RESERVATIONS FORECASTING IN AIRLINE YIELD MANAGEMENT

BY

JOAO SA

Eng. Aer. Instituto Tecnologico de Aeronautica
(1980)

SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF

MASTER OF SCIENCE IN AERONAUTICS AND ASTRONAUTICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1987

Signature of the Author: ______________________________
Department of Aeronautics and Astronautics
February, 1987

Certified by : ______________________________
Prof. Robert W. Simpson
Thesis Supervisor

Accepted by : ______________________________
Prof. Harold Y. Wachman
Chairman, Department Graduate Committee

RESERVATIONS FORECASTING

IN AIRLINE YIELD MANAGEMENT

BY

JOAO SA

Submited to the Department of Aeronautics and Astronautics on February, 1987 in partial fulfillment of the requirements of the degree of Master of Science in Aeronautics and Astronautics.

ABSTRACT

This thesis shows the application of Regression Analysis in reservations forecasting in airline yield management.

The first three chapters highlight the need for yield management and the automation of seat inventory control. The seat inventory control problem is related to the determination of an optimal allocation of seats among the various fare classes being offered in a flight so as to maximize revenues. In order to determine such optimal seat allocation, forecasts of final bookings need to be made.

Forecasting alternatives are presented in this thesis. An example of application of Time Series Analysis is given as an alternative in providing such forecasts. Results obtained via Time Series Analysis were not encouraging enough in providing acceptable estimates.

Regression Analysis is also presented as a forecasting tool. Although regression models were developed for each market, a generalized model structure was thought to be preferable in view of the reduction of modeling efforts, data handling and model specification, that are need for forecasting final bookings for all markets/flights/classes. A general structure model is presented in this thesis as the result of the search for structural behavior across markets and flights.

Regression Analysis results are presented for a set of five citypairs, one flight in each directional market, i.e. ten flights in total. These results evidenced that a general structure model via regression analysis can indeed be used in the forecasting module of an automated seat inventory control system, and thus provide better estimates of final bookings when compared to Time Series Analysis or historical averages.

Thesis Supervisor : Robert W. Simpson
Title: Professor of Aeronautics and Astronautics

ACKNOWLEDGMENTS

I would like to thank Prof. Robert Simpson for the opportunity of working in this project. He read numerous draft versions of this thesis and helped me to clarify and express the ideas presented here.

I am also grateful to Peter Belobaba, who helped me since the begining of my participation in the project.

I would also like to express my satisfaction in working at FTL, where in so little time I made good friends. I certainly will miss them : Ray Ausrotas, Mike Fisher, Jan Van Cothem, Peter Belobaba, Peter Appel and Francisco Salas-Rocha.

I am deeply grateful to Angela, whose dedication, understanding, caring and love have certainly made life at MIT more bearable.

to Angela

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

The Airline Industry, since its very inception, has experienced many changes that have constantly challenged it and contributed to intensify operations and improve management of the today's diversified air transportation markets.

The introduction of jet aircraft in commercial operation required airlines to quickly adapt and respond to the "new equipment". Markets that were served with over than a day's flight could then be reached within the same day. The industry experienced rapid growth in passenger traffic. The combination of new equipment and traffic growth favored competition , which made some airlines operate more efficiently both in operational and managerial standpoints, while some other airlines that could not cope with these changes experienced financial problems, and eventually went out of business.

Technological advances have dominated the scenario of innovations in the airline industry. New aircraft guidance systems, new aircraft with more fuel efficient engines, and new navigational aids can be cited as examples of technological innovations that have, one way or the other, changed the Airline Industry.

Today, the U.S. Airline Industry experiences a rather different source of innovation.

Managerial innovations caused by the deregulation of the Airline Industry have dominated the industry scenario in the last years. The "fare war" that immediately followed the Airline Deregulation Act in 1978, was just the beginning. Price competition among airlines became a vivid reality.

Price competition has quickly evolved to price-availability competition . Seats allocated to lower fares are capacity controlled,and there are limitations or restrictions associated to low fare seats.

The airline product, a seat in a flight from A to B, has now two dimensions : fare and restriction. The emergence of a vast set of airline products has generated the need of a sophisticated decision support system devoted to the development/management of price-availability-restriction policies. Such management decisions are central to the permanence of an airline in the marketplace, and they have changed the airline industry.

CHAPTER TWO

THE NEED FOR YIELD MANAGEMENT

The deregulation of the U.S. Airline Industry launched the industry into a new era. The change from a regulated to a "free" market caused radical modifications in the airline industry. The objective of this chapter is to highlight the different environments that airlines were subject to in these two periods. A brief comparison between these periods is presented in the next two sections of this chapter. The third section highlights the need for an Yield Management System.

2.1 THE AIRLINE INDUSTRY BEFORE DEREGULATION

Before deregulation, US airlines were controlled by a government body: the Civil Aeronautics Board-CAB. The decision of either flying or expanding existing services in any given market, was subjected to the approval of the CAB. Fares were calculated by using the mileage-based Standard Industry Fare Level, adopted by the CAB. Fare levels and structure were, therefore, fixed. Markets were regulated and controlled. Marketing decisions of airlines were dependent of the CAB.

Fare discounting was, nevertheless, practiced in the U.S. Airline Industry before deregulation. The concept at that time was that the revenue of a flight could be increased by offering the unsold seats to a new and different segment of passengers.

This new market segment consisted of some few passengers that would only travel at a discount fare. The operating cost associated to this "new segment" was considered as being minimal, since the full fare passengers would already have absorbed most of the operating costs. Incremental costs would typically involve reservations, ticketing, baggage handling and on board meal service for these "additional" passengers.

Few seats were sold at discount fare to these few passengers. The majority of passengers would still fly at regular full fares. Revenue was not diverted from the airline's "regular" passengers and a new and different market , with low yield passengers, was created. Yield is defined as the revenue per passenger-mile of traffic carried by an airline.

A limited set of "discount" tickets was also available to some passengers meeting pre-determined criteria. For instance, senior citizens could buy airline tickets at a lower price. There were discounts associated with bulk travel, being either on a family basis or on a group basis. Thus, some discount tickets were available for those who would meet these pre-determined criteria.

A step towards a more complex discounting practice was then observed with the introduction of "red eye" flights (i.e. late night flight services). These flights/seats at a discount were available to anyone willing and able to travel at late hours. This new discounting practice differed from the existing in the sense that passenger would not need to meet/have pre-determined qualifications (age,organizations) , nor need to travel in "bulk" (family or group). Anyone could buy a ticket on these flights .

Seat inventory management was not needed in the regular flights of an airline. The available set of alternatives (airline product) available to passengers was relatively small :

(1) first and regular coach class seats;

(2) limited possibilities of "discount" tickets in regular flights;

(3) few special flights at discount fares.

The management of flight revenue was a relatively straightforward task. Once the potential of sales of full fare passengers was estimated for a given market, the remaining seats were made available for low yield passengers. Profit maximization was strongly related to +maximization of flight loads. No special attention or routine was used for controlling the seats sold to discount fare passengers. Revenue maximization was achieved by filling up planes with as many revenue passengers as possible. Regular full fare passengers accounted for most of the revenue. The remainder or unsold seats were sold to few passengers, at a discount fare. A flight with as much revenue passengers as possible was thought to be reaching revenue optimality

For the profit maximizing airline, revenue and cost would have to be taken into account, since revenue maximization does not always lead to profit maximization. Nevertheless, the dominant criterion was that a flight with high load factor was the sign of good business.

2.2 THE AIRLINE INDUSTRY AFTER DEREGULATION

The deregulation of the U.S. Airline Industry marked the beginning of a new era. With deregulation, U.S. airlines were allowed to enter or leave any domestic market, increase or reduce existing services. It was up to the management staff of an airline to fully decide where and when services should be offered. If market A was not considered as profitable as market B, an airline could decided to offer services only in market B. Any airline could offer services in market B. No government approval was needed.

Fares were also deregulated so the industry experienced a real change. Today, it is the airline who determines how much should it charge in a given market/class.

As a consequence, new airlines were created and more airlines started to fly in traditionally profitable markets. Unprofitable markets experienced either a reduction in the level of service or they were abandoned. Fare competition was inevitable in busy markets. Competition was increased as result of the "free market" era.

The marketplace, formerly regulated and subject to limitations, gave place to a "free" and strongly competitive market. With the freedom to enter or leave any market, airlines started to increase competition in profitable markets, by offering more seats/flights at lower and lower prices. The advent of low-cost/low-yield new entrant carriers led to a fare war.

A fare war immediately followed the free entry market era. Airlines were forced, once again, to react and adapt to a new scenario : stiff competition and low fares, with a high level of diversity.

Passengers who used to fly at full fare, because few alternatives were available, started taking advantage of these lower fares. They benefited from the fare reduction by flying at more competitive prices. Demand increased as a response to low fares. New markets were even created because of some extremely low fares. As a consequence, to fly at discount fares became a common practice, and today, only few passengers fly at nominal full fares.

Established airlines needed to remain or be competitive. They needed to compete with new entrant carriers, and to offer/match some low fares, and yet they also needed to avoid fare diversion, which happens when a potential high yield passenger takes advantage of a low fare seat. But, most importantly, they needed to maintain, or even recover profitability.

The creation of a very complex fare structure was the response of the airline management to the fare war and low-cost/low-yield carriers challenge. By offering a more complex set of services with differential pricing, airlines were able to maintain regular/traditional passengers, attract low fare passengers, maintain a competitive image in the market and remain profitable.

Restrictions and limitations were attached to some fares and, as a general rule, the cheaper the fare gets, the more restrictions/limitations it has. By discriminating passengers, offering different fares, with associated different restrictions, an airline can differentiate passengers in respect to price, and pursue profit maximization.

Evidently, the new fare structure could not coexist with the cost structure that was in effect at that time, specially for old and established airlines. Cost levels were no longer compatible with fare/revenue levels. Low fares must be followed by low costs. It would have been impossible to survive in the low fare market without drastic changes in the cost structure. As a consequence, airlines were forced to reduce cost.

2.3 THE NEED FOR YIELD MANAGEMENT

Today, every single class/passenger is an important part in the apportioning of flight costs. The concept of "incremental" cost associated to lower fare passengers no longer exists. Management of flight revenue became a complex task. The optimal combination of passengers and fares is now the what airlines aim at. A flight need not to be at its maximum load factor, but rather the overall product seats/fares sold needs to be maximized: a "revenue load factor" maximization problem.

Management decisions related to how many seats to sell at what price, became crucial for airlines. From the complexity derived from managing such decisions has emerged the need of Yield Management. An Yield Management System is, therefore, a decision support tool designed to help an airline to determine how many seats should be sold, given the price levels. Finding such answers is central to the permanence of the airline in the market place.

The passenger load factor of a flight is no longer a proxy to infer flight revenue performance. The maximization of the flight load has been replaced with flight revevue.

Profit maximization is now related to the degree of success that an airline achieves in selling the right number of seats to the right number of passenger so as to maximize revenue, while at the same time keeping the associated costs down.

The revenue maximization problem has now two major components: seats and fares. Average yields can be increased by either increasing price levels or reducing the proportion of seats sold in the lowest fare product categories. In both cases, the potential of sales in each and every fare level has to be estimated so as to allocate the optimal number of seats to each fare class, maximizing the revenue of a flight.

Pricing and seat inventory control represent, therefore, the core of Yield Management.

"While pricing is clearly an important component of yield management, no one airline can influence it's own revenue through pricing, without taking the reactions of its competitors into account. Revenue increases resulting from pricing actions are possible only when all of the major competitors in a market agree implicitly to follow a price leader."[1]

Prices are published in airline guides and displayed in reservation systems that are available to everyone. An airline can always monitor and follow its competitors price changes.

Seat inventory control, on the other hand, is a logistical component of yield management that is entirely under control of each individual airline. It is an in-house component that only the airline itself knows and controls, with the exception of some few one-price-only low-cost/low-yield airlines. As a consequence, airlines do not know the seat inventory management decisions of the competitor airlines.

While pricing is not fully dependent on the decision of one airline alone, seat/class allocation is. Through seat inventory control, airlines have the potential of managing revenue from a flight on a departure by departure basis, which would be far more difficult to replicate via pricing.

Today, the need for an Yield Management System is evident. Airlines can no longer be profitable in the marketplace without it. The competition is strong. An efficient Yield Management System is, therefore, very important to an airline.

CHAPTER THREE

THE AUTOMATION OF SEAT INVENTORY CONTROL

The seat inventory control problem is related to the determination of an optimal (revenue maximizing) allocation of seats on the aircraft among the various fare classes being offered in a flight. Other management decisions, such as capacity allocation, equipment utilization, aircraft routing cannot be dissociated from the seat inventory control problem. These decisions interact with each other, and as a consequence, the seat inventory control problem has to either use them as input, or interact with them.

3.1 SEAT INVENTORY CONTROL

A flight leg seat inventory control approach is commonly used in the industry. On a flight leg basis, the aircraft seating capacity is divided among classes with the objective of revenue maximization on that flight leg.

Although the flight leg approach might not lead to revenue maximization over the whole flight and/or the entire network of an airline, because it maximizes flight leg revenues, its simplicity makes it very attractive.

With a flight leg approach, true origin and destination of passengers are not taken into account. All passengers flying the same class are treated equally. The deficiency of this approach is that seat allotment decisions on a leg basis do not assure revenue maximization over the whole flight.

A more coherent approach should consider origins and destinations - O&D, when allocating seats to different classes for a flight. The O&D approach becomes complex when one considers the possible combinations of a multiple leg flight. The increase of hub and spoke operations by the airlines has certainly added further difficulty. The number of possible O&D combinations can become unmanageable and, as a consequence, this approach can hardly be pursued.

In order to determine how many seats should be allocated to each class on a future flight departure, the airline has to have estimates of expected loads in each and every class, as well as the associated average revenues. With these two parameters for each class, the airline can then figure out the best seat/class allocation that maximizes flight revenues.

Expected loads have to be transformed in expected bookings. Because of large incidence of no-shows in some markets, substantial analysis has to be devoted in determining the number of seats to be assigned to each class. Overbooking analysis determines the level of total bookings for a flight that will minimize the total of the costs associated with denied boarding of passengers and the costs in the lost revenues associated with no-shows and unsold seats.

The estimates of expected loads, by class and by flight, together with estimates of what will happen on the boarding day, (i.e. no-shows, go-shows, upgrades, etc.) are then used to generate booking thresholds for each class, on a given flight leg. These limits can then be input in the reservations system of the airline and bookings can be monitored.

3.2 SEAT ALLOCATION MODELS - AN OVERVIEW.

The seat inventory control system is a decision support system that provides inputs to the reservation system of an airline, helping the yield management analyst to determine booking limits or thresholds. In the case of a flight leg based system, limits are given by class and by flight leg.

Several alternatives have been proposed as core routine of a seat inventory control system. This routine, apart from determination of overbooking policies, are intended to determine whether to accept or reject a request for a seat (reservation) according to the fare paid.

Littlewood [2], in 1972, suggested a seat allocation routine, probabilistic based, that maximizes total expected flight revenues, using a "marginal seat" model. In his routine, a low-yield passenger paying a lower fare f2 should be accepted as long as the expected revenue from selling all S1 seats to passengers paying the higher fare f1 is less than f2.

That is , if $f2 > f1 \cdot P(S1)$ take f2 passengers,

where:

$f1$ = higher fare;

$f2$ = lower fare ;

$P(S1)$ = probability of selling S1 seats at f1 fare.

The aircraft seating capacity (C) , in this simple case, is divided in two compartments. One with S1 seats, calculated with the probability distribution function (pdf) of f1 passengers, and fares (f1 and f2). The other compartment takes the remaining seats, that is (C - S1). Note that the pdf of f2 was not needed in the seat allotment decision process.

Mayer [3], in 1976, suggested a two class seat allocation model that would utilize dynamic programming (DP) as a framework. He suggested a simple model to be used to determine initial seat allotments, and that a multi-period DP-based model should be used to modify initial limits, taking into account bookings already made.

The set of assumptions he made in deriving his model was:

(1) no cancellations;

(2) total loss of rejected bookings (no vertical shift);

(3) in each booking period, low fare passengers make reservations before high fare passengers;

(4) the demand in each period is independent of the actual demand in all previous periods.

He concluded that the initial seat allotment did not benefit from a DP-based approach. He suggested Littlewood's model to set initial allotment. From there on, a model that permits corrective action (reallotment), such as a DP-based one, should be used.

Buhr [4] contributed to the marginal probabilistic approach in 1982, suggesting a seat allotment model for a two leg flight (A to B to C). His model was based on the expected "residual" revenue, defined as the revenue from allocating an additional seat to passenger flying from A to B as the product of the average fare from A to B, times the probability of selling more than x seats in the A/B market, or :

$$E_{AB}(x) = P_{AB}(x) \cdot R_{AB} \quad ;$$

where,

$E_{AB}(x)$ = expected residual revenue ;

$P_{AB}(x)$ = probability of getting more that x passengers in the A/B market ; and

R_{AB} = average fare the A/B market.

For the two leg flight, the demand for each O&D market were assumed as independent. For each of the three markets, expected residual revenues were calculated and seat allotment decisions were taken based on such revenues, that is :

$$| E_{AC}(x) - (E_{AB}(y) + E_{BC}(y)) | \dashrightarrow \min ,$$

where y is the seat capacity allocated to local AB and BC passengers.

An extension to the simple marginal seat allotment model, proposed by Belobaba [5], handles multiple fare classes and flight with multiple legs. Given the expected bookings in a fare class i, the expected revenue for this class is given by :

$$E(Ri) = fi \cdot bi(Si)$$

where fi is the net fare or the yield to the airline from a passenger booked in class i, and bi is the expected bookings, given a seat allotment Si.

The expected marginal revenue for class i (EMSRi) is defined as the increase in revenue when the seat allotment is increased by one seat, i. e. Si+1 seats. EMSRi is, therefore, calculated with the following expression:

$$EMSRi = fi \cdot P[\ ri > Si\ ],$$

where ri is the total reservations made on class i.

Given a natural ranking of fares f1 > f2 > f3 > f4 etc. , in order to maximize flight revenue, the reservation process should be able to discriminate bookings, giving priority to passengers that contribute most with revenue. Although the assessment of the probability of having Si passengers at the fare level fi is made, priority should always be given to higher yield passengers. This leads to a nested version of authorized booking limits, where the authorized booking limit for a given higher fare class overlaps with the authorized booking limit for all subsequent lower fares. For example, in the case of a three class flight, where f1 > f2 > f3, the authorized limit (AU) for each class should be as:

$$AU1 = C \quad ;$$

$$AU2 = C - S1 \quad ;$$

$$AU3 = C - S1 - S2 \quad ;$$

where

C = seating capacity of the aircraft;

S1 = number of seats protected for class 1;

S2 = number of seats protected for class 2;

r1 and r2 represent reservations made in class 1 and 2.

A protection level is calculated for each class in order to achieve this priority. The protection level for a class in the minimum number of reservations that are accepted in that fare class and that must be protected from lower fare classes. In the above example, S1 seats are always protected from f2 and f3 passengers. Likewise, S2 seats are always protected from f3 passengers. No protection level for f3, of course. Note that nothing prevents passengers paying higher fare from taking up low fare seats. In the example, up to C passengers paying f1 fare can make reservations, and up to to (C-S1) for passengers paying f2.

3.3 RESERVATIONS FORECASTING MODULE

Central to any model briefly presented here, is the ability of knowing the fare and the probability distribution function for each class i, in every market.

Airlines can obtain average fare figures by sampling tickets on a class i, for a given market. Several problems arise when averaging fares within a class. Within each class there is a multitude of fare codes, and very often they are not well structured. It is very common to observe price overlapping between adjacent fare classes.

The average fare , calculated by class, may not be representative, if too many fare codes exists. Not only prices vary but also restrictions change.

Another problem associated with fare averaging is the pro-rating of fares. When a passenger buys a ticket from A to C, and the service he gets is a one-stop flight A to B to C, the apportioning of fares in legs AB and BC will tend to drive down the AB and BC average fare calculations. With the increase of the airline hub-and-spoke operation , the pro-rating problem tends to be widespread. If the seat allocation routine were O&D based, this problem would not exist.

Due to this high degree of non-homogeneity it is really difficult to define what is the "average service" and to estimate its probability distribution function. Fortunately, several fare data analysis performed together with this thesis show that although there is a high degree of non-homogeinity within each fare class, the overall class result tends to exhibit a stable pattern, as far average prices are concerned, especially for higher fare classes.

The reason for this result is that within each class there is always a "dominant fare code". The majority of passengers fly with dominant fare code tickets. The remaining passengers use other fare code tickets (within the same class) but with different proportions each time. The dominant fare code tends to be less predominant as the fare gets lower. For the lower fare class, passengers are more disperse. Therefore, the fare aggregation within the same class should be more meaningful and yield better result than for lower fare classes, provided there is no radical changes in fare levels. This may result in a better ability in forecasting reservations for higher fares than for lower ones.

The seat inventory control routine will need a forecast element that provides (1) initial estimates of final loads, and (2) updates on such estimates, as the reservation process for a flight is under way.

These two inputs, namely fare and expected bookings, are then used by an automated booking limit routine. The automated booking limit system calculates reservations thresholds, by class and by flight. The promptness of the automated booking limit routine in providing such thresholds is dependent on the reservation system itself. There are reservation systems that start taking reservation as early as one year in advance. The usefulness of generating booking thresholds so early in time are, of course, questionable. As a general rule, reservation systems have some form of booking limit introduced at least 6 weeks before departure. From there on, booking thresholds limit the number of seats made available in each class, in every market.

The reservation forecasting module of an Automated Seat Inventory Control System is primarily intended to provide the dynamic booking limit adjustment routine with estimates of expected bookings for individual future flights. A seat allocation routine will then use these estimates of expected bookings to calculate how many seats should be allocated/protected for each upper fare class.

The first step in reservations forecasting involves initial estimates of final bookings, well in advance of flight departures. These estimates are typically needed for each future flight, up to 90 days out, and are used to set initial authorized booking limits. Sophisticated forecasting models are of little use, and outweighed by the error associated with the forecast produced for such large time interval. As a consequence, a simplistic but conservative approach is thought as being the most appropriate and effective.

These initial estimates need to be improved later in the bookings process as more information (data) on the specific flight for which more accurate forecasts are needed.

A simple forecasting model is suggested for the initial estimation of final bookings. It consists of moving average process that is sensitive to day of week variation only. That is to say, for instance, that a 8-week average is used to describe or estimate final bookings for a given flight (e.g. flight F1), on a specific day of week (e.g. Monday).

Although no information on actual bookings on hand for future flights is ever used, nor additional adjustments are made for cyclic or seasonal variations other that on weekly basis, the implicit assumption of this simple approach is that a small sample of final demand for recent flights will be representative of the demand for the same flights in the near future.

The final step in the development of a forecasting module is to improve the estimates of bookings to come, over those strictly based on recent historical averages. These new estimates are used to re-calculate expected revenues ,and again the seat allocation routine is used to update allotments.

CHAPTER FOUR

EXPLORATORY DATA ANALYSIS

An exploratory data analysis is designed to give the forecaster more insight into the variable (s)he is trying to forecast. Trends in daily booking levels, variations across markets, and seasonalities are among the characteristics the forecaster searches. Reservations data are extremely confidential and, very reluctantly, airlines make them available. The exploratory data analysis presented in this thesis represents a moment of rare opportunity in which actual and recent data was available. As a consequence, extensive data analysis are presented here.

4.1 DATA SAMPLE DESCRIPTION

A sample of five city-pairs was selected for data statistical analysis and hypothesis testing. The sample included a variety of market types and stage lengths . One short, one short-medium, two medium and one long haul markets were included in the sample. One of the medium-haul markets was a Canadian market.

A total of 28 flights were included in the sample. Some flights did not operate throughout the whole sample period. At least two flights were operating in any given month, for any of the five markets. The sample period was from January, 1986 through June, 1986. The Airline Industry registered no major abnormality during the sample period. Therefore, it is expected that the data set provides a normal picture of what happens in the first half of an year.

Data was collected from the actual database of an existing US airline. Table 4.01 shows the markets and flights selected. In order to maintain confidentiality of the data presented in this thesis, single capital letters were assigned to markets ,and single digit numbers to flights. Distances and flight times were rounded.

For instance, the flight F1 in the A/B market departs at 09:00 am. The distance flown is approximately 500 miles, and the aircraft type is a B73S. Additional market characteristics are also given in Table 4.01.

TABLE 4.01

DATA SAMPLE
GENERAL CHARACTERISTICS

MARKET	FLT NUMBER	DISTANCE (MILES)	FLIGHT TIME (HH:MM)	DEPART TIME (HH:MM)	AIRCRAFT TYPE	MARKET CHARACTERISTICS
A/B	F1	500	1:30	09:00a	B735	Short-medium haul hub feeder with a minimum of two daily flights each way, and with high load factors.
	F2			01:50p		
	F3			07:19p		
B/A	F1	500	1:30	10:00a	B735	
	F2			02:00p		
	F3			06:15p/08:30p		
C/D	F1	300	0:55	08:30a	B725	Short-haul hub feeder, very stable business market, with a minimum of three caily flights, with consistently high load factors.
	F2			12:00n	B735	
	F3			03:25p	B725	
	F4			06:50p	B733	
D/C	F1	300	0:55	09:15a	B735	
	F2			11:15a	B725	
	F3			05:10p/06:15	B733	
	F4			09:30p	B725	
E/F	F1	2000	5:00	07:50a	B725	Long-haul hub leg, with high load factors, with two daily flights each way.
	F2			05:00a		
F/E	F1	2000	4:15	10:10a	B725	
	F2			05:00p		
G/H	F1	1200	2:30	06:35p/12:25	B725	Medium-haul leg with an average load factor and a good mix of traffic. Two daily flights , each direction.
	F2			09:45a/09:45		
H/G	F1	1200	2:30	08:15a	B725	
	F2			05:00p		
I/J	F1	750	1:40	11:10a	B725	Medium-hub feed ,with high load factor, different fare structure/mix. A Canadian market, with two daily flights each direction.
	F2			05:10p		
	F3			09:30p		
J/I	F1	750	1:40	07:40a	B725	
	F2			01:40p		
	F3			05:50p		

SOURCE :Official Airline Guide, North American Edition, 1986.

Table 4.02 shows how many days a given flight operated, throughout the sample period, on a monthly basis. Some flights started operation only June, e.g. flight F2 in the C/D market. Some flights operated throughout the whole sample, e.g. flight F1 in the A/B market. Within the sample period some flight ceased operations. Sometimes, a new flight was created, departing at the same time as the old one, e.g. flight F2 in the G/H market, or the new flight departed between one and two hours later, e.g. flight F3 in the D/C markets. In both cases, the old and the new flights were considered as the same flight.

TABLE 4.02

DATA SAMPLE
MARKETS & FLIGHTS

MARKET	FLT NUMBER	OBSERVATIONS JAN.	FEB.	MAR.	APR.	MAY	JUN.
A/B	F1	31	28	31	30	31	30
	F2	30	28	31	30	31	30
	F3	0	0	0	0	0	30
B/A	F1	31	28	31	30	31	30
	F2	0	0	0	0	0	30
	F3	31	28	29	30	31	30
C/D	F1	30	28	31	30	31	30
	F2	0	0	0	0	0	30
	F3	31	28	31	30	31	30
	F4	31	28	31	30	31	30
D/C	F1	0	0	0	0	0	30
	F2	30	28	31	30	31	30
	F3	31	28	31	30	31	30
	F4	31	28	31	30	31	30
E/F	F1	31	28	31	30	31	30
	F2	31	28	31	30	31	30
F/E	F1	31	28	31	30	31	30
	F2	31	27	31	30	31	30
G/H	F1	26	24	29	26	26	30
	F2	30	27	30	29	31	30
H/G	F1	26	24	29	26	27	30
	F2	30	27	30	29	31	30
I/J	F1	[illegible]	28	31	30	31	30
	F2	31	28	31	0	0	30
	F3	31	28	31	29	30	30
J/I	F1	31	28	31	30	31	30
	F2	28	28	29	30	31	30
	F3	0	0	0	0	0	30

The markets in the sample exhibited high levels of bookings on boarding day. Table 4.03 shows reservations load factors , defined here as total reservations on the boarding day, divided by the seating capacity of the aircraft assigned to that flight. From now on the term load factor will be loosely used meaning not the actual load factor, which is calculated with departure loads in the passenger cabin, but rather the reservations load factor already defined. It can be observed that there were months in which the average load factor was greater than 100%, which means that in the average flights were overbooked. This is the case of markets A/B, B/A, D/C, E/F, and F/E.

It is interesting to observe, still on Table 4.03 the change in the performance of reservations caused by the introduction of a new flight in the market. For the A/B market, the third flight introduced in June, a night flight, exhibited a reservations load factor that ranked second. In the opposite direction, market B/A, the third flight exhibited the highest reservations load factor in the month. Reservations load factors in the other flights, in June, were bellow average, with the exception of flight F1 in the A/B market. This result suggests that some of demand generated by the new flight might be the result of diversion of "regular" passengers from other flights.

The new flights in the C/D and D/C markets, on the other hand, exhibited very low load factors, and one may also speculate about passenger diversion. There was an overall reduction in load factors in the month of June. If the reduction of load factors was a consequence of a seasonality in the market demand then passenger diversion cannot solely justify the observed reduction in demand levels.

For the J/I market, the result was very different. Flight F3 showed a reservations load factor that ranked second, but the overall market behavior suggests that reservations demand was indeed increased in the market. In the I/J case, there were three flights from January to March. When the third flight came back in operation in June, reservation levels were brought back to normal levels.

TABLE 4.03 RESERVATIONS LOAD FACTOR ON BOARDING DAY
ALL CLASSES

MARKET	FLT NUMBER	AVERAGE LOAD FACTOR JAN.	FEB	MAR.	APR.	MAY	JUN.
A/B	F1	74	80	101	89	98	95
	F2	89	88	99	90	90	61
	F3	0	0	0	0	0	71
B/A	F1	73	84	97	84	85	54
	F2	0	0	0	0	0	82
	F3	80	88	100	94	107	72
C/D	F1	97	96	91	84	84	81
	F2	0	0	0	0	0	47
	F3	89	86	85	76	78	57
	F4	48	48	50	31	41	41
D/C	F1	0	0	0	0	0	37
	F2	87	89	86	72	78	70
	F3	84	92	100	81	86	70
	F4	49	58	65	53	57	66
E/F	F1	77	86	104	60	78	101
	F2	86	97	102	79	89	101
F/E	F1	93	99	111	96	105	114
	F2	68	74	90	61	62	91
G/H	F1	41	39	71	58	70	30
	F2	56	51	80	77	93	87
H/G	F1	49	47	77	65	83	64
	F2	50	51	80	74	95	77
I/J	F1	58	61	65	66	74	82
	F2	66	76	72	0	0	49
	F3	80	87	79	94	84	59
J/I	F1	67	82	82	92	85	57
	F2	60	81	80	59	64	74
	F3	0	0	0	0	0	63

Seating capacity : B735 -115
(all classes) B725 -148
B733 -128

Table 4.04 shows reservation averages for a flight on the boarding day. Reservations were totaled, for all classes, and then averages were calculated by month. The result is presented on table 4.04 . For instance, the average in January for flight F1 in the C/D market was 144 reservations, for the month of January. This table presents the intensity in bookings for each market analyzed. market. The C/D and D/C markets exhibited a high reservation activity. In this example, a total of six flights, eight in June only, one can also observe that the high level on bookings did not vary too much from month to month, exception made only in June, when flights were added. The same stability pattern is also observed in the other markets.

TABLE 4.04 REVERVATIONS ON BOARDING DAY (ALL CLASSES)

MARKET	FLT NUMBER	AVERAGE TOTAL BOOKINGS JAN.	FEB.	MAR.	APR.	MAY	JUN.
A/B	F1	85	92	116	102	113	109
	F2	102	101	114	103	103	70
	F3	0	0	0	0	0	82
B/A	F1	84	97	111	97	98	62
	F2	0	0	0	0	0	94
	F3	92	101	115	108	123	83
C/D	F1	144	142	134	124	124	120
	F2	0	0	0	0	0	54
	F3	132	127	126	113	116	85
	F4	61	61	64	40	52	53
D/C	F1	0	0	0	0	0	43
	F2	129	132	127	106	115	104
	F3	108	118	128	104	110	90
	F4	72	86	96	79	84	97
E/F	F1	114	128	154	89	116	150
	F2	128	143	151	117	131	150
F/E	F1	137	146	165	142	155	168
	F2	100	110	133	90	92	134
G/H	F1	60	58	105	86	104	44
	F2	83	75	119	114	137	129
H/G	F1	72	70	114	96	123	95
	F2	74	76	119	109	140	114
I/J	F1	86	91	96	97	110	121
	F2	97	112	106	0	0	72
	F3	118	129	117	139	124	87
J/I	F1	99	121	122	136	126	84
	F2	89	120	119	88	95	109
	F3	0	0	0	0	0	93

Table 4.05 shows averages in the reservations for a flight on the boarding day. The objective of this table is to show the contribution of different classes in the flight load. First class was also included in the calculation of percentages, and the total coach compartment contribution is shown on the last column. The contribution of the first class is, on the average, less than 5% .

For the business market, the Y contribution in the flight load would be expected to be a little higher than the average. It was indeed observed in the E/F & F/E data that participation of Y class was slightly higher in this business market.

In the I/J & J/I market, the contribution of Y class was the highest. It happens to be the canadian market. A large proportion of non-restricted Y seats were sold in this market. In the other markets the Y contribution was less than 8%, in the average. The results are as expected. While the expected typical contribution of Y class is less than 10%, when there is a strong business component in this market. Canadian markets usually exhibit different behavior when compared to US markets.

The participation of Q class observed in the sample was very high. In the average, excluding the business and the Canadian market, Q class participation was 42.27% .

TABLE 4.05

REVERVATIONS ON BOARDING DAY
BOOKINGS BY CLASS
(% TOTAL)

MARKET	FLT	CLASS Y	M	B	Q	TOTAL COACH
A/B	F1	7.00	22.83	24.83	42.83	97.50
	F2	7.33	19.67	26.50	43.33	96.83
	F3	7.00	22.00	19.00	51.00	99.00
B/A	F1	6.83	22.00	22.33	46.17	97.33
	F2	5.00	27.00	16.00	50.00	98.00
	F3	7.83	22.67	23.33	43.50	97.33
C/D	F1	6.00	28.17	24.83	37.83	96.83
	F2	9.00	24.00	27.00	38.00	98.00
	F3	4.33	31.67	24.67	36.17	96.83
	F4	2.17	34.83	19.67	41.33	98.00
D/C	F1	6.00	24.00	21.00	46.00	97.00
	F2	5.17	30.67	24.83	36.67	97.33
	F3	3.50	32.83	23.00	37.83	97.17
	F4	5.17	23.50	26.33	42.50	97.50
E/F	F1	15.00	13.83	27.33	36.33	92.50
	F2	16.33	11.50	28.67	35.83	92.33
F/E	F1	16.67	14.83	29.17	32.00	92.67
	F2	9.50	13.50	31.67	35.17	89.83
G/H	F1	5.17	16.00	17.17	57.67	96.00
	F2	3.83	17.00	22.50	53.17	96.50
H/G	F1	4.83	16.50	17.17	58.17	96.67
	F2	4.33	18.00	20.83	52.67	95.83
I/J	F1	35.33	33.00	8.50	18.17	95.00
	F2	35.25	28.00	10.50	21.25	95.00
	F3	26.17	18.00	19.50	32.00	95.67
J/I	F1	29.00	28.67	10.50	27.67	95.83
	F2	39.33	25.50	9.83	19.83	94.50
	F3	6.00	25.00	13.00	53.00	97.00

Table 4.06 shows class participation in the flight load, on a monthly basis. Table 4.07 shows the authorized booking limits that were imposed to the reservation system, by flight, by class and by month. The reference level was the Y class limit, which was always greater than the coach class seating capacity. One could observe the "high" authorized level assigned to M-class, exception made only to the Canadian I/J & J/I markets. In the business markets, i. e. C/D & D/C, the authorized limit was slightly lower. Authorized bookings limits are, in this case, nested. Considering the seat allocation routine proposed by Belobaba, one could calculate the average protection level assigned to each upper fare class. For instance, for flight F1 in the A/B market, the average protection level for Y class was 7%, in the month of January. The M protection level was 15% (22% - 7%),and for the B class it was 33% (55%-22%).

TABLE 4.06

REVERVATIONS ON BOARDING DAY
BOOKINGS BY CLASS
(x TOTAL)

MARKE	FLT NUMBER	Y BOOKINGS (x) JAN.	FEB	MAR.	APR.	MAY	JUN.	M BOOKINGS (%) JAN.	FEB.	MAR.	APR.	MAY	JUN.	B BOOKINGS (x) JAN.	FEB.	MAR.	APR.	MAY	JUN.	Q BOOKINGS (x) JAN.	FEB.	MAR.	APR.	MAY	JUN.
A/B	F1	8	8	5	6	6	9	31	21	18	21	22	24	24	27	26	21	24	27	36	42	48	49	44	38
	F2	11	7	4	7	10	5	23	19	18	20	19	19	26	29	29	29	24	22	38	42	46	40	43	51
	F3	0	0	0	0	0	7	0	0	0	0	0	22	0	0	0	0	0	19	0	0	0	0	0	51
B/A	F1	7	8	6	8	6	6	27	21	19	20	21	24	23	24	27	21	18	21	41	45	45	48	52	46
	F2	0	0	0	0	0	5	0	0	0	0	0	27	0	0	0	0	0	16	0	0	0	0	0	50
	F3	10	9	6	6	7	9	27	19	19	24	24	23	25	27	26	24	18	20	37	42	47	42	47	46
C/D	F1	7	7	5	4	5	8	30	26	26	31	25	31	27	27	28	22	24	21	34	37	39	39	41	37
	F2	0	0	0	0	0	4	0	0	0	0	0	22	0	0	0	0	0	20	0	0	0	0	0	52
	F3	5	6	4	4	3	4	36	30	28	33	26	37	24	27	26	24	24	23	34	34	39	34	42	34
	F4	3	2	2	2	1	3	40	40	33	41	16	39	13	22	23	22	17	21	43	34	41	32	62	36
D/C	F1	0	0	0	0	0	3	0	0	0	0	0	43	0	0	0	0	0	13	0	0	0	0	0	39
	F2	4	5	6	7	4	5	37	34	30	36	24	23	23	26	26	19	24	31	35	32	36	34	44	39
	F3	4	3	4	3	3	4	40	35	30	35	19	38	19	27	23	24	22	23	36	33	41	34	50	33
	F4	5	5	4	4	6	7	28	25	22	25	16	25	26	31	29	25	22	25	39	38	43	42	51	42
E/F	F1	13	8	11	16	20	22	7	9	17	12	19	19	23	32	32	24	24	29	50	44	34	38	28	24
	F2	13	12	14	16	19	24	5	8	13	12	16	15	26	31	30	26	28	31	[illegible]	41	36	38	28	23
F/E	F1	15	11	14	17	18	25	9	12	16	16	18	18	26	32	30	29	29	29	41	39	33	30	27	22
	F2	8	7	9	8	10	15	5	8	13	13	21	21	29	33	31	31	32	34	48	41	38	36	26	22
G/H	F1	10	5	4	5	4	3	42	13	7	8	14	12	17	29	20	20	11	6	28	49	67	61	66	75
	F2	6	4	3	3	3	4	39	9	5	6	14	29	18	31	28	19	17	22	34	52	61	68	62	42
H/G	F1	7	4	3	4	5	6	41	11	7	5	10	25	16	25	18	14	17	13	32	56	70	73	65	53
	F2	7	5	3	3	4	4	42	10	7	7	14	28	17	33	23	17	16	19	29	48	64	69	61	45
I/J	F1	35	35	33	38	41	30	32	28	36	25	34	43	16	10	8	7	4	6	12	23	19	23	15	17
	F2	33	33	37	0	0	38	27	29	30	0	0	26	18	9	8	0	0	7	16	25	19	0	0	25
	F3	20	21	24	24	30	38	18	14	17	17	17	25	20	19	19	22	23	14	39	42	37	31	24	19
J/I	F1	26	25	27	27	32	37	33	27	31	22	27	32	7	9	8	18	12	9	32	35	29	28	23	19
	F2	40	36	36	46	46	32	26	21	29	20	22	35	15	12	14	5	4	9	14	27	16	21	20	21
	F3	0	0	0	0	0	43	0	0	0	0	0	20	0	0	0	0	0	5	0	0	0	0	0	28

TABLE 4.07

AUTHORIZED BOOKINGS LEVELS
(YAU as 100%)

MARKET	FLT	Y	M LEVELS JAN.	FEB.	MAR.	APR.	MAY	JUN.	B LEVELS JAN.	FEB.	MAR.	APR.	MAY	JUN.	Q LEVELS JAN.	FEB.	MAR.	APR.	MAY	JUN.
A/B	F1	100	93	98	99	92	92	72	78	83	84	78	68	54	45	53	48	39	35	26
	F2	100	93	95	97	96	70	85	78	77	80	76	61	67	46	50	44	33	33	35
	F3	100	0	0	0	0	0	88	0	0	0	0	0	75	0	0	0	0	0	38
B/A	F1	100	92	94	94	92	67	88	78	80	80	76	66	66	49	47	40	36	37	35
	F2	100	0	0	0	0	0	93	0	0	0	0	0	77	0	0	0	0	0	44
	F3	100	94	94	94	86	93	86	81	80	80	65	70	66	49	59	43	31	38	32
C/D	F1	100	83	83	83	85	89	78	65	62	62	51	62	41	38	50	28	25	30	19
	F2	100	0	0	0	0	0	93	0	0	0	0	0	64	0	0	0	0	0	35
	F3	100	86	84	83	84	86	84	71	62	60	59	57	53	42	37	30	28	32	29
	F4	100	95	91	91	90	89	92	79	76	76	69	65	71	52	59	54	44	43	45
D/C	F1	100	0	0	0	0	0	93	0	0	0	0	0	82	0	0	0	0	0	40
	F2	100	86	87	83	70	91	82	70	65	62	39	69	55	44	33	25	15	32	28
	F3	100	89	90	84	78	82	83	71	65	60	46	61	62	40	39	28	23	32	30
	F4	100	84	91	91	77	92	91	72	73	72	57	68	67	51	44	39	32	42	41
E/F	F1	100	84	92	92	92	92	71	73	73	65	70	67	48	44	35	27	37	37	23
	F2	100	83	87	83	89	87	67	67	67	58	66	62	40	44	44	25	34	30	17
F/E	F1	100	86	92	92	90	90	75	58	70	65	63	60	43	30	41	25	28	29	16
	F2	100	82	86	83	89	91	80	72	72	66	68	72	50	53	43	30	38	41	33
G/H	F1	100	93	93	92	93	94	92	78	75	73	73	78	71	59	59	53	47	55	41
	F2	100	89	92	92	93	94	92	82	76	75	74	76	55	52	49	45	47	48	20
H/G	F1	100	92	91	92	93	94	85	77	74	74	73	77	63	49	54	51	53	55	31
	F2	100	94	92	92	93	94	82	80	74	72	73	77	61	50	46	45	49	50	30
I/J	F1	100	78	84	84	83	85	78	40	59	61	58	58	49	23	19	23	26	27	21
	F2	100	79	84	82	0	0	84	51	51	43	0	0	59	25	13	20	0	0	32
	F3	100	87	78	78	77	80	82	62	62	56	52	66	53	32	25	17	19	31	24
J/I	F1	100	92	85	83	90	88	88	66	65	55	72	66	66	37	31	19	47	48	48
	F2	100	75	82	81	84	84	88	51	57	47	53	56	67	26	23	13	20	23	39
	F3	100	0	0	0	0	0	89	0	0	0	0	0	71	0	0	0	0	0	42

Table 4.08 shows bookout analysis performed for the flights in the sample. Columns with heading " # DAYS " indicate the actual number of days in which the class was closed. Columns with " (%) " heading indicates the percentage of days, during the whole sample period in which the class was closed. As a function of the nested authorized booking limit reservation system, the following equations describe a bookout in a given class:

Y is closed if:

YRES + MRES + BRES + QRES > YAU.

M is closed if:

YRES + MRES + BRES + QRES > YAU , OR

MRES + BRES + QRES > MAU.

B is closed if:

YRES + MRES + BRES + QRES > YAU ;OR

MRES + BRES + QRES > MAU ;OR

BRES + QRES > BAU.

Q is closed if:

YRES + MRES + BRES + QRES > YAU ;OR

MRES + BRES + QRES > MAU ;OR

BRES + QRES > BAU ;OR

QRES > QAU.

A consequence of the bookout equations in the previous page, is the following relation :

QCLOSED > BCLOSED > MCLOSED > YCLOSED.

One can observe that this relation is not observed in the flight F1 in the B/A market. Flight F2 in the A/B market exhibit a high level of bookout for all classes. M was closed in 18.33% of the flight. The high percentage in Y bookout is likely to be the consequence of bookout in other class, rather than in Y alone. In the opposite market, B/A, flight F2 consistently exhibit the lowest bookout levels, for all classes.

TABLE 4.08 BOOKOUT ANALYSIS

MARKET	FLT	Y # DAYS	Y (%)	M # DAYS	M (%)	B # DAYS	B (%)	Q # DAYS	Q (%)	# FLTS
A/B	F1	14	7.73	24	13.26	34	18.78	48	26.52	181
	F2	32	17.68	33	18.23	44	24.31	71	39.23	180
	F3	0	0.00	0	0.00	0	0.00	2	1.10	30
B/A	F1	12	6.63	22	12.15	15	8.29	39	21.55	181
	F2	1	0.55	3	1.66	0	0.00	4	2.21	30
	F3	12	6.63	13	7.18	25	13.81	49	27.07	179
C/D	F1	5	2.76	39	21.55	76	41.99	80	44.20	180
	F2	0	0.00	0	0.00	1	0.55	2	1.10	30
	F3	2	1.10	14	7.73	47	25.97	43	23.76	181
	F4	0	0.00	2	1.10	4	2.21	4	2.21	181
D/C	F1	0	0.00	0	0.00	0	0.00	0	0.00	30
	F2	7	3.87	20	11.05	38	20.99	57	31.49	180
	F3	9	4.97	31	17.13	49	27.07	50	27.62	181
	F4	0	0.00	3	1.66	10	5.52	17	9.39	181
E/F	F1	12	6.63	13	7.18	46	25.41	36	19.89	181
	F2	8	4.42	22	12.15	98	54.14	59	32.60	181
F/E	F1	11	6.08	12	6.63	82	45.30	86	47.51	181
	F2	4	2.21	17	9.39	47	25.97	35	19.34	180
G/H	F1	0	0.00	0	0.00	6	3.31	11	6.08	161
	F2	0	0.00	4	2.21	16	8.84	62	34.25	177
H/G	F1	5	2.76	10	5.52	26	14.36	22	12.15	162
	F2	0	0.00	5	2.76	18	9.94	39	21.55	177
I/J	F1	9	4.97	9	4.97	15	8.29	39	21.55	179
	F2	3	1.66	6	3.31	17	9.39	31	17.13	120
	F3	5	2.76	16	8.84	46	25.41	87	48.07	179
J/I	F1	6	3.31	8	4.42	10	5.52	23	12.71	181
	F2	1	0.55	9	4.97	34	18.78	44	24.31	176
	F3	0	0.00	0	0.00	0	0.00	1	0.55	30

For the remaining markets, the same behavior is observed. At least one flight for each directional market exhibited a high level of bookouts. The remaining flights showed moderate to low bookout statistics.

The flights that exhibited high level of bookouts were:

MARKET	FLIGHT
A/B	F2
B/A	F1 & F3
C/D	F1
D/C	F2 & F3
E/F	F2
F/E	F1
G/H	F2
H/G	F1
I/J	F3
J/I	F2

The flights in the following list did not exhibit high level of bookouts. They were :

MARKET	FLIGHT
A/B	F1
B/A	F2
C/D	F4
D/C	F3
E/F	F1
F/E	F2
G/H	F1
H/G	F2
I/J	F1
J/I	F1

Statistical analysis on these flights above should produce results that are likely to be more representative of the market behavior than the original set of 28 flights because they have exhibited low level of bookout. There is not a simple routine for bookout correction, i.e. how to estimate what would have been the reservations demand given that no seat limitation was imposed to a flight. As a consequence, major attention will be given to the flights that did not bookout. Therefore, statistical analysis, distribution plotting and the demand analysis presented in this thesis will only show results for these flights.

4.2 DISTRIBUTION ANALYSIS

The objective in this analysis was to produce and examine distribution plots of reservations by fare class for the markets and flights in the sample, specially those which exhibit low levels of bookout. The first reservation data retrieved for any flight refers to a period that corresponds to 28 days before departure. From there on, snapshots were taken for every seven days. That is to say that each flight will be analyzed in 5 seven-days' periods, from day 28 to boarding day, that is, periods are: T28, T21, T14, T7 and TBD.

Analyses were made in terms of reservations made for the M-class, up to a particular period , or bookings-on-hand, and the expected number of reservation still to come,or bookings-to-come. This is to say, that on day 21, we should have data analysis for both bookings-on-hand and bookings-to-come. The objective is to observe these two related but distinct variables in terms of shapes, means, and standard deviation. Exception will be made for the Canadian markets, where the analyzed class will be Y.

Correlation analysis performed in bookings-on-hand and final bookings indicated that the correlation between these two variables , in the majority of cases, was very low. This means that bookings-on-hand should not exhibit good explanatory power when forecasting final bookings via bookings-to-come. This result corroborates the conclusion arrived by Littlewood [1]:

> "The subsequent arriving passengers
> can be regarded as independent of
> the booked load".

Figures 4.01 through 4.10 show distribution plots for the flights/markets selected for data analysis.

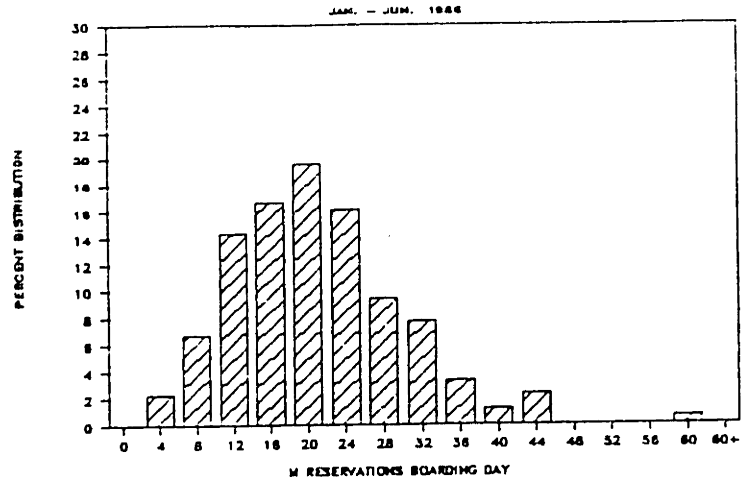
Figure 4.01 shows distribution plots for flight F1 in the A/B market. The shape of the distribution observed for final bookings, i.e. reservations on boarding day, resembles the bell shape of a normal distribution. An increase in skewness is observed for bookings on hand as it gets further from boarding day. On the other hand, the shape of bookings-to-come plots, in any period, resemble that of a normal distribution. Bookings-on-hand exhibits increase in standard deviation, going from 6.00 to 7.87, whereas bookings-to-come is the reverse, going from 8.80 to 6.42 . A side by side plot comparison in given in figure 4.01 . The small table on the top right corner of figure 4.01 shows statistical analysis for this flight.

The notation used in these figures is :

FiMBD = reservations on boarding day for flight i, M class;

FiMt = total reservation made up to day t, for flight i, M class, (t = 7, 14, 21 and 28);

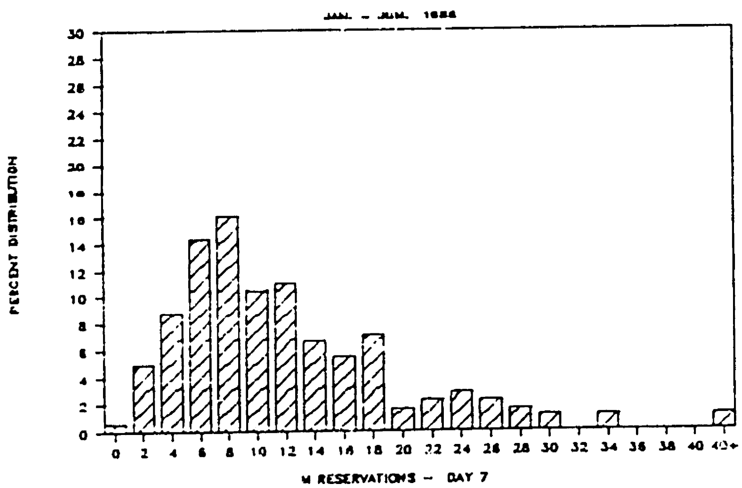
FiMt_BD = bookings-to-come for flight i, from day t to boarding day, M class, (t = 7, 14, 21 and 28).

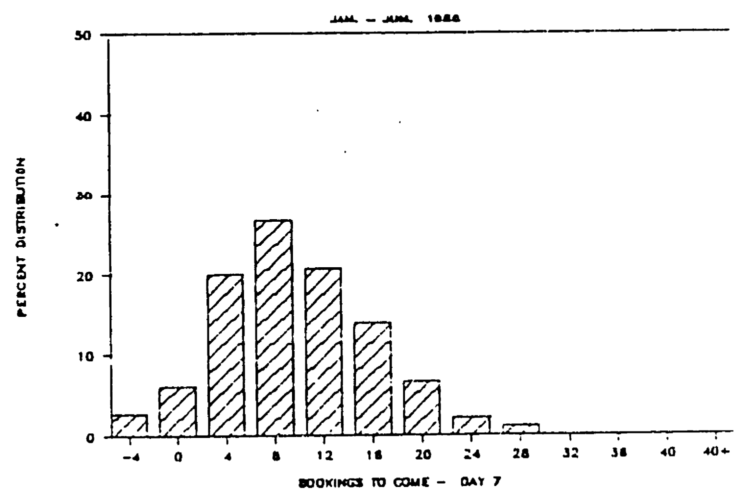


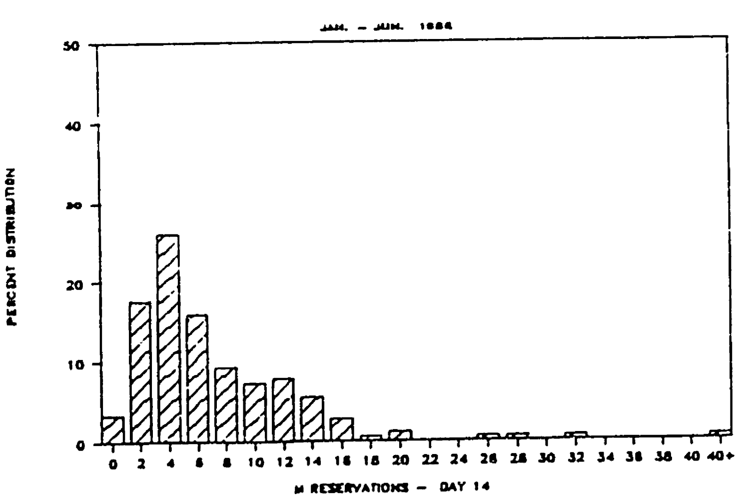
Sample set to -> ALL
Descriptive Statistics - 180 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F1M80	19.406	8.8908	.79593	4.1779	3.000	57.00
F1M7	11.250	7.8746	1.4836	6.1708	.0000	49.00
F1M14	6.6944	6.8749	4.4066	35.232	.0000	67.00
F1M21	4.2556	6.2919	6.2318	57.496	.0000	67.00
F1M29	3.1444	5.9936	7.1261	69.900	.0000	66.00
F1M7_80	8.1556	6.4252	.13393	3.3533	-13.00	27.00
F1M14_80	12.711	8.1826	-.54475	6.8530	-31.00	37.00
F1M21_80	15.150	8.7379	-.33213	6.0647	-31.00	41.00
F1M28_80	16.261	8.7971	-.27681	6.7125	-30.00	42.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







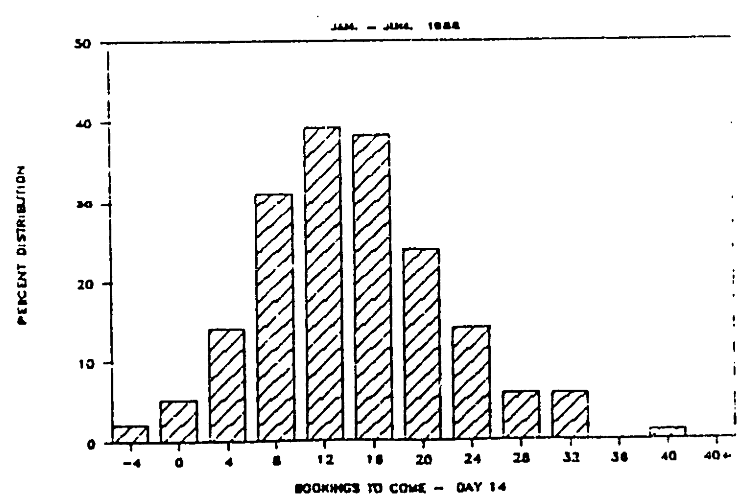
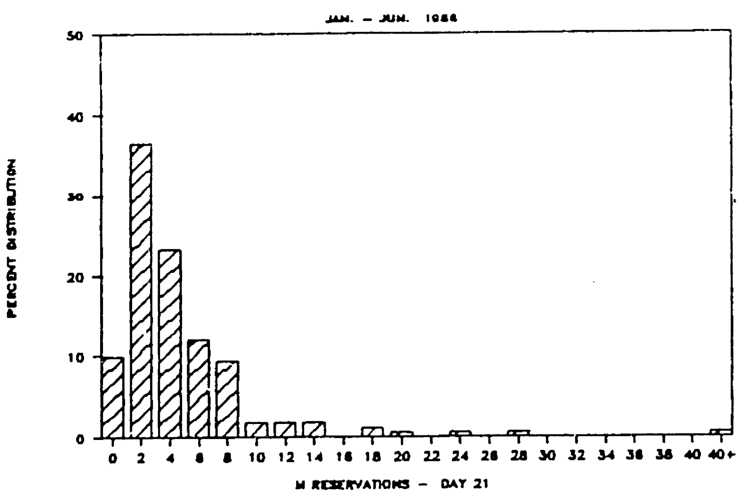
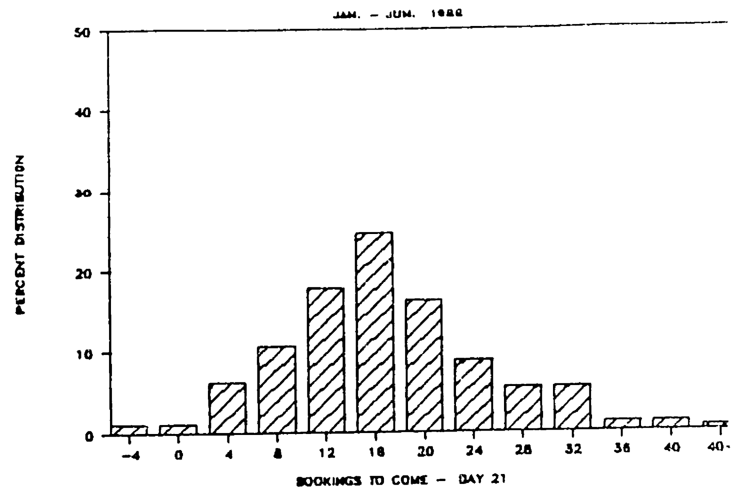
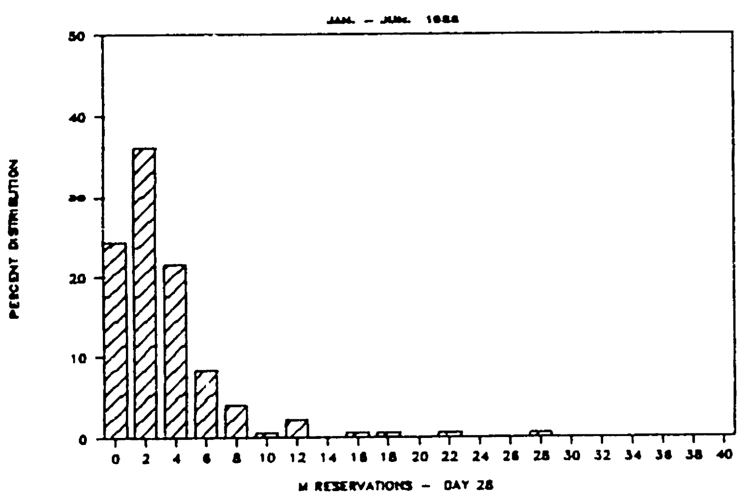


Figure 4.01 : Distribution Plots

M-class Flight F1 Market A/B







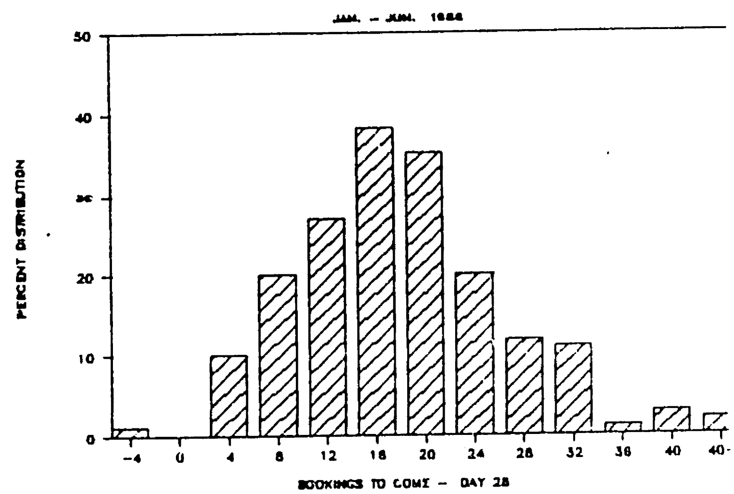
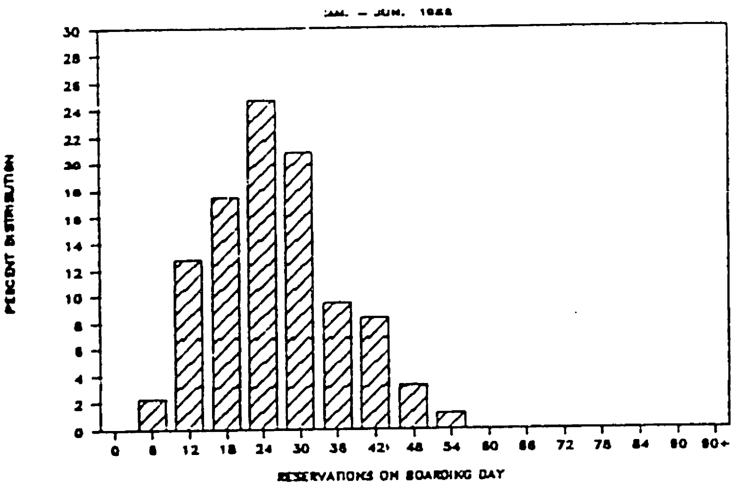


Figure 4.01 (cont.) : Distribution Plots

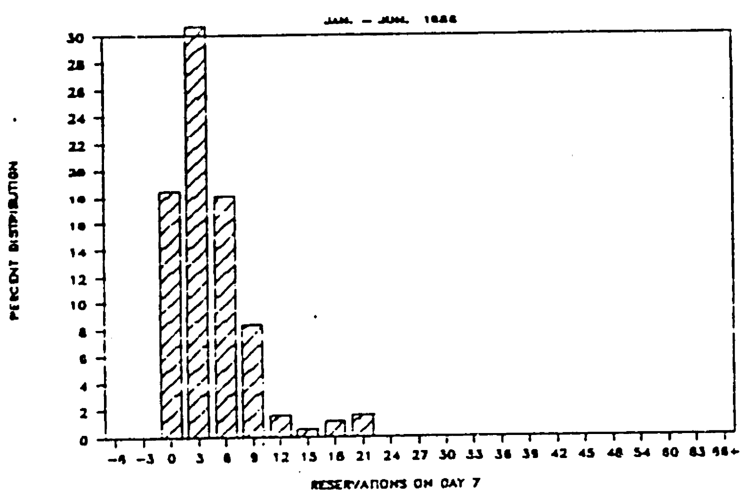
M-class Flight F1 Market A/B

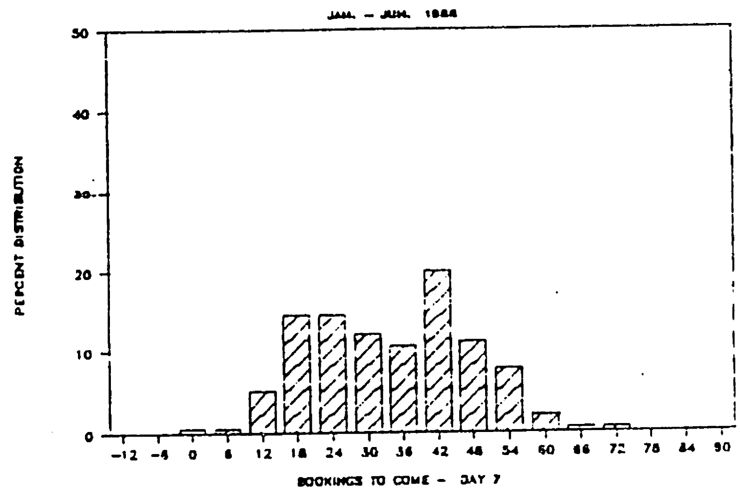


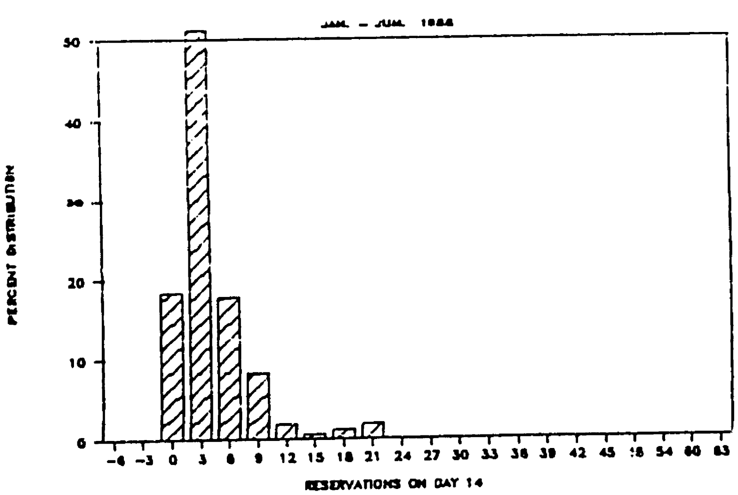
Sample set to -> ALL
Descriptive Statistics - 181 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F2MBD	23.337	10.383	.34722	2.7745	.0000	52.00
F2M7	13.011	8.2050	1.1963	4.9249	.0000	49.00
F2M14	6.8398	5.4896	1.7008	7.4439	.0000	35.00
F2M21	4.6961	4.5694	2.0665	8.9319	.0000	30.00
F2M28	3.1934	3.7150	2.2790	9.3770	.0000	21.00
F2M7_BD	10.326	7.9302	.44086	3.2102	-9.000	34.00
F2M14_BD	16.497	9.7049	.45437	3.0684	-6.000	46.00
F2M21_BD	18.641	10.121	.27984	2.8918	-2.000	46.00
F2M28_BD	20.144	10.361	.33458	2.8315	.0000	49.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







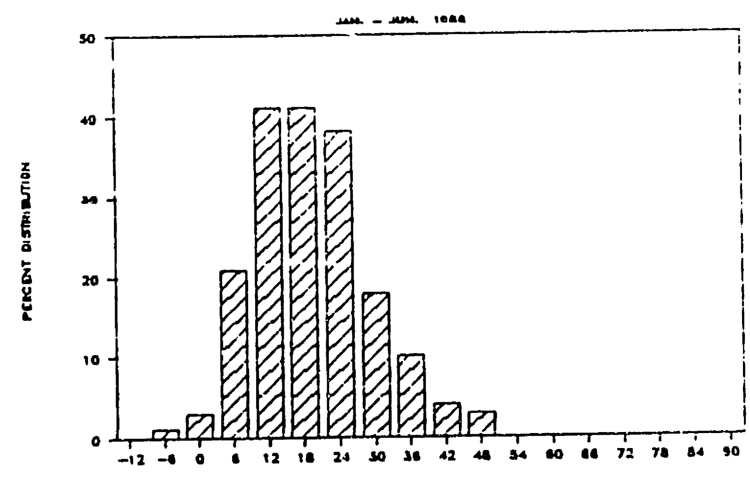


Figure 4.02 : Distribution Plots

M-class Flight F2 Market B/A

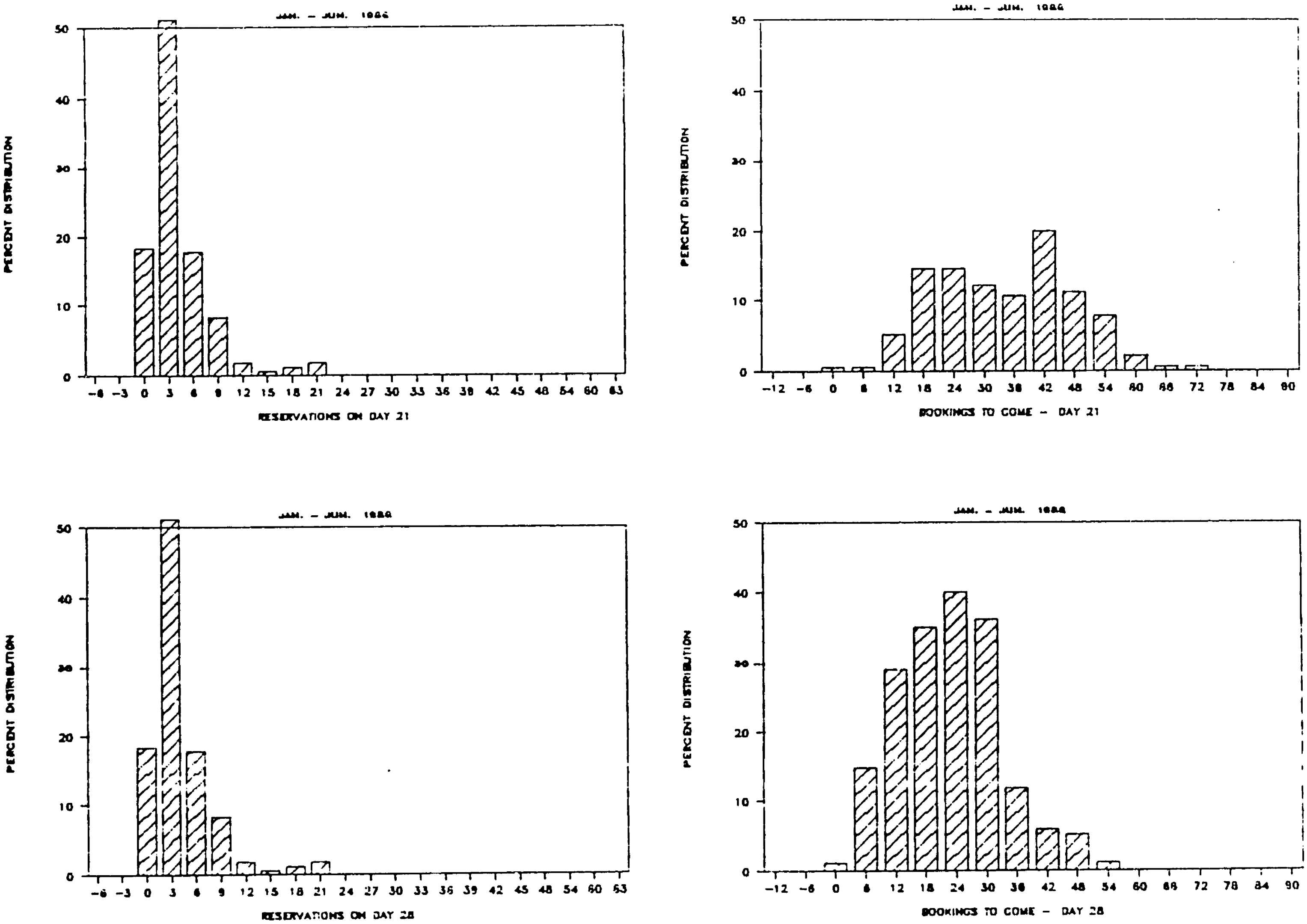
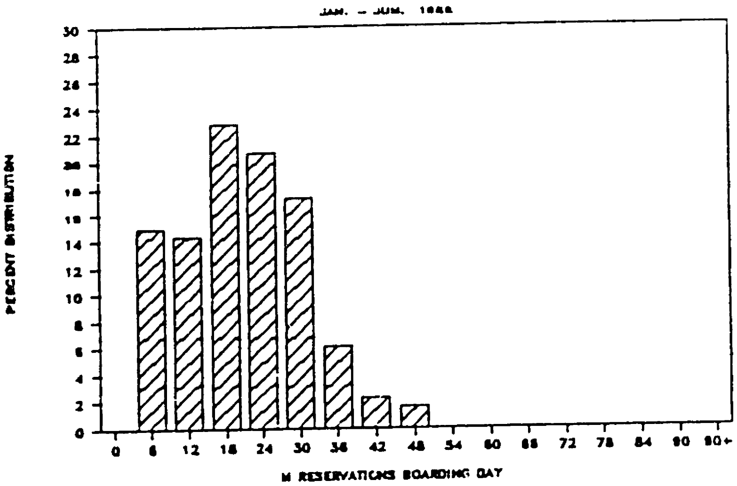


Figure 4.02 (cont.) : Distribution Plots

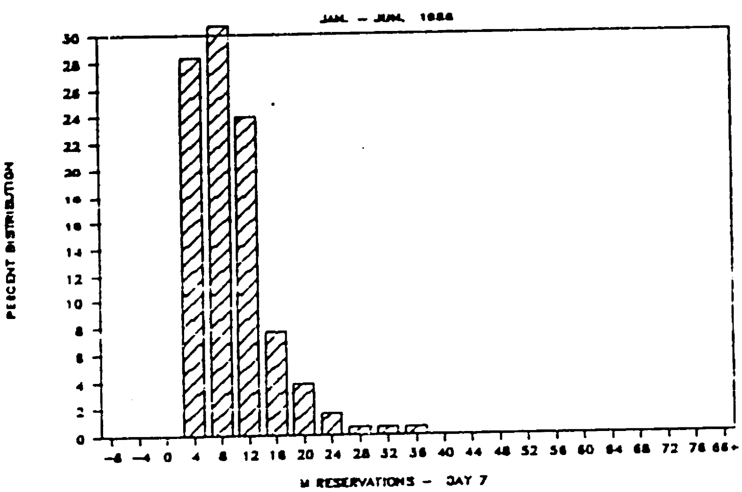
M-class Flight F2 Market B/A

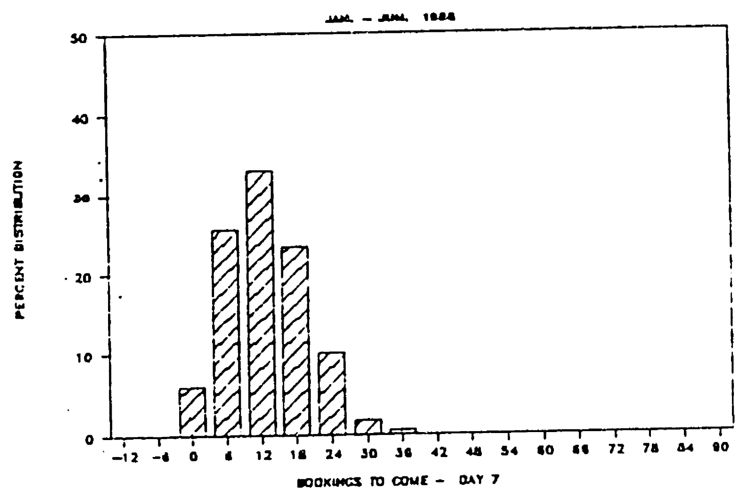


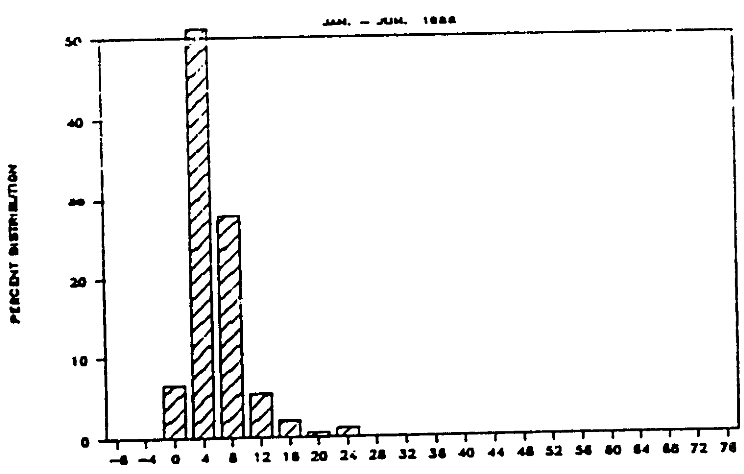
Sample set to -> ALL
Descriptive Statistics - 180 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F4MBD	18.267	10.010	.33665	2.6426	1.000	48.00
F4M7	8.0667	5.4717	1.5856	6.8878	1.000	34.00
F4M14	4.3667	3.6615	1.8727	7.9251	.0000	21.00
F4M21	2.6556	2.9997	2.2730	10.068	.0000	19.00
F4M28	1.5667	1.9292	1.8897	7.1777	.0000	11.00
F4M7_BD	10.200	6.8784	.36585	2.6981	-3.000	32.00
F4M14_BD	13.900	8.5907	.15795	2.4534	-7.000	36.00
F4M21_BD	15.611	8.9470	.16705	2.2709	-1.000	38.00
F4M28_BD	16.700	9.5729	.31096	2.4013	.0000	43.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







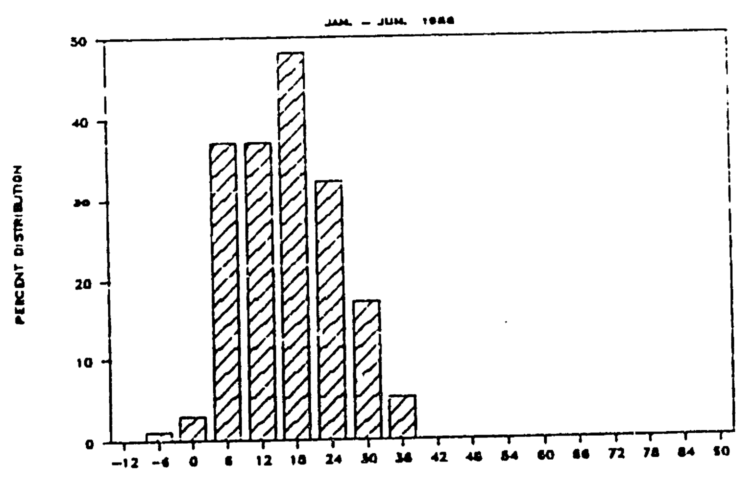


Figure 4.03 : Distribution Plots

M-class Flight F4 Market C/D

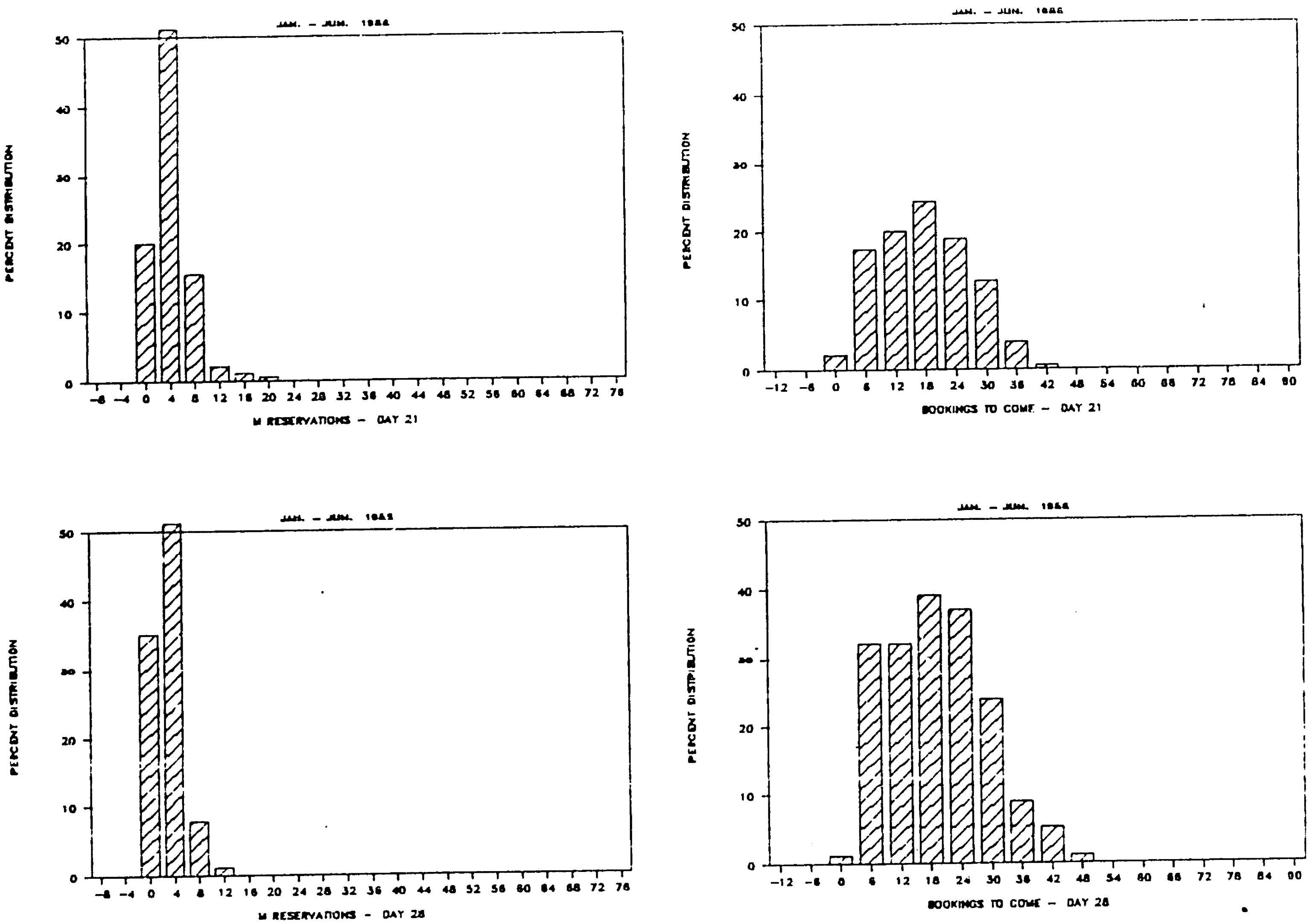
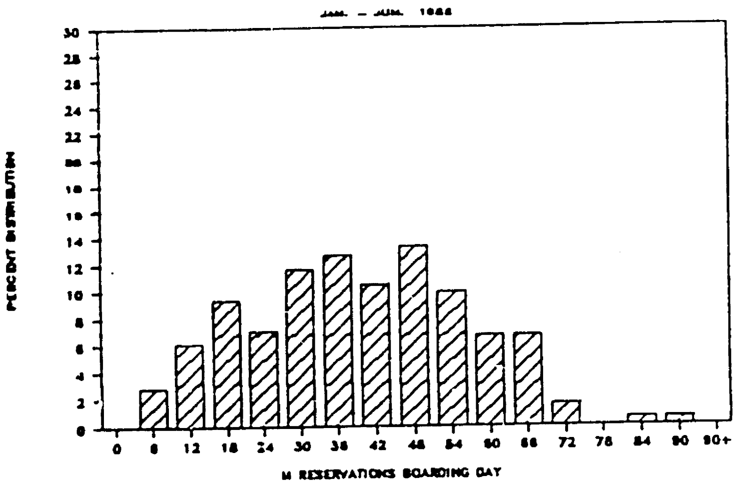


Figure 4.03 (cont.) : Distribution Plots

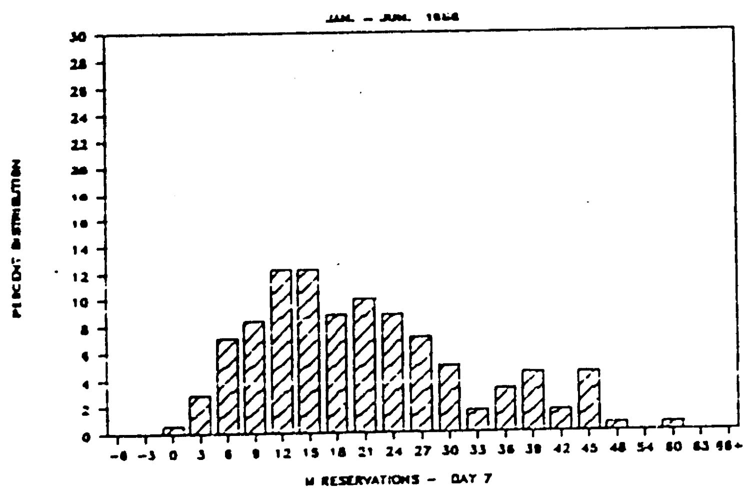
M-class Flight F4 Market C/D

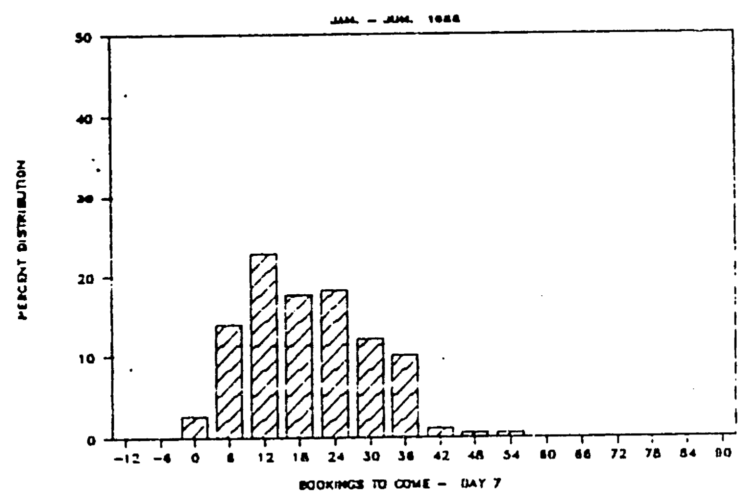


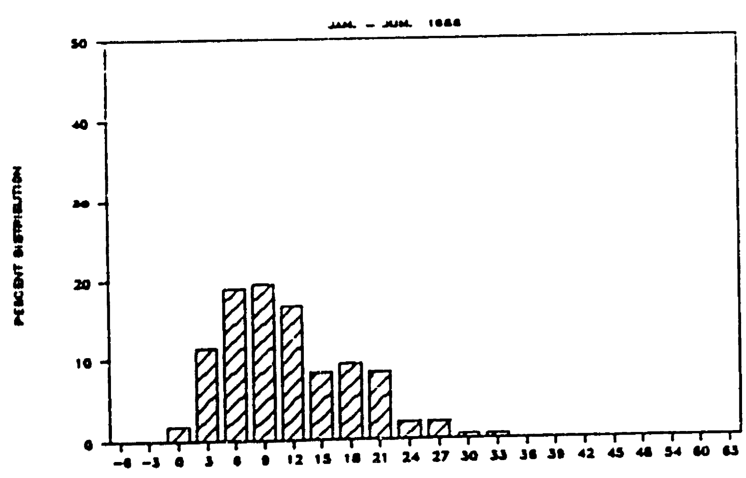
Sample set to -> ALL
Descriptive Statistics - 181 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F3MBD	36.564	17.241	.11119	2.4484	4.000	85.00
F3M7	19.564	11.384	.68521	2.8519	.0000	56.00
F3M14	10.370	6.4945	.72749	3.0515	.0000	32.00
F3M21	6.1050	4.0968	.94272	3.6303	.0000	20.00
F3M28	3.7459	3.1940	1.2185	4.5497	.0000	16.00
F3M7_BD	17.000	10.423	.40856	2.7033	-4.000	52.00
F3M14_BD	26.193	14.137	.30349	2.6268	-1.000	69.00
F3M21_BD	30.459	15.682	.18637	2.5519	1.000	79.00
F3M28_BD	32.818	16.119	.12118	2.4817	.0000	81.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







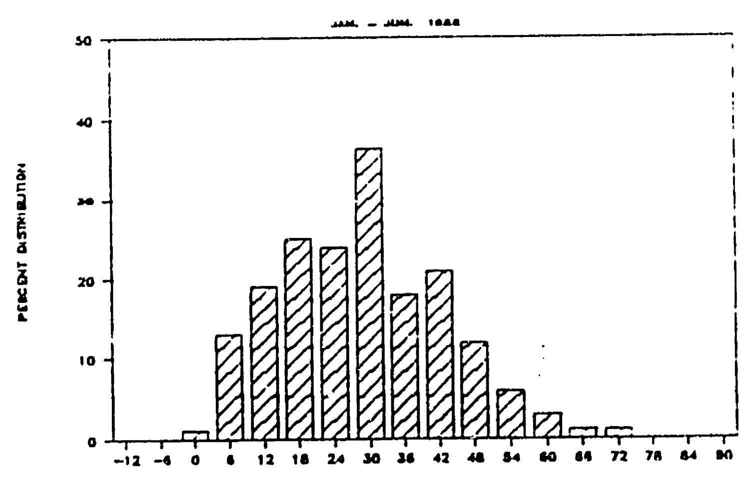


Figure 4.04 : Distribution Plots

M-class Flight F3 Market D/C

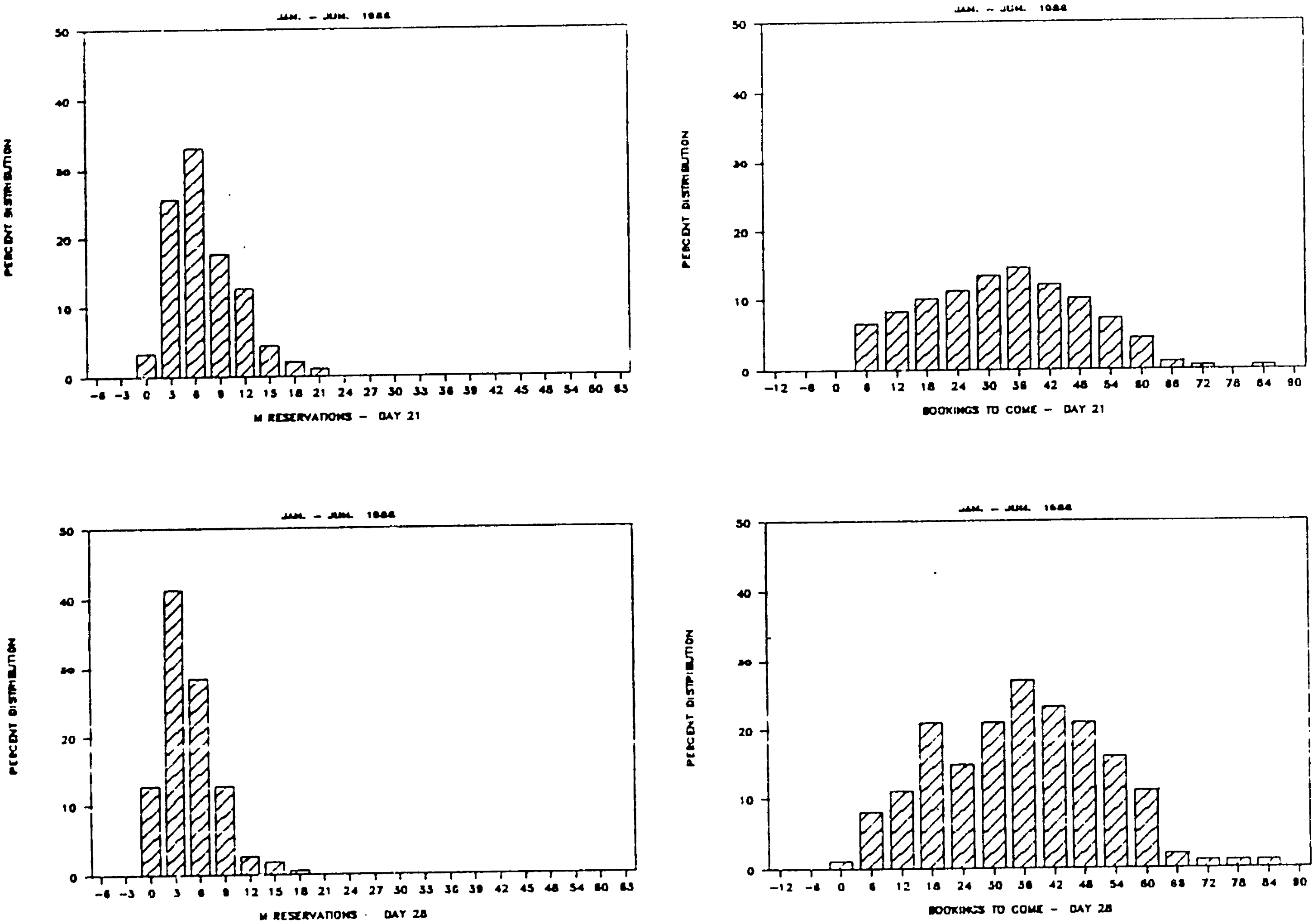
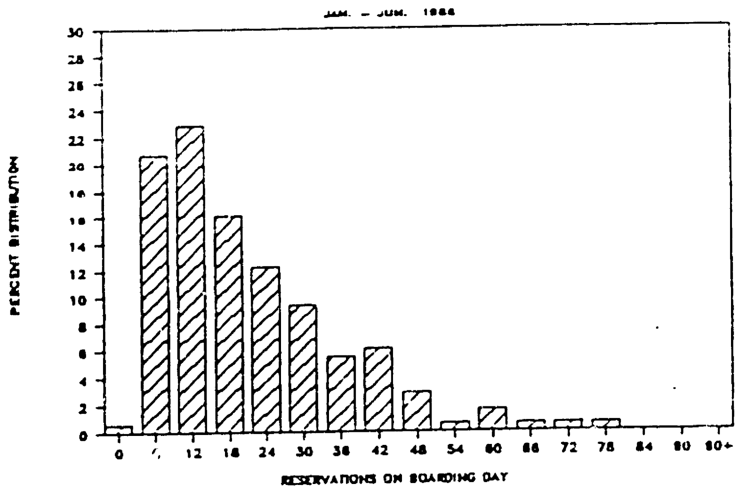


Figure 4.04 (cont.) : Distribution Plots

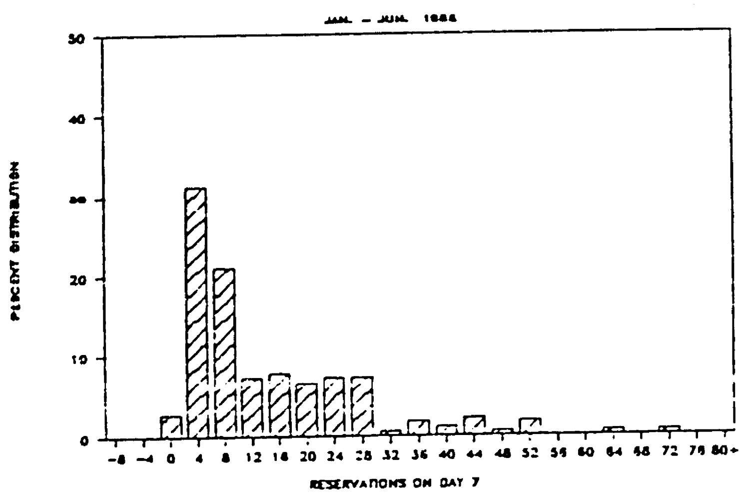
M-class Flight F3 Market D/C

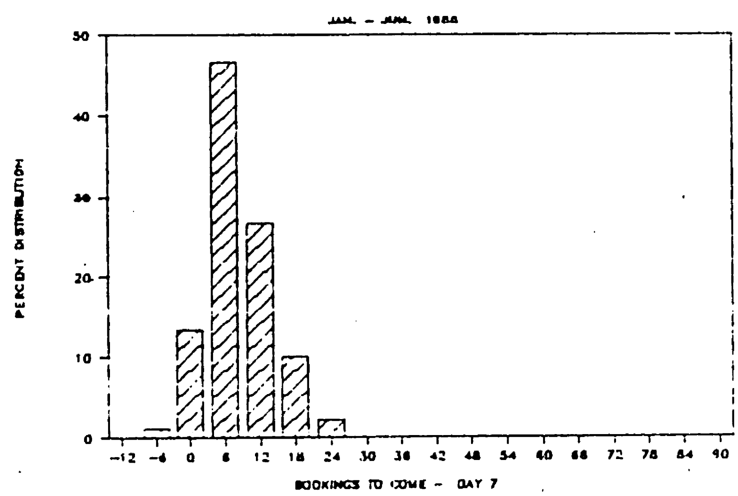


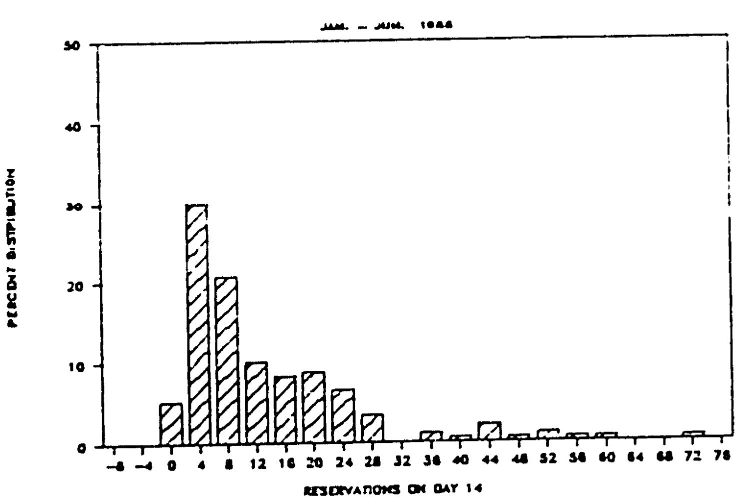
Sample set to -> ALL
Descriptive Statistics - 181 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F1M80	18.492	14.380	1.2502	4.4513	.0000	73.00
F1M7	12.752	12.965	1.5975	6.2956	.0000	72.00
F1M14	11.657	12.295	1.9275	7.2190	.0000	69.00
F1M21	10.249	11.475	2.0563	7.5916	.0000	59.00
F1M29	9.5912	11.792	2.2476	8.6765	.0000	64.00
F1M7_80	5.7293	5.5446	.56191	3.0394	-9.000	22.00
F1M14_80	6.8343	7.0691	.73960	3.6576	-10.00	34.00
F1M21_80	8.2431	8.8980	.38958	3.6611	-17.00	42.00
F1M29_80	8.9006	10.246	.18590	6.0279	-36.00	53.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







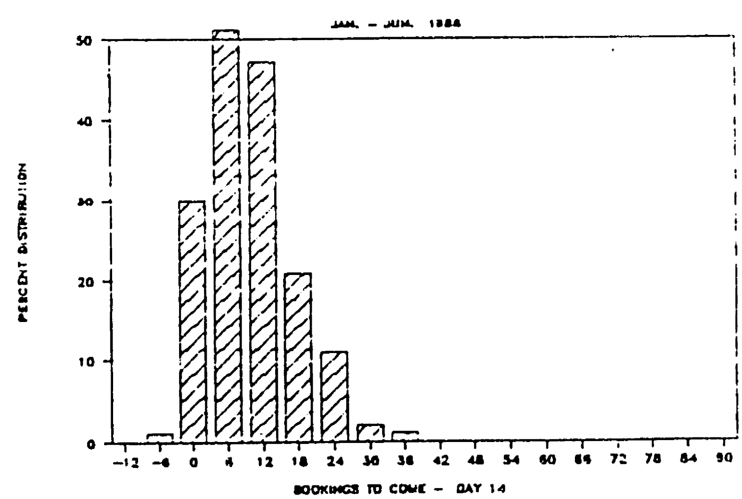


Figure 4.05 : Distribution Plots

M-class Flight F1 Market E/F

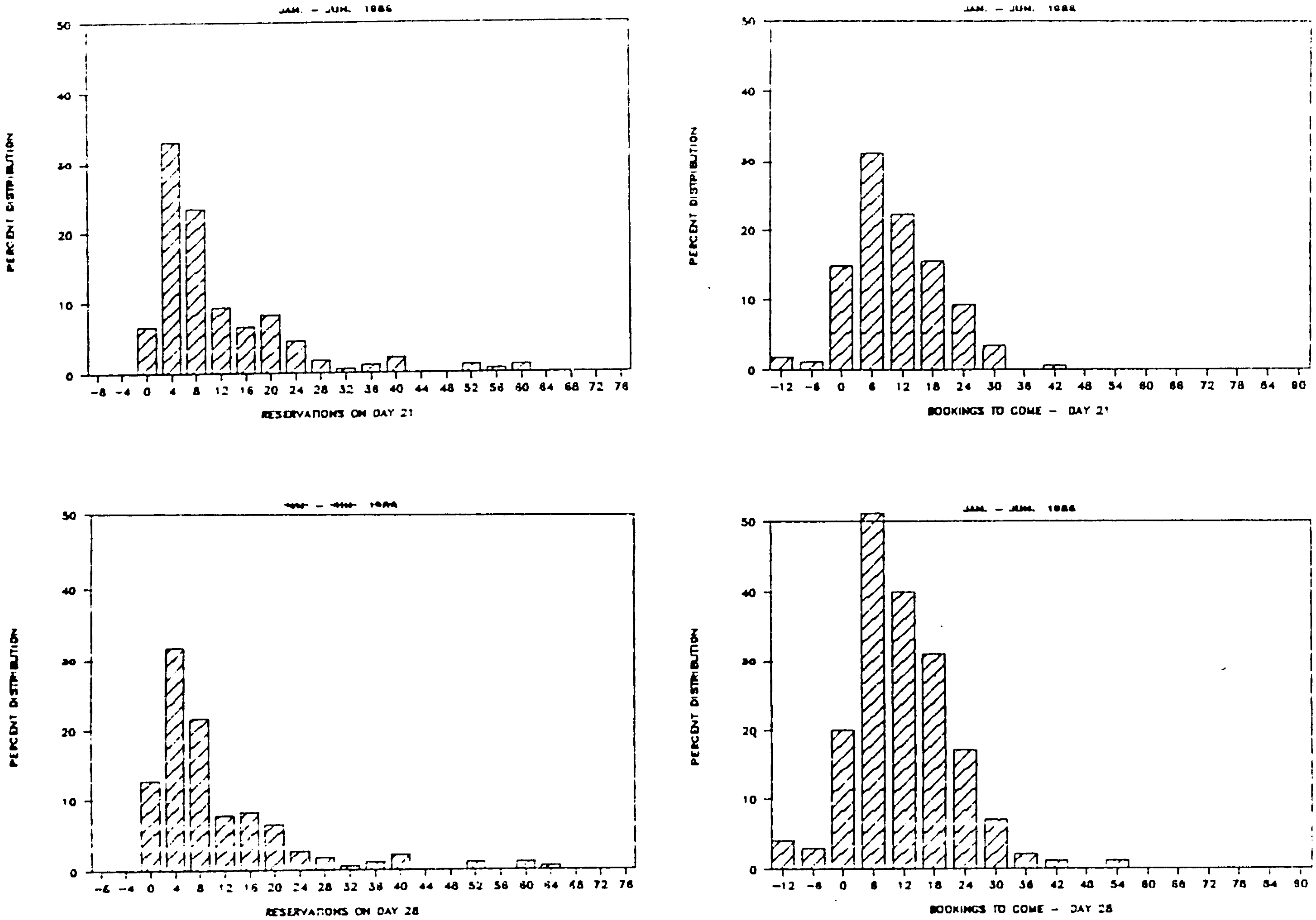
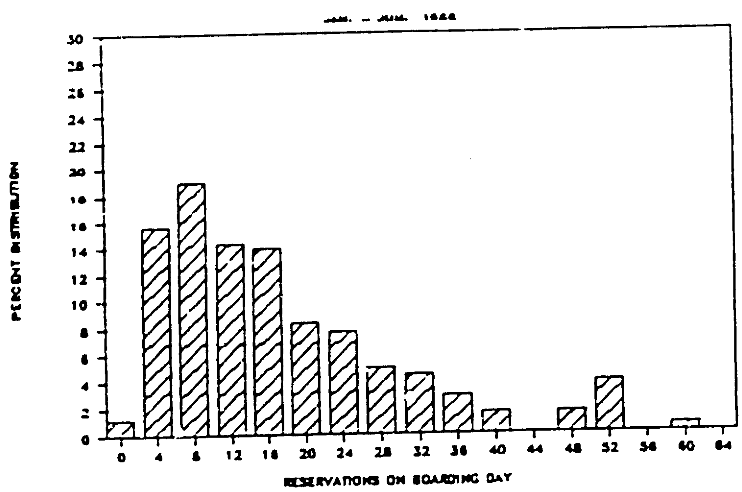


Figure 4.05 (cont.) : Distribution Plots

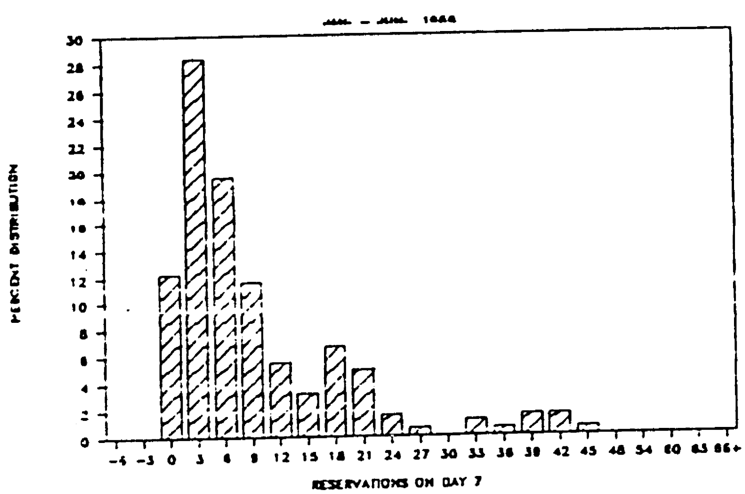
M-class Flight F1 Market E/F

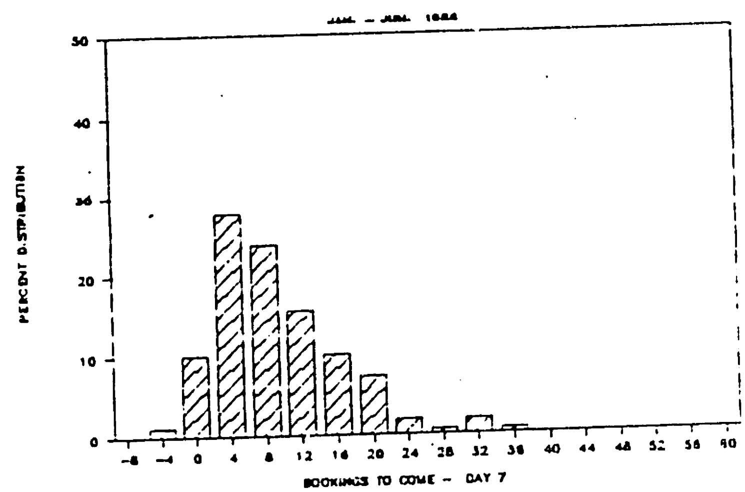


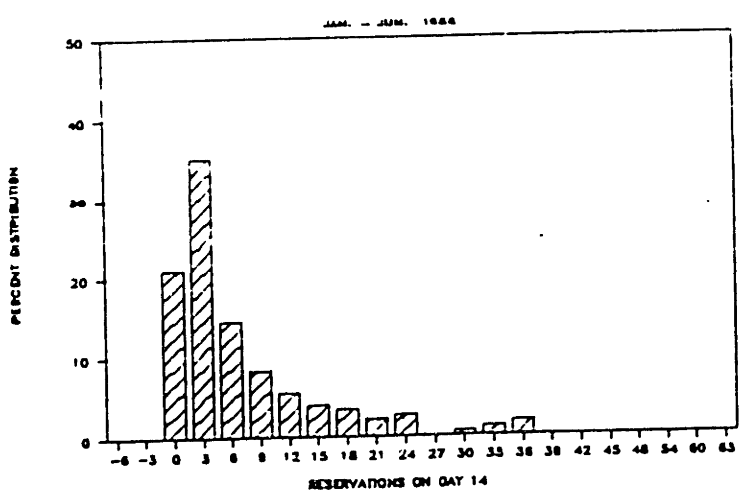
Sample set to -> ALL
Descriptive Statistics - 180 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F2MBD	15.811	12.703	1.2455	4.1136	.0000	57.00
F2M7	8.2333	9.4951	1.9791	6.3571	.0000	43.00
F2M14	5.9056	7.7594	1.9319	6.3720	.0000	35.00
F2M21	5.2778	7.4437	2.3957	10.002	.0000	47.00
F2M28	4.3389	6.8071	2.6277	11.693	.0000	41.00
F2M7_BD	7.5778	6.9956	.98957	4.1794	-6.000	33.00
F2M14_BD	9.9056	8.8701	.81249	4.0811	-15.00	40.00
F2M21_BD	10.533	9.6094	.74549	4.0999	-19.00	42.00
F2M28_BD	11.472	10.194	.64110	4.5303	-25.00	44.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







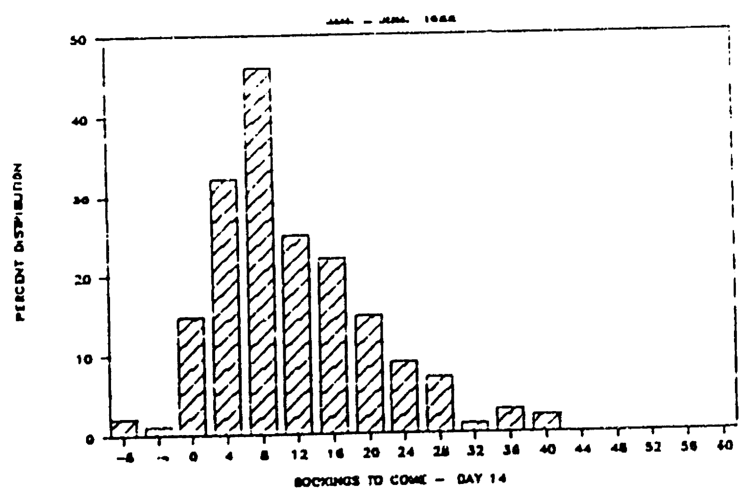


Figure 4.06 : Distribution Plots

M-class Flight F2 Market F/E

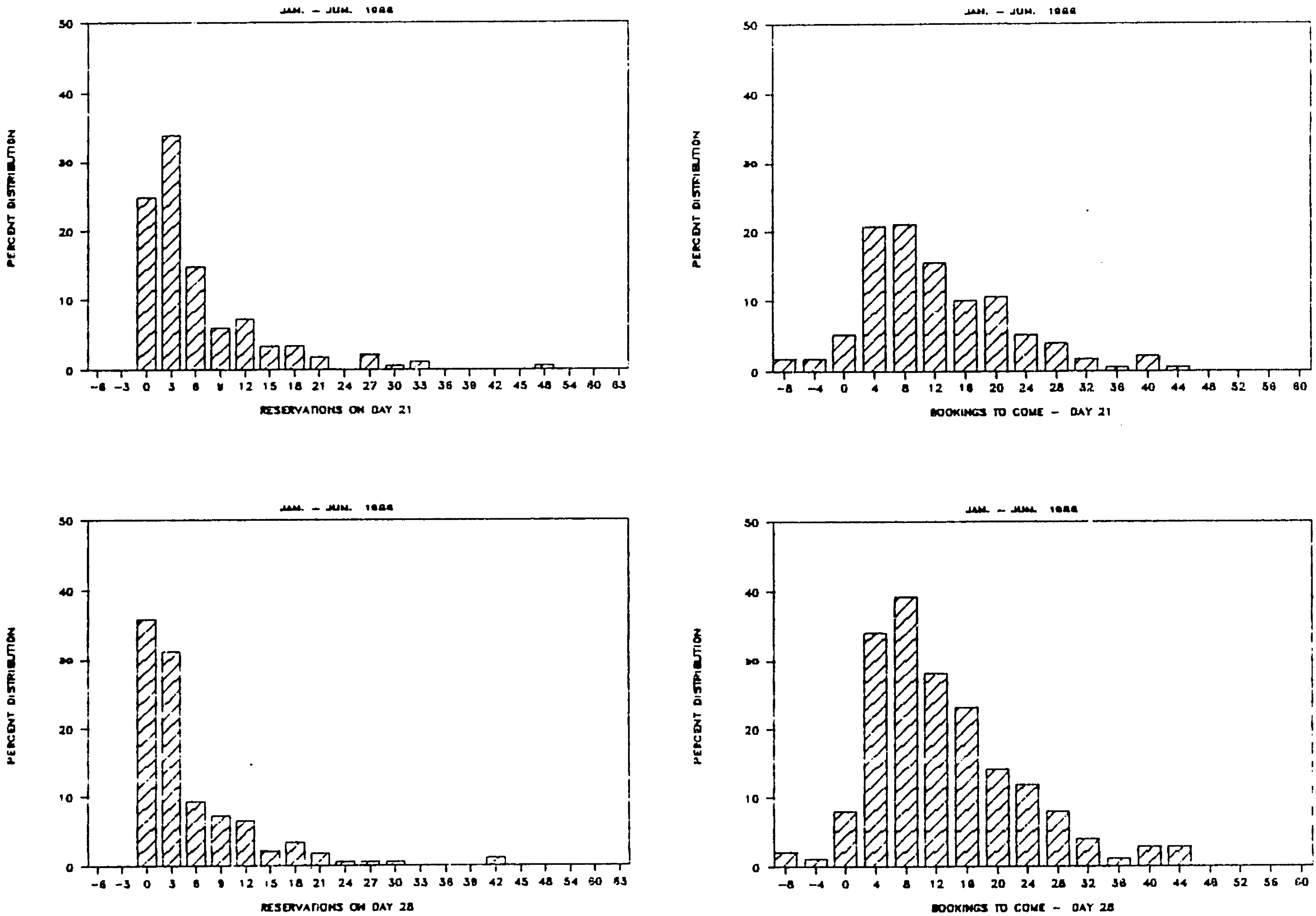
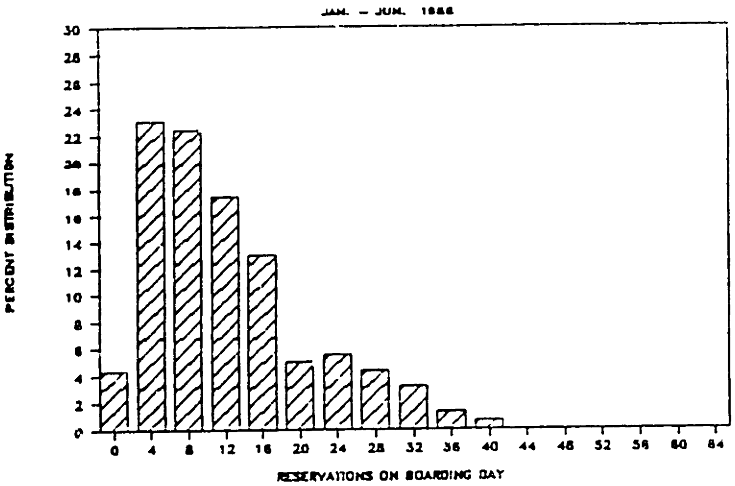


Figure 4.06 (cont.) : Distribution Plots

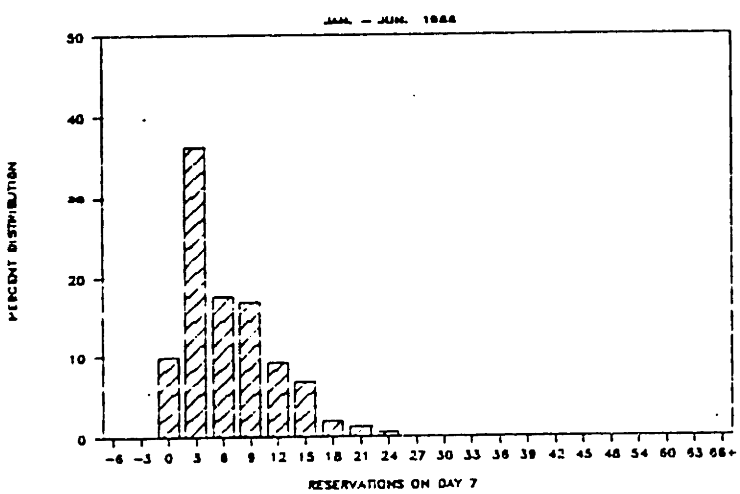
M-class Flight F2 Market F/E

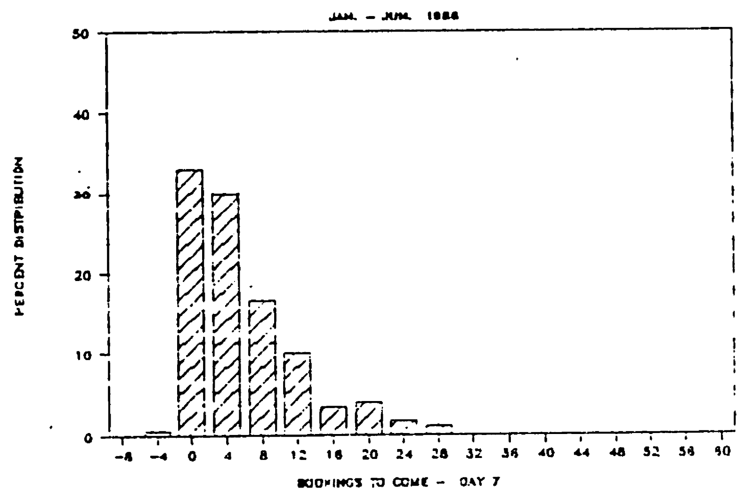


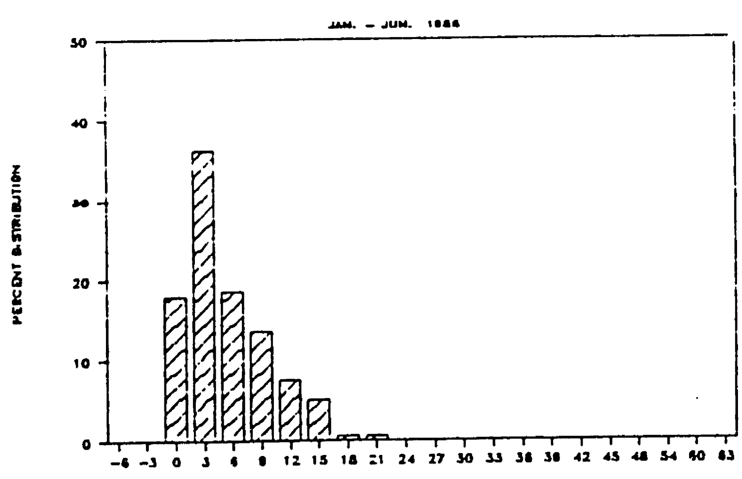
Sample set tc -> ALL
Descriptive Statistics - 177 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F1M90	9.5876	8.7441	1.0562	3.5150	.0000	39.00
F1M7	5.0339	4.8464	1.1587	3.9874	.0000	22.00
F1M14	4.0113	4.2707	1.2575	4.0886	.0000	20.00
F1M21	3.3955	4.0748	1.5439	5.0282	.0000	20.00
F1M28	2.5028	3.5099	1.9954	7.3824	.0000	19.00
F1M7_BD	4.5537	6.0367	1.3625	4.6258	-5.000	26.00
F1M14_BD	5.5763	6.9818	1.1649	3.9173	-5.000	29.00
F1M21_BD	6.1921	7.3002	1.0470	3.6200	-7.000	30.00
F1M28_BD	7.0347	7.7738	1.0703	3.6601	-6.000	32.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







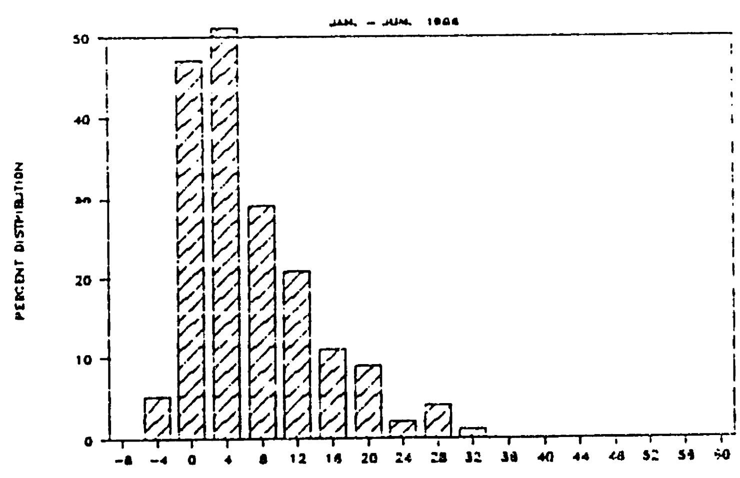


Figure 4.07 : Distribution Plots

M-class Flight F1 Market G/H

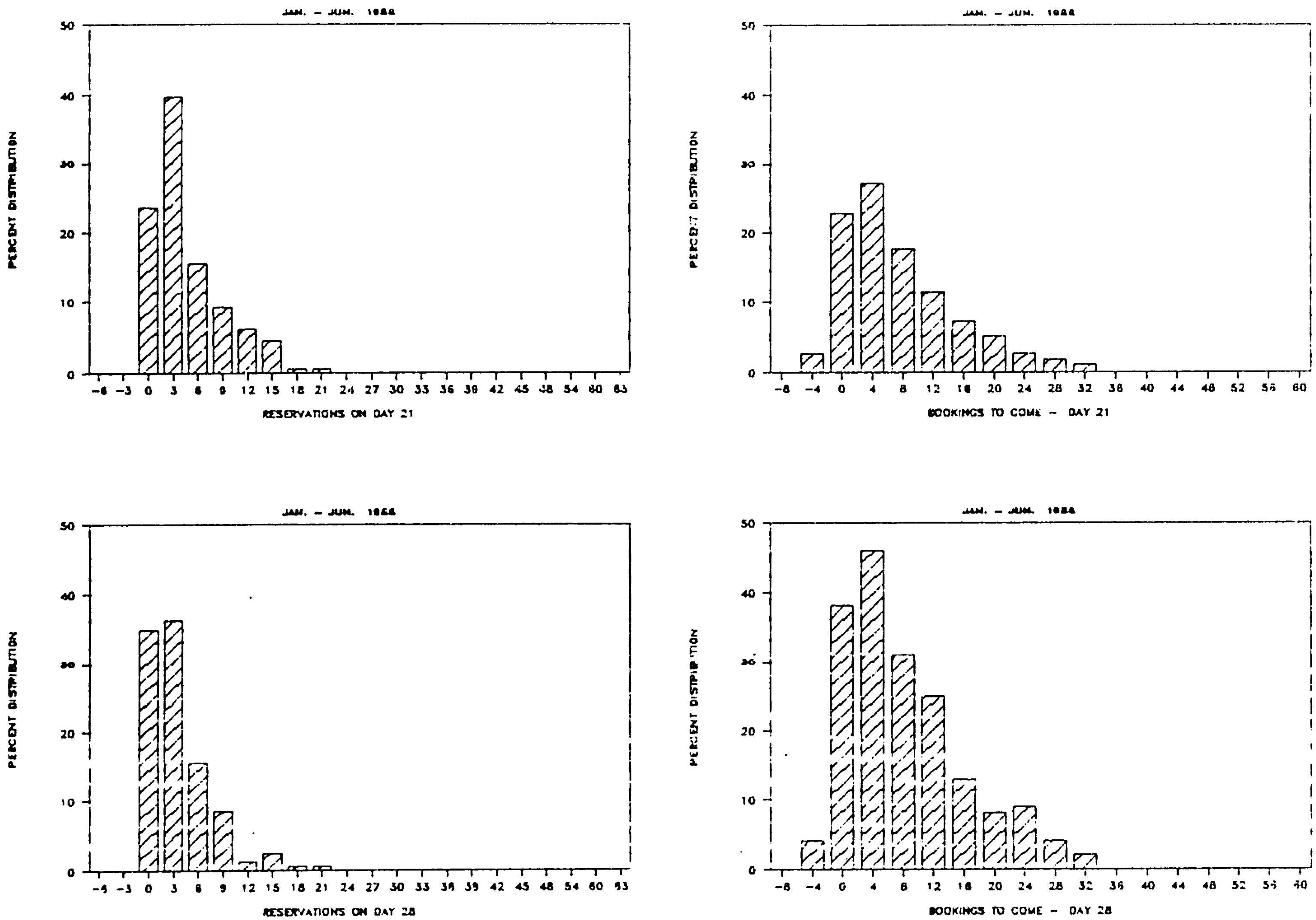
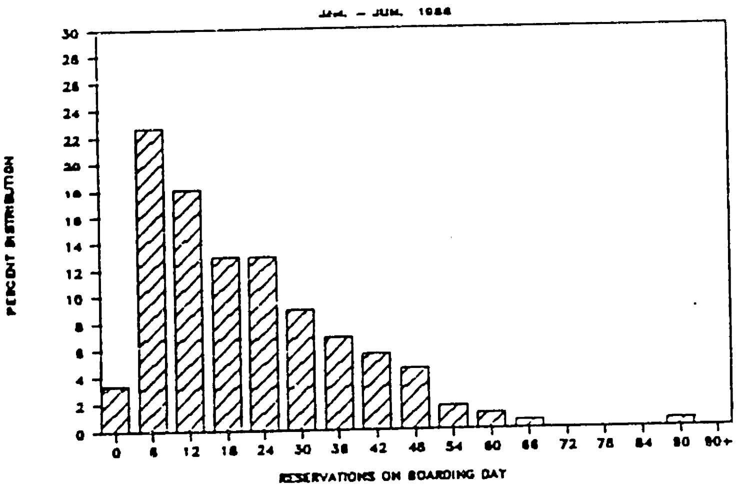


Figure 4.07 (cont.) : Distribution Plots

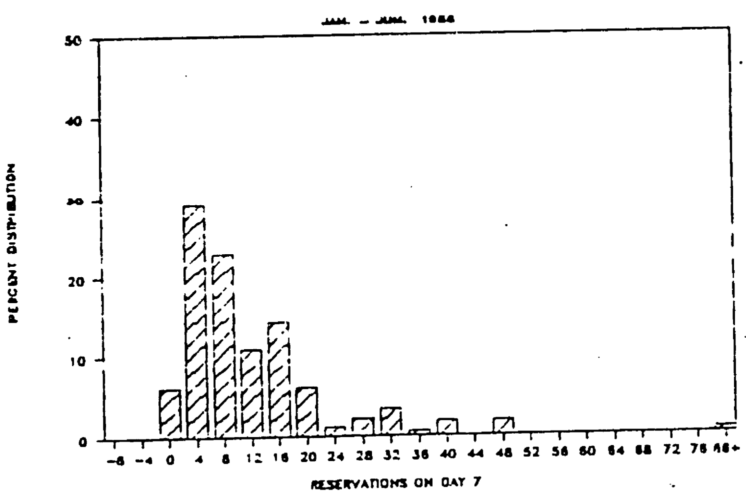
M-class Flight F1 Market G/H

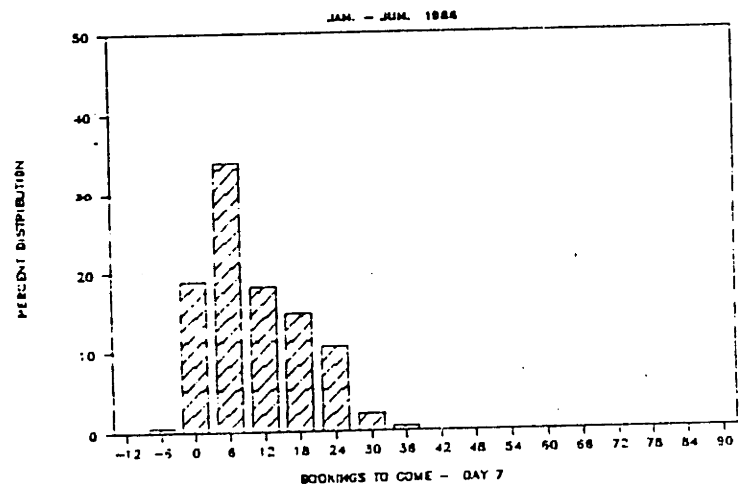


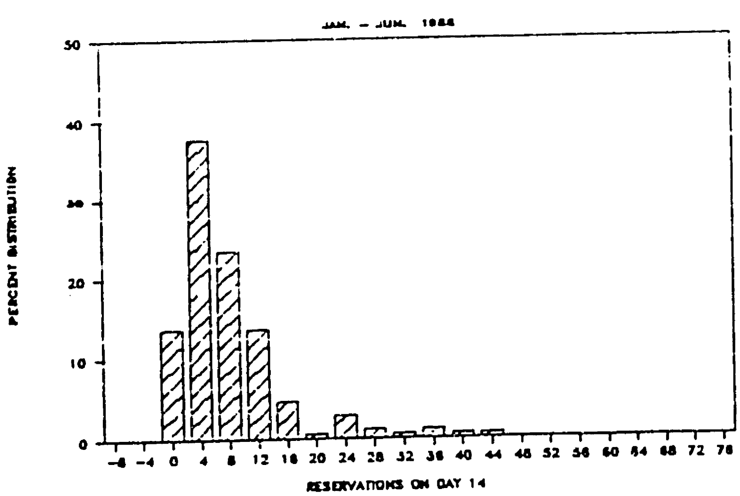
Sample set to -> ALL
Cescriptive Statistics - 177 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F2M8D	18.458	15.468	1.1536	4.6206	.0000	88.00
F2M7	10.350	11.427	2.7474	15.177	.0000	88.00
F2M14	6.9209	9.7558	4.4785	32.974	.0000	90.00
F2M21	5.1582	8.8175	5.2529	42.639	.0000	86.00
F2M28	3.8023	7.8935	7.1373	69.124	.0000	86.00
F2M7_8D	8.1073	8.0032	.75510	2.8912	-8.000	35.00
F2M14_8D	11.537	10.723	.78253	2.7486	-3.000	43.00
F2M21_8D	13.299	12.200	.65665	3.1032	-23.00	52.00
F2M28_8D	14.655	12.555	.79869	2.9498	-2.000	56.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







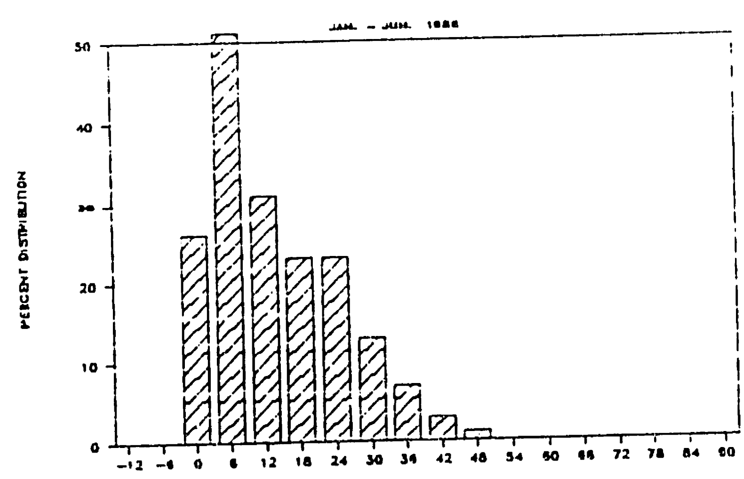


Figure 4.08 : Distribution Plots

M-class Flight F2 Market H/G

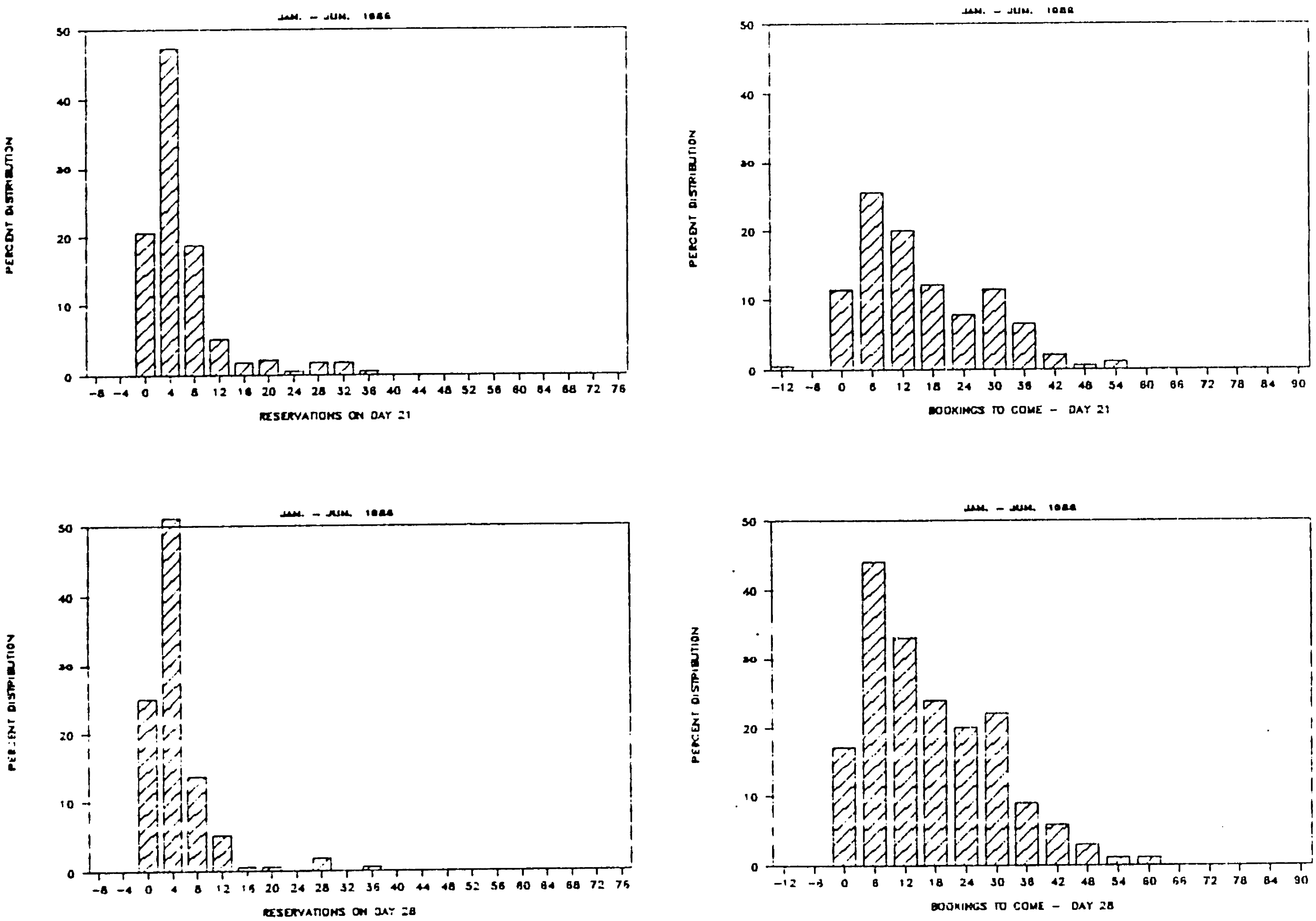
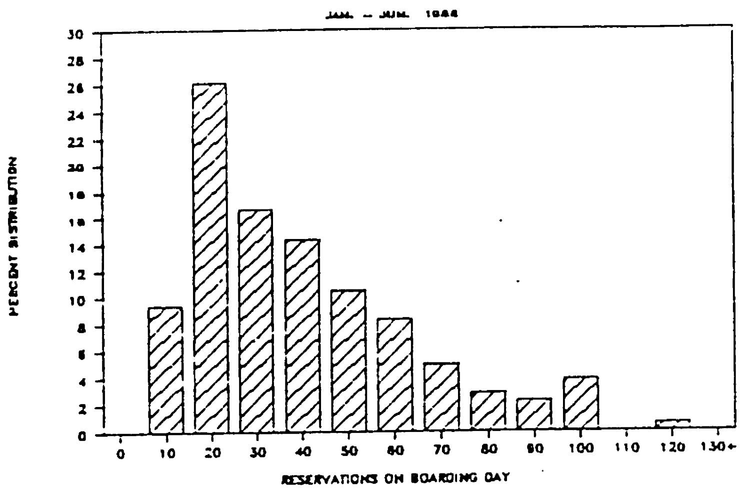


Figure 4.08 (cont.) : Distribution Plots

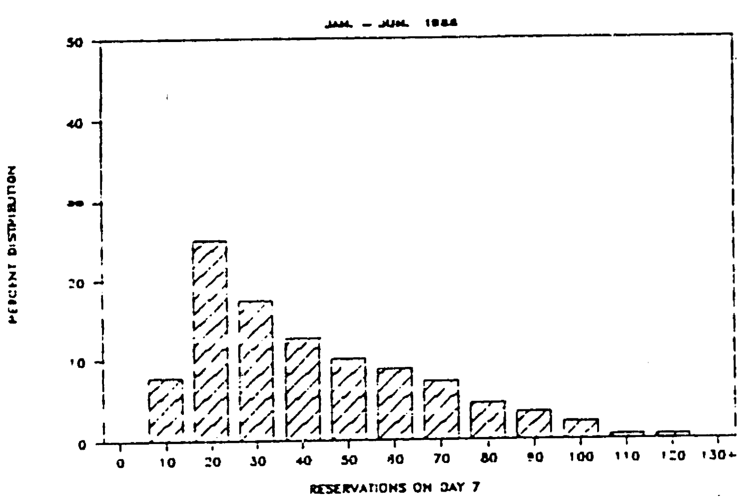
M-class Flight F2 Market H/G

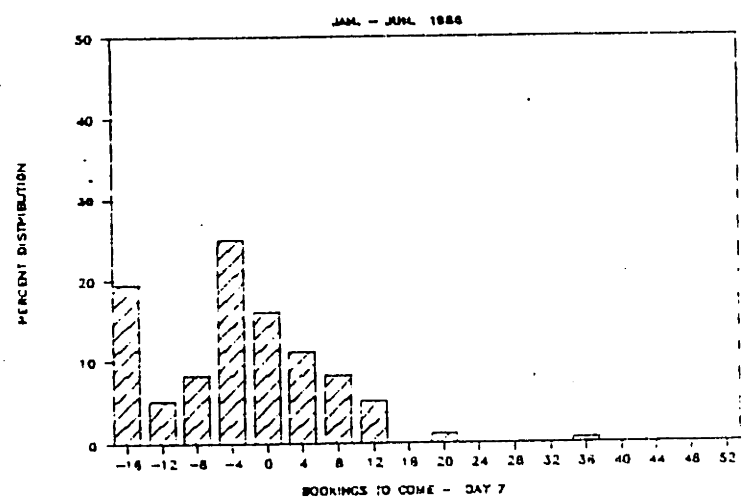


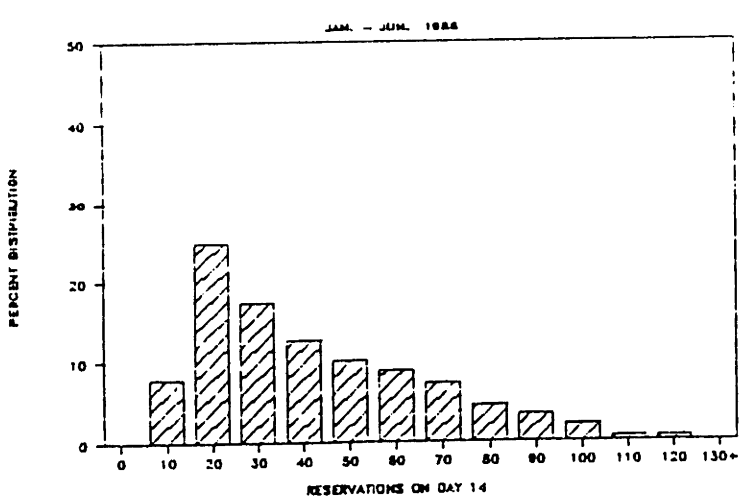
Sample set to -> ALL
Descriptive Statistics - 181 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F1MBD	33.414	18.656	1.3495	5.3380	2.000	112.0
F1M7	36.414	20.924	1.1379	4.1376	6.000	112.0
F1M14	35.315	22.328	1.1341	4.0560	5.000	116.0
F1M21	32.713	23.651	1.2593	4.3340	3.000	119.0
F1M28	30.276	24.502	1.4118	4.7802	1.000	119.0
F1M7_BD	-3.0000	6.9873	-2.4831	13.880	-49.00	9.000
F1M14_BD	-1.9006	9.4717	-.32929	3.9236	-32.00	33.00
F1M21_BD	.70166	11.645	-.29316	4.0552	-32.00	38.00
F1M28_BD	3.1381	12.390	-.22423	3.6612	-33.00	39.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







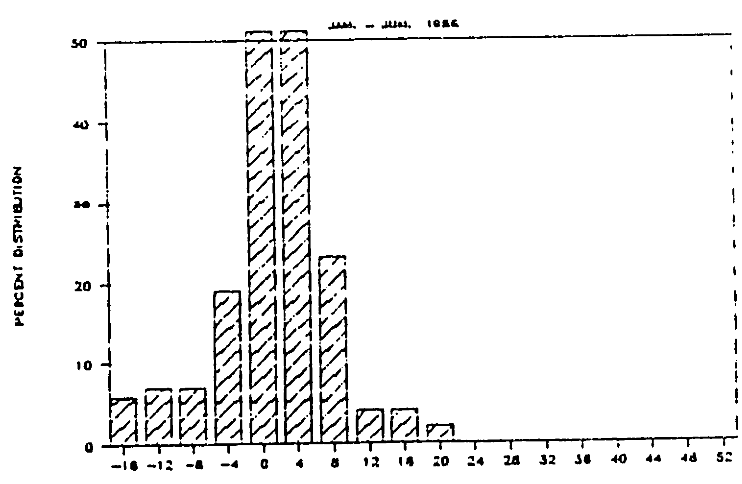


Figure 4.09 : Distribution Plots

Y-class Flight F1 Market I/J

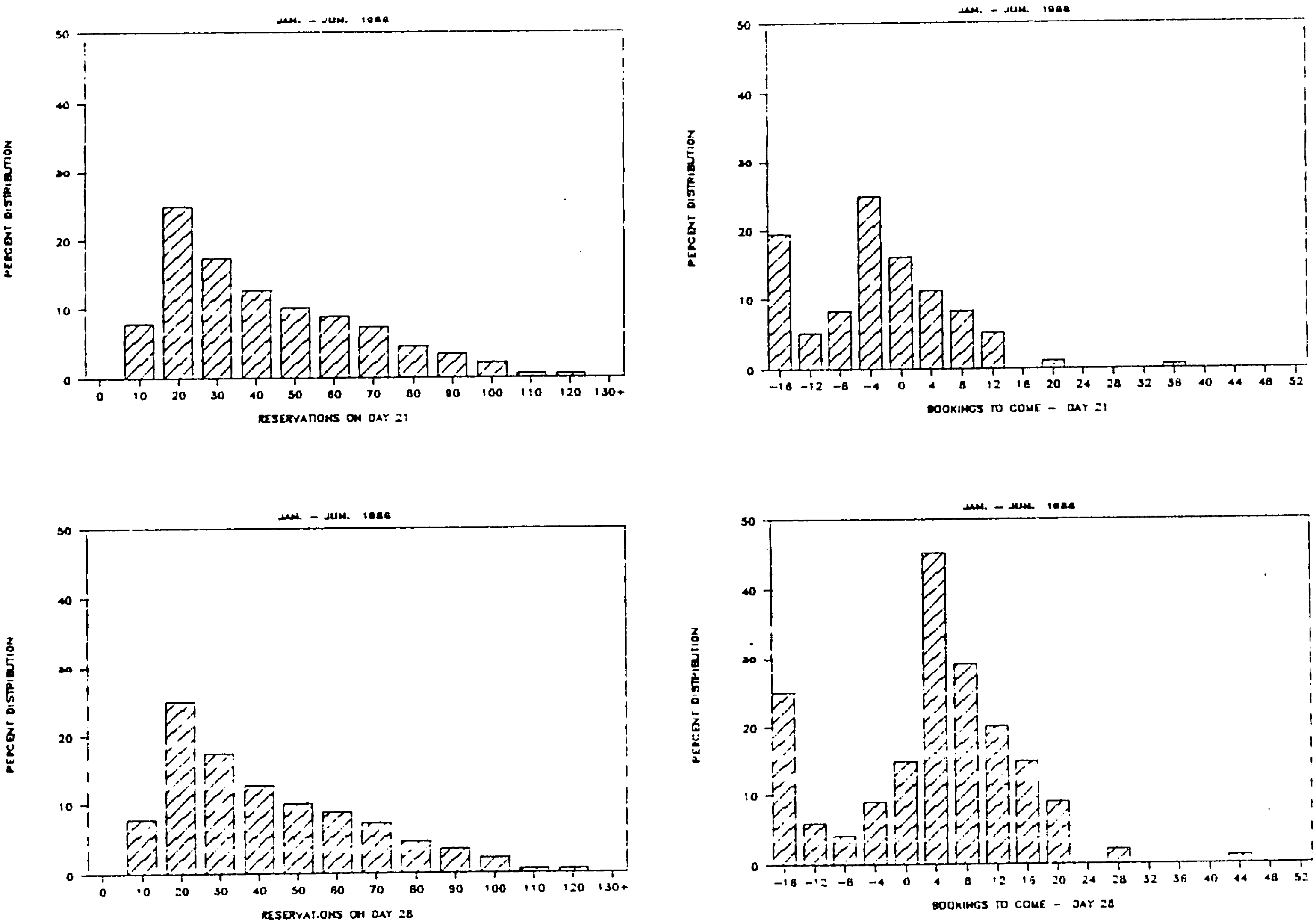
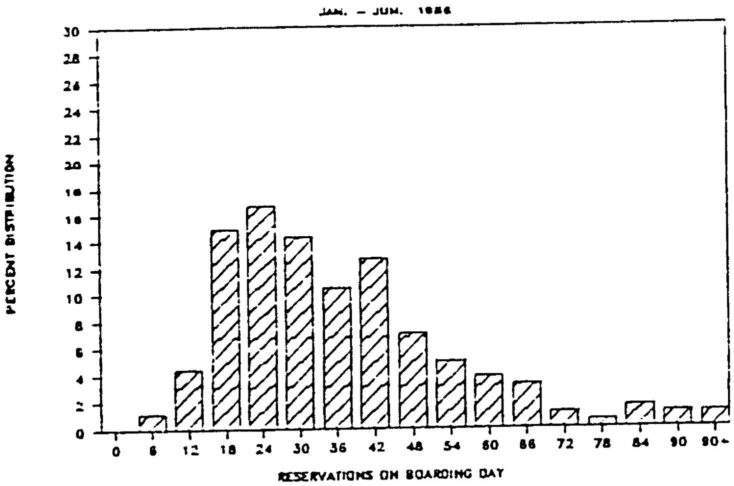


Figure 4.09 (cont.) : Distribution Plots

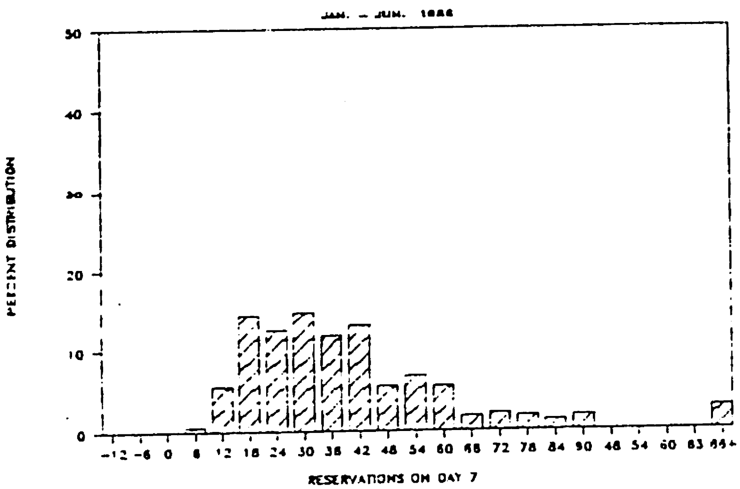
Y-class Flight F1 Market I/J

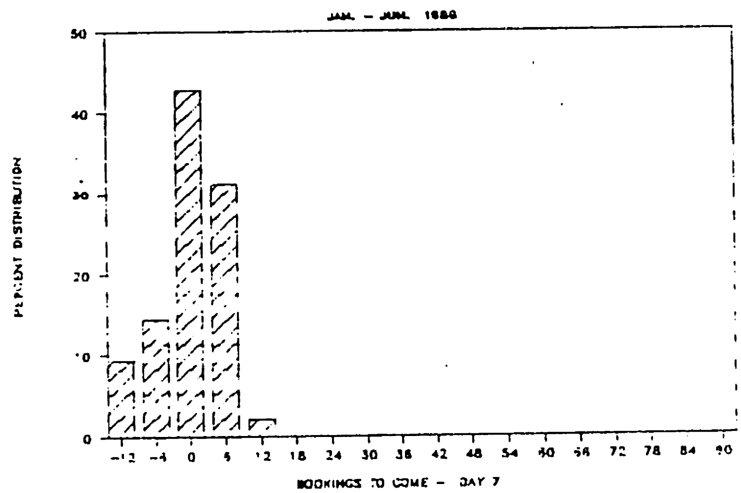


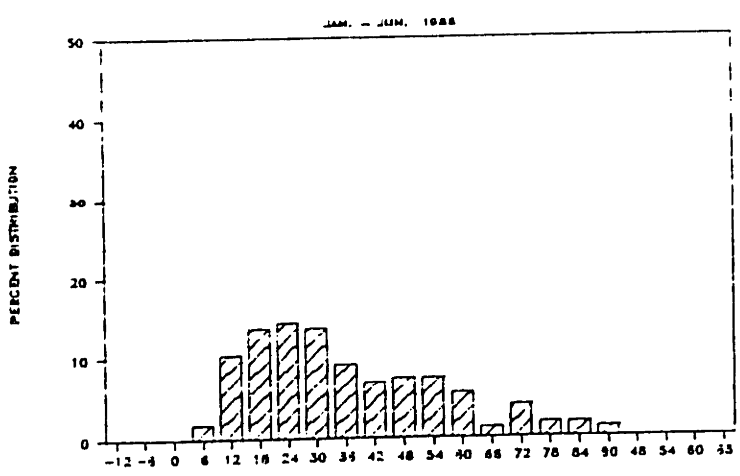
Sample set to -> ALL
Descriptive Statistics - 179 observations used.

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
F1MBD	35.073	23.250	1.0442	3.4700	3.000	111.0
F1M7	36.514	24.254	.96106	3.2470	4.000	115.0
F1M14	35.225	25.853	1.0526	3.5919	1.000	123.0
F1M21	34.453	27.356	1.1066	3.7430	1.000	123.0
F1M28	34.520	30.476	1.0933	3.5124	.0000	133.0
F1M7_BD	-1.4413	4.4026	-1.0050	5.5798	-18.00	10.00
F1M14_BD	-.16201	6.6790	-.75446	5.3971	-27.00	17.00
F1M21_BD	.61453	9.9621	-.65930	5.9584	-40.00	39.00
F1M28_BD	.55307	14.710	-1.3151	5.8150	-59.00	44.00

(Skewness = m3/s**3; Kurtosis = m4/s**4)







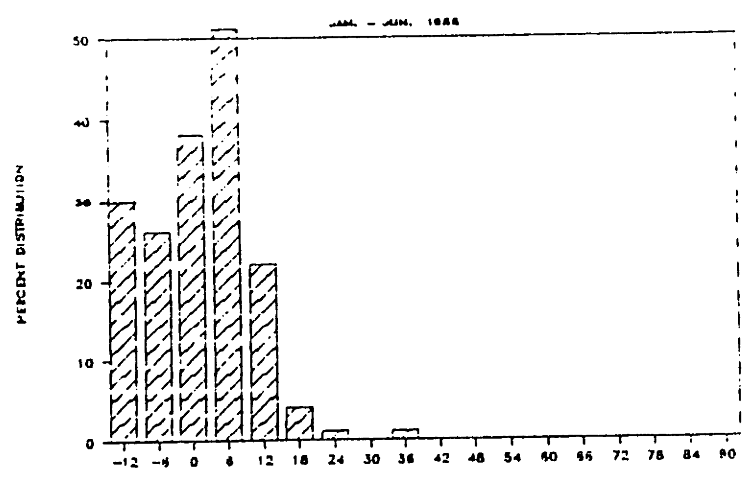


Figure 4.10 : Distribution Plots

Y-class Flight F1 Market J/I

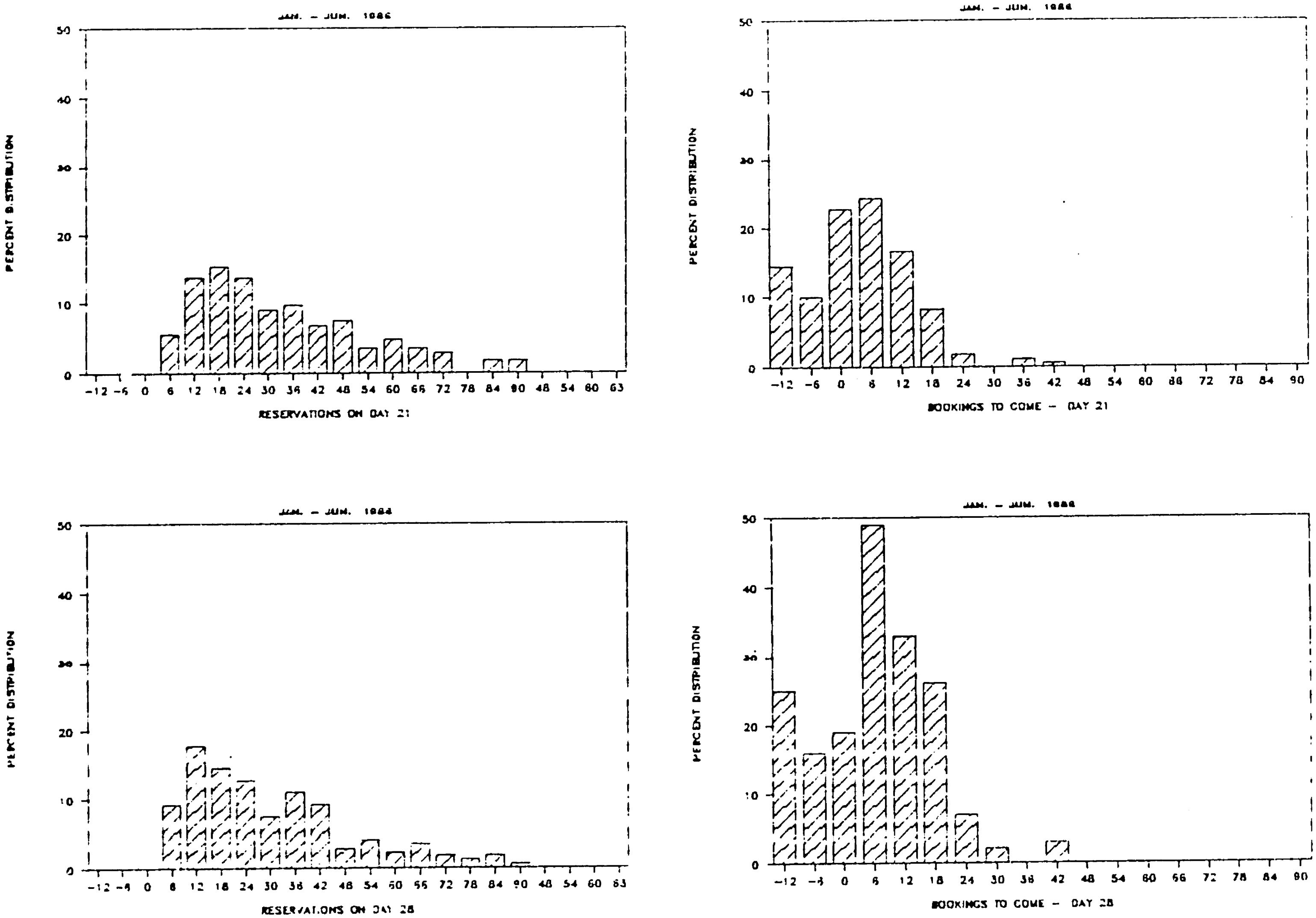


Figure 4.10 (cont.) : Distribution Plots

Y-class Flight F1 Market J/I

CHAPTER FIVE

RESERVATIONS FORECASTING

Forecasts are made because they assist the decision making process from the analysis of policy, activity, or plan to the timing and implementation of an action, program, or strategy [7]. The need to forecast final reservation requests, particular to this thesis, represents a key ingredient of the decision making process in a Yield Management System. Regardless of the core algorithm that the Seat Inventory Control routine uses, forecasts of final reservation requests are needed.

In general, the various forecasting methods can be divided into three broad categories: quantitative or scientific, qualitative or judgmental, and decision analysis, which is a combination of the first two methods.

Quantitative methods have been used more often, and have gained a wide acceptance for several reasons. They rely heavily on the existence and use of historical data, and to a large extent, on the perpetuation of past behavior.

Forecasting methods are based on cause-effect relationships, statistical analysis or simulation methods, which in turn provide a description of the underlying process one is trying to understand, explain and make forecasts.

Qualitative methods, on the other hand, rely on the "intuition and/or experience" of the forecaster, and are, as a consequence, dependent on the forecaster ability in describing the process. Subjective by nature, this class of forecast "models" tend to be used when very little information is available, or when there is an inherent inability of modelling in an objective fashion.

Decision analysis, the remaining category, is a combination of both quantitative and qualitative methods. In this category, assumptions on some unknown parameters are made, and using quantitative methods, forecasts are made.

Given the ability of any airline to provide a vast set of historical data , quantitative methods can be applied to fulfill the forecast need of the Yield Management System. Therefore, quantitative methods in forecasting final bookings for a flight are explored in this thesis.

5.1 ALTERNATIVES IN FORECASTING

Time Series Analysis are heavily based on statistical behavior. The model/parameters estimated in Time Series Analysis do not have specific meaning. Models in this class can be used as a forecast tool, but they do not try to explain the underlying nature of the process. Models are developed based on a statistical basis only.

In this class of models, we presume to know nothing about the real world causal relationships that affect the variable to be forecasted. Instead, past behavior of a time series is examined in order to infer something about its future behavior. Ratio analysis, trend projection, moving averages, spectral analysis, and Box_Jenkins' ARIMA modelling can be cited as examples of Time Series Analysis methods.

The second class of Quantitative methods, namely Causal Methods, concentrates on models that can be expressed in equation form, relating variables quantitatively. Data are then used to estimate parameters of the equations, and theoretical relationships are tested statistically.

In single equation regression models, the variable under study is explained by a single function (linear or nonlinear) of explanatory variables. The equation will often be time-dependent (i.e. a time index will appear explicitly in the model), so that one can predict the response over time of the variable under study.

5.2 TIME SERIES ANALYSIS

Time Series Analysis presumes that the series to be forecasted has been generated by a stochastic process with a structure that can be characterized or described. In other words, a times series model provides a description of the random nature of the stochastic process that generated the sample under study. The description is given in terms of how the randomness is embodied in the process, and not in terms of cause-effect relationships.

Because time series analysis require a large deal of statistical analysis, it provides better estimates and it is more sophisticated than simple extrapolations. Simple extrapolation, such as trend analysis do not account for the fact that a time series is the result of a stochastic process.

With Box-Jenkins' Auto Regressive Integrated Moving Average models , known as ARIMA models [8], one can describe time series process, by using autoregressive and moving average components. A constant may also be included in the model. Model parameters can be estimated for the original series, or for the time series differentiated i times. Parameters are chosen for the terms included in the model in such a way that it minimizes the sum of square differences between actual time series and fitted time series. Model parameters, again, convey no special mean.

ARIMA models were developed and estimated for flight F1 in the A/B market, as an example of application and to observe model fitting results. The data used for modelling was final bookings for M class for this flight, for the six month sample. Table 5.01 shows fitting results for an ARIMA(3,0,2) model. The equation fitted was:

$$AR3 \;.\; MBD(t) = C + MA2 \;.\; r(t)$$

where :

$AR3 = (1 + AR(1).B + AR(2).B^{2} + AR(3).B^{3})$

B = backward shift operator, defined as $B^{n} [X(t)]= X(t-n)$;

MBD = final reservations, M-class;

$MA2 = (1 + MA(1).B + MA(2).B^{2})$;

$r(t)$= residual at time t ;

and (to be determined by model fitting)

C = constant ;

AR(i) = coefficient i calculated for the Moving Average components ;

MA(i) = coefficient i calculated for the Auto-Regressive components.

The model estimated is statistically accepted, because the calculated chi-square test statistic on first 20 residual autocorrelations is 17.91, which is meaningful at least at a confidence level of .90 , chi-square(15, .90) = 22.3 . Three from five parameters do not exhibit acceptable t statistics. The estimated white noise variance is 77.56, which corresponds to a standard error of regression of 8.81 . From figure 4.01 one can observe that the standard deviation of the time series variable is 8.89 . No much explanatory power was gained with this ARIMA model.

```
ITERATION 4:  RESIDUAL SUM OF SQUARES ..... 13739
ITERATION 5:  RESIDUAL SUM OF SQUARES ..... 13434.9
ITERATION 6:  RESIDUAL SUM OF SQUARES ..... 13346.9
```

SUMMARY OF FITTED MODEL

parameter	estimate	stnd.error	t-value	prob(>|t|)
AR (1)	1.00765	.14360	7.01705	.00000
AR (2)	-.11036	.17264	-.63922	.52353
AR (3)	.02083	.08176	.25474	.79923
MA (1)	1.07407	.14845	7.23533	.00000
MA (2)	-.22857	.15545	-1.47036	.14329
MEAN	22.04925	1.33373	16.53206	.00000
CONSTANT	1.87547			

ESTIMATED WHITE NOISE VARIANCE = 77.5587 WITH 172 DEGREES OF FREEDOM.
CHI-SQUARE TEST STATISTIC ON FIRST 20 RESIDUAL AUTOCORRELATIONS = 17.9051

Table 5.01 Time Series Analysis

ARIMA (3,0,2)

Reservations on Boarding Day

M-class   Fight F1   Market A/B

Another time series model was estimated for the original series differenced once, which corresponds to the series D(t) = MBD(t) - MBD(t-1). A week seasonality was introduced in the model . Table 5.02 shows fitting results for a seasonal ARIMA(0,1,4), with length of 7 days. The equation that represents the model is :

$$D(t) = C + MA4 \,.\, r(t)$$

where :

$D(t) = MBD(t) - MBD(t-1)$;

MBD = final reservations, M-class;

$$MA4 = 1 + MA(7).B^{7} + MA(14).B^{14} + MA(21).B^{21} + + MA(28).B^{28}$$

$r(t)$ = residual at time t ;

and (to be determined by model fitting)

C = constant ;

MA(i) = coefficient i calculated for the Auto-Regressive components.

The calculated chi-square statistic is 17.72 (< 22.3), which means that the model can be accepted. Estimated white noise variance, this time, was higher than before: 108.23 , which means a standard error of regression of 10.40.

```
ITERATION 1:  RESIDUAL SUM OF SQUARES ..... 16433.8
ITERATION 2:  RESIDUAL SUM OF SQUARES ..... 15282
ITERATION 3:  RESIDUAL SUM OF SQUARES ..... 15260.9
```

SUMMARY OF FITTED MODEL

parameter	estimate	stnd.error	t-value	prob(>|t|)
SAR(7)	-.70067	.08339	-8.40285	.00000
SAR(14)	-.43076	.10223	-4.21363	.00004
SAR(21)	-.24550	.10321	-2.37660	.01872
SAR(28)	-.32319	.09446	-3.42135	.00082
MEAN	-.10325	.30574	-.33770	.73609
CONSTANT	-.36782			

MODEL FITTED TO SEASONAL DIFFERENCES OF ORDER 1 WITH SEASONAL LENGTH = 7
ESTIMATED WHITE NOISE VARIANCE = 108.233 WITH 141 DEGREES OF FREEDOM.
CHI-SQUARE TEST STATISTIC ON FIRST 20 RESIDUAL AUTOCORRELATIONS = 17.7191

Table 5.02 Time Series Analysis

ARIMA (0,1,4)

Reservations on Boarding Day

M-class   Fight F1   Market A/B

The level of error in each model presented is not related to poor model specification. Instead, they reflect the variability of the parameter we are modelling. Although reservations on boarding day for a flight/class exhibit time related patterns, such as day of week seasonality, they cannot be used alone to describe the random process associated with bookings. Furthermore, given the need to make forecast for a flight departure, say 28 days ahead, the error associated with the forecast will sharply increase as the time interval increases.

These two examples are illustrative of the randomness that is present in reservations data. Rather than showing how to use ARIMA models, they serve to illustrate the difficulty posed to the forecaster in modelling and the seemingly disadvantage of time series models. Moreover, because no structural behavior is associated with any time series model, model specification becomes extremely time consuming. No clear cut approach can be developed in modelling, not in an reasonable fashion, when one uses time series models. Just too many forecaster's interventions are needed. Time series analysis were applied in the remaining markets, and the results obtained were similar.For the above reasons, the use of time series analysis in reservation forecasting becomes unattractive, although with only two model examples a forecaster should never discard a forecasting method.

One should note that in the above presented models, only reservations data on boarding day were used. No other available data, such as reservation made 28 days before departure for the same flight, or reservation for the same class in the same day, or even other flight reservations data, were ever used. That leads to the next step which is the use of Regression Analysis in reservations forecasting.

5.3 REGRESSION ANALYSIS

The use of Regression Analysis in reservations forecasting implicitly leads to the presumption that one knows something about causal relationships that are relevant and influences booking patterns for a given flight. One may assume that there is a relationship between booking levels in a directional market, for instance. Cause-effect relationships can also be tested among different classes in the same flight/market. One could argue that some passengers that made reservation on a full Y class did so because they did not find a seat in the "M compartment". Correlation among classes, among flights and between markets are few examples that one could test in developing a regression model.

Given the need to provide estimates of final seat requests, for a given flight, for a given class, the forecaster can launch himself in model building, knowing that he has at his disposal a large database. The data that is usually available includes :

(1) historical data from past operations of the same flight, for all classes, which can be retrieved from the airline reservation system;

(2) reservation data for the flight itself;

(3) current and past applicable fares;

(4) changes in schedules/airlines in the market;

(5) level of service variables, such as a new flight was introduced in the market;

(6) time related information, e.g. as flight will depart on a Friday morning, Thanksgiving week;

(7) socio-economic variables .

Historical data from past flight operation can be retrieved for all classes, and flight build up patterns can be derived from the database. Data can be retrieved and analyzed in a seven days intervals, for instance. That is to say that for the M class of flight F1, market A/B, reservations data can be retrieved for the boarding day (MBD), 7 (M7), 14 (M14) ,21 (M21) days before flight departure and so forth.

Current and past fares are also available. A regression analysis model with micro/macro-economic variables should increase the explanatory power of regression models. By including such variables, one could infer about consequences of fare changes. It is very unfortunate that, in this thesis, fare data was not so readily available. Some problems were detected in the fare database that made them of little use. A "pollution" problem was particularly detected in the highest fare class, with exception only of the I/J & J/I markets, which makes them unusable as far statistical analysis and demand modelling are concerned.

Because reservations were modeled on a leg basis, not on a O&D basis, the absence of a fare variable does not implicitly represent a loss in the model explanatory power. If O&D forecasts are needed, then fare variables should inevitably be incorporated, as well other socio-economic variables, e.g income level, population ,etc.

Variables type (4) and (5) cannot be explicitly introduced in the model. An example of a methodology for determining the relationship between air transportation demand and the level of service has been proposed by Eriksen, Scalea and Taneja [8]. In essence, a level of service index is created and used in regression analysis models. The level of service index generated is a nondimensional generalized trip time scaled from zero to one, which takes into account not only the number of flight, but also the number of intermediate stops, direct or connecting service, speed of aircraft, and most important, the matching of departure schedules to time variability of demand. By the same token that fares cannot be introduced in the model, (only in O&D models), this index cannot be applied to the regression models developed in this thesis.

Although some variables above mentioned are not explicitly considered in the regression models presented in this thesis, some structural behavior are implicitly assumed here. For instance, fare elasticities are expected to increase as one moves from a potential full fare passenger to the lowest fare potential passenger. Correlations are expected to be higher in adjacent classes in comparison with extreme classes, for instance , or, in other words, reservations on the lowest fare class are not expected to be an important explanatory variable in the regression model of the highest fare class. A coherent fare structure, as well associated set of restriction, are assumed in each and every market analyzed. As the models presented here are the result of a search of a structural behavior among classes and even markets, those implicit assumptions are of very importance.

Under these circumstances, the models that are presented here can be called as booking performance models, in the sense that they search for structural behavior and or performance of booking for a given class, on a given flight. The search of a structural behavior across markets leads to the formulation of a general structure model, which is a model that is expected to hold across flights and markets. The rational in developing this kind of model is that one expects to find a minimal set of explantory variables that is able to describe reservations as a function of these variables. In building the so called general structure model, causal-effect relationships are examined between bookings-on-hand and bookings-to-come, among adjacent fare classes, among flights and classes, and time related variables, such as week-of-year and day-of-week seasonal variables.

One may argue that there is loss in model precision when a general structure is used. It is true that there are losses involved in adopting a general structure model. A model that is market and flight specific is by all means better than a generalized one. Because it is specified and fitted with the data of one flight only, it tends to show better fitting results. Building models that are market/flight specific may be a time consuming task, especially when one considers the large number of flights/markets served by an airline. One may introduce a variable that is relevant for the flight F1 in the A/B market, whereas the same variable may not be of relevance in the flight F1 in the E/F market.

Therefore, building models that are flight/market specific has some practical disadvantages when one considers that a vast set of forecasting models has to to be generated for an airline. When a new flight is introduced in the A/B market, for instance, a general structure model,that was fitted for the A/B market is likely to yield better forecast than the one particularly specified and fitted to any flight in the market. One also has to have in mind that schedule changes happen very often in the airline industry, and the ability of building flight specific models is sharply reduced. It would require a much higher stability in schedules that is observed in the airline industry.

Therefore, if a general structure model can be developed and the incurred precision losses are not relevant, those models are preferred. With a general structure model approach, forecasts of final bookings can be made more efficiently without spending too much time in modeling.

Another positive consequence of a general structure model is the associated reduction in data handling routines that are required for model fitting. Moreover, with a market specific approach, one could run into the problem of having to re-specify the model as time goes by, because the model could also be specific for a determined period/season of the year, or even worse, a model dependent on the data set used for fitting.

The regression models presented in this thesis are the result of a search of a general structure model. Several model specifications and variables were tested for each market, and pooled together in a time consuming regression analysis effort. Variables that showed statistical adherence in most of all markets were kept.

The general structure model developed here, includes a long term cyclic (seasonal variation) component, recent historical data, as well as actual bookings. The minimum cycle is a week, and the model is sensitive to day-of-week variation.

The variables used in the general structure model to forecast bookings to come, for flight **Fi** in M class, for a given market, from t days before departure, (Mt_BD), are as follows:

ONE	-	a constant or base booking level;
DAYS	-	day of week dummy variables, (**MO**,**TU**,**WE**,**TH**,**FR**, and **SA** relative to **SU**);
Mt	-	bookings-on-hand, on day t, M-class;
INDEX	-	week of year non-dimensionalized index for traffic levels and growth through the major hub of the airline;
S5MAiB	-	historical average of bookings made in M-class , between day t and departure, for the most five recent departures of the same flight **Fi** ;
MTt	-	total bookings on hand ,for all future flights in the same directional market t days before departure.

Long-term cycles and market growth are expected to be captured by the **INDEX** variable, while **S5MAiB** should be sensible to shorter term trends.

Regression analysis results are shown from Tables 5.03 through 5.12 .

Table 5.03 shows fitting results of the general structure model applied to market A/B, flight F1. Ten explanatory variables are included in the model. The model also includes a constant term. The model is fitted in a subset of the original data set. The subset used was from observation #35 to observation #181, which means a total of 147 observations. The sample size was reduced to 147 because of the S5MA variable. S5MA is a 5-week lag average, and consequently the first non-trivial S5MA is for observation #35 (35 = 7 . 5).

The degree of freedom for the F-statistic is therefore equals to 10 (the number of explanatory variables) in the numerator, and equals to 136 (number of observations minus the number of parameters to be estimated, or 136=147-10-1),in the denominator. Therefore, the critical value of the F-Statistic (10,136), at 95% level of confidence is 1.91. All runs exhibit a higher F level, which means that the models are accepted.

The adjusted R-squared , or R-bar squared, for all runs is characterized by a low value. One has to remember that the dependent variable is the result of the difference of final bookings and bookings-on-hand. Model fitting for differenced variables will always exhibit low R squared statistics. This explain, to some extent, why R square is relatively low for each model run.

In this example there was a distinct behavior for Mondays, Fridays, Thursdays, and Saturdays. They were statistically different from the base day, which were Sundays. The Bookings-on-hand and INDEX variables were significant in all runs : in any run, the t statistic of coefficients were greater than the critical value t(136)=1.98.

This market/flight is an example of model fitting results that is expected for the general structure model. The INDEX variable is expected to be significant and capture week-of-year seasonality. A "local" seasonality is also expected to be captured by day-of-week dummy variables. It was indeed possible for this flight/market to detect some different behavior for some dummy variables.

TABLE 5.03

REGRESSION ANALYSIS
SUMMARY

MARKET A/B   FLIGHT   F1

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	15.56		14.44		12.01		7.56	
STD. DEV.	8.32		8.24		7.72		6.09	
STD. ERROR OF REGRESSION	6.38		6.34		6.09		5.35	
R SQUARED	0.45		0.45		0.42		0.28	
R-BAR SQUARED	0.41		0.41		0.38		0.23	
F-STATISTIC (10,136)	11.17		11.01		9.79		5.29	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	34.56	6.57	33.87	6.46	31.61	6.21	24.04	5.19
MO	-3.94	-1.98	-3.86	-1.96	-2.85	-1.51	-1.36	-0.87
TU	0.41	0.19	0.45	0.21	1.34	0.66	2.14	1.21
WE	1.6	0.81	1.35	0.68	1.94	1.01	1.37	0.82
TH	3.12	1.34	3.31	1.42	3.39	1.52	2.28	1.18
FR	5.91	2.93	5.32	2.65	4.54	2.32	1.36	0.75
SA	-5.11	-2.52	-5.14	-2.55	-3.81	-1.94	-20.9	-1.21
Mt	-0.51	-4.23	-0.49	-4.13	-0.42	-3.95	-0.29	-3.18
INDEX	-0.11	-2.23	-0.09	-2.11	-0.11	-2.31	-0.08	-2.18
S5MAt	-0.32	-2.15	-0.32	-2.16	-0.23	-1.62	-0.13	-1.04
MTt	-0.15	-1.97	-0.15	-1.94	-0.16	-2.25	-0.09	-1.57

Table 5.04 shows fitting results for flight F2, in the B/A market. This is an example of a flight in the reverse direction of the one shown in the previous example.

In this example, F-statistics were acceptable for all model runs. R-bar squared although almost constant was also low across runs. The day-of-week variables exhibited an interesting behavior. While only Fridays and Saturdays variables were statistically significant in the day_28 run, all days but Mondays were significant in the day_7 run. The bookings-on-hand variable did not behave consistently for all runs. It showed an almost acceptable t_statistics in day_28 and day_7 runs while in between the t values were not acceptable. The INDEX variable exhibited a very stable behavior, although not all t-statistics were acceptable.

This market/flight is an example of a mixed statistical behavior for the general structure model. Although some coefficients of the variables could not be statistically determined, the overall model performance was acceptable. When compared to Table 5.03, one can observe that relatively similar results were obtained for flights in both direction of the markets defined by the citypair A&B.

TABLE 5.04

REGRESSION ANALYSIS
SUMMARY

MARKET B/A FLIGHT F2

MODEL RUN (DAY)	:	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	:	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	:	20.46		18.77		16.61		1021	
STD. DEV.	:	10.45		10.22		9.78		8.01	
STD. ERROR OF REGRESSION	:	9.03		8.59		8.21		6.81	
R SQUARED	:	0.31		0.34		0.34		0.32	
R-BAR SQUARED	:	0.25		0.29		0.29		0.27	
F-STATISTIC (10,136)	:	5.86		6.99		7.02		6.55	
VARIABLES	:	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	:	24.31	2.69	26.83	3.11	2659	3.24	22.31	3.29
MO	:	1.71	0.58	1.37	0.49	1.16	0.44	0.05	0.02
TU	:	-2.15	-0.75	-2.49	-0.91	-2.97	-1.14	-3.66	-1.71
WE	:	-1.24	-0.43	-2.09	-0.76	-4.32	-1.61	-7.01	-3.15
TH	:	-7.53	-2.35	-9.59	-3.11	-9.37	-3.21	-8.93	-3.73
FR	:	-2.01	0.66	-3.43	-1.19	-4.37	-1.59	-7.31	-3.17
SA	:	-5.69	-1.85	-5.02	-1.71	-4.26	1.51	-4.78	-2.04
Mt	:	-0.42	-1.41	-0.12	-0.56	0.006	0.03	-0.16	-1.49
INDEX	:	-0.07	-0.91	-0.11	-1.37	-0.11	-1.52	-0.08	-1.38
SSMAt	:	0.38	3.01	0.37	3.07	0.35	3.06	0.19	4294
MTt	:	-0.08	-0.68	-0.11	-1.21	-0.16	-1.88	-0.02	-0.33

Table 5.05 shows model fitting results for flight F4 in the C/D market. R-bar statistics are little higher for this flight. In this example all day-of-week variables were different from the base day, Sundays. Thursday was constantly the busiest day in any week, in any model run. The bookings-on-hand variable was not significant in any model run. The INDEX and S5MA variables were constantly significant across model runs.

The general structure model behave as expected. Fitting improvement was observed, when compared to previous flights/markets.

Table 5.06 shows model fitting results for flight F3 in the D/C market. R-bar squared statistics were improved. All day-of-week variables were significant, in any model run. Bookings-on-hand were not significant again. The INDEX and S5MA variables were siginificant in most of the runs.

The overall model behavior was similar to the C/D market example. That is, all dummy variables were significant in the C&D citypair markets. The INDEX and S5MA variables did contribute to the general structure model. The remaining variables did not explicitly improve model statistics.

TABLE 5.05

REGRESSION ANALYSIS
SUMMARY

MARKET C/D FLIGHT F4

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	15.93		15.11		13.57		9.81	
STD. DEV.	9.25		8.84		8.42		6.79	
STD. ERROR OF REGRESSION	6.82		6.63		6.21		5.38	
R SQUARED	0.49		0.47		0.49		0.41	
R-BAR SQUARED	0.45		0.43		0.45		0.37	
F-STATISTIC (10,136)	13.12		12.28		13.2		9.6	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	-18.83	-2.95	-18.99	-3.09	-17.68	-3.04	-18.18	-3.64
MO	5.49	2.33	6.34	2.74	4.81	2.24	3.78	1.98
TU	6.63	2.67	6.48	2.69	5.08	2.23	4.98	2.53
WE	9.38	3.44	9.39	3.56	6.91	2.74	5.42	2.45
TH	11.68	3.94	11.04	3.82	9.42	3.45	7.21	2.97
FR	7.92	2.81	8.11	2.93	6.03	2.33	4.87	2.08
SA	6.81	2.62	6.78	2.61	4.51	1.85	3.35	1.49
Mt	-0.04	-0.09	0.11	0.38	0.09	0.45	0.03	0.22
INDEX	0.21	3.05	0.21	3.35	0.17	2.93	0.18	3.69
SSMAt	0.34	3.12	0.31	2.91	0.31	3.01	0.19	2.23
MTt	0.25	1.91	0.08	1.01	0.12	1.91	0.04	0.87

TABLE 5.06

REGRESSION ANALYSIS
SUMMARY

MARKET D/C FLIGHT F3

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	31.78		29.44		25.33		16.63	
STD. DEV.	15.66		15.24		13.71		10.33	
STD. ERROR OF REGRESSION	10.04		10.08		9.67		7.88	
R SQUARED	0.61		0.59		0.53		0.45	
R-BAR SQUARED	0.58		0.56		0.51		0.42	
F-STATISTIC (10,136)	21.91		19.76		15.73		11.48	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	-5.21	-0.56	-4.47	-0.46	-3.51	-0.38	-3.53	-0.47
MO	2.08	0.52	2.88	0.87	4.38	1.35	5.84	2.24
TU	8.15	2.24	9.13	2.47	9.95	2.72	9.67	3.27
WE	14.19	3.46	13.91	3.34	14.11	3.41	11.51	3.46
TH	11.84	3.02	11.91	3.02	10.94	2.84	6.21	1.94
FR	10.31	2.84	10.41	2.89	6.49	1.82	0.39	0.13
SA	-7.64	-2.17	-7.66	-2.14	-7.13	-2.06	-5.73	-1.97
Mt	0.23	0.53	-0.18	-0.62	-0.27	-1.16	-0.11	-0.83
INDEX	0.19	2.09	0.17	1.89	0.14	1.64	0.11	1.63
SSMAt	0.41	3.53	0.38	3.31	0.28	2.44	0.11	1.25
MTt	-0.29	-1.27	-0.12	-0.66	0.07	0.53	0.51	0.73

Table 5.07 shows model fitting results for flight F1 in the E/F market. R-bar squared statistics are very low. They range from 0.19 to 0.16 . Nevertheless, all F_statistics were acceptable. The critical value is F(10,136)=1.91, at a 95% confidence interval. As it gets closer to the departure day the more significant the day-of-week dummy variables are. While the INDEX and S5MA variables were not significant to the model runs, the Mt and MTt variables exhibited acceptable t-statistics. It is an example of the opposite behavior so far observed.

Table 5.08 shows model fitting results for flight F2 in the F/E market. R-bar squared statistics are a little higher than in the previous example. All F_statistics were also acceptable. Statistical significance was marginally observed for day-of-week variables. The Mt variable, bookings-on-hand was not significant in any model run. The INDEX and S5MA were significant in this example. Total bookings-on-hand (MTt) was not so significant as in the previous example.

These two examples illustrate the distinct behavior expected for the (INDEX & S5MA) variables vs. (Mt & MTt) variables. When the first two are significant, the others are not, and vice-versa. It explains why both sets are included in the model, together with the fact that no one can a priori predict which two will be significant.

TABLE 5.07

REGRESSION ANALYSIS
SUMMARY

MARKET E/F FLIGHT F1

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	10.61		9.71		8.08		6.63	
STD. DEV.	10.46		9.07		7.12		5.67	
STD. ERROR OF REGRESSION	9.4		8.45		6.81		5.21	
R SQUARED	0.24		0.19		0.15		0.21	
R-BAR SQUARED	0.19		0.13		0.09		0.16	
F-STATISTIC (10,136)	4.45		3.21		2.41		3.73	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	-5.58	-1.01	-2.11	-0.42	1.71	0.42	4.93	1.57
MO	1.41	0.46	2.73	1.01	3.26	1.48	3.76	2.23
TU	-0.85	-0.29	1.67	0.62	3.45	1.61	4.66	2.81
WE	1.35	0.46	1.65	0.62	2.42	1.15	3.88	2.39
TH	4.54	1.55	5.98	2.26	6.27	2.95	7.06	4.31
FR	2.79	0.91	1.99	0.72	3.01	1.36	2.61	1.52
SA	4.05	1.34	2.51	0.92	1.89	0.87	1.68	1.01
Mt	-0.67	-4.35	-0.51	-3.79	-0.21	-2.10	-0.14	-2.06
INDEX	0.13	2.74	0.07	1.59	0.01	0.22	-0.04	-1.36
S5MAt	0.13	0.97	0.15	1.27	0.16	1.67	0.07	0.99
MTt	0.39	2.92	0.35	3.32	0.16	2.22	0.14	2.77

TABLE 5.08

REGRESSION ANALYSIS
SUMMARY

MARKET F/E FLIGHT F2

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	13.81		12.68		11.81		8.96	
STD. DEV.	9.45		9.02		8.44		6.81	
STD. ERROR OF REGRESSION	7.83		7.94		7.64		6.48	
R SQUARED	0.35		0.28		0.23		0.15	
R-BAR SQUARED	0.31		0.22		0.17		0.19	
F-STATISTIC (10,135)	7.57		5.17		4.16		2.48	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	-13.58	-2.55	-11.93	-2.184	-8.24	-1.54	-3.52	-0.77
MO	2.51	1.01	3.28	1.31	3.91	1.66	3.01	1.44
TU	4.61	1.83	4.06	1.56	4.31	1.68	4.24	1.89
WE	3.18	1.29	3.36	1.33	4.07	1.65	3.56	1.64
TH	3.55	1.42	3.13	1.21	3.51	1.38	1.79	0.81
FR	1.71	0.68	0.48	0.18	2.11	0.82	2.41	1.11
SA	3.57	1.31	3.06	1.08	3.82	1.39	4.87	2.13
Mt	0.04	0.26	0.02	-0.13	-0.04	-0.29	-0.14	-1.54
INDEX	0.16	3.08	0.15	2.81	0.11	2.01	0.05	1.25
S5MAt	0.61	5.43	0.53	4.74	0.46	4.17	0.29	3.11
MTt	0.08	1.11	0.04	0.54	0.06	0.89	0.07	1.42

Table 5.09 shows model fitting results for flight F1 in the G/H market. R-bar squared statistics are again low. F_statistics are acceptable, and in the day_7 run, it reaches the minimum so far observed, 1.92 . In this example, only the INDEX variable is statistically significant. One possible reason is the high variation observed for the dependent variable. While means are extremely low, ranging from 6.65 to 3.97, standard deviations are relatively high, ranging from 6.7 to 5.3, respectively. As a consequence, model performance is reduced. Nevertheless, the standard error of the regression was always smaller than the standard deviation of the dependent variable.

Table 5.10 shows model fitting results for flight F2, in the H/G market. Although R-bar squared are extremely low, one can observe the distinct model behavior. In this example, all day-of-week variables were significant. The INDEX and S5MA variables were also significant, while MTt and Mt showed bad t_statistics. Again, only two variables were significant.

The flight F1, market H/G, example illustrates that even for a flight with very small load the general structure model can be used.

TABLE 5.09

REGRESSION ANALYSIS
SUMMARY

MARKET G/H FLIGHT F1

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	6.04		5.46		4.88		3.97	
STD. DEV.	6.65		6.41		6.12		5.25	
STD. ERROR OF REGRESSION	5.56		5.47		5.57		5.07	
R SQUARED	0.35		0.32		0.23		0.14	
R-BAR SQUARED	0.31		0.27		0.17		0.06	
F-STATISTIC (10,136)	6.37		5.68		3.58		1.92	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	-5.35	-1.45	-4.22	-1.16	-4.88	-1.32	-3.02	-0.89
MO	-1.61	-0.91	-1.39	-0.81	-1.41	-0.79	-1.68	-1.14
TU	-0.54	-0.31	-1.26	-0.71	-1.06	-0.59	-1.38	-0.85
WE	-2.05	-1.16	-2.61	-1.49	-2.65	-1.49	-2.27	-1.41
TH	-0.81	-0.45	-2.05	-1.17	-1.53	-0.86	-1.78	-1.11
FR	4.31	2.45	3.54	2.03	3.31	1.83	0.96	0.57
SA	-1.31	-0.53	-0.43	-0.17	-1.31	-0.52	-2.03	-0.91
Mt	-0.18	-0.83	0.02	0.11	-0.24	-1.17	-0.11	-0.79
INDEX	0.12	4.35	0.11	4.13	0.11	3.88	0.08	3.13
SSMAt	0.13	1.48	0.13	1.61	0.09	1.01	0.02	0.21
MTt	-0.51	-3.44	-0.61	-3.54	-0.23	-1.67	-0.07	-0.74

TABLE 5.10

REGRESSION ANALYSIS
SUMMARY

MARKET H/G FLIGHT F2

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	M28_BD		M21_BD		M14_BD		M7_BD	
MEAN	12.51		11.26		9.88		7.13	
STD. DEV.	11.58		11.21		9.77		7.45	
STD. ERROR OF REGRESSION	10.38		10.44		9.14		7.02	
R SQUARED	0.25		0.19		0.18		0.17	
R-BAR SQUARED	0.19		0.13		0.12		0.11	
F-STATISTIC (10,132)	4.46		3.18		3.01		2.81	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	5.69	1.42	6.14	1.52	4.22	1.18	3.45	1.27
MO	-12.27	-3.21	-11.61	-3.01	-9.23	-2.71	-7.03	-2.67
TU	-13.01	-3.38	-12.85	-3.31	-10.81	-3.16	-7.59	-2.92
WE	-14.72	-3.85	-15.14	-3.95	-12.01	-3.56	-7293	-3.04
TH	-12.92	-3.33	-12.01	-3.08	-9.55	-2.78	-6.52	-2.45
FR	-13.64	-3.56	-12.46	-3.21	-10.97	-3.23	-8.58	-3.27
SA	-11.01	-2.88	-10.77	-2.81	-10.61	-3.15	-8.87	-3.42
Mt	-0.14	-0.71	-0.26	-1.33	-0.21	-1.14	-0.07	-0.64
INDEX	0.09	2.81	0.09	2.64	0.08	2.51	0.06	2.61
SSMAt	0.59	4.56	0.45	3.37	0.35	2.95	0.19	2.11
MTt	0.15	0.89	0.22	1.32	0.23	1.51	0.14	1.44

Table 5.11 shows model fitting results for flight F1, in the I/J market, for the Y-class. The same general structure is applied to the Y-class in the Canadian market. R-bar squared statistics are low as it were in the case of the M-class, for domestic U.S. markets. In the day_7 run, all model statistics dropped significantly, and the model exhibited a distinct behavior: only two variables were significant. Nevertheless, F_statistics were acceptable for all runs. The day-of-week dummy variables exhibited the expected behavior, that is for some days (e.g. TU or SA) different behavior from the base day was observed, i.e. t_statistics were significant. Bookings-on-hand (Yt) were significant for all model runs, but day_7 run. The INDEX variable was marginally accepted in some runs, while the remaining variables were not.

Table 5.12 shows model fitting results for flight F1 in the J/I market, for the Y-class. R-bar squared statistics were a little higher that in the previous example. In the day_28 model run it was 0.41 and it dropped to 0.16 on the day 7. The day-of-week variables did not exhibit good t_statistics. No "local" seasonality could be picked up by the model. The Yt variable was significant in the first two model runs, while YTt was not. The INDEX variable was significant in all model runs. S5MA was not significant.

TABLE 5.11

REGRESSION ANALYSIS
SUMMARY

MARKET I/J FLIGHT F1

MODEL RUN (DAY)	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	Y28_BD		Y21_BD		Y14_BD		Y7_BD	
MEAN	22.98		21.36		18.48		13.11	
STD. DEV.	13.36		12.21		10.48		8.19	
STD. ERROR OF REGRESSION	11.3		10.77		9.76		7.98	
R SQUARED	0.33		0.27		0.18		0.11	
R-BAR SQUARED	0.28		0.22		0.12		0.04	
F-STATISTIC (10,136)	6.81		5.13		3.18		2.81	
VARIABLES	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	48.22	5.29	45.5	5.15	35.56	4.29	25.51	3.62
MO	-5.24	-1.28	-3.34	-0.86	-0.41	-0.11	1.22	0.41
TU	-10.99	-2.98	-8.91	-2.52	-5.22	-1.59	-0.56	-0.21
WE	-3.93	-1.01	-1.97	-0.53	1.19	0.34	2.01	0.71
TH	-4.64	-1.29	-3.87	-1.12	-1.89	-0.61	0.37	0.14
FR	5.91	1.65	6.01	1.75	6.49	2.09	3.79	1.45
SA	-6.29	-1.79	-6.36	-1.91	-5.49	-1.81	-3.42	-1.35
Yt	-0.72	-5.69	-0.51	-3.96	-0.27	-2.21	-0.01	-0.02
INDEX	-0.11	-1.64	-0.11	-1.49	-0.11	-1.86	-0.08	-1.89
S5MAt	-0.15	-1.25	-0.15	-1.29	-0.09	-0.84	-0.05	-0.64
MTt	0.03	0.46	0.01	0.21	0.02	0.37	-0.05	-1.01

TABLE 5.12

REGRESSION ANALYSIS
SUMMARY

MARKET J/I FLIGHT F1

MODEL RUN (DAY)	:	t=28		t=21		t=14		t=7	
DEPENDENT VARIABLE	:	Y28_BD		Y21_BD		Y14_BD		Y7_BD	
MEAN	:	20.98		15.42		17.01		11.36	
STD. DEV.	:	13.87		10.34		8.29		6.62	
STD. ERROR OF REGRESSION	:	10.74		9.53		7.81		6.04	
R SQUARED	:	0.44		0.21		0.17		0.22	
R-BAR SQUARED	:	0.41		0.15		0.11		0.16	
F-STATISTIC (10,134)	:	10.61		3.54		2.81		3.88	
VARIABLES	:	value	t stat	value	t stat	value	t stat	value	t stat
CONSTANT	:	40.03	6.84	31.89	5.92	23.67	5.16	18.41	5.04
MO	:	-3.59	-1.07	-3.37	-1.13	-2.13	-0.87	0.29	0.15
TU	:	-4.79	-1.39	-0.81	-0.25	2.34	0.91	2.47	1.24
WE	:	-0.84	-0.25	2.82	0.91	5.96	2.32	6.92	3.45
TH	:	-1.99	-0.58	0.99	0.32	3.21	1.26	3.17	1.61
FR	:	1.43	0.42	3.42	1.13	6.15	2.44	5.27	2.73
SA	:	-0.41	-0.12	0.01	0.01	2.37	0.94	0.71	0.35
Yt	:	-0.89	-6.22	-0.47	-3.52	-0.11	-0.94	0.06	0.81
INDEX	:	-0.15	-3.27	-0.14	-3.51	-0.11	-2.91	-0.06	-2.11
S5MAt	:	0.24	1.82	0.18	1.55	0.05	0.48	-0.06	-0.81
YTt	:	0.07	0.92	0.11	1.58	0.03	0.59	-0.08	-1.81

The statistical adherence of the variables considered in the generalized model vary from market to market. For instance, the bookings-on-hand variable was statistically significant in the directional A/B market, but in the other direction it was not. In the majority of markets, the INDEX and S5MA variables contributed to the model. MTt and Mt variables seem to pick up explanatory power when INDEX and S5MA variables can not. Therefore, they work together in an almost exclusive basis. Nevertheless, all four are kept in the general structure model because one can never predict what variables will be significant. In general, the model was able to pick up day-of-week seasonality: day-of-week dummy variables were significant in most of the cases. A reduction in the number of explanatory variables included in the set of variables selected for the general structure model causes noticeable reduction of Durbin-Watson statistics to values that are not acceptable, which in turn means the presence of serial correlation. This result leads to the conclusion that either variables can only be added to this minimal set, or carefully replaced.

In all markets, the general structure model outperformed simple estimates of bookings-to-come based on local historical averages. The application of the model in a Y-class, as in the case of the Canadian market, yielded equivalent fitting results, which may suggest that the model can be adapted and applied to the Y-class.

CHAPTER SIX

CONCLUSION

6.1 SUMMARY

The forecasting module of an Automated Seat Inventory Control System is intended to provide the dynamic booking limit adjustment routine with estimates of expected bookings for individual future flights. A seat allocation routine will then use these estimates of expected bookings to calculate how many seats should be protected for each upper fare class, in addition to bookings already on hand.

The initial work in forecasting involved models that derived direct estimates of final bookings. Bookings by fare class, on the day of departure was the variable we wanted to forecast - the dependent variable. Cause-effect relationships between this variable and a set of explanatory (independent variable) as well qualitative and quantitative time series behavior.

Further examination of these cause-effect relationships, together with statistical analysis of historical reservations data, have indicated that a focus on bookings-to-come would be a better approach. The key factor to this conclusion was an observed correlation between bookings on hand and final bookings, obtained across the data set selected for hypothesis testing. With the focus shifted to bookings-to-come, final bookings were indirectly estimated as the sum of actual bookings on hand and the estimated bookings to come.

A simple forecasting model is suggested for the initial estimation of final bookings. It consists of moving average process that is sensitive to day of week variation only. That is to say, for instance, that a 8-week average is used to describe or estimate final bookings for a given flight (e.g. flight F1), on a specific day of week (e.g. Monday). Although no information on actual bookings on hand for future flights are ever used, nor additional adjustments are made for cyclic or seasonal variations other that on weekly basis, the implicit assumption of this simple approach is that a small sample of final demand for recent flights will be representative of the demand for the same flights in the near future.

The first step in reservations forecasting involves initial estimates of final bookings well in advance of flight departures. These estimates can be improved later in the bookings process as more information (data) on the specific flight for which more accurate forecasts are needed.

The final step in the development of a forecasting module is to improve the estimates of bookings to come, over those strictly based on recent historical averages. The closer it gets to departure day, the better is to improve forecasts of final bookings. As a general rule, better forecasts of final bookings, via bookings-to-come, are obtained using regression analysis as the 28 days before departure day threshold is surpassed.

The models tested in this thesis ranged from analytical models (Time Series Analysis and Regression Analysis) to non conventional models (Bookings Curves and ad-hoc methods). Results obtained via Time Series Analysis (Box and Jenkins' ARIMA models) were not encouraging enough in providing better estimates, when compared to results obtained via Regression Analysis or even simple historical averages. Any improvements were far outweighed by elaborated data handling routines that would have to be used to fit the models. Non-conventional methods required too many "tuning" interventions by the forecaster, which is not helpful if an automated routine is to be developed, and again their results did not improve over Regression Analysis. As a result, effort was concentrated on Regression Analysis.

The first step was to develop market specific models. Models were formulated and generated for the markets selected in the data sample. The search was for a specific structure (model specification) that yielded better and better model fitting results. At this level, model structure was considered as independent of direction. That is to say, for instance, that the same model structure for the A/B market, should also hold for the B/A market, although estimated coefficients were allowed to be directionally sensitive.

Although it was possible to develop models that were specific to markets, a general structure model was thought to be preferable in view of the associated reduction in specific data handling routines that would be required for model fitting. As model generalization involves losses in forecast precision caused by aggregation of markets, this approach was preferred because these losses were not large enough to distort forecasting results. All forecasts produced with the general structure approach were consistently better (less variable) than simple historical average from a sample of recent flights.

A proposal of a general structure model is then presented and tested in this thesis. The general structure model proposed here includes a long term cyclic (seasonal) component , short term cyclic and trend components calculated over recent historical data, and the model is sensitive to day of week variations. Apart from showing "good" model fitting statistics, this model structure demonstrates how Regression Analysis can be used in a forecasting module of an Automated Booking Limit System, and thus provide improved estimates of bookings to come.

6.2 TOPICS FOR FURTHER RESEARCH

The general structure model was developed in this thesis for the M-class only.

As the typical airline has, at least, four different classes, models need also to be developed and tested for the remaining classes. For the upper fare class, in this case Y-class, a similar model structure can be applied.

The expected set of build-up curves for the Y class is rather similar to the set of M-class, and generally speaking, flights start to heavily build up during the last week, before flight departure. Therefore, a bookings-to-come approach, with the reference on the boarding day can also be used. Almost no "supply limitation" is also expected to occur, since Y authorized booking levels are usually greater than the total coach seating capacity.

As one moves to lower fare classes, both build up behavior (build-up curves) and supply limitations change. For the B-class, for instance, the build-up curve reaches its peak at least a week before the flight departure, say on day 14, and from there on a period of cancellation is expected to be observed. The same phenomena is also observed in the lowest fare class, although, as far as seat inventory control routines are concerned, forecasting models will not be developed. This build-up phenomena can be taken into account by changing the reference day to day 14, as in the case of above mentioned example for the B-class.

On the other hand, the supply limitation problem, caused by bookout phenomena, (the class was closed due to either a low authorized booking limit, or by a large number of reservations made in other classes), needs to be carefully addressed.

A single equation regression model can no longer be applied. Instead, a multi-equation regression model needs to developed, using a simultaneous equation system approach. Now, supply variables, such as authorized levels for a given class need to be explicitly taken into account. For instance, cause-effect relationships such as, for a given fight, for a given class, there was a cutoff in the flight build up because the authorized level for the class itself was reached (low authorized booking limit ---> low demand observed ---> change in flight statistics) should be investigated.

BIBLIOGRAPHY

[1] Belobaba, Peter : "Air Travel Demand and Airline Seat Inventory Control", forthcoming Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[2] Littlewood, Kenneth : Forecasting and Control of Passenger Bookings". AGIFORS Proceedings 1972.

[3] Mayer, Michael : "Seat Allocation, or a Simple Model of Seat Allocation via Sophisticated Ones." AGIFORS Proceedings 1976.

[4] Buhr, Jorn : " Optimal Sales Limit for 2-Sector Flights". AGIFORS 1982.

[5] Belobaba , Peter : "Development of Airline Yield Management Models - Past Work and Future Directions. Unpublished paper. Flight Transportation Laboratory. Massachusetts Institute of Technology, 1985.

[6] Littlewood, Kenneth : Opus cited, page 101.

[7] Taneja, Nawal : "Airline Traffic Forecasting". Lexington Books. Lexington, MA.

[8] Box, George and Jenkins, Gwilym : "Time Series Analysis -Forecasting and Control". Holden-Day. Oakland, CA.

[9] Pindick, Robert and Rubinfeld, Daniel : "Econometric Models & Economic Forecasts". McGraw Hill. 1981.

[10]Wells, Alexander : "Air Transportation : A Management Perspective" . Wadsworth Publishing Co.Belmont,CA. 1984.

[11]Eriksen, S. : "Demand Models for U.S. Domestic Air Passenger Markets". FTL Report R78-2. Massachusetts Institute of Technology, Cambridge, MA. 1978.

[12]Eriksen, S. & Liu, E. : "Effect of Fare and Travel Time on the Demand for Domestic Air Transportation". FTL Report R79-2. Massachusetts Institute of Technology, Cambridge, MA. 1979.