Journal of Air Transportation World Wide Vol. 5, No. 2 – 2000

# A FUZZY APPROACH TO OVERBOOKING IN AIR TRANSPORTATION

**Matteo Ignaccolo**
**and**
**Giuseppe Inturri**
**Catania, Italy**

## ABSTRACT

A high load factor is important for airlines trying to maximise their profits without alienating customers. The loss of revenue caused by empty seats cannot be recaptured. The aim of this paper is to propose a method that minimises the unused seats and the denied boarding at the same time for every single flight. This can be achieved by monitoring the booking process during the days before the departure and by using an Inference Fuzzy System as an easy decision support system to assist the revenue management analysts.

## INTRODUCTION

A flight, like most services, is produced by an airline company while supplying and cannot therefore be stored. If an aircraft takes off with some empty seats, there is a loss of revenue that cannot be recaptured.

The marginal revenue of an extra passenger occupying a seat which otherwise would have not been sold, is very large, while the additional supported costs are very small. For this reason it is very important for the airlines to reach a high load factor of the aircraft.

The problem is that even if a flight is sold out, that is, the aircraft capacity matches exactly the number of booked seats, it is almost sure that the aircraft will leave the gate with some empty seats. This happens because some passengers do not appear to claim their seats the day of the departure and some cancel their reservation too late to allow the company to sell the seats again.

Matteo Ignaccolo is a researcher at the University of Catania, Istituto Strade Ferrovie Aeroporti in Catania Italy.

Giuseppe Inturri earned his Ph.D. at the University of Palermo. He is a professor at the Istituto Strade Ferrovie Aeroporti in Catania Italy.

To reduce these effects most airlines overbook their scheduled flights to a certain extent in order to compensate for no-shows. As a consequence, some passengers are sometimes left behind or bumped as a result. By bumping passengers from an oversold aircraft, an airline can incur costs ranging from nothing, if the excess passengers can be rebooked with the same airline on a later flight that day, to meals, hotel rooms, vouchers for free flights, and the cost of transportation on another airline, not considering the potential loss of customer goodwill.

Overbooking and automated reservation systems are today an important chapter of the yield management, which has become a basic tool for the survival of the airlines in the air transport market, increasing today more and more in competitiveness and complexity. It has been evaluated that in the period from 1989 to 1992 American Airlines have saved through yield management about 50 percent more than its net profit for the same period (Davis, 1994).

Generally airlines accept reservation requests up to a booking limit, if the number of initial reservations is less than the booking limit, and decline the reservation requests otherwise.

As the number of no-shows is a stochastic variable, it is possible that the passengers that show up are more than the available seats for the flight, thus producing the opposite problem of the seat spoilage, that is, a number of denied boarding. These may be voluntary, if a passenger with a confirmed reservation accepts some kind of refund to abdicate the flight (money, hotel accommodation, meals, etc.), otherwise is an involuntary denied boarding, causing damages to the company image and additional costs.

Selling more seats than the aircraft's capacity might be seen as an incorrect behaviour, but the airlines sustain that without the balancing factor of an overbooking policy, the load factors of the flight would be lower than the actual one, thus producing an inevitable increase in the average fares.

The problem we want to face in this paper is what kind of booking policy should an airline adopt in the days before the departure in order to reduce the double risk of empty seats and denied boarding. In other words the company should establish what is the optimal authorisation level at any given time before the take-off, that is, the optimum number of reservations to be accepted.

The aim of this paper is to propose a method which minimises the spoiled seats and the denied boarding at the same time for every single flight. This can be achieved by monitoring the booking process during the days before the departure and using an Inference Fuzzy System as an easy decision support system to assist the revenue management analysts. This allows an understanding of any unusual event or action taken by competitors for each flight from the opening of the reservations to the take-off.

## REVIEW OF THE EXISTING MODELS

Several models had been proposed in these last four decades based on different approaches to match the objectives of the airline companies.

The cost minimisation model (Beckmann, 1958 and Kosten 1960 in Holm, 1995) finds the optimal authorisation level as the one which determines the minimum expected total cost of overbooking, calculated as the sum of the cost due to denied boarding (that increases with the number of accepted bookings) and the spoilage due to empty seats (that is reduced instead).

In 1961, Thompson proposed a model to limit the probability of denied boardings calculated as the area of a standard normal distribution of the number of show-up passengers exceeding the aircraft capacity (Holm, 1995).

A similar approach was used by Taylor (1962), which takes into account the ratio of denied boarding over the number of booked passengers as a constraint not to be overcome, while Rothstein & Stone (1967) maximises the expected revenue of the flight under the limit of an acceptable pre-set risk of denied boarding.

The model made by Gerbracht (1979) for Continental Airlines selects the optimum level of booking to maximise the expected net revenue as a result of the revenue obtained from the passengers actually carried, and also the penalty arising from the number of passengers with denied boarding. Since the number of no-shows varies randomly for each flight, if the probability distribution of no-shows is given, the statistical expected net revenue can be maximised. As it is much more expensive to have a denied boarding than to spoil an empty seat, the optimum booking levels are always shifted toward low overbooking values with regard of the average no-shows.

Alstrup (1986) considers different booking policy for different classes of passengers (and fares) and models the booking process as a non-homogenous Markovian chain. The aim is to find the optimal level of booking to be adopted for each class and for each time interval.

Andersson (1989) treats the case of an aircraft with a flexible cabin divided into fare classes, with different priority in respect of denied boardings.

Dunleavy (1994) uses a classical probabilistic overbooking approach to the determination of the marginal fare for a group and the uncostraining of origin-destination fare level data within a seamless, bid-price environment.

Reviewing the different overbooking models used by the airlines, we note that they try to arrange a compromise between the aim of maximising the net revenues with the need of assuring a more competitive level of service, avoiding the denied boarding as much as possible. However most of the proposed models are focused on determining the expected number of no-shows in terms of probability distributions.

The ratio of denied boardings per 1,000 boarded passengers is often used as an overbooking performance index. Although the airlines are interested in keeping a fixed level of service, the basic aim is the revenue maximisation. This is why it is becoming critical to monitor the booking process in real-time in order to counterbalance changes and shifts from the expected values of the process output variables (i.e., the mean show-up rate). In fact the show-up rate is probabilistic, therefore uncertain, and besides is the aggregate result of the available historical data. So a perfect hit on every flight cannot be achieved on a probabilistic base. This why often the airlines allow the intervention of a booking analyst that overrides the automated system's overbooking advice in order to embody common feeling and human judgement in unusual situations.

Whatever is the method adopted, we believe that for a given flight the limits of the authorisation level cannot be evaluated through static considerations owing to the tightly dynamic nature of the booking process, which requires a continuous check and change of these limits, in order to suit the unpredictable passenger behaviour, which becomes more and more changeable as the day of the take-off approaches.

Fuzzy Logic represents a very promising mathematical approach to model a process characterised by subjectivity, uncertainty and imprecision. The linguistic information expressed by a booking analyst is a subjective knowledge which can hardly be incorporated in a classical mathematical model. In the next section the basic fuzzy logic theory assumptions are presented.

## WHY USE FUZZY LOGIC

### Introduction

Why should the fuzzy logic be applied to perform an optimal booking policy for a flight?

If we ask a revenue management analyst how he settles the level of booking authorisation to be adopted in the days before the aircraft take-off, he probably would say that if he finds a low booking level, he makes the decision to authorise a level of reservations which is more than the aircraft's capacity to compensate for the expected no-show passenger. If we ask him what he means by a low booking level, he could say that this depends on many factors, such as the type of flight, the season, the ratio between business passenger and leisure ones, but anyway, less than 50 percent of the aircraft capacity ten days before departure might be seen as a low booking level. The question is if he will use a different overbooking policy with a booking level of 51 percent. Actually he thinks that 50 percent is a limit for unequivocally saying that an over-sale of seats must be done, but for a lower level as well as for a higher booking level, an overbooking of seats must be accepted.

In other words we see how this kind of problem requires that the variables controlling the system must shift from a mathematical and deterministic formalism to a linguistic representation based on fuzzy sets.

Actually fuzzy systems suit very well in modelling non-linear systems. The nature of fuzzy rules and the relationship between fuzzy sets of different shapes provides a powerful capability for the description of a system whose complexity makes traditional expert system, mathematical, and statistical approach very difficult.

The problem is now to manage the experience of the expert and to transform it in a set of inference fuzzy rules expressing the dynamics of the system we want to model. As Lotfi Zadeh (1973) said, “when the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance become almost mutually exclusive characteristics (Cox, 1994, p.2)”. The basic idea underlying the *Fuzzy Logic* is that when we try to describe a system by a traditional model we use mathematical variables, which represent the state of the system as existing or not existing. If we represent the state of the system in terms of *fuzzy sets*, and not in terms of discrete symbols and numbers, we can obtain a representation of the system closer to human reasoning and the transition from a system state to the next is more gradual.

## Fuzzy Sets and Membership Functions

According to *Fuzzy Logic*, when a system is characterised by an incomplete knowledge, the hypothesis is not only true or false, but are true or false by a certainty factor.

The Fuzzy Set is a function indicating to what degree (between 0 and 1) the value of a variable belongs to the set. A degree of zero means that the value is not in the set, while a degree of one means that the value is completely representative of the set. A membership function maps to what degree of confidence each value belongs to the fuzzy set. It is important to outline that the degree of confidence we are talking about is not to be interpreted as a probability but as a degree of truth, that is, a measure of compatibility of the value of a variable with an approximate set, and not the occurring frequency of that value.

Formally, if *X* is a set of elements indicated as *x*, a fuzzy set ŧ of *X* is a set of paired values as shown below:

$b=[(x,\mu_b(x) : x\in X]$

$\mu_ŧ(x)$ is called membership function and it associates a degree of confidence m to each value of *x* in b. For instance the curve of Figure 1 can be seen as the degree of membership of each value of the booked seats of an aircraft to the set "High booking level". In this example, 50 percent and 150 percent are the limits of the so called interval of confidence.

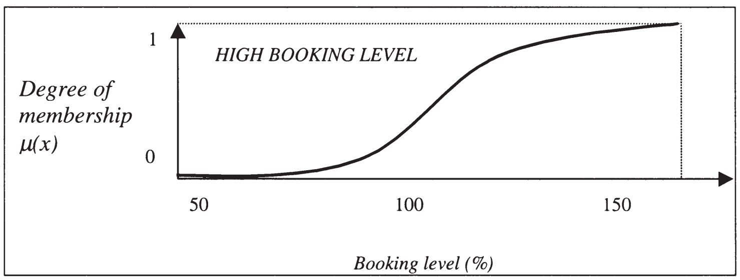


**Figure 1. Fuzzy set of the booking level as percent of the aircraft capacity.**

Using the Fuzzy Set Theory it is possible to approximate the behaviour of complex and non-linear systems, which otherwise would require a high level of computational resources. At the same time it is possible to have a model of the system very close to the human way of reasoning and to the way experts themselves think about the decision process, while many traditional expert and decision support systems lose fast comprehension as the complexity of the system increases because they persist in applying dichotomised rules with artificial and crisp boundaries.

We are talking of approximate (or possibilistic) reasoning, that is the way the experts think. So trying to perform an optimal booking policy during the period elapsing from the opening of the reservations of a flight till the departure day, a revenue management analyst would give us suggestions such as: *if the booking level is low, and the rate of cancellation is high, then the number of no-show passengers will be quite high*. Actually, the expert of the problem shows a knowledge of the system through concepts without a well defined pattern, based on his sensations, experience and intuitions, more than on precise data. Now the fact is that fuzzy systems are able to directly manage these kind of imprecise recommendations, reducing the distance that lies between the idea expressed by an expert and the one coded in a conventional model.

Another basic difference between a conventional expert system and a fuzzy system is that the former has a series of statements which are executed serially and is carried out with algorithms that reduce the number of rules examined, while the second has a parallel processing and activates all the rules at same time.

## Fuzzy Rules

The *fuzzy rules* are the building blocks of a fuzzy system. A fuzzy rule is a conditional proposition that settles a link between the fuzzy sets. Each rule is appraised for its degree of truth and shares to the final output set.

The proposition has the general form,

*if w is Z then x is Y*

where *w* and *x* are scalar values and *Z* and *Y* are linguistic variables, i.e., fuzzy sets.

In this example *w* is the "process state", while *x* is the "control action".

The meaning of the statement is then,

*x is a member of (the fuzzy set) Y to the degree that w is a member of (the fuzzy set) Z.*

The final solution fuzzy space is created by the collection of correlated fuzzy propositions, called rules of inference, each contributing with its degree of truth.

The main methods of inference used in fuzzy systems are the *min-max method* and the *fuzzy additive method*.

### *The min-max rules of implication*

By this method, the contribution of the antecedent part to the consequent fuzzy region is restricted to the minimum, that it, to the smaller value of the grades of inputs, while the final output region is obtained as a maximum, that it, by summing the fuzzy sets region corresponding to each rule.

### *The fuzzy additive rules of implication*

The fuzzy additive compositional operation is a slightly different approach as the output fuzzy region is bounded by [1,0], so that the result of any addition cannot exceed the maximum truth value of a fuzzy set.

Both methods reduce the level of truth of the output fuzzy region activated by the relevant rule of inference

### *Methods of decomposition and defuzzification*

Using the general rules of inference, the evaluation of a proposition produces one fuzzy set associated with each model solution variable. To find the actual scalar value representing the solution, the method of defuzzification is used. It is the final step of the fuzzy reasoning. As shown in Figure 2, this is obtained through an aggregation process that produces the final fuzzy regions, which have to be decomposed using one of the defuzzification methods.

There are different defuzzification functions, some computing the centroid of the output sets, some averaging the maximum points of the output sets. However, each of them inevitably results a compromise between the need to find a single point outcome and the loss of information that such process produces, by reducing to a single dimension the output region solution.

## FUZZY INFERENCE SYSTEM

### Building a Fuzzy Inference System

There are five main steps that must to be carried out to build a Fuzzy Inference System (FIS):

1. to choose the system variables (control variables for the input and solution variable for the output);

2. to define the fuzzy sets (number, shape and confidence intervals of the membership functions);
3. to write the relationships between the input and the output (inference rules);
4. to defuzzify to get the value of the solution;
5. to run a simulation of the model.

***Choose the system variables***

One of the most difficult parts to achieving a good formulation of the problem is identifying the data which influence the operation of the system and those which represent the output value of the model.

In this paper an overbooking fuzzy model has been constructed by selecting as control variables (input) the following:

- the booking level (BL) at a given time, that is, the difference between the total number of people who had booked a seat from the opening of the reservation period and the one who had cancelled it (in percent of the aircraft capacity); and
- the rate of cancellation (CR) at a given time, that is, the ratio between the number of people who had cancelled their reservation and those who had booked.

The number of no-show passengers (NS) in percent of the aircraft capacity is assumed as the solution variable (output).

A scheme of the proposed model is shown in Figure 3.

**Figure 3. Fuzzy Inference System.**

### *Define the Fuzzy Sets*

The shape of the fuzzy set is quite important, but most models do not show a very wide sensitivity to it. Triangular, trapezoid or bell curves are often used. Neural networks models have been used to find natural membership functions in the data and thus automatically creating fuzzy surfaces.

It is convenient to use a wide and elastic domain rather than a restrictive one.

To obtain a smooth and continuous control of the output variable a suitable degree of overlap of each fuzzy set should be assured.

### *Write the Inference Rules*

The rules that activate the same solution fuzzy set are grouped together. The application of a rule of inference that gets the shape of the consequent (output fuzzy set) as a result of the implication of the antecedent is reported in Figure 4. The implication form used is a minimum function, called an implication of Mamdani.

**Figure 4. Application of an inference rule.**

### *Aggregation of the Rules*

The application of each rule determines an adjusted fuzzy set of the consequent part. The final conclusion is then derived by summing the fuzzy sets of the conclusion of each rule, by a process called *determining MAX* (maximum) deriving from the application of the inference rules

### *Defuzzification*

This step selects the expected crisp value of the solution (output) from the fuzzy region resulting from the aggregation of the fuzzy sets each activated by all the rules applied in parallel.

There are several methods of defuzzification, but the most widely used is the method of the centroid, where the abscissa of the centre of gravity of the output fuzzy set region represents the balance point of the solution.

## THE AIRLINE BOOKING PROCESS

The booking process, from an airline company point of view, is rather complex. From a microeconomic point of view it is an economic interaction between the consumer (the potential air traveller) who tries to maximise his utility function under some given factors (travel dates, price, service and restrictions) and the airline trying to maximise its profit.

In the weeks before the departure many reservations are made for each type of fare. As the time of departure approaches some cancellations are added to the new reservations. Moreover at the day of departure there are additional complications due to travellers who show up without a reservation (go-show), travellers who fail to show-up (no-show) and travellers who are inserted in a waiting list. Furthermore there are many external factors which affect the booking process, such as different fare levels for each class, flight frequency, season or type of aircraft.

When the spaces corresponding to a certain fare class are filled, the request of travel is denied, but the airline (or the reservation agent) can try to recapture the traveller on a different class or on a different flight in the requested fare class. Nevertheless the actual number of boarded people depends also on the level of authorisation which has been adopted during the booking process.

A typical flow chart of a booking process is shown in Figure 5.

The result of the economic interaction between potential customers and the airline is a certain number of reservations and cancellations in each class on each flight.

Without any specific mathematical effort, a Fuzzy Inference System is able to incorporate all these factors, affecting the problem, as they are perceived by an expert (marketing specialist).

The booking process is divided into N time intervals of unequal length, that it, the duration each interval decreases as the departure date approaches.

Most airlines keep a record of some data describing the evolution of this process. A large number of such intervals is computationally impractical, while a small number allows no adjustment for differences between forecast and actual bookings as the booking history for each flight develops. Alitalia Airline holds an historical flight database where the booking process is photographed by 13 pictures. Pictures with their relevant time intervals are indicated in Table 1. Time intervals have a

**Table 1. Pictures of the Flight, as Used by Alitalia Airline Booking Process**

| *Picture* | *Days to Departure* |
|---|---|
| 1 | 342-90 |
| 2 | 89-60 |
| 3 | 59-43 |
| 4 | 42-23 |
| 5 | 22-13 |
| 6 | 12-7 |
| 7 | 6-5 |
| 8 | 4-4 |
| 9 | 3-3 |
| 10 | 2-2 |
| 11 | 1-1 |
| 12 | 0-0 |
| 13 | check in |

decreasing width. This is why airlines need to improve the monitoring resolution of the booking process, because, as the day of departure approaches, the possibility of managing the variability of the process is reduced.

In details, Alitalia reservation data contain company, flight number, origin and destination, day of the week, type of aircraft, compartment (i.e. top, business, economy), booking class, picture number (from 1 to 13), event code, date of departure, and value of the event.

The events recorded for each picture are

B = actual booked passengers, that it, the difference between reservations and cancellations;

C = cancelled passengers;

N = No-Show at departure, booked at the relevant picture; and

G = Go-Show are the total passengers appearing at the departure (picture 13) without reservation.

The effective number of boarded passengers is the minimum between the physical compartment capacity and the term (B+G-N).

*A typical example of average historical booking flight data is shown in Figure 6*.

The Booking Level is the cumulative sum of B relevant for each picture. The Cancellation Rate is the ratio of the cumulative sum of the total cancelled seats to the reserved ones. The falling down of the booking curve at the 13$^{th}$ picture is the effect of passengers who do not show up at the day of departure.

**Figure 6. Average of historical booking database.**

The shape of the booking curve for a specific class on a given flight depends on several factors. In fact, it is important how early before take-off the reservations are made. Leisure travellers usually book early, while business men late. Furthermore it is important to know that a large amount of cancellations occurs as a consequence of discouraging penalty for lower fare classes or as a consequence of the high cancellation rates and no-show rates observed for the higher fares.

The slope of the curve is steeper near a restriction expiration for lower fares and near the very last few days for higher fares. The booking limits tend to flatten the curve and the airline loses information about the shape of the real curve based on unconstrained demand

The fact that, occasionally in the example shown, the average overbooked seats coincide with the average no-show level is scarcely meaningful, as it might be the average result of flights with many denied boardings and flights with many empty seats. The goal for a effective forecasting policy is to get the “perfect fit” for each flight.

## THE FUZZY INFERENCE OVERBOOKING MODEL

Following the Fuzzy Inference System concepts, a Fuzzy Inference Overbooking Model has been built. The experience coming from the historical data of a flight reservation process that has been incorporated to construct the membership functions and writing the rules as previously discussed. The No-Show Level as a function of the Cancellation Rate and as a function of the Booking Level are plotted respectively in Figure 7 and in Figure 8, for a set of historical data of a typical booking process. The chart’s data include 33 flights and 12 pictures per flight. All data are

**Figure 7. No-Show Level as a function of the Cancellation Rate.**

**Figure 8. No-Show Level as a function of the Booking Level.**

referred to the Roma-New York Alitalia link and range from April to December 1993.

As shown by these charts, there is a substantial growth trend of the No-Show Level both with the Booking Level and with the Cancellation Rate. The number of no-show, as percentage of the cabin capacity, seem to be a function of how much the cabin is engaged and how much passengers tend to reject their reservation. This means that the evolution of the booking process depends fundamentally on the state of the process, described by the Booking Level and by the Cancellation Rate, while the dependency from the time is weak.

For the modelling of the system the following input control variables has been chosen: the booking level (BL) at any time before departure, as the total reservations made up to that time minus the total cancellations (in percent of the aircraft capacity); and the cancellation rate (CR) at any time before departure, as the ratio of the number of people who had cancelled their reservation to those who had booked (C/B for each picture).

The number of no-show passenger (NS) in percent of the aircraft capacity is the solution variable (output).

Hereby follow the nine rules of the Inference Fuzzy System, as they could be suggested by a booking process expert:

1. If (Cancellation Rate is Low) and (Booking Level is Low) then (No-Shows Level is Low)
2. If (Cancellation Rate is Low) and (Booking Level is Medium) then (No-Shows Level is Low)
3. If (Cancellation Rate is Low) and (Booking Level is High) then (No-Shows Level is Medium)

4. If (Cancellation Rate is Medium) and (Booking Level is Low) then (No-Shows Level is Low)
5. If (Cancellation Rate is Medium) and (Booking Level is Medium) then (No-Shows Level is High)
6. If (Cancellation Rate is Medium) and (Booking Level is High) then (No-Shows Level is High)
7. If (Cancellation Rate is High) and (Booking Level is Low) then (No-Shows Level is Medium)
8. If (Cancellation Rate is High) and (Booking Level is Medium) then (No-Shows Level is High)
9. If (Cancellation Rate is High) and (Booking Level is High) then (No-Shows Level is High)

In Figure 9 the input and output fuzzy sets activated by the parallel action of each rule with the corresponding aggregated fuzzy regions are shown, while the final solution is obtained as defuzzification with the centroid method. It can see how a Cancellation Rate of 75 percent and a Booking Level of 120 percent do not activate the rules from 1 to 4 and rule 7, while the remaining four rules contribute to the final result. Only the fuzzy set high in the No-Show level is activated in the output variable and applying the defuzzification procedure we obtain the final crisp result which is a No-Show level of 14.4 percent. All the percentages in the horizontal axes are referred to the aircraft seat capacity, while the vertical axes indicates the level of activation of the relevant rule.

**Figure 9. Application of the FIS rules.**

Finally Figure 10 shows the surface representing a three dimensional view of the overbooking model. It represents the outcome of the application of the FIS rules for each combination of the input variables.

The model has been constructed using the software *Fuzzy Logic Toolbox*, which is an extension of the MATLAB software application.

Using this fuzzy model an optimal booking policy can be adopted, by dynamically modifying the booking limit for the reservations that can be authorised in each time interval of the booking process. The authorisation level to be adopted for each picture is the sum of the cabin capacity and the number of no-show as calculated by the fuzzy model.

## NEURO-ADAPTIVE FUZZY INFERENCE SYSTEMS

As already said, the fuzzy sets (number, shape, range and overlapping) and the fuzzy rules are built by the co-operation of an expert and a fuzzy engineer, which traduces the experience into the fuzzy model. Otherwise it is possible to automate the process using a procedure based on the neural networks, such as the ANFIS function, contained in the Fuzzy Logic Toolbox of MATLAB. This is a Neuro-Adaptive Fuzzy Inference System essentially constituted of a fuzzy inference system, whose rules and membership functions are derived by a back-propagation algorithm based on some collection of input-output data. By this way the fuzzy system is able to learn from the example data, applying some optimisation routines to reduce the error between the data and the fuzzy system output.

To carry out this learning procedure a fuzzy inference system has to be specified, or alternatively, if no supposition can be made on how the initial membership functions should be, it is possible to use the command *genfis1*, which will examine the training data set and then generate a FIS matrix based on the given numbers and types of membership functions. The membership functions of the input variables are uniformly distributed in the range of the training data. Of course this procedure requires that a large amount of historical data are available.

A Neuro Adaptive Fuzzy Inference System has been built using the Alitalia reservation database for the flight Rome-New York in 1993. A simple program has been written in the internal MATLAB language to demonstrate that a fuzzy inference system can be adopted to simulate the booking process of a flight. Two main aspects can be pointed out: it can be easily and rapidly built; and it is a good approximation of the intrinsic complexity of the problem.

The program is reported in the Appendix.

As it can be seen in Figure 11, using a set of training data, ANFIS is able to approximate the booking process, showing a clear growth in the No-Show Level for increasing values of the Booking Level and of the Cancellation Rate.

## CONCLUSIONS

A fuzzy approach to the overbooking problem in air transportation has been presented.

The aim is to show that a complex system, such as the booking process, can be better controlled in terms of fuzzy sets than crisp numbers and mathematical models.

The underlying idea is that the notion of high booking level or low no-show level may change from day-to-day, flight-to-flight, airlines-to-airlines, season-to-season, but the logic is always the same and is contained in the inference rules. Therefore the method can be easily tuned just shifting the fuzzy sets averages or the intervals of confidence.

It has been shown the capability of the function ANFIS, contained in the Fuzzy Logic Toolbox of MATLAB, as a simple instrument to build an adaptive fuzzy inference system. When you try to approximate a function with an adaptive fuzzy inference system, there are several parameters that you can vary, some relevant to the fuzzy system, such as number and shape of the membership functions, or the method of inference and defuzzification, some relevant to the training method, such as the number or the sequence of the training data. It would be worthwhile to carry out some analysis to find out which is the best configuration of these parameters to obtain the best approximation.

The model has been built considering only a fare class of passengers, while in reality it would be better to extend the forecast of total bookings for each fare class.

A problem that should be investigated with more detail is the consequence of setting limits on the number of seats that can be sold. As a result, the airline companies can only evaluate the accepted demand, while no observation can be made on the demand that was turned away.

## REFERENCES

Alstrup, J. (1986). *Booking policy for flights with two types of passengers*, European Journal of Operation Research, vol. 27.

Andersson, S. E.(1989). *Departure overbooking when the denied boarding cost is non-linea*, AGIFORS Symposium, Chicago, USA.

Cox, E. (1994). *The Fuzzy Systems Handbook*, AP Professional.

Davis, P. (1994). *Airline Ties Profitability Yield to management*, SIAM News, Vol. 27, No. 5, May/June 1994

Dunleavy, H. (1994). *Probability density function approach to the determination of overbooking, marginal ODF fares and ODF level passenger unconstraining*, IATA 6th Yield management Conference, Barcelona.

Gerbracht, R. (1970). *Optimum overbooking levels*, AGIFORS Symposium, 1979

Holm, C. (1995). *Airline Overbooking performance measurement*, MIT Flight Transportation Lab Report, Cambridge, MA.

Rothstein, M. J. & Stone, A. W. (1967). *Passenger Booking Levels*, Proceedings of the 7th AGIFORS Symposium, 1967

Taylor, C. J.(1962). *The determination of the Passenger Booking Level*, Proceedings of the 2th AGIFORS Symposium, 1962

Zadeh, L. A. (1973). *Outline of a new approach to the analysis of complex systems and decision processes*, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 (1).

## APPENDIX

```
% function overb(dati,numMf,epochs,outputfile)
%
% dati = input matrix + output (last column)
% numMf = number di membership functions for the input (i.e. [3 2 4])
% epochs = number of iterations
% outputfile = name of the output file with .fis extension
%
function overb(dati,numMf,epochs,outputfile)
    data=dati;
    y=data(:,size(data,2));
    NumInput = size(data,2) - 1;
    TrainData = data;
    NumMfs = numMf;
    MfType = str2mat('trapmf');
    for i=1:NumInput-1,
        MfType = [MfType' str2mat('trapmf')']';
    end
    NumEpochs = epochs;
    StepSize = 0.1;
    InputFismat = genfis1(TrainData, NumMfs, MfType);
    close all;
    for i = 1:NumInput;
        subplot(NumInput, 1, i);
        plotmf(InputFismat,  'input', i);
        xlabel(['input ' num2str(i) ' (' MfType(i, :) ')']);
    end
    title('Initial fuzzy sets');
    OutputFismat = anfis(TrainData, InputFismat, [NumEpochs nan StepSize]);
    yy = evalfis(data(:,1:NumInput), OutputFismat);
    figure;
    plot(1:size(y,1),y,'o',1:size(y,1),yy,'x');
    legend('real','simulated');
    title('Real system vs. simulated system');
    figure;
    for i = 1:NumInput;
        subplot(NumInput, 1, i);
        plotmf(OutputFismat, 'input', i);
        xlabel(['input ' num2str(i) ' (' MfType(i, :) ')']);
    end
    title('Final fuzzy sets');
    writefis(OutputFismat,outputfile);
end
```