENT

This paper is a part of the PhD thesis of Ahmet Z. Bulum in the Graduate Education Institute, Karabuk University, Karabuk 2021 [Advisor: Assoc. Prof. Dr. Muharrem Dugenci, Co-Advisor: Dr. Mumtaz Ipek].

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

[1] J. I. McGill and G. J. van Ryzin, "Revenue Management: Research Overview and Prospects," *Transportation Science*, vol. 33, no. 2, pp. 233–256, May 1999, https://doi.org/10.1287/trsc.33.2.233.

[2] R. Phillips, *Pricing and Revenue Optimization*. Stanford, CA, USA: Stanford Business Books, 2005.

[3] S. L. Brumelle and J. I. McGill, "Airline Seat Allocation with Multiple Nested Fare Classes," *Operations Research*, vol. 41, no. 1, pp. 127–137, Feb. 1993, https://doi.org/10.1287/opre.41.1.127.

[4] R. E. Curry, "Optimal Airline Seat Allocation with Fare Classes Nested by Origins and Destinations," *Transportation Science*, vol. 24, no. 3, pp. 193–204, Aug. 1990, https://doi.org/10.1287/trsc.24.3.193.

[5] A. Ciancimino, G. Inzerillo, S. Lucidi, and L. Palagi, "A Mathematical Programming Approach for the Solution of the Railway Yield Management Problem," *Transportation Science*, vol. 33, no. 2, pp. 168–181, May 1999, https://doi.org/10.1287/trsc.33.2.168.

[6] H. Luo, L. Nie, and Z. He, "Modeling of multi-train seat inventory control based on revenue management," in *International Conference on Logistics, Informatics and Service Sciences*, Sydney, NSW, Australia, Jul. 2016, pp. 1–6, https://doi.org/10.1109/LISS.2016.7854371.

[7] Z. Xie, W. Zhu, and L. Zheng, "A Dynamic Railway Seat Allocation Problem," in *Industrial and Systems Engineering Research Conference*, Puerto Rico, PR, USA, Dec. 2013, pp. 3480–3489.

[8] Y. Bao, J. Liu, M. Ma, and L. Meng, "Seat inventory control methods for Chinese passenger railways," *Journal of Central South University*, vol. 21, no. 4, pp. 1672–1682, Apr. 2014, https://doi.org/10.1007/s11771-014-2109-y.

[9] P. Hetrakul and C. Cirillo, "A latent class choice based model system for railway optimal pricing and seat allocation," *Transportation Research Part E: Logistics and Transportation Review*, vol. 61, pp. 68–83, Jan. 2014, https://doi.org/10.1016/j.tre.2013.10.005.

[10] A. Armstrong and J. Meissner, "Railway Revenue Management: Overview and Models," 2010. https://eprints.lancs.ac.uk/id/eprint/49017.

[11] W. Yuan and L. Nie, "Optimization of seat allocation with fixed prices: An application of railway revenue management in China," *PLOS ONE*, vol. 15, no. 4, Apr. 2020, Art. no. e0231706, https://doi.org/10.1371/journal.pone.0231706.

[12] C. Cizaire and P. Belobaba, "Joint optimization of airline pricing and fare class seat allocation," *Journal of Revenue and Pricing Management*, vol. 12, no. 1, pp. 83–93, Jan. 2013, https://doi.org/10.1057/rpm.2012.27.

[13] G. Dutta and D. P. Marodia, "Comparison of forecasting techniques in revenue management for a national railway in an emerging Asian economy," *International Journal of Revenue Management*, vol. 8, no. 2, pp. 130–152, Jan. 2015, https://doi.org/10.1504/IJRM.2015.070000.

[14] X. Wang, H. Wang, and X. Zhang, "Stochastic seat allocation models for passenger rail transportation under customer choice," *Transportation Research Part E: Logistics and Transportation Review*, vol. 96, pp. 95–112, Dec. 2016, https://doi.org/10.1016/j.tre.2016.10.003.

[15] H. Jafarzadeh, N. Moradinasab, and M. Elyasi, "An Enhanced Genetic Algorithm for the Generalized Traveling Salesman Problem," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2260–2265, Dec. 2017, https://doi.org/10.48084/etasr.1570.

[16] T. Kara and M. C. Savas, "Design and Simulation of a Decentralized Railway Traffic Control System," *Engineering, Technology & Applied Science Research*, vol. 6, no. 2, pp. 945–951, Apr. 2016, https://doi.org/10.48084/etasr.631.

[17] J. Zheng, J. Liu, and D. B. Clarke, "Ticket Fare Optimization for China's High-Speed Railway Based on Passenger Choice Behavior," *Discrete Dynamics in Nature and Society*, vol. 2017, Feb. 2017, Art. no. e6237642, https://doi.org/10.1155/2017/6237642.

[18] J. Zheng and J. Liu, "The Research on Ticket Fare Optimization for China's High-Speed Train," *Mathematical Problems in Engineering*, vol. 2016, Aug. 2016, Art. no. e5073053, https://doi.org/10.1155/2016/5073053.

[19] A. Nikseresht and K. Ziarati, "A Demand Estimation Algorithm for Inventory Management Systems Using Censored Data," *Engineering, Technology & Applied Science Research*, vol. 7, no. 6, pp. 2215–2221, Dec. 2017, https://doi.org/10.48084/etasr.1517.

[20] Z. Xiaoqiang, M. Lang, and Z. Jin, "Dynamic pricing for passenger groups of high-speed rail transportation," *Journal of Rail Transport Planning & Management*, vol. 6, no. 4, pp. 346–356, Jan. 2017, https://doi.org/10.1016/j.jrtpm.2017.01.001.

[21] M. Qin, Y. Li, and G. Che, "Railway passenger ticket pricing policy portfolio," in *International Conference on Logistics, Informatics and Service Sciences*, Sydney, NSW, Australia, Jul. 2016, pp. 1–5, https://doi.org/10.1109/LISS.2016.7854432.

[22] G. Dutta and P. Ghosh, "A passenger revenue management system (RMS) for a National Railway in an Emerging Asian Economy," *Journal of Revenue and Pricing Management*, vol. 11, no. 5, pp. 487–499, Sep. 2012, https://doi.org/10.1057/rpm.2012.10.

[23] K. T. Talluri and G. J. van Ryzin, *The Theory and Practice of Revenue Management*. Boston, MA, USA: Springer, 2005.

[24] P.-S. You, "An efficient computational approach for railway booking problems," *European Journal of Operational Research*, vol. 185, no. 2, pp. 811–824, Mar. 2008, https://doi.org/10.1016/j.ejor.2006.12.049.

[25] X. Jiang, X. Chen, L. Zhang, and R. Zhang, "Dynamic Demand Forecasting and Ticket Assignment for High-Speed Rail Revenue Management in China," *Transportation Research Record*, vol. 2475, no. 1, pp. 37–45, Jan. 2015, https://doi.org/10.3141/2475-05.

[26] M. Riss, J. Cote, and G. Savard, "A new revenue optimization tool for high-speed railway: finding the right equilibrium between revenue growth and commercial objectives," in *8th World Congress on Railway Research*, Seoul, Korea, Dec. 2008.

[27] P. Hetrakul, "Discrete choice models for revenue management," Ph.D. dissertation, University of Maryland, College Park, MD, USA, 2012.

[28] S. S. Azadeh, M. Hosseinalifam, and G. Savard, "The impact of customer behavior models on revenue management systems," *Computational Management Science*, vol. 12, no. 1, pp. 99–109, Jan. 2015, https://doi.org/10.1007/s10287-014-0204-z.

[29] W. L. Cooper, T. Homem-de-Mello, and A. J. Kleywegt, "Models of the Spiral-Down Effect in Revenue Management," *Operations Research*, vol. 54, no. 5, pp. 968–987, Oct. 2006, https://doi.org/10.1287/opre.1060.0304.

[30] S.-W. Kim, "The impact of customer buying behavior on the optimal allocation decisions," *International Journal of Production Economics*, vol. 163, pp. 71–88, May 2015, https://doi.org/10.1016/j.ijpe.2015.02.009.

[31] J. G. Wilson, C. K. Anderson, and S.-W. Kim, "Optimal booking limits in the presence of strategic consumer behavior," *International Transactions in Operational Research*, vol. 13, no. 2, pp. 99–110, 2006, https://doi.org/10.1111/j.1475-3995.2006.00537.x.

[32] A. Sen and A. X. Zhang, "The newsboy problem with multiple demand classes," *IIE Transactions*, vol. 31, no. 5, pp. 431–444, May 1999, https://doi.org/10.1023/A:1007549223664.

[33] P. P. Belobaba and L. R. Weatherford, "Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations," *Decision Sciences*, vol. 27, no. 2, pp. 343–363, 1996, https://doi.org/10.1111/j.1540-5915.1996.tb00856.x.

[34] I. S. Yeoman and U. McMahon-Beattie, "The turning points of revenue management: a brief history of future evolution," *Journal of Tourism Futures*, vol. 3, no. 1, pp. 66–72, Jan. 2017, https://doi.org/10.1108/JTF-11-2016-0040.

[35] B. Vinod, "Evolution of yield management in travel," *Journal of Revenue and Pricing Management*, vol. 15, no. 3, pp. 203–211, Jul. 2016, https://doi.org/10.1057/rpm.2016.15.

[36] A. K. Strauss, R. Klein, and C. Steinhardt, "A review of choice-based revenue management: Theory and methods," *European Journal of Operational Research*, vol. 271, no. 2, pp. 375–387, Dec. 2018, https://doi.org/10.1016/j.ejor.2018.01.011.

[37] M.-E. A. Martinez, M.-A. G. Borja, and J.-A. M. Jimenez, "Yield Management As A Pricing Mechanism," *Review of Business Information Systems*, vol. 15, no. 5, pp. 51–60, Sep. 2011, https://doi.org/10.19030/rbis.v15i5.6018.

[38] R. K. Roy, *Design of Experiments Using The Taguchi Approach: 16 Steps to Product and Process Improvement*. New York, NY, USA: Wiley, 2001.