Developing an Overbooking Fuzzy-Based Mathematical Optimization Model for Multi-Leg Flights

Al Bazi, A, Uney, E & Abu Monshar, A

Original citation:
Al Bazi, A, Uney, E & Abu Monshar, A 2020, 'Developing an Overbooking Fuzzy-Based Mathematical Optimization Model for Multi-Leg Flights' Transportation Research Procedia, vol. 43, pp. 165-177
https://dx.doi.org/10.1016/j.trpro.2019.12.031

DOI 10.1016/j.trpro.2019.12.031
ISSN 2352-1465
ESSN 2352-1465

Publisher: Elsevier

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 43 (2019) 165–177

Transportation Research Procedia

www.elsevier.com/locate/procedia

8th International Conference on Air Transport – INAIR 2019
GLOBAL TRENDS IN AVIATION

Developing an Overbooking Fuzzy-Based Mathematical Optimization Model for Multi-Leg Flights

Ammar Al-Bazi[a,*], Emre Üney[b], Anees Abu-Monshar[a]

[a] *School of Mechanical, Aerospace and Automotive Engineering Coventry University, Coventry, UK*
[b] *Roketsan Missles Inc., Istanbul, Turkey*

Abstract

Overbooking is one of the most vital revenue management practices that is used in the airline industry. Identification of an overbooking level is a challenging task due to the uncertainties associated with external factors, such as demand for tickets, and inappropriate overbooking levels which may cause revenue losses as well as loss of reputation and customer loyalty. Therefore, the aim of this paper is to propose a fuzzy linear programming model and Genetic Algorithms (GAs) to maximize the overall revenue of a large-scale multi-leg flight network by minimizing the number of empty seats and the number of denied passengers.
A fuzzy logic technique is used for modeling the fuzzy demand on overbooking flight tickets and a metaheuristics-based GA technique is adopted to solve large-scale multi-leg flights problem.
As part of model verification, the proposed GA is applied to solve a small multi-leg flight linear programming model with a fuzzified demand factor. In addition, experimentation with large-scale problems with different input parameters' settings such as penalty rate, show-up rate and demand level are also conducted to understand the behavior of the developed model.
The validation results show that the proposed GA produces almost identical results to those in a small-scale multi-leg flight problem. In addition, the performance of the large-scale multi-leg flight network represented by a number of KPIs including total booking, denied passengers and net-overbooking profit towards changing these input parameters will also be revealed.

Peer-review under responsibility of the scientific committee of the 8th International Conference on Air Transport – INAIR 2019, GLOBAL TRENDS IN AVIATION

Keywords: revenue management; overbooking; airline networks; fuzzy demand; genatic algorithm

* Corresponding author. Tel.: +0-247-765-9182
E-mail address: aa8535@coventry.ac.uk

2352-1465 © 2019 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 8th International Conference on Air Transport – INAIR 2019, GLOBAL TRENDS IN AVIATION
10.1016/j.trpro.2019.12.031

1. Introduction

Revenue Management (RM) is a field of management that aims at maximizing the revenues of an organization by selling the available products to the customers at the right time and price by developing revenue-optimal strategies of customer and product selection (Cross 1997). Overbooking is one of the most vital revenue management practices that is currently used in many industries including the airline industry. It allows airlines to increase revenues and aircraft utilization and, also, generates economic benefits for passengers by increasing the number of available seats and reducing the cost of air travel (Klophaus and Pölt 2010). Research indicates that companies that exploit overbooking practice may increase their revenues by 2% to 5% (Wangenheim and Bayón 2007). On the other hand, inappropriate overbooking strategies not only cause revenue losses in the short term but also have long term consequences for the airline companies such as loss of reputation and customer loyalty. A good example of reputation loss and its negative social effect was seen on United Airlines when its flight attendant forcibly removed a paying customer (Ma et al. 2019).

Given the significant volume of literature in the field of revenue management that has been produced after the liberalization of the airline industry in 1978, overbooking is one of the oldest strategies that were exploited even before this period (McGill and van Ryzin 1999). Different studies proposed various approaches for revenue management considering overbooking at the leg level (Sumida and Topaloglu 2019) or at the network level (Mou, Li, and Li 2019) or at a competitive market level (Alavi, Sy, and Ivanov 2019). The network problem is naturally more complex since the seat allocation in one leg affects other legs in the network but in many cases addressing the problem at the network level may be more profitable.

The novelty of this paper is that it presents a mathematical optimization model, based on fuzzy demand for the multiple-flight network overbooking problem. The purpose is to maximize the net profit of the flight network by minimizing the number of empty seats and the number of denied passengers considering vague and uncertain demand conditions. Knowledge is elicited after running a number of numerical experiments, in which reasonable assumptions are set.

The remainder of this paper is organized as follows: Section 2 reviews the related literature in airline revenue management. The problem of study is stated in Section 3. The fuzzy demand mathematical programming model is presented in Section 4. The proposed Genetic Algorithm (GA) is shown in Section 5. Experimentation with both small-sized and large-sized problem are shown in Section 6. Finally, conclusions are drawn in Section 7.

2. Previous literature on multiple-leg revenue management model with overbooking

This section provides studies that are related to multiple-leg airline revenue management overbooking problems. In this context, Bertsimas and Popescu (2003) investigated the network revenue management problem with certain equivalent control policy to overcome errors of bid-price method. However, Monte Carlo fashion was used to model uncertainty in demand rather than a defined stochastic function. Karaesmen and van Ryzin (2004a) investigated the static overbooking problem with multiple reservation and inventory classes with the possible substitution of different fare classes. In another study, Karaesmen and van Ryzin (2004b) developed a two-stage optimization model for the dynamic airline network overbooking problem where the first stage deals with overbooking, given the probabilistic nature of cancellations, to find the overbooking limit and, accordingly, the second stage allocates capacities for different inventory classes and revenues from the flight network. Although compensation paid to the denied passengers and the goodwill costs are considered, penalties for cancellations and no-shows are not. Bertsimas and de Boer (2005) aimed to find the booking limits for an airline network considering the dynamic and stochastic nature of the reservations and the complexity of booking control policies in a network environment by estimating the expected revenue of any booking limit policy. However, they considered a simplified demand model and cancelations were not considered. Siddappa, Rosenberger and Chen (2008) determined overbooking levels for different flights in an airline network with an aim of maximizing the total profit from the whole flight network. However, the probability of passengers showing up is assumed to be constant which makes the model rather unrealistic. Gosavi, Ozkaya and Kahraman (2007) addressed the inventory allocation problem of airlines for single-leg, multi-leg or network problems considering random cancellations and overbooking. As for the demand modeling, these methods usually require only a ‘black box’ estimate of the demand, however, it would be desirable to have an insight about demand uncertainty in practice. Kunnumkal and Topaloglu (2008) captured the uncertainty of requests for an itinerary given their probabilistic arrivals as well as the uncertainty in no-shows in a multi-leg overbooking airline problem. Although requests are modeled as stochastic, the requests do not fully capture the uncertain behavior of the demand in which fuzzy logic can be utilized. Erdelyi and Topaloglu (2009) investigated the problems of capacity allocation and overbooking in airline networks. Their study offers a viable solution for the overbooking problem in an airline network, however, the model under study is deterministic and omits the probabilistic elements of the booking process. Hjorth et al. (2018) considered the overbooking of seats problem for an airline network under a different fare structure called fare families, that are groups of highly similar products (fares), reduced from a larger fare classes problem. Demand is modeled as deterministic and dependent on the fare family, cancellations and overbooking. Soleymanifar (2019) investigated the problem of airline network overbooking by dealing with booking request decisions to maximize revenues. Request arrivals, cancellations and no-shows are simulated and modeled stochastically apart from the fully uncertain behavior of the demand in terms of its total quantity. Mou, Li and Li (2019) tackled the airline capacity allocation problem with overbooking but considered uncertainty in demand (in terms of distribution), show-up and cancelations as random variables where the latest two were based on historical data.

However, few attempts in literature addressed the multiple-leg flight overbooking problem. Furthermore, researchers addressed the problem by utilizing deterministic and probabilistic model; however, the use of fuzzy logic theory for modeling uncertainty, such as demand, is not well investigated.

3. Problem Statement

With the increase in hub-and-spoke network operations in the airline industry, the network effects in airline revenue management have become more important. There are many studies highlighting that considering the overbooking problem at a network level may significantly increase the profits of the airlines (Williamson 1992, McGill and van Ryzin 1999, De Boer, Freling and Piersma 2002).

In larger networks, there may be more than one hub and many spokes. Moreover, each itinerary may have many different fare classes. Hence hundreds of itineraries and products can be defined. In a network with N spokes, the number of itineraries is calculated as N (N+1). If the airline offers M classes on each flight the number of products is calculated as M (N) (N+1). Fig. 1 shows a larger airline network with one hub and three spokes. Therefore, there are 12 itineraries in this network.

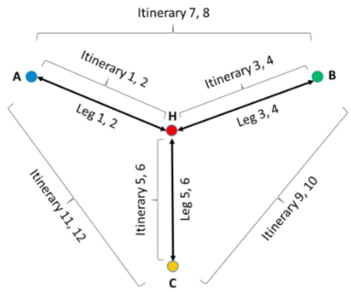

Fig. 1. Overbooking problem for an airline network with a hub and three spokes.

Assuming that the airline offers two fare classes on each flight the number of products becomes 24. The overbooking problem can be solved either at the leg level or the network level. At the leg level, the aim is to find the number of tickets to be sold for each fare class. However, at the network level, the aim is to determine the number of bookings requests to be accepted for any product (Gosavi, Ozkaya and Kahraman 2007). The network problem is naturally more complex due to the fact that seat allocation in one leg affects other legs in the network. Hence an airline operating flights in a network type organization should ideally address the problem at the network level.

4. The fuzzy-based mathematical optimization methodology

The proposed methodology consists of a fuzzy logic technique which is used to model the fuzzy non-deterministic features of demand on overbooking flight tickets, while the linear programming technique is used to imitate the overbooking problem described in section 3. A meta-heuristics approach represented by GA is also introduced for solving large-scale optimization problems.

In this methodology, GA is integrated with the developed fuzzy-based mathematical model for a reasonably good local optimal solution of the considered large-scale multi-leg flight networks. The function of the GA optimization is to provide the developed fuzzy-based model with high quality inputs of feasible parameters of decision variables. The proposed parameters are then evaluated using the proposed fuzzy-based model. This process continues in terms of iterations to evaluate the suggested numbers of allocated seats and denied boarding by the developed fuzzy-based model, while GA evaluates the performance of the resultant allocation of these parameters, and based on this, adjust the parameters of the decision variables using GA operators and select the most promising ones for further evaluation. The results of the evaluation in terms of revenues, O/B costs and net profits are captured per iteration, and the evaluation process continues until optimal/ near optimal solution is obtained.

5. The fuzzy demand overbooking mathematical programming model

In this section, the proposed Overbooking Fuzzy Mathematical Programming model is developed as an extension to the model previously introduced by Bertsimas and Popescu (2003) to accommodate fuzzy non-deterministic demand. Unlike probabilistic functions, the membership function of a fuzzy set involves subjectivity since human involvement is needed to obtain the fuzzy sets. The membership function of the future demand can be estimated based on the experience and intuition of the decision maker by considering historical data and present demand (Teodorović 1998). Although there are various types of fuzzy numbers defined in the literature, triangular fuzzy numbers are one of the simplest and most widely used types of fuzzy numbers that may be used to define uncertainty, mainly for the reason that fuzzy calculations are relatively easier for the triangular type of fuzzy numbers, which are preferred in this study to define uncertainties in demand probabilities.

The following notation is defined in the developed mathematical model:

- l: number of a leg in the network
- i: number of an itinerary in the network
- j: number of an itinerary fare-class combination
- OD: origin and destination pair
- ODF: origin-destination-fare class combination
- f_j: airfare if itinerary j
- C_l: capacity of the leg l
- q_j: show up rate for ODFj at the departure
- a_{lj}: resource coefficient of the ODFj
- θ_j: penalty cost of denied boarding for itinerary j
- $\tilde{D}_j$: fuzzy number denoting the demand for itinerary j
- z_j: number of allocated seats for itinerary j
- ω_j: number of denied boarding for itinerary j

The mathematical model ‘OB_FMP’ is formulated based on the assumptions as listed below:

- Only one airline is considered with three different fare classes
- No refund given for cancelations and no-shows
- Demand uncertainty is the same for all itineraries
- Triangular membership function is used to define uncertainties in demand probabilities
- Show-up is determined by a deterministic variable q_j
- Penalty cost for denied boarding is given by θ_j= airfare + X% of the airfare, where X is either 10 or 20.

The proposed fuzzy optimization formulation of the problem is shown below

$$\text{Maximize} \quad \sum_{j\in\mathcal{J}} f_j z_j - \sum_{j\in\mathcal{J}} \theta_j \omega_j \tag{1}$$

$$\text{Subject to} \quad \sum_{j\in\mathcal{J}} a_{lj}[q_j z_j - \omega_j] \le C_l \qquad l \in \mathcal{L} \text{ (for l=1,....,m)} \tag{2}$$

$$z_j \leq \tilde{D}_j \qquad j \in \mathcal{J} \text{ (for } j=1,\ldots,n) \tag{3}$$

$$\omega_j - q_j z_j \leq 0 \qquad j \in \mathcal{J} \text{ (for } j=1,\ldots,n) \tag{4}$$

$$z_j, \omega_j \geq 0 \qquad j \in \mathcal{J} \text{ (for } j=1,\ldots,n) \tag{5}$$

The objective function in (1) aims at maximizing the revenues from the allocated seats of each itinerary while taking into consideration the penalties resulting from the denied boarding, making z_j and ω_j as decision variables. Constraint (2) of the fuzzy model above ensures that the number of passengers allowed boarding is not greater than the leg capacities. Constraint (3) defines that the upper bound of the allocated seats to ODFj is equal to the fuzzy demand of that product so that the accepted bookings do not exceed booking requests. The fuzzy demand follows a triangular membership function of (d_{1j}, d_{2j}, d_{3j}). Constraint (4) ensures that the number of denied passengers does not exceed the number of passengers that show up at the departure time. Finally, constraints (5) are the non-negativity constraints.

Bellman and Zadeh (1970) introduced a method to deal with the fuzzy parameters by converting the fuzzy model into an equivalent crisp problem where the level of satisfaction of an objective function and a constraint can be defined by λ such that:

$$\mu_G(z_j) \geq \lambda \tag{6}$$

$$\mu_{C_n}(z_j) \geq \lambda \tag{7}$$

where $\mu_G(z_j)$ and $\mu_{C_n}(z_j)$ are membership functions of the goal and the nth constraint respectively and $0 \leq \lambda \leq 1$.

As the demand values ranges between d_{1_j} and d_{3_j}, two crisp problems are considered where the objective function takes values that are between P_1 and P_2. Let $P_l = \min(P_1, P_2)$ and $P_u = \max(P_1, P_2)$, then P_l *and* P_u are equal to the lower and upper bounds of the acceptable profit, respectively. The fuzzy set of optimal values for the acceptable profit is defined as follows:

$$\mu_G(z_j) = \begin{cases} 0, & if\ (\Sigma f_j z_j - \Sigma \theta_j \omega_j) < P_l \\ \dfrac{(\Sigma f_j z_j - \Sigma \theta_j \omega_j - P_l)}{(P_u - P_l)}, & if\ P_l \leq (\Sigma f_j z_j - \Sigma \theta_j \omega_j) < P_u \\ 1, & if\ (\Sigma f_j z_j - \Sigma \theta_j \omega_j) \geq P_u \end{cases} \tag{8}$$

The fuzzy set of the nth constraint is defined as follows:

$$\mu_{C_n}(z_j) = \begin{cases} 1, & if\ z_j < d_{1_j} \\ (d_{3_j} - z_j)/(d_{3_j} - d_{1_j}), & if\ d_{1_j} \leq z_j < d_{3_j} \\ 0, & if\ z_j \geq d_{3_j} \end{cases} \tag{9}$$

From equation (8), the inequality from (6) and $\mu_G(z_j)$ membership function shape adapted from (Teodorović 1998), the inequality can be transformed into:

$$\sum_{j \in \mathcal{J}} f_j z_j - \sum_{j \in \mathcal{J}} \theta_j \omega_j \geq P_l + \lambda(P_u - P_l) \tag{10}$$

Similarly, the fuzzy demand constraint is converted into a crisp constraint using the inequality in (7), equation (9) as well as the adapted shape of the demand membership function (Teodorović 1998) to get:

$$z_j \leq d_{1_j} + (1 - \lambda)(d_{3_j} - d_{1_j}) \tag{11}$$

The new objective function becomes equal to the level of satisfaction denoted by λ. The objective function (1) of the original fuzzy problem is transformed into a constraint formulated by equation (10) while the demand constraint (8) is transformed into a fuzzy constraint formulated by equation (11). The resultant crisp problem can be written as follows:

$$\text{Maximize} \quad \lambda \tag{12}$$

$$\text{Subject to} \quad \sum_{j\in\mathcal{J}} f_j z_j - \sum_{j\in\mathcal{J}} \theta_j \omega_j \geq P_l + \lambda(P_u - P_l) \qquad j \in \mathcal{J}\ (for\ j=1,\ldots,n) \tag{13}$$

$$\sum_{j\in\mathcal{J}} a_{lj}[q_j z_j - \omega_j] \leq C_l \qquad l \in \mathcal{L}\ (for\ l=1,\ldots,m) \tag{14}$$

$$z_j \leq d_{1j} + (1-\lambda)(d_{3j} - d_{1j}) \qquad j \in \mathcal{J}\ (for\ j=1,\ldots,n) \tag{15}$$

$$\omega_j - q_j z_j \leq 0 \qquad j \in \mathcal{J}\ (for\ j=1,\ldots,n) \tag{16}$$

$$z_j, \omega_j \geq 0 \qquad j \in \mathcal{J}\ (for\ j=1,\ldots,n) \tag{17}$$

The objective function $\max(\lambda)$ of the developed model shows the aspiration of maximizing the level of satisfaction. On the other hand, the objective function of the original deterministic problem is transformed into constraint (13) in the developed model which states that we can achieve a greater than satisfactory level of profit with a satisfaction level equal to λ. Constraint (14) ensures that the number of passengers allowed boarding is not greater than the leg capacities. Constraint (15) is the fuzzy demand constraint that defines the upper bound of the seat allocation of ODF_j is equal to the crisp value of the fuzzy demand so that the accepted bookings do not exceed booking requests. Constraint (16) ensures that the number of denied passengers does not exceed the number of passengers that show up at the departure time. Finally, constraint (17) is the non-negativity constraints.

6. Genetic Algorithm based metaheuristic optimization

The investigated overbooking multi-leg airline problem is relatively a large-scale problem and hence, solving it using traditional techniques requires an exponential runtime or an excessive amount of memory and hence, a meta-heuristics optimization technique such as GA is used for a reasonably good local optimal solution of the proposed fuzzy model.

In our mathematical model (12-17), the three decision variables are satisfaction level, number of assigned seats for ODF_j and number of denied passengers for ODF_j denoted by λ, z_j and ω_j, respectively. It should be noted that λ is a scalar variable whereas z_j and ω_j are column vectors. Therefore, these variables are transformed into row vectors when being encoded in the chromosome string. A representation of the chromosome structure used for the developed mathematical model is given in Fig. 2 below.

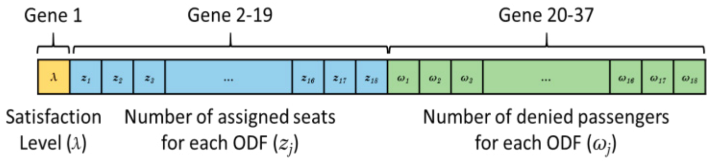

Fig. 2. Structure of the chromosome string used for the solution.

The chromosome string mainly comprises the three decision variables of the optimization model. The first gene represents the satisfaction level λ, which is a variable present in both the objective function and the constraints. The genes 2-19 are allocated to z_j, where each gene represents the number of assigned seats for each of the 18 different ODFs offered in the network. The final 18 genes in the chromosome string represent the number of denied passengers for each ODF denoted by ω_j in the developed model. The Genetic Algorithm basically creates a randomly generated population that comprises a number of chromosomes structured in the way shown in Fig. 2 and enables the population to evolve so as to search and obtain the optimal solution that represents the variables of the optimization model. A traditional GA including generation of population, selection, crossover and mutation is then applied.

7. Numerical Examples

In this section, small and large-sized problems are investigated. The small problem is adapted from literature and modified to accommodate fuzzy demand. This is used to validate the developed model including its solution approach-based GA. A moderately large-sized domestic flight network is then attempt and GA is applied.

7.1. The small-sized airline network overbooking problem

In order to validate and test the performance of the developed model, the numerical data set given by De Boer (1999) for a sample airline network is adopted. The demand parameters of the original problem are slightly changed according to our experimental analysis requirements. The lower and upper demand values that we use in our analyses, shown in Table 1 below, are calculated as ±10% of the Expected Demand value $E(D_j)$.

Table 1. Upper and lower bounds of the fuzzy demand.

# OD	OD Pair	Fare Class 3		Fare Class 2		Fare Class 1	
		d1j	d3j	d1j	d3j	d1j	d3j
1	AB	56	69	45	55	34	41
2	AC	45	55	28	34	23	28
3	AD	34	41	27	33	23	28
4	BC	34	41	23	28	23	28
5	BD	34	41	23	28	23	28
6	CD	56	69	45	55	34	41

In practice, there may be significant differences in demand for certain itineraries in an airline network based on seasonality or other external factors. However, in order to simplify the problem, it is assumed that uncertainties in demand are the same for all the itineraries. It should also be noted that, since a static solution approach is applied in the solutions, it will not be wrong to consider cancellations together with no shows. Hence, a combined show-up rate of q_j=0.75 is assumed for the overbooking optimization models with stochastic show-up probabilities in this study.

It is worth mentioning that the actual show-up rate is defined as the percentage of the demand booked that shows up at departure. This actual rate is obtained by the total number of passengers that show up at the time of departure for a given flight number divided by the total allocated seats for all the passengers. Predictions of the show-up rate could also be estimated using statistical distributions (Popescu et al. 2006). Actual demand data required to calculate this rate can be obtained from sales results accumulated in the seat reservation system of an airline, and a low-cost airline could be selected as the main source of such data. Future airlines seat reservations could be modelled and predicted using forecasting models (Varedi M. 2010).

In order to test the behavior of the developed model, the problem settings are varied, and the behavior of the fuzzy mathematical model is observed. Firstly, with the aim of testing different satisfaction levels, the upper and lower bounds (P_u and P_l) of the expected profit parameters of the model are varied as shown in Table 2. This is analogous to a revenue management analyst arbitrarily defining acceptable profit levels and calculating the satisfaction level and optimal booking levels (Teodorović 1998).

Table 2. Tested values for upper and lower bounds for expected profit.

Expected Profit Level	EP 1	EP 2	EP 3	EP 4	EP 5	EP 6	EP 7	EP 8	EP 9	EP 10
P_l	70,000	75,000	80,000	85,000	90,000	95,000	100,000	105,000	110,000	115,000
P_u	75,000	80,000	85,000	90,000	95,000	100,000	105,000	110,000	115,000	120,000

EP 9 is chosen because this is the case where we can observe a relatively high number of denied boarding while achieving a 100% satisfaction level. This allows for a better comparison between the two methods.

A test scenario is considered for the numerical analyses. The penalty costs for denied boarding, the airline refunds the initial airfare for the concerned itinerary, is considered 20% of the airfare as a compensation and calculated as f_j x 120%, where f_j is the fare for itinerary j. On the other hand, the demand level is increased by 20% for all the ODFs (d_{2j} x 120%), where d_{2j} is the average demand for itinerary j. The demand level is increase for the purpose of testing the model under increased demand conditions.

7.1.1. GA Parameters tuning for small sized problem

The configuration of the GA parameters mainly affects the best fitness of the objective function and how fast the GA converges to the optimal solution. Therefore, experiments are performed to find out the best GA configuration including population size, cross-over and mutation fractions. The experimental settings for the parameters are shown in Table 3 below. Mutation rate is experimented with both 1% and 5%. GA Creation function and Mutation function are set as uniform while selection function is set as stochastic uniform.

Table 3. GA parameter settings.

GA Settings/ Runs	1	2	3	4	5	6	7	8	9
Population Size	10	10	10	20	20	20	40	40	40
Crossover Fraction (%)	55	70	85	55	70	85	55	70	85

The obtained values for the satisfaction level at the end of each GA generation under 9 different parameter settings are illustrated in Fig. 3 (a) below. For GA runs with smaller size populations, the best fit values for the satisfaction level converge to values less than one, although the trend in the generation results suggest that if the GA is allowed to run for more generations, they may eventually reach one. On the other hand, GA generations with larger populations tend to converge to the optimum value more quickly and give results closer to one. Additionally, it can be seen that, although they both have a population size of 40, setting 9 converges to the optimal value faster than setting 8 as it has a higher crossover rate. Therefore, it can be concluded that setting 9 is the best configuration in terms of GA performance, which is the configuration with the largest population size and highest crossover fraction. The obtained values for the satisfaction level at the end of each GA generation under 9 different parameter settings are illustrated in Fig. 3.

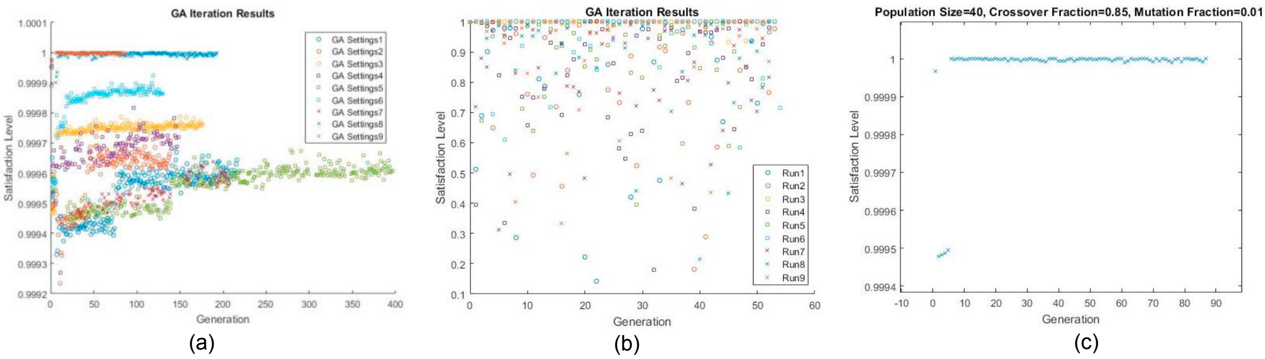

Fig. 3. (a) GA generations for each parameter settings (b) GA runs with mutation rate of 5% (c) Best GA run.

In order to observe the effect of mutation fraction on the performance of the GA, the mutation rate is set to 5%. Fig. 3 (b) above shows the behavior of the GA optimization for the increased mutation rate of 5%. GA settings/runs 1-9 have exactly the same settings as Settings 1-9 outlined in Table 32 except for the increased mutation fraction. By comparing the results shown in Fig. 3 (a) and (b), it can be seen that the increase in the mutation rate increases the randomness of the GA iterations. In addition, for the low mutation rate, increasing the crossover rate improves the optimization performance. On the other hand, for increased mutation rate, the performance reduces as the crossover rate increases. Based on these runs, it shows that the mutation and crossover are inversely proportional for this problem. Although there are very small differences between the final values, the satisfaction level diverges from one as the crossover rate increases under increased mutation conditions and hence the worst results are obtained from the runs with high mutation and high crossover fractions. Out of all the 18 runs so far, the best run is shown in Fig. 3 (c) above. All things considered, for the proposed GA based optimization method, it seems reasonable to use the GA setting 9 given in Table 3 with a mutation rate of 1%.

7.1.2. Validation of the Proposed Fuzzy based GA Optimization Model

In this section, MATHLAB is used to provide an exact solution to the fuzzy mathematical programming small-sized problem. This solution is then compared with the one obtained by using GA to validate the developed GA model. The validated model can then be adopted to tackle large-sized problems as presented in the next section. Table 4 shows the ODF level outputs of the test scenario at EP 9 using the mathematical and the proposed GA model.

Table 4. Optimization results by MP and GA.

Test Scenario					MP Optimization		GA Optimization	
ODF	Class	Fare	Penalty	Demand	Allowed Bookings	Denied Boarding	Allowed Bookings	Denied Boarding
AB	3	75	90	68	68	10	48	0
AB	2	125	150	54	54	0	54	0
AB	1	250	300	41	41	0	41	0
AC	3	130	156	54	0	0	20	12
AC	2	170	204	34	34	2	34	0
AC	1	400	480	27	27	0	27	0
AD	3	200	240	41	41	31	41	31
AD	2	320	384	32	32	0	32	0
AD	1	460	552	27	27	0	27	0
BC	3	100	120	41	41	0	27	1
BC	2	150	180	27	27	0	27	0
BC	1	330	396	27	27	0	27	0
BD	3	160	192	41	5	3	33	18
BD	2	200	240	27	27	0	27	0
BD	1	420	504	27	27	0	27	0
CD	3	80	96	68	59	0	50	0
CD	2	110	132	54	54	0	54	0
CD	1	235	282	41	41	0	41	0

It can be noticed that in both methods Mathematical Programming (MP) and Genetic Algorithm (GA), only passengers on lower fare classes are denied boarding as bumping passenger from lower fare classes leads to lower penalties. Also, both methods deny most bookings on longest route of AD as it comprises more flight legs than any other route in the network and a denied boarding on this route releases a seat on all the three flight legs in the network.

The output of these two approaches are represented and compared in Table 5. GA obviously has allowed slightly more bookings and caused a relatively higher denied passenger boarding. GA has generated more revenues, resulting in more overbooking costs. The extra generated revenues and the additional costs incurred are nearly the same, resulting in net profit that is very close to the profit obtained by mathematical modeling. The net profit from both methods is slightly more than the EP level because the optimization models work with non-integer numbers to calculate the allocated seats for each ODF and the results are rounded later at the end.

Table 5. Optimization KPIs for MP against Tuned GA.

Optimization Output	MP	GA
Satisfaction level	1	1
Total Bookings	632	637
Denied Boarding	46	62
Revenues	124,435	127,895
O/B Costs	9,324	12,888
Net Profit	115,111	115,007

Although an explicit limit is not set on the number of denied boarding in this study, in practice, a high number of denied boarding may cause loss of reputation and a decrease in revenues in the long run for the airline company. GA shows a rational behavior in how it allocates seats for different ODFs and which reservations it decides to deny boarding.

7.2. GA applied to a large-sized problem

For the large sized airline network overbooking problem, a randomly generated network that comprises 16 nodes is considered. The nodes are connected by 15 flight legs that are assumed to have equal capacities of 200 seats. The airline company serves 120 OD pairs over the entire network and three fare classes on each flight. Hence there are 360 different ODF combinations offered over the network. In practice, this corresponds to a moderately large-sized domestic flight network (Chen, Günther, and Johnson 1998).

For the numerical analyses, the network is first modeled for a single fare class and then extended to comprise 3 fare classes. Therefore, the first 120 ODFs in Table 6 represent the different OD pair combinations in the network for Fare Class 3, the next 120 ODFs represent the OD pairs for Fare Class 2 and the final 120 ODFs represent the OD pair combinations for Fare Class 1.

It is assumed that there is more demand for lower fare classes than the highest fare class and the demand values are randomly generated by the algorithm. The penalty for denied boarding is assumed to be equal to (airfare + 10% of the airfare). The parameters for the fare and demand settings are shown in the 6 below.

Table 6. Demand and fare settings for the large sized network.

ODF number	Upper and Lower Bounds of Demand	Show-up Rate	Fare Class	Airfare	Penalty Cost
1-120	40-50	90%	3	100	110
121-240	30-40	90%	2	200	220
241-360	20-30	90%	1	300	330

The best configuration for the GA is again chosen after experimentation with the GA parameters, in a similar way to that followed for the small sized network problem. Different combinations of population size, crossover fraction and mutation rate parameters of the GA are tested and the best configuration is concluded as shown in Table 7. The settings are similar to the ones used in the small sized network problem, however, this time population size is increased as the chromosome size has increased.

Table 7. GA configuration for the large sized network problem.

Parameter	Population Size	Crossover Fraction	Creation Function	Selection Function	Mutation Function	Mutation Rate
Value/ Definition	300	0.80	Uniform	Stochastic Uniform	Uniform	0.01

Fig. 4 represents solution results for the considered large sized network after implementing the developed fuzzy linear programming and GA optimization. Fig. 4. below shows the optimization results for the considered large sized network.

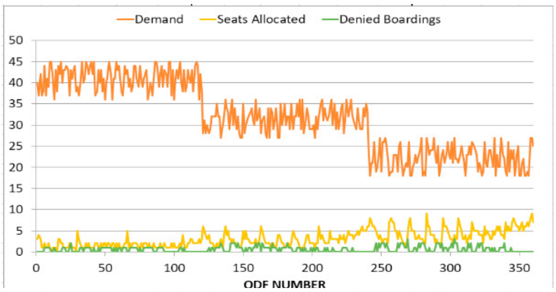

Fig. 4. GA optimization results for a large-sized network.

Fig. 4. shows that as the expected profit level increases, more bookings are allowed, and more passengers are denied since the overbooking level increases. When the demand is increased fewer bookings are allowed because it is possible to fill the available seat capacities with lower overbooking levels under increased demand conditions. When a passenger needs to be denied boarding, it is more likely that an itinerary with a longer route is picked by the algorithm because this way more seats become available as longer routes comprise more flight legs. When the EP level is low (and when the demand is higher than the available capacity of course), the number of denied passengers is increased. This is because the EP can be achieved with fewer passengers on board.

7.2.1. Experimentations

Seven different scenarios are considered. In each scenario, only one or two parameters from the original scenario are changed and the results are compared with those of the original problem. The penalty costs for denied boarding are increased to 20% in scenario 1 while keeping the demand, show-up rate and the expected profit level unchanged. In scenario 2, show-up probability for all fare classes is reduced to 80%. Demand for all ODFs in scenario 3 are reduced by 10% compared to the original problem. Both of the penalty costs and the expected profit are increase in scenario 4 by 20% and 25%, respectively. The last three scenarios experiment with different fare pricing settings. Scenario 5 has an unchanged EP level compared to the original problem, whereas the EP level is increased 25% in scenario 6. In scenario 7, the demand and EP level are increased by 20% and 25%, respectively. The changed input parameters in all scenarios are summarized in Table 8.

Table 8. Values of the parameters in test scenarios.

Scenario	Fare Settings	Penalty	Show-up Rate	Demand	EP Level
Original Problem	1	f_j x 110%	90%	d_{2j}	200,000
Scenario 1	1	**f_j x 120%**	90%	d_{2j}	200,000
Scenario 2	1	f_j x 110%	**80%**	d_{2j}	200,000
Scenario 3	1	f_j x 110%	90%	**d_{2j} x 90%**	200,000
Scenario 4	1	**f_j x 120%**	90%	d_{2j}	**250,000**
Scenario 5	**2**	f_j x 110%	90%	d_{2j}	200,000
Scenario 6	**2**	f_j x 110%	90%	d_{2j}	**250,000**
Scenario 7	**2**	f_j x 110%	90%	**d_{2j} x 120%**	**250,000**

* f_j=Initial airfare for itinerary j of and d_{2j}=Initial average demand for itinerary j

This kind of experimentation is similar to a revenue management analyst testing the overbooking algorithm under changing practical conditions due to seasonality or other reasons to find the optimal profit levels, therefore, show-up rates and demand are varied in order to mimic such seasonality. Penalty is increased to show the model behavior if the denied boarding cost has risen. The expected profit level is also increased to check the possibility of generating more profit while maintaining the denied boarding level.

In order to test the behavior of the overbooking optimization model, the inputs to the problem are varied which are the fare classes prices, the penalty cost θ_j as percentage of the fare f_j, show-up probability q_j, average demand

d_{2j} and the expected profit level. The fare prices are varied in two settings, the first fixes the prices for each fare class while the other bases the price on the travelled distance of the itinerary; prices are compared in Table 9 below. The behavior of the fuzzy based GA optimization model is observed.

Table 9. Fare settings for Scenarios (1-7).

Number of Legs	Fare Setting-1 Original Problem & Scenarios (1-4)			Fare Setting-2 Scenarios (5-7)		
	Fare Class 3	Fare Class 2	Fare Class 1	Fare Class 3	Fare Class 2	Fare Class 1
1	100	200	300	60	120	180
2	100	200	300	80	160	240
3	100	200	300	100	200	300
4	100	200	300	120	240	360
5	100	200	300	140	280	420
6	100	200	300	160	320	480
7	100	200	300	180	360	540
8	100	200	300	200	400	600
9	100	200	300	220	440	660
10	100	200	300	240	480	720
11	100	200	300	260	520	780
12	100	200	300	280	560	840
13	100	200	300	300	600	900
14	100	200	300	320	640	960
15	100	200	300	340	680	1020

Table 10 below compares the results of the original problem to all scenarios in terms of the KPIs.

Table 10. Optimization results of the original problem and each scenario.

KPIs	Original Problem	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7
Total Bookings	1,098	1,116	1,136	1,135	1,282	1,100	1,319	1,289
Denied Boarding	225	221	259	251	223	288	342	331
Revenues	255,800	259,700	263,100	261,400	308,098	333,781	393,752	388,792
O/B Costs	53,240	56,760	61,600	59,950	57,985	132,945	143,416	138,749
Net Profit	202,560	202,940	201,500	201,450	250,113	200,836	250,336	250,042

The original problem allowed for 1,098 bookings, making 255,800 in revenues, of which 20% were denied boarding which resulted in total penalty costs of 53,240. The net profit would be around 202,560. As mentioned earlier, the net profit in all scenarios are slightly more than the EP level due to the rounding of non-integer numbers that calculates the seat allocation. With penalty increases as in scenario 1, the O/B costs have increased; however, the model allowed marginally more bookings for more revenue to maintain the same EP level. With fewer show-up rates as in scenario 2, the model allowed 38 more bookings, compared to the original problem; however, the resulting O/B costs were high enough to decrease the profit by 1,000. Highly similar outcomes to scenario 2 occur when the demand was decreased by 10% in scenario 3. In scenario 4, since the EP level is higher, the optimization algorithm allows more bookings to fill the empty seats on the empty flights and achieve the targeted profit level. As a result, the load factor over the network increases but the number of denied passengers does not change considerably.

The Fare Setting-2 increases the average ticket price for the products offered over the network. Hence, the algorithm can reach the expected profit levels more easily. The optimization results show that the changed fare setting does not affect the number of accepted bookings significantly in scenario 5, though the GA algorithm tends to deny more passengers boarding the plane. This is because the airline can achieve the EP level of 200,000 by carrying fewer passengers. On the other hand, when the EP level is raised in scenario 6, the number of allowed bookings and denied passengers increase. As a result, both the overbooking costs and the total revenues generated over the network increase significantly. Finally, looking at the results of scenario 7, compared with the Original Problem, it shows that the optimization model inclines to allow fewer bookings as the demand level increases because the airline is more likely to fill the available seat capacities with lower overbooking levels under increased demand conditions. The outputs of Scenarios (5-7) show that the distance-based fare setting does not change the character of the optimization algorithm. However, the overbooking level is directly related to the ticket prices and hence the GA decides on the number of accepted bookings and overbooked passengers to comply with the expected profit level.

8. Conclusion and Future Work

This study proposed a mathematical model for an overbooking multi-leg airline problem under fuzzy demand conditions and proposed a Genetic Algorithm (GA) model aiming to solve large-sized problems efficiently. The GA was implemented on a large-sized problem tested under different problem settings by changing fare pricing, denied passenger penalties, show-up rate, demand and the expected profit level. Results indicated that the GA model gives rational results when applied on the large sized network problem and adapts to the different problem settings. This paper proves that the proposed GA can be utilized for solving different sizes of the fuzzy overbooking optimization problems since they have shown its suitability in solving such problems efficiently.

As for future work, the model can be extended to represent both the show-up rate and the future demand. In addition, a booking control policy, such as service level policy, may be included in the optimization model to control the number of denied passengers.

References

Alavi F.F., Sy M., 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. https://doi.org/10.1016/j.tre.2019.07.001

Bellman, R.E., Zadeh, L.A., 1970. Decision-Making in a Fuzzy Environment. Management Science 17, B-141.

Bertsimas, D., de Boer, S., 2005. Simulation-Based Booking Limits for Airline Revenue Management. Operations Research 53, 90–106.

Bertsimas, D., Popescu, I., 2003. Revenue Management in a Dynamic Network Environment. Transportation Science 37, 257–277.

Chen, V. C., Günther, D., and Johnson, E. L. (1998) A Markov decision problem based approach to the airline YM problem. In YM Problem, Working Paper, the Logistics Institute, Georgia Institute of Technology

Cross, R.G., 1997. Revenue management: hard-core tactics for market domination, 1st ed., New York: Broadway Books.

De Boer, S.V., 1999. Mathematical programming in airline seat inventory control (Master's thesis). Department of Mathematics, Vrije Universiteit, Amsterdam, the Netherlands.

De Boer, S.V., Freling, R. & Piersma, N. (2002) 'Mathematical programming for network revenue management revisited'. European Journal of Operational Research 137 (1), 72

Erdelyi, A., Topaloglu, H., 2009. A Dynamic Programming Decomposition Method for Making Overbooking Decisions over an Airline Network. INFORMS Journal on Computing 22, 443–456.

Gosavi, A., Ozkaya, E., Kahraman, A.F., 2007. Simulation optimization for revenue management of airlines with cancellations and overbooking. OR Spectrum 29, 21–38.

Hjorth, R., Fiig, T., Larsen, J., Bondoux, N., 2018. Joint overbooking and seat allocation for fare families. Journal of Revenue and Pricing Management 17, 436–452.

Karaesmen, I., van Ryzin, G., 2004a. Overbooking with Substitutable Inventory Classes. Operations Research 52, 83–104.

Karaesmen, I., van Ryzin, G., 2004b. Coordinating overbooking and capacity control decisions on a network. Unveröffentlichtes Arbeitspapier, Graduate School of Business, Columbia University, New York.

Klophaus, R., Pölt, S., 2010. Airline Overbooking Considering Passengers' Willingness to Pay. Trier University of Applied Sciences.

Kunnumkal, S., Topaloglu, H., 2008. A tractable revenue management model for capacity allocation and overbooking over an airline network. Flexible Services and Manufacturing Journal 20, 125–147.

Ma, J., Tse, Y.K., Wang, X., Zhang, M., 2019. Examining customer perception and behaviour through social media research – An empirical study of the United Airlines overbooking crisis. Transportation Research Part E: Logistics and Transportation Review 127, 192–205. https://doi.org/10.1016/j.tre.2019.05.004

McGill, J.I., van Ryzin, G.J., 1999. Revenue Management: Research Overview and Prospects. Transportation Science 33, 233–256.

Mou, D., Li, W., Li, J., 2019. A network revenue management model with capacity allocation and overbooking. Soft Computing.

Popescu A., Keskinocak P., Johnson E. LaDue M., Kasilingam R. (2006). "Estimating Air-Cargo Overbooking Based on a Discrete Show-Up-Rate Distribution". INFORMS, Vol. 36, No. 3, May–June 2006, pp. 248–258.

Siddappa, S., Rosenberger, J.M., Chen, V.C.P., 2008. Optimizing airline overbooking using a hybrid gradient approach and statistical modeling. Journal of Revenue and Pricing Management 7, 207–218.

Soleymanifar, R., 2019. An LP based approximate dynamic programming model to address airline overbooking under cancellation, refund and no-show. Yugoslav Journal of Operations Research 29, 113–132.

Sumida, M., Topaloglu, H., 2019. An Approximation Algorithm for Capacity Allocation Over a Single Flight Leg with Fare-Locking. INFORMS Journal on Computing 31, 83–99.

Teodorović, D., 1998. Airline network seat inventory control: a fuzzy set theory approach. Transportation Planning and Technology 22, 47–72.

Varedi M. (2010). "Forecasting seat sales in passenger airlines: introducing the round-trip model". MSc thesis, Waterloo, Ontario, Canada.

Wangenheim, F. v., Bayón, T., 2007. Behavioral Consequences of Overbooking Service Capacity. Journal of Marketing 71, 36–47.

Williamson, E.L., 1992. Airline network seat inventory control: Methodologies and revenue impacts (Ph.D. thesis). Massachusetts Institute of Technology.