Airline overbooking in the multi-class case

J Coughlan

Dublin Institute of Technology, Dublin, Ireland

This paper presents an airline overbooking model at a class level for one service compartment–cabin. Class level demand data is used to determine the number of bookings that can be taken for each class. The model is optimised through the use of multi-dimensional search routines. The control level model developed is tested with data supplied by Ireland's national airline, Aer Lingus. The model shows a significant improvement over previous methods employed by Aer Lingus and was subsequently adopted by the airline.

Keywords: air transport; yield management; overbooking

Introduction

Airlines operate their passenger services on the basis of advance reservations. The air travel market is very competitive in Ireland, especially on the Ireland–UK routes and new electronic systems of booking enable the customer to make multiple reservations with different airlines. The airline may find that fewer passengers show up for a flight than are expected from historical patterns. Traditionally, the level of passengers who book but do not show up is estimated from historical data, and the flight is overbooked in order to maximize revenue and reduce the possibility of flying with an empty seat.

Overbooking is the practice of intentionally selling more reservations for a flight than there are actual seats on an aircraft.[1] This reduces the number of vacant seats but may also result in the rejection of passengers with confirmed bookings which costs the airline in terms of compensation payouts. Also in a period of intensifying competition, the loss of goodwill from this activity is not welcome.

Overbooking is an aspect of yield management that is also used in other service businesses[2,3] and falls under the general area of Perishable Asset Revenue Management (PARM).[4] It has been described as economically inefficient and is often thought to be brought about by operational imperfections in demand forecasting.[5] In the single service compartment case, for example, the economy cabin, the airline is selling what is essentially the same seat for different fares. This is also the case in the other service compartments. Overbooking historically was carried out at service compartment (cabin) level.[6] This was achieved by aggregating class level data and determining an overbooking level at cabin level. Intuitively, individuals who book in different classes should behave differently in terms of

booking patterns and cancellations and this causes distortion in the aggregation to cabin level. The solution is to use class level data and determine an overbooking level for each class.

In the first section of this paper, the choice of model will be justified, the assumptions presented and the model developed. The solution methods will be presented, the starting heuristic developed and the Aer Lingus model explained, and finally, concluding remarks and future research directions will be discussed.

The overbooking model

Historically, overbooking models have taken an incremental control approach.[7–13] This allows a maximum number of additional reservations to be taken based on the reservations already taken for that flight and the historical patterns for the flight. However, several airlines use a level control approach in which reservations are accepted until the total number of reservations exceed specified levels (authorisations).[11] This paper takes a level control approach. This required more detailed data about reservation and cancellation behavior, although this data is now being collected by the airlines. For the purposes of this paper, the data was made available by Ireland's national airline Aer Lingus. The advantage of a level control model is that for a given length of time, the level control approach responds faster to rapid changes in reservations.[11]

A revenue maximisation model was used. This involves the calculation of the estimated number of bookings who will show up in each class based on the bookings received and the historical data on the bookings that failed to show up in that class. Making a number of assumptions about the distributions of how customers book and 'no show', the expected number of 'show ups' can be calculated at class level. This is then used to set an overbooking level for each

Correspondence: J Coughlan, Dublin Institute of Technology, Mountjoy Square, Dublin 1, Ireland.

class, called the 'authorization'. If more passengers show up than there are available seats, the extra passengers will be denied boarding and the airline will incur a compensation cost. Also, if the service compartment (cabin) is not full after the passengers who booked have boarded, then passengers who turned up without a booking will be boarded. These are known as 'go shows'. It follows that, in one cabin, 'denied boardings' and boarded 'go shows' cannot occur together.

Assumptions

A general taxonomy of yield management is presented in Weatherfield and Bodily[4] and this paper uses these general assumptions. However there are some specific assumptions which are required for this model.

1. *The model does not employ nesting of classes.* Most airlines use fare nesting in their yield management models. Classes are ordered according to fare and nesting allows the yield analyst to offer a prospective passenger the opportunity of booking a seat in a higher fare class than they originally requested if the requested fare class is full and there is availability in the higher fare class. Ireland is experiencing an unprecedented degree of competition in the air travel market, particularly the Dublin–UK market segment. In this competitive environment, if a particular fare class is not available, the individual making the booking may go to a competitor airline.
2. *The different classes book over the same period.* This is a simplification. Many airlines require different fare classes to have different booking periods, whereas some classes are allowed to book all the way through a booking period. The case of purely sequential booking based on rising fares per class was considered but was felt that this was not how customers actually booked.[14] The model could be extended to look at this aspect.
3. The number of passengers seeking bookings is assumed to be a normally distributed random variable. It is reasonable to assume that the receipt of requests for bookings over any given period will follow a Poisson distribution. In the case, where booking periods are long (assumption 2), it is reasonable to approximate the number of requests for bookings by the normal distribution.
4. The 'No Show' rate does not vary with time and is independent of the number of bookings in that class. The 'No Show' rate for those passengers who book in the last few days of the booking period is assumed to be equal to the rate for those who book at the start of the period. This is termed the 'fortgetfulness property'.[7]
5. The number of 'no shows' from any given booking level is binomially distributed. The probability of x 'No Shows', out of b bookings, is assumed to be binomial.

That is,

$$P(X = x) = \binom{b}{x} r^x (1 - r)^{b-x}$$

where p is the probability that any booking will result in a 'No Show', namely the No Show rate.

6. The number of 'go shows' in any class is independent of the number of 'show ups' in that class. This assumption is based on the independent of 'Go Shows' and booking demand. It can be argued that the two are dependent, but in general it appears that the assumption of independence is reasonable.
7. The probability of a booking resulting in a 'No Show' is independent of whether that booking is part of a group. Group bookings in many cases will come not from actual groups, but from a collection of individuals booked by an agent. This means that group identification is in itself a major issue.[15]

Notation and terms

The following are used in the development of the model:

n = number of classes in the service compartment

c = total number of seats available in the service compartment (cabin)

Let t_i, the demand for class i, be an independent normally distributed random variable with cumulative distribution $T(\ )$.

Let b_i, the number of bookings in class i, be an independent normally distributed random variable with p.d.f. $p(b_l)$ mean: b_i and standard deviation F_{bi}. Let s_i, the number of 'Show Ups', be an independent normally distributed random variable with p.d.f. $y(s_i)$, means: s_i, standard deviation Fs_i and cumulative distribution $S(\ )$.

Let r_i, be the no show rate in class i

Let g, the number of 'Go Shows', be an independent normally distributed random variable with p.d.f. $q(g)$, mean: g, standard deviation Fg and cumulative distribution $Q(\ )$

f_i = fare in class i

a_i = authorisation in class i. The authorisation is the maximum number of bookings allowed for each class

d = number of 'Show Ups' who are denied boarding

h = number of 'Go Shows' who are permitted to board

m = denied boarding cost

s_c = number of 'Show Ups' at cabin level

Development of function

The revenue is generated from the number of 'show ups' in each class multiplied by the fare in that class. The 'show

ups' are calculated as follows, the number of bookings, b_i, is

$$b_i = \begin{cases} t_i & \text{if } 0 \leqslant t_i \leqslant a_i \\ a_i & \text{if } t_i > a_i \end{cases}$$

If, $E(b_I)$ represents the expected bookings in class i, it follows that

$$E(b_I = \int_0^{a_i} t T_i(t) dt + a_i \int_{a_i}^{\infty} T_i(t) dt$$

Having now calculated the expected bookings in each class from the demand in each class, an expression was developed for 'Show Ups'. The number of 'show ups' is the number of bookings in a class multiplied by the 'No Show' rate

$$E(s_i) = E(b_i) \times (1.0 - r_i)$$

It follows that the expected revenue is

$$\text{rev} = \sum_{i=1}^{n} \left[\left(\int_0^{a_i} t T_i(t) dt + a_i \int_{a_i}^{\infty} T_i(t) dt \right) \times (1.0 - r_i) \times f_i \right]$$

However the capacity of the service compartment (cabin) has to be taken into account. The total number of 'show ups' must be calculated and checked against the capacity to see if there are too many 'show ups' (denied boarding situation) or if there are empty seats ('Go Shows' may be boarded). The number of 'show ups' at cabin level is the sum of the class level 'show ups'

$$s_c = \sum_{i=1}^{n} s_i$$

The expression for the number denied boarding, if any, is as follows

$$d = \begin{cases} s_c - c & \text{if } s_c < c \\ 0 & \text{otherwise} \end{cases}$$

The expected number of 'denied boardings' is as follows

$$E(d) = \int_0^{s_c} (s_c - c) S_c(s) ds$$

Since those passengers who were denied boarding were included in the revenue function above, their contribution to the expected revenue must be subtracted. Also for each of these passengers, a figure representing denied boarding compensation, m, should be subtracted.

Because denied boarding occurs at the cabin level and passengers pay different fares for seats, the calculation of the contribution of the denied boarding cannot be made exactly. In order to combat this, a weighted average fare, v, is calculated. Once a booking shows up they get a seat in the cabin, after the cabin is full, the remainder of the bookings that show up are denied boarding. This model does not take the fact that there could be available seats in another cabin, as it is a single cabin model only. It could be

assumed that those passengers who booked last are more likely to be late for the flight and therefore the highest fare is the fare to use. It is also possible that any other passenger may be late. Therefore the contribution to be subtracted is the weighted average of the fares of the 'show ups'.

$$\text{rev} = \sum_{i=1}^{a_i} t T_i(t) dt + a_i \int_{a_i}^{\infty} T_i(t) dt) \times (1.0 - r_i) \times f_i]$$
$$- \left(\int_0^{s_c} (s_c - c) S_c(s) ds \times (m + v) \right)$$

The situation may also arise that the total 'show ups' is less than capacity and there will be empty seats. If there are 'Go Shows', then they will be allowed to board. As was the case with 'denied boardings', 'go shows' are calculated at cabin level. It may be assumed that 'go shows' pay the highest fare, but this is not always the case. 'Go Shows' may be individuals with open tickets (which may be at a number of fares) or may be individuals with 'stand-by' tickets (which are relatively inexpensive). To overcome these possibilities, a weighted average fare was calculated. Provided there are empty seats on the plane (namely, $c - s_c > 0$) then,

$$h = \begin{cases} g & \text{if } 0 \leqslant g \leqslant (c - s_c) \\ c - s_c & \text{if } g > (c - s_c) \end{cases}$$

The expected number of boarded go shows, $E(h)$, for a given number of 'Show Ups' is:

$$E(h) = \int_0^{c-s_c} g Q_c(g) dg + (c - s_c) \int_{c-s_c}^{\infty} Q_c(g) dg$$

This must be integrated over all possible 'Show Ups' in order to calculate the expected boarded 'go shows' for the cabin

$$bgs = \int_0^c \left(\int_0^{c-s_c} g Q_c(g) dg + (c - s_c) \int_{c-s_c}^{\infty} Q_c(g) dg \right) S_c(s) ds$$

combining to give,

$$\text{rev} = \sum_{i=1}^{n} \left[\left(\int_0^{a_i} t T_i(t) dt + a_i \int_{a_i}^{\infty} T_i(t) dt \right) \times (1.0 - r_i) \times f_i \right]$$
$$- \left(\int_0^{s_c} (s_c - c) S_c(s) ds \times (m + v) \right)$$
$$+ \left[\int_0^c \left(\int_0^{c-s_c} g Q_c(g) dg (c - s_c) \int_{c-s_c}^{\infty} Q_c(g) dg \right) S_c(s) ds \right]$$

Solution methods

Having developed the revenue function, a number of optimisation techniques were available. Derivatives could be sought and either by developing an iterative solution procedure or by use of a package, the optima found. The calculation of derivatives in this case is quite complicated and a simpler approach was sought.

Direct search methods were investigated. These methods use the value of the function to find a maximum (or minimum) and are very robust although these could not guarantee optimality, due to the relatively flat areas found in the function (see Figure 1). Two methods were chosen to be tested, Hooke and Jeeves[16] and Nelder and Mead.[16]

Due to the flat areas of the function, a starting heuristic was sought in order to speed up the solution time. A number of different heuristics were developed and tested, the final one chosen was the following:

Step one

A cabin-level overbooking rate, w, was calculated by using the historical demand and no show data for that flight, as follows

$$w = \frac{\sum_{i=1}^{n}(t_i \times (1.0 + r_i))}{\sum_{i=1}^{n} t_i}$$

Using this method, a class with large demand but a low 'no show' rate would not dominate and neither would a class with low demand but a large 'no show' rate.

Step two

The above rate is used to find an authorisation (total number of bookings to accept) at cabin level.

Step three

The cabin level authorisation is allocated across the classes using the Fare Mix Algorithm. The algorithm is based on the principle that since $\text{EMSR}_i(j)$ is the additional revenue

Figure 1 Graph of a sample two class case.

that is expected to accrue when the jth seat is allocated to fare class i, then

$$\text{EMSR}_i(j) = f_i \times T(t_i > j)$$

The EMSR of the jth seat in the ith fare class is the seat fare multiplied by the probability of there being more than j requests for seats in that class. The algorithm goes through the seats in the cabin on an iterative basis and at each stage allocates the seat to the class which shows the greatest EMSR.

The Aer Lingus model

At the time of the development of the revenue model above, Aer Lingus were using the following heuristic procedure for overbooking. Their approach was based on a cabin-level overbooking strategy. Class-level data was collected and aggregated to cabin-level. The procedure was as follows

Step one

Starting with a cabin-level authorisation of capacity, the expected 'show ups', 'denied boardings' and empty seats were calculated.

Step two

The total cost of flying with this authorisation was calculated using the fares from the 'show ups', an empty seat cost and a denied boarding cost.

Step three

The cabin-level authorisation was incremented and steps one and two repeated until the difference between the new and old cost was less than a specified tolerance.

The data

The data for testing the revenue function developed was sourced from Ireland's national airline, Aer Lingus. The data was from single-leg flights and was adjusted for special events. Special events occur during the year which change the demand patterns for flights. Day of week effects, sporting events, national and school holidays are all examples of special events which change the demand patterns. The data was also adjusted for seasonality.

Analysis of results

Extensive trials were carried out using real data from Aer Lingus. The solution procedures employed were tested to see if they were reaching the global optimum. As can be
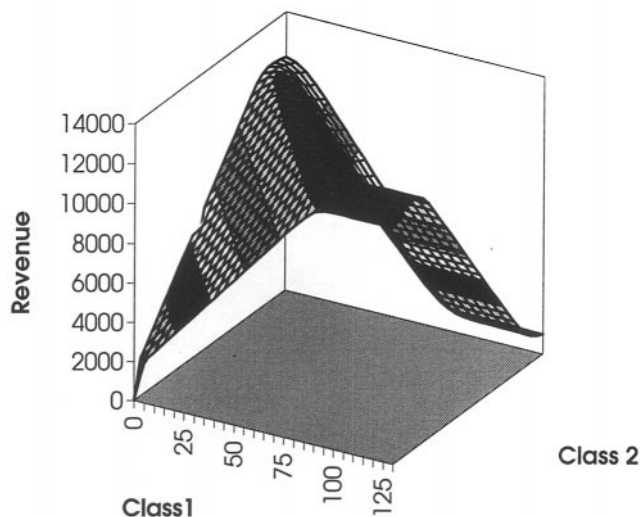
seen from Figure 1, the solution area for the model in the two class case shows flat areas, and also the model has many local optima close to the global optimum. This led to the development of the starting heuristic developed in the previous section.

Through complete enumeration of the model for a number of test cases of varying dimensions (no. of classes), it was found that the model did not achieve the global optimum in approximately 15% of cases. An ancillary result of this test was that complete enumeration was not a viable solution strategy due to the computation time involved.

The two solution procedures were employed separately and together to attempt to achieve the global optimum. When used individually the method of Hooke and Jeeves[16] outperformed Nelder and Mead,[16] but not by a significant amount. The solution procedure was then applied to the model in sequence and although this led to better results, the increase in computation time was not justified for the magnitude of the increase achieved.

The model was tested for different levels of denied boarding cost. The authorisations (allowed number of bookings) per class were found to be more conservative as the denied boardings cost increased.

The data contained in Table 1 is used to illustrate the model developed. Table 2 shows the solution of the model using each of the procedures employed. Upon further investigation of this case, it can be seen that the direct search methods are authorising higher bookings in the classes with higher fares and reducing the authorisations for lower fare classes. Further testing bore out this observation and it can be concluded that the Aer Lingus model was more conservative than the model developed in this paper. The average potential revenue improvement from the use of this model was in the order of 1–2%.

Conclusions

The Aer Lingus model was found to be conservative in comparison to the model developed. The model developed also required estimation of one less parameter than the Aer Lingus model because it implicitly brings in the cost of an empty seat, whereas the Aer Lingus model has to explicitly determine this parameter.

The solution procedures employed do not guarantee optimality, an area which future research will address. The model uses cabin level data for the 'boarded go shows' and the 'denied boardings'. This is to be expected because these activities occur at cabin level rather than at class level. The model does not allow for 'nesting' of classes within the cabin and future research will concentrate this aspect of the problem. In conclusion, the expected revenue model is currently in use by Aer Lingus for overbooking. It achieves increases in revenue of the order of 1–2% on average over the Aer Lingus model.

References

1 Smith B, Leimkuhler J and Darrow R (1992). Yield management at American airlines. *Interfaces* **22**: 8–31.
2 Bitran G and Mondschein S (1995). An application of yield management to the Hotel Industry considering multiple day stays. *Opns Res* **43**: 427–443.
3 Rothstein M (1974). Hotel overbooking as a Markovian sequential decision process. *Dec Sci* **5**: 389–404.
4 Weatherfield L and Bodily S (1992). A taxonomy and research overview of perishable-asset revenue management: yield management, overbooking and pricing. *Opns Res* **40**: 831–844.
5 Botimer T (1996). Efficiency considerations in airline pricing and yield management. *Trans Res A* **30**: 307–317.
6 Rothstein M (1985). OR and the airline overbooking problem. *Opns Res* **33**: 237–248.
7 Thompson H (1961). Statistical problems in airline reservation control. *Opl Res Q* **12**: 167–185.
8 Schlifer R and Yardi Y (1975). An airline overbooking policy. *Trans Sci* **9**: 101–114.
9 Alstrup J, Boas S, Madsen O and Vidal R (1986). Booking policy for flights with two types of passengers. *Eur J Opl Res* **27**: 274–288.
10 Belobaba P (1989). Application of a probabilistic decision model to airline seat inventory control. *Opns Res* **37**: 183–197.

Table 1   Test data for three class case, cabin capacity=50

	Class 1	Class 2	Class 3
Demand	15	25	50
No show rate	20%	10%	5%
Fare	160	125	112

Table 2   Solution of a sample three class case, cabin capacity=50

Methods	Authorisations			Total	Exp. Rev.	Time (s)
	Class 1	Class 2	Class 3			
Nelder and Mead	20	22	14	56	6508.14	0.096
Hooke and Jeeves	20	27	12	59	6518.72	0.168
Start heuristic	15	22	18	55	6441.92	0.059
Aer Lingus	15	22	17	54	6427.77	0.063

11 Alstrup J *et al* (1989). Booking control increases profit at Scandinavian airline. *Interfaces* **19**: 10–19.

12 Chatwin R (1996). Optimal control of continuous-time terminal-value birth-and-death processes and airline overbooking. *Naval Res Logist* **43**: 159–168.

13 Brumelle S and McGill J (1993). Airline seat allocation with multiple nested fare classes. *Opns Res* **41**: 127–137.

14 Robinson L (1995). Optimal and approximate control policies for airline booking with sequential nonmonotonic fare classes. *Opns Res* **43**: 252–263.

15 O'Connor B (1991). An evaluation of quantitative seat inventory models in airline yield management, Management Science Thesis, University College Dublin.

16 Bunday B (1985). *Basic Optimisation Methods*. Edward Arnold: London.