NASA/CR-1999-208987

A Method for Forecasting Commercial Air Traffic Schedule in the Future

Dou Long, David Lee, Eric Gaier, Jesse Johnson, and Peter Kostiuk
Logistics Management Institute, McLean, Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NAS2-14361

January 1999

Available from:

NASA Center for AeroSpace Information (CASI)
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 605-6000

Contents

FIGURES

TABLES

Chapter 1
Introduction

BACKGROUND

This report describes a series of models that will predict the air carriers' operations in the future as demand for air travel increases more rapidly than the ability of the National Airspace System (NAS) to accommodate flights in today's patterns. The work advances the discussion of NAS capacity beyond the level of recent dire forecasts that we are going to see massive flight delays in the United States in about a decade [1,2,3,8].

Those huge delays will not actually happen. They *would* occur if air carriers attempted to meet all future demand by simply increasing the number of flights while maintaining the same scheduling practices and other operating methods in use now. The airlines will certainly change their operating practices long before massive delays develop [2].

This report is concerned with predicting what those changes might be, and their likely economic implications. It is about the carriers' strategic response—not their tactical one—to operation disruptions caused by capacity-related flight delays.

In the *1998 Current Market Outlook*, Boeing recognizes the capacity limitations on the air traffic growth and says that the carriers will find ways to increase the traffic despite those limitations [4]. Specifically, Boeing suggests that the airlines take the following actions to avoid traffic congestion:

- Avoid congested hubs and gateways
- Use secondary airport metropolitan areas
- Move flights to off-peak time
- Broaden the range of departure times
- Shift short-haul flights to long-haul flights.
- Increase airplane size, especially for the short-haul flights.

While Boeing's list gives the airlines ways to grow the market under the capacity limitation, we also expect results from the severe congestion predicted. Those results are as follows:

- Increased flight block times or more padding in flight schedules
- Creation of new hub airports
- More slot-controlled airports.

This report intends to address how the increased air traffic and limited NAS capacity will modify the measures listed by Boeing and ourselves. All the listed measures are related to airlines' service schedules, which are described by the origin and destination airports, departure and arrival times, aircraft types, and operators. Airport operations are defined as the total number of departures and arrivals at an airport in a given time period, e.g., an hour. Given the flight schedules of all airlines, one can get the operations for any airport at any time by simply aggregating the appropriate flights.

A completed flight from the origin gate to the destination gate requires many air traffic control (ATC) services, each of which may cause delays. Delays may be caused by shortages in taxi capacity on the ground; runway capacities, holding aircraft either on the ground or in the air; terminal radar approach control (TRACON) capacity in the air near the airport; or en route ATC capacities. Since the overwhelming majority of flight delays in the United States will be caused by the airports, either from runway or taxiway shortages, here we model only the impacts of limited airport capacities on the air carriers' operations [1,2,8].

ORGANIZATION

In the next chapter, we present some empirical evidence that airlines have stretched their flight block times to counterbalance increased delays, and we present our statistical models relating block time and airport operation.

In Chapter 3, we construct the unconstrained future air traffic schedule based on the airport operations forecast. Chapter 4 is devoted to the modeling of constrained air traffic demand forecast, and we outline our general modeling approach, delineate the inputs and components of our model, and present the forecast of air traffic schedule in 2007.

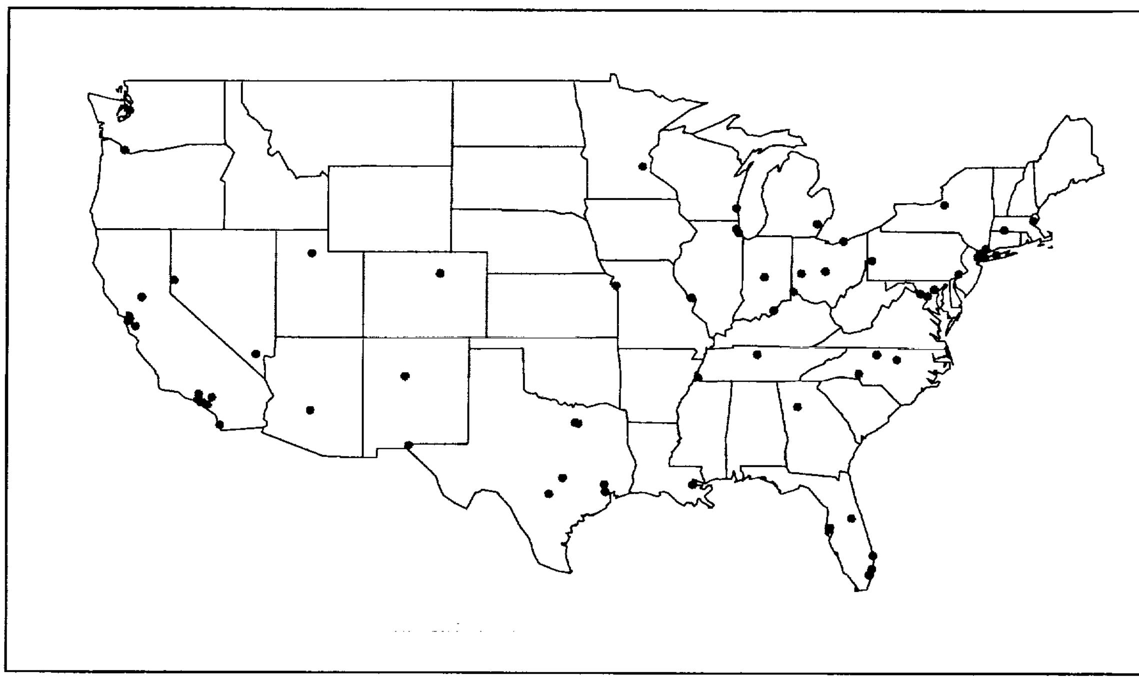
In Chapter 5, we model the high-level economic effects of changing air carrier operations, using the Air Carrier Investment Model (ACIM) of the Aviation System Analysis Capability (ASAC) model suite, based on the delay figures produced by LMINET, a queueing network model of the NAS. Appendix A provides a detailed description of the newest development in LMINET to compute flight delays. Appendix B gives analytic details for Chapter 5.

Part of the material in Chapters 3 and 5, although developed for this task, has been presented in another report sponsored by NASA [2]. It is included for the sake of completeness.

Chapter 2
Modeling Block Times

Throughout this report, we will often refer to the Federal Aviation Administration (FAA) large hub airports and the LMINET airports. The FAA classifies an airport as a large hub if it has more than 1 percent of domestic enplanements. Currently, there are 29 large hub airports,[1] which account for 68.1 percent of total domestic enplanements [5] There are 64 LMINET airports,[2] which make up a superset of the FAA's 57 pacing airports (a set of airports that the FAA has used to study flight operations in the NAS). The 64 LMINET airports contribute 84.9 percent of total domestic enplanements and about 85 percent of total domestic operations as reported by the Department of Transportation's (DOT's) T-100 data. The locations of the LMINET airports are depicted in Figure 2-1.

Figure 2-1. LMINET Airports



[1] The 29 FAA large hub airports are ORD, ATL, LAX, DFW, SFO, MIA, DEN, JFK, DTW, PHX, LAS, EWR, STL, MSP, BOS, IAH, MCO, SEA, HNL, CLT, LGA, PIT, SLC, PHL, CVG, DCA, SAN, BWI, and TPA.

[2] The 64 LMINET airports are ABQ, ATL, AUS, BDL, BNA, BOS, BUR, BWI, CLE, CLT, CMH, CVG, DAL, DAY, DCA, DEN, DFW, DTW, ELP, EWR, FLL, GSO, HOU, HPN, IAD, IAH, IND, ISP, JFK, LAS, LAX, LGA, LGB, MCI, MCO, MDW, MEM, MIA, MKE, MSP, MSY, OAK, ONT, ORD, PBI, PDX, PHL, PHX, PIT, RDU, RNO, SAN, SAT, SDF, SEA, SFO, SJC, SLC, SMF, SNA, STL, SYR, TEB, and TPA.

Since most of the flight delays happen at these busy airports, whether represented by the FAA's 29 or LMINET's 64 airports, restricting our study to them will filter out unhelpful noise in the results due to operations at the smaller airports and will not materially affect our conclusions.

REPORTED FLIGHT DELAYS

The Airlines Service Quality Performance (ASQP) data published by DOT is the information source used in this chapter. Any airline with more than 1 percent of domestic enplanements is required to participate in ASQP, which accounts for about 90 percent of total domestic operations.[3]

In the ASQP database, the scheduled departure and arrival times as well as the actual departure and arrival times of a flight are recorded. Since January 1995, the scheduled wheels-off and wheels-on times as well as the actual wheels-off and wheels-on times are also recorded. The scheduled times are based on the Official Airline Guide (OAG), when available; otherwise they are based on the times shown in the Computer Reservation System (CRS). ASQP keeps the information for every flight operated by the reporting airlines (if not canceled or delayed for mechanical reasons). ASAC had ASQP data for 1993 and 1995, and has expanded to include data for 1996 and 1997 since this study began.

The block time of a flight, either scheduled or reported, is the gate-to-gate time from departure to arrival. The flight delay time is the difference between the scheduled and actual block times, which, in contrast with the actual delay, does not include delays in departing from the gate. Flight delay, therefore, captures the delays due to the limited ATC or airport capacities. A flight delay is reported as zero if the actual block time is smaller than the scheduled block time.

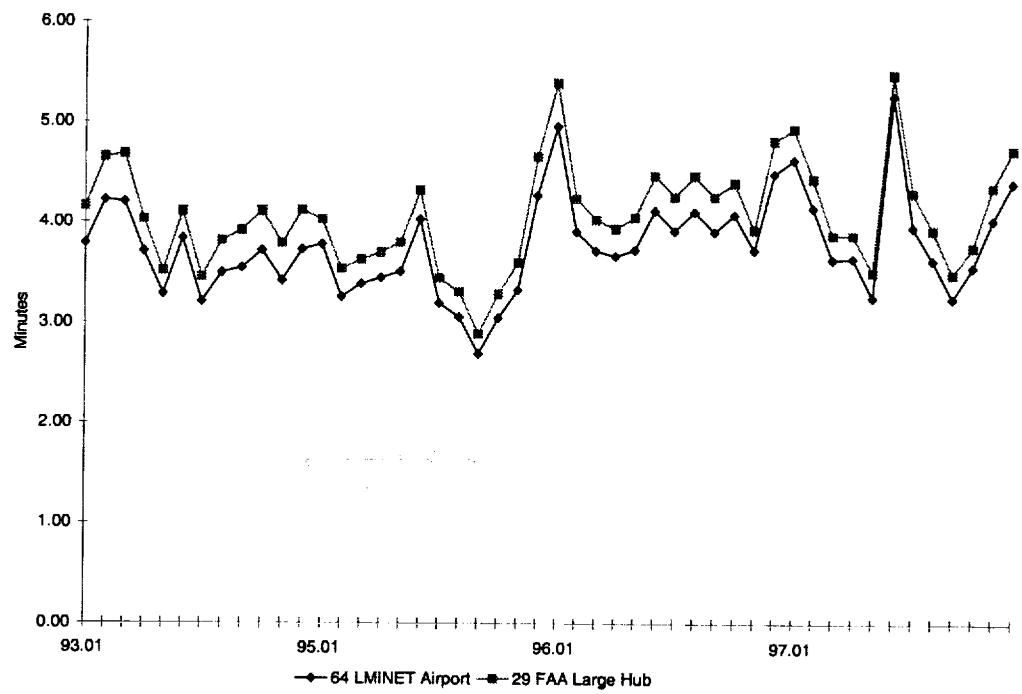
BLOCK TIME STRETCH

Figure 2-2 shows the monthly average flight delays for the 64 LMINET airports as well as the 29 FAA large hub airports.

One observation from this chart is that there are larger average delays in the 29 FAA large hub airports than in the 64 LMINET airports. This should not surprise us since the hub airports are generally busier than others, and they generally have more intense "banks" of arrivals and departures.

[3] Currently, 10 airlines report in ASQP. They are United Airlines, American Airlines, Delta Airlines, Northwest Airlines, Continental Airlines, US Airways, TWA, Southwest Airlines, America West Airlines, and Alaska Airlines.

Figure 2-2. Monthly Average Flight Delays



Note: The abscissas have the form YY.MM, where YY denotes the year and MM, the month. Thus, 93.02 indicates February 1993, 95.06 indicates June 1995, and so on.

Figures 2-3 and 2-4 depict the monthly average flight delays for the Atlanta Hartsfield International Airport (ATL) and the Dallas/Fort Worth International Airport (DFW), respectively, by each scheduled arrival hour. For example, the curves under the symbol associated with _7 are for the monthly average flight delays if the scheduled arrival time is between 7:00 and 8:00 a.m. local time. Although we find delay patterns across a day—e.g., delays at certain times are always higher than those at other times—trends for each individual delay series are much weaker.

The reported flight delays stayed essentially constant from 1993 to 1997, as evidenced by Figures 2-2, 2-3, and 2-4, while traffic steadily increased. It is a familiar result of queuing theory that the increased traffic will inevitably increase the flight delays. The relatively constant reported delays thus lead us to the conjecture that the airlines have stretched the block times of their schedules to compensate for the increased delays. This section explores that possibility.

We compared the published OAG block times of April 1993 and April 1997 for flights among the 29 FAA large hub cities. The reason we confine the data to the 29 large hub cities is because they enjoy the most stable service in terms of the cities served, either as origin or destination, flight frequencies, and flight equipment.

Figure 2-3. Monthly Average Flight Delay Minutes to ATL by Arrival Time

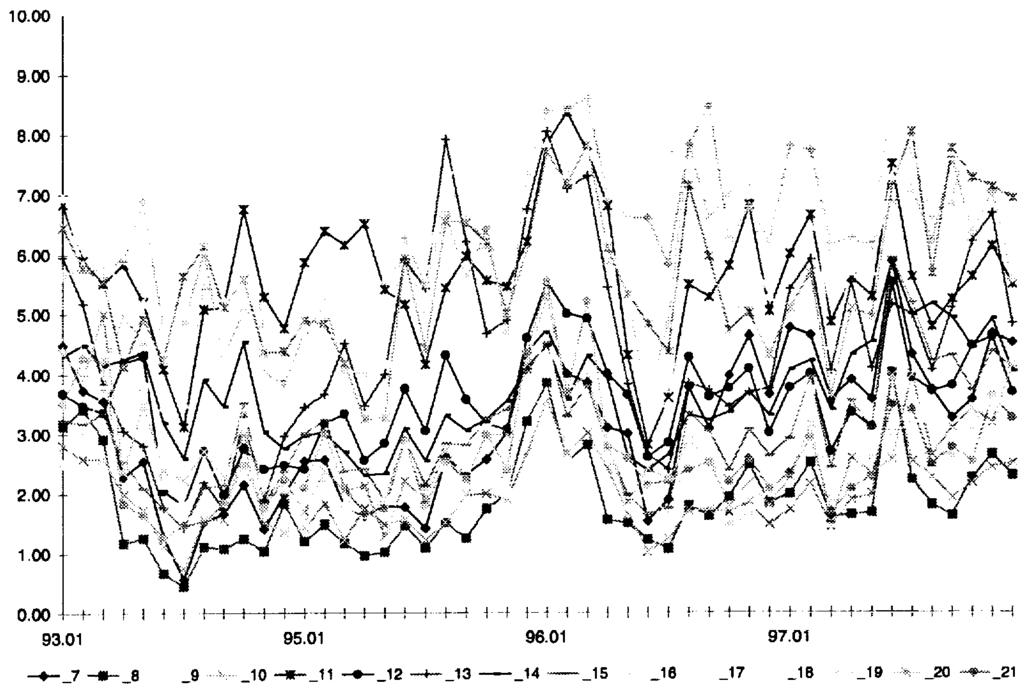
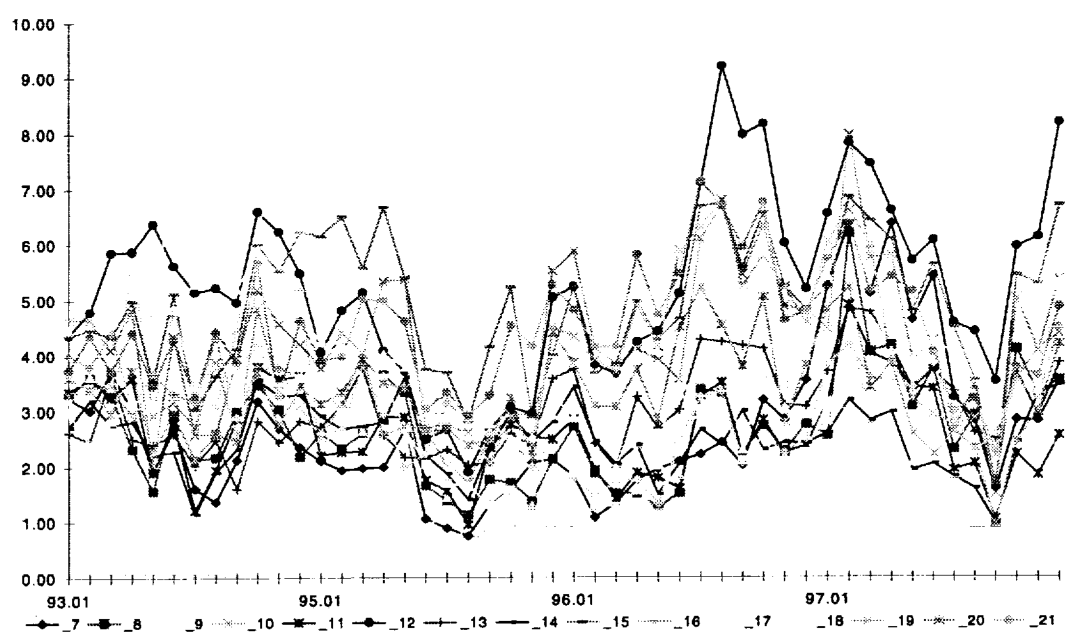


Figure 2-4. Monthly Average Flight Delay Minutes to DFW by Arrival Time



From 1993 to 1997, the operations in the 29 large hub cities increased 10.6 percent, while the average block times, weighted by the flight frequency, increased from 134.64 minutes to 140.61 minutes. However, much of the difference in the block times is due to new service (origin/destination pairs), which cannot be compared. If we focus on the head-to-head comparison, the block time increases by 1.25 minutes from 1993 to 1997.

Besides demand, the other factor determining the delays is capacity. During the period from 1993 to 1997, all the airports had relatively stable capacities, except Denver. That terminal commenced operations at the new Denver International Airport in January 1995. We have to exclude all the flights leaving from or arriving to Denver in order to have a fair comparison. If Denver is excluded, then the block time increases by 1.61 minutes from 1993 to 1997. This block time increase is much greater than the delay increase implied by the linear trend for the data of Figure 2-2, which is about 0.1 minute.

CORRELATION OF PUBLISHED BLOCK TIME AND AIRPORT OPERATIONS

In the previous section, we showed that block times are increasing, which appears to explain why reported delays held almost constant while operations increased. In this section, we explore the relation between increased block time and airport operations, through rank statistics.

Two pairs of random variables (X_1, Y_1) and (X_2, Y_2) are said to be *concordant* if $X_1<X_2$ and $Y_1<Y_2$ or if $X_1>X_2$ and $Y_1>Y_2$; and *discordant* if $X_1<X_2$ and $Y_1>Y_2$ or if $X_1>X_2$ and $Y_1<Y_2$. Then Kendall's τ statistic can be defined as the difference between the probabilities of concordance and discordance for two independent pairs (X_1, Y_1) and (X_2, Y_2),

$$\tau = Pr\{(X_1-X_2)(Y_1-Y_2)>0\} - Pr\{(X_1-X_2)(Y_1-Y_2)<0\}. \qquad \text{[Eq. 2-1]}$$

Like Kendall's τ, Spearman's ρ is defined as the difference between the probabilities of concordance and discordance. Let (X_1, Y_1), (X_2, Y_2), and (X_3, Y_3) be three independent pairs of random variables. Spearman's ρ is defined as the probability of the concordance of the pairs of (X_1, Y_1) and (X_2, Y_3),

$$\rho = Pr\{(X_1-X_2)(Y_1-Y_3)>0\}. \qquad \text{[Eq. 2-2]}$$

One can immediately see that both Kendall's τ and Spearman's ρ are what are called rank statistics, i.e., they will not change—unlike such moment statistics as the commonly used Pearson's second central moment—under any monotonic transformation preserving the ranks of the sampled data. Values of both Kendall's τ and Spearman's ρ are confined to the interval of [-1, 1], and -1 and 1 are attained when $X=-Y$ or $X=Y$, respectively.

The two order statistics are actually closely related to one another in that they are bounded by one another for any pair of random variables. The actual computations for the Kendall's τ and Spearman's ρ follow more convenient schemes than our definitions, and they are readily available from many popular statistical packages.

A random variable $U \in [0, 1]$ is said to be the probability transformation of another random variable, X, if U is the probability of X, or specifically $u = sup\{x: Pr(X \leq x) = u\}$. It is known that Spearman's ρ between two random variables (X, Y) is the same as the standard deviation of their probability transformations (U, V), which is also Pearson's σ.

Armed with the rank statistics, we are now ready to study the dependence between the airport operations and the published block times. The purpose of this study is to see whether the air carriers take the airport operations into consideration when assigning block times in schedule. Roughly speaking, the positive correlations can be interpreted as the block time being more likely to be lengthened if the airport operations are high. The increased block time is the carriers' response to the delays associated with high airport operations.

We define the operations of an airport in 1 hour as the sum of departures and arrivals in that hour from the OAG. We first calculate the correlations of the scheduled block time and the airport operations for the same origin-destination (O-D) pair, same airline, and same equipment type. Then these correlations are averaged, by weight of number of flights, for each departure airport and arrival airport. This gives us two sets of statistics: correlations of departure airport operations and the block time, and correlations of arrival airport operations and the block time. For most airports, whether it is departure block time or arrival block time, we see positive correlation measures, as shown in Figures 2-5 and 2-6.

Although Spearman's ρ and Kendall's τ are well defined for the same O-D pair, same airline, and same equipment type, their weighted measures are not since they involve different samples. However, since Spearman's ρ and Kendall's τ are rank statistics, defined as the probabilities that the two random variables move in the same direction, then weighting them makes sense as computing the overall probability of such tendency. After the probability transformation, all subsamples form a single large sample, where the rank statistics remain since the transformation is monotonic. The weighting of Pearson's σ does not have any theoretical interpretation since it depends on the *scale* of the random variable deviations, which is why we cannot use it to get the overall correlations although it is more widely used.

Figure 2-5. Correlations of Published Block Time and Arrival Airport Operations

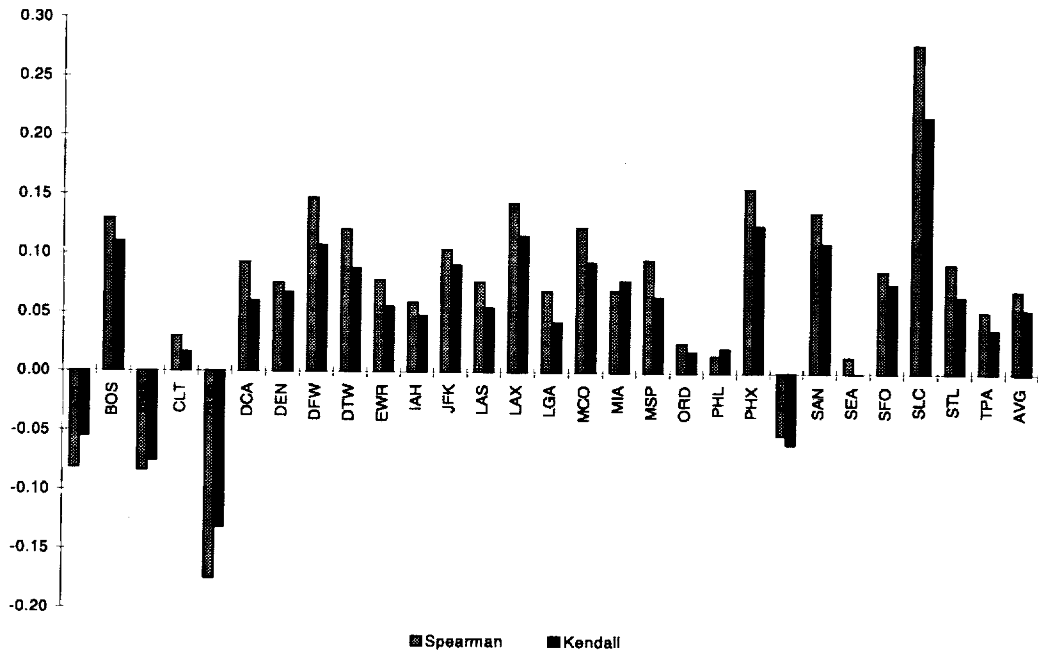
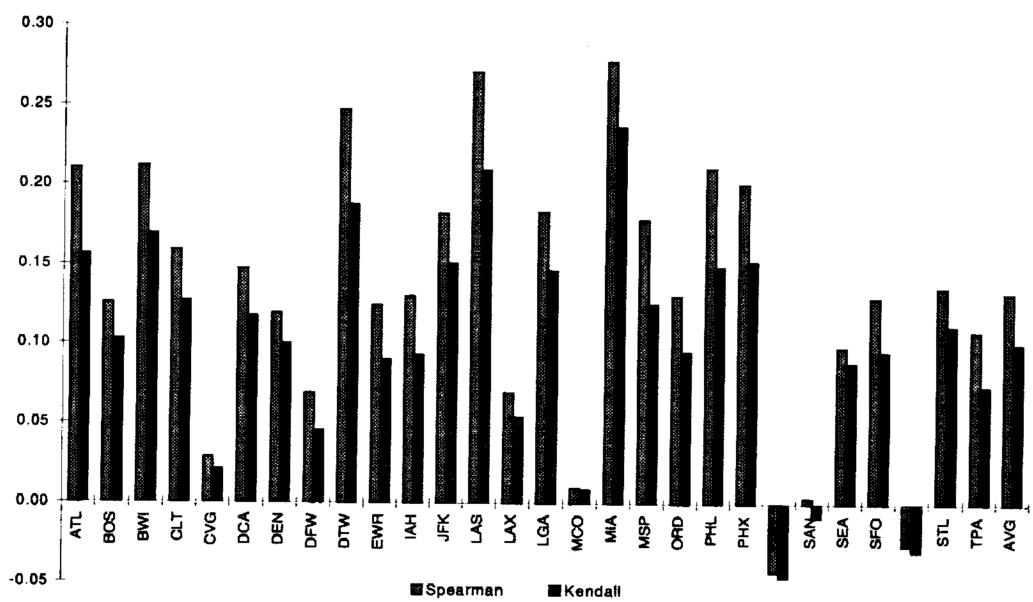


Figure 2-6. Correlations of Published Block Time and Departure Airport Operations



Most airports have positive correlations, which means there is more chance that airlines will increase the block time with the increased airport operations. Taking the 29 FAA large airports all together, the overall Kendall's τ between the arrival airport operations and block time is 0.05, and it is 0.10 between the departure airport operations and the block time. Taking the 29 FAA large airports all together, the overall Spearman's ρ between the arrival airport operations and the block time is 0.07, and it is 0.15 between the departure airport operations and the block time. These statistics indicate that the airlines pay more attention to the delays caused by congestion in departure airports than to the delays caused by the congestion in arrival airports.

We must caveat our conclusion by noting that a flight record is not used if there are *no* multiple flights during a day for the same O-D, same airline, and same equipment. We believe this omission of flights is random and hence has no effect on the correlation measures.

A General Linear Model of Published Block Time

In the previous section, we showed that the carriers' block times are positively correlated with airport operations. In this section, we develop a quantitative model to help us to understand how much padding in the published block time is due to delays from airport congestion.

Conceptually, the published block time can be decomposed into the following components:

- Ground movement time in the departure airport
- Airborne time
- Ground movement time in the arrival airport
- Schedule padding.

In addition to the departure and arrival airports' geographic locations, which determine the distance flown, block time depends on the equipment, which determines the speed of travel. Ground movement times are airline specific because different carriers use different clustered gates.

Padding is directly related to delays, which are determined by congestion of airport operations. The congestion measures that we used are the utilization ratios at the two airports. The arrival utilization ratio, ρ, is the quotient of arrival demand and arrival capacity, and the departure utilization ratio, σ, is the quotient of departure demand and departure capacity.

Guided by these ideas, we made our general linear model of block time as

$$B=T_{ij\Omega}+D_{i\alpha}+A_{j\alpha}+\beta_1\sigma_i(t_d)+\beta_2\rho_j(t_a). \qquad \text{[Eq. 2-3]}$$

In Equation 2-3, *B* is published block time. Roman indices, which identify airports, range from 1 to 29. Uppercase Greek indices, which identify equipment, range from 1 to the total number of aircraft considered. Lowercase Greek indices identify the airline, and they range from 1 to the number of airlines considered.

Thus, the right side of Equation 2-3 shows the model of block time from origin airport *i* to destination airport *j*, by airline α, using equipment Ω, departing at time t_d, and arriving at time t_a.

Obviously our linear model has a large number of adjustable constants. We considered the flights among the 29 FAA large hub airports operated by the 10 major airlines with and all the equipment types. Not all combinations of origin, destination, and equipment actually occur, however, in the data for April 8, 1997. In fact, the data required us to consider only 2,105 distinct values of the $T_{ij\Omega}$ and 238 distinct values of Δ_{ia} and $A_{j\alpha}$, respectively. Since the number of records is about 7,000 and our model has 2,583 adjustable parameters, our regression analysis has a large number of degrees of freedom.

We selected the FAA's 29 large airports to analyze, since they are busy and are operated close to capacity, so that delays are pronounced and may well be incorporated into the schedule. Including a less busy airport in the sample would introduce noise to the parameter estimation.

We took the hourly departure and arrival capacities for an airport from the FAA's Performance Monitoring Analysis Capability (PMAC) system for the 57 pacing airports in visual meteorological conditions (VMC). An airport operating under the visual flight rule has the maximum capacities that can be utilized, and we believe that the airlines design their flight schedules based on VMC. Table 2-1 shows the capacities that we used.

We used the SAS statistics package to calibrate our linear model. Based on the SAS estimate of this general linear model, we have

$$padding = 5.18 \times \sigma_d + 3.82 \times \rho_a. \qquad \text{[Eq. 2-4]}$$

The average origin airport utilization ratio and the average destination airport utilization ratio, weighted by the flight frequency, are 0.74 and 0.71, respectively. This means that the average padding incorporated in the block times for the flights among the 29 FAA large hub airports is $5.18 \times 0.74 + 3.82 \times 0.71 = 6.55$ minutes.

Table 2-1. FAA Large Airport Capacities Under VMC

Airport	Arrival	Departure
ATL	90	100
BOS	57	57
BWI	39	39
CLT	52	52
CVG	52	52
DCA	40	40
DFW	90	90
DTW	63	63
EWR	49	49
IAH	72	69
JFK	46	46
LAS	59	59
LAX	72	72
LGA	38	38
MCO	53	53
MIA	59	59
MSP	53	53
ORD	78	78
PHL	45	45
PHX	51	51
PIT	81	81
SAN	32	32
SEA	48	48
SFO	53	53
SLC	59	59
STL	58	58
TPA	50	50

Data source: FAA PMAC system.

Note: DEN and HNL are part of the 29 FAA large airports but are not part of the 57 pacing airports. Therefore, they are not used in the general linear model estimation.

The average padding in the block times would be increased by 0.65 minute from 1993 to 1997 based on this model, since the operations increased by 10.8 percent. This is not far from our empirical estimate of 1.25 minutes. One has to bear in mind that this comparison is based on the assumptions that the airlines will equally increase the operations among the airports and among the hours, and that the airlines would have the same policy in deciding how much padding to put into the block time. Nevertheless, Equation 2-4 gives a reasonable estimate of the paddings that airlines incorporate in their scheduled block times in response to airport congestion.

Another observation from Equation 2-4 is that the airlines pay more attention to the congestion in the departure airport than to the arrival airport, which is a reaffirmation of the finding in the previous section based on correlation figures.

Obviously, there are many ways to construct the general linear model. We have tried some others, including some nonlinear components like the square of the utilization ratios, and they all yield similar results. We keep the current one since it is the simplest, and yet it captures the essence of the effect.

Many airports have the flexibility to trade the arrival and departure capacities as modeled by our LMINET airport capacity models, so the airport utilization ratios based on FAA's airport capacity figures are approximate. On the other hand, using the LMINET airport capacity models will involve the policy decision of allocating the capacities from the Pareto frontier, which will make the utilization ratio subject to this policy decision. FAA's capacity figures are chosen since (1) they are more official, (2) the model itself is on the aggregate level so that it is not intended to be accurate nor does it need accurate data input, and (3) we are not extremely concerned with the accuracy of the model; the magnitude of the estimate is good enough for our purpose.

An airport-specific padding model may seem more appealing. However, we could not carry out this modeling effort due to the very large amount of data records needed. With available data, we cannot get significant parameter estimates for all airports.

Chapter 3
Modeling Unconstrained Air Traffic Demand in the Future

This chapter concerns forecasting unconstrained air traffic demand in the future when the delay and congestion due to limited air traffic capacities are not binding constraints on air traffic growth.

The term *air traffic demand* is only loosely defined. It can mean anything from aircraft operations to passenger enplanements on different aggregate levels, but we are primarily interested in the *schedule* of the air travel.

Specifically, a schedule is a planned service from the origin airport to the destination airport, leaving at a certain time and arriving at a certain time, operated by an air carrier using certain equipment. For the present we ignore the operator and the equipment of the schedule, which we will deal with in the subsequent section "Air Traffic Schedule in the Future."

The models developed in this report are for the long term, when the effects of the capacity shortage will be more pronounced. Our previous studies show that the NAS will be severely congested in about a decade if nothing is done to augment its capacities [1,2,3,8] We choose the year 2007 as the "future" for this study.

If we discretize time by dividing a day into a certain number of epochs, e.g., hour-long epochs, then a schedule of the NAS in a day will be represented by a matrix, $\{s_{ijkl}\}$, where $i, j \in I = \{0,1,\ldots,64\}$, and $k, l \in K = \{0,1,\ldots,23\}$. Here i and j are the indices of the airports in the LMINET, where 0 represents an out-of-network airport; and k and l are the time indices of the departure and arrival, respectively. The schedule, s_{ijkl}, is the number of flights from airport i to airport j that depart in the epoch k and arrive in the epoch l.

FAA Terminal Area Forecast

The Terminal Area Forecast (TAF), published annually by the FAA, has several data tables for the total annual enplanements, operations, and various FAA workload measures for the set of airports and control towers that the FAA tracks. Each table has several columns to give more detailed information, e.g., the enplanements can be domestic or international. For the most recent TAF, released in February 1998, the data from 1976 to 1996 are the annual totals reported by the airport control towers, while the data from 1997 to 2010 are the FAA's predicted values.

The Airport Council International sponsors the annual FAA Commercial Aviation Forecast Conference every year, and 1998 is its 23rd. The FAA not only updates its TAF every year, it also improves the forecast's methods constantly. The TAF has become the *de facto* official aviation demand forecast. In this report, we are interested in the TAF of operations for the LMINET airports.

The FAA derives forecasted operations in the TAF in the following way: [5]

- It forecasts the enplanements based on outputs of socioeconomic models, such as gross domestic product (GDP) and demographic growth rates, with due consideration of originating traffic and connection traffic. Each major airport has its own specific models.
- It forecasts the load factors to and from each airport based on the demand, fare yield, and airlines cost.
- It forecasts the average number of seats per aircraft for arrivals and departures at the airport.
- It divides the forecasted enplanement by the forecasted load factor and by the forecasted average number of seats per aircraft to get forecasted operations.

In deriving the forecasts, flight delays due to traffic congestion are never explicitly considered. Implicitly, the TAF assumes that airport and ATC capacities will grow to meet the potential demand.

Table 3-1 shows the FAA's values for total enplanements and total operations at the LMINET airports for 1996 and 2007, and their ratios of commercial and general aviation (GA) operations in 2007 and 1996, based on the information from FAA's TAF. We used the total of air carrier, air taxi, and itinerant GA in the TAF as the airport operations measure. Air carrier and air taxi are the operations of scheduled air transport service corresponding to the OAG; air taxi data are for aircraft with less than 60 seats, which are typical of commuter operations.

One can see that all the airports will enjoy positive total and commercial operations growth from 1996 to 2007, but there are many airports with negative GA operations growth from 1996 to 2007—the GA traffic ratio is less than one. This may imply that the commercial traffic growth will be at the expense of GA operations. For all airports reported in FAA's TAF but not included in the 64 LMINET airports, the aggregate commercial and GA traffic in 2007 will be 1.149 and 1.031, respectively, times their 1996 levels.

Table 3-1. Annual Enplanements (Millions) and Operations (Thousands) at LMINET Airports

	Enplanements		Operations		Ratio of traffic	
Airport	1996	2007	1996	2007	Commercial	GA
BOS	12.3	16.0	462	509	1.088	0.906
BDL	2.7	4.1	151	181	1.242	0.971
HPN	0.5	0.9	153	160	1.271	0.905
ISP	0.6	0.9	109	117	1.258	0.951
TEB	0.0	0.0	189	189	1.000	1.000
LGA	10.3	13.8	342	381	1.109	0.822
JFK	15.0	20.7	360	397	1.109	0.823
EWR	14.2	20.7	443	561	1.229	0.906
PHL	9.1	14.8	401	509	1.269	0.821
BWI	6.6	10.3	260	338	1.301	0.906
DCA	7.2	8.6	305	318	1.055	0.948
IAD	6.0	9.7	323	397	1.240	1.000
GSO	1.4	2.5	138	170	1.300	0.937
RDU	3.1	4.8	217	256	1.240	0.906
CLT	10.7	15.6	454	563	1.239	0.906
ATL	30.7	41.4	770	916	1.177	0.906
MCO	11.8	22.6	337	502	1.507	1.000
PBI	2.8	3.9	182	196	1.164	0.951
FLL	5.2	9.4	234	304	1.359	1.000
MIA	16.1	27.4	540	694	1.336	0.905
TPA	6.2	9.2	269	331	1.257	0.906
MSY	4.2	5.9	162	190	1.197	0.906
MEM	4.6	6.2	358	467	1.279	1.094
BNA	3.4	5.4	222	260	1.195	0.904
SDF	1.8	2.9	168	215	1.267	1.000
CVG	8.8	16.9	392	613	1.507	0.906
DAY	1.0	1.0	143	160	1.159	1.000
CMH	3.1	5.3	185	238	1.290	0.952
IND	3.5	5.8	230	305	1.390	0.906
CLE	5.4	8.6	287	373	1.294	0.906
DTW	15.0	24.7	530	708	1.343	1.035
PIT	10.1	14.4	438	536	1.205	1.166
SYR	1.0	1.3	122	148	1.239	1.000
MKE	2.7	4.3	187	239	1.268	1.000
ORD	32.2	43.2	906	1039	1.180	0.863
MDW	4.5	6.4	251	297	1.268	0.906
STL	13.5	20.5	511	637	1.240	0.906
IAH	11.9	20.0	391	566	1.420	0.906
HOU	4.0	5.3	252	287	1.226	1.000
AUS	2.8	4.6	203	245	1.340	1.105

Table 3-1. Annual Enplanements (Millions) and Operations (Thousands) at LMINET Airports (Continued)

Airport	Enplanements		Operations		Ratio of traffic	
	1996	2007	1196	2007	Commercial	GA
SAT	3.3	5.5	238	293	1.340	1.105
DAL	3.5	5.2	219	264	1.388	0.906
DFW	27.4	43.7	869	1234	1.38	1.034
MSP	13.4	20.8	478	615	1.282	0.905
MCI	5.0	7.1	195	244	1.22	0.906
DEN	15.2	20.6	453	553	1.211	0.906
ABQ	3.2	5.1	173	217	1.339	0.952
ELP	1.8	2.8	122	125	1.131	0.906
PHX	14.6	24.2	531	698	1.356	0.906
SLC	9.8	15.5	369	491	1.372	1.000
LAS	14.3	26.1	445	637	1.474	1.000
SAN	6.8	10.4	238	309	1.305	0.823
SNA	3.6	6.4	369	483	1.378	1.233
LGB	0.2	0.4	263	312	1.425	1.151
LAX	28.2	41.9	761	947	1.232	0.842
BUR	2.5	4.3	180	222	1.352	1.048
ONT	3.2	4.6	149	177	1.222	0.878
RNO	3.0	5.4	144	189	1.403	0.906
SMF	3.5	5.7	145	201	1.382	1.000
OAK	4.8	7.8	400	494	1.271	1.119
SFO	18.3	29.4	426	562	1.299	1.000
SJC	4.8	8.0	210	258	1.394	0.905
PDX	6.1	10.2	290	384	1.351	0.851
SEA	11.7	17.5	397	503	1.242	1.000

Tables 3-2 and 3-3 compare LMINET to the network for operations and enplanements, respectively. Figures 3-1 and 3-2 graphically depict the LMINET airport annual operations and enplanements for 1996 through 2017.

Table 3-2. LMINET Airports Versus the Total Operations (Millions)

		Operations			Growth rate (%)	
	Count	1996	2000	2010	1996–2000	2000–2010
Large hubs	29	13.6	14.9	18.3	2.37	2.04
Medium hubs	42	9.2	9.9	11.6	2.04	1.53
Small hubs	67	8.2	8.6	9.3	1.24	0.74
Non-hub towers	305	30.9	31.8	33.5	0.70	0.54
Total	443	61.9	65.3	72.7	1.35	1.08
LMINET airports	64	20.7	22.6	27.3	2.20	1.91

Table 3-3. LMINET Airports Versus the Total Enplanements (Millions)

		Enplanements			Growth rate (%)	
	Count	1996	2000	2010	1996–2000	2000–2010
Large hubs[a]	29	412.6	490.1	684.3	4.40	3.39
Medium hubs[b,c]	42	135.7	163.6	237.9	4.79	3.81
Small hubs[d]	67	41.6	48.8	67.5	4.08	3.30
Non hub towers	273	15.5	17.6	22.2	3.18	2.38
Total	411	605.5	720.2	1,012.0	4.43	3.46
LMI airports	64	514.0	613.0	863.0	4.50	3.50
Share (%)	—	85.0	85.1	85.3	—	—

Source: Department of Transportation, *Terminal Area Forecasts, Fiscal Years 1997—2010,* Report No. FAA-APO-97-7, Federal Aviation Administration, Office of Aviation Policy and Plans, Statistics and Forecast Branch, Washington, DC, October 1997.

[a] > 1% of total enplanement.

[b] > 0.25% of total enplanement.

[c] The 42 medium hub airports are ABQ, ANC, AUS, BDL, BNA, BUF, BUR, CLE, CMH, COS, DAL, ELP, FLL, GEG, HOU, IAD, IND, JAX, MCI, MDW, MEM, MKE, MSY, OAK, OGG, OKC, OMA, ONT, PBI, PDX, RDU, RNO, RSW, SAT, SDF, SJC, SJU, SMF, SNA, TUL, TUS, and GUM.

[d] > 0.05% of total enplanement.

Figure 3-1. Total LMINET Airport Annual Operations (Millions)

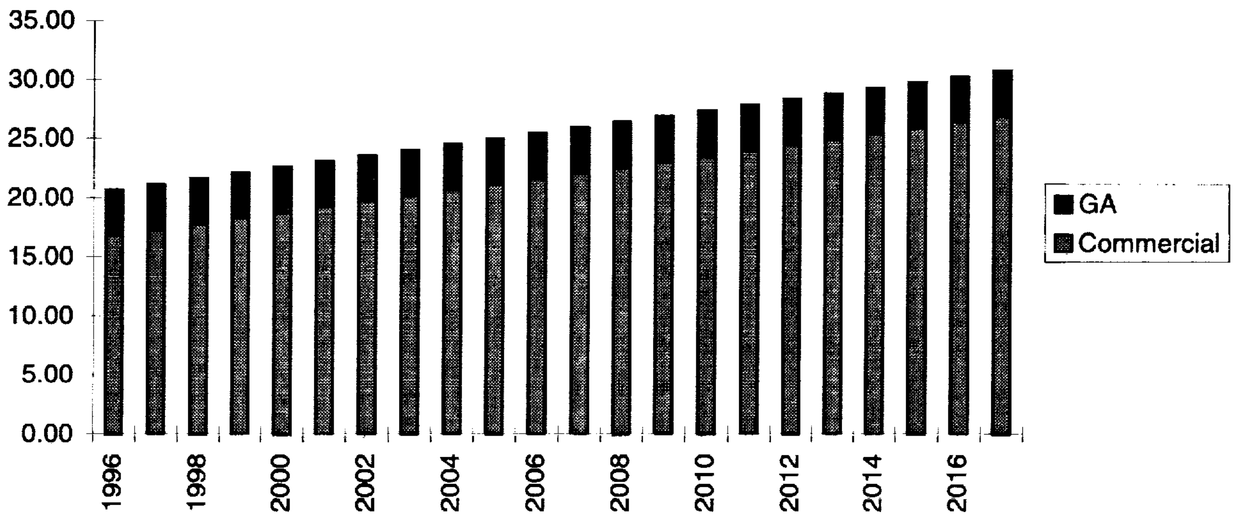
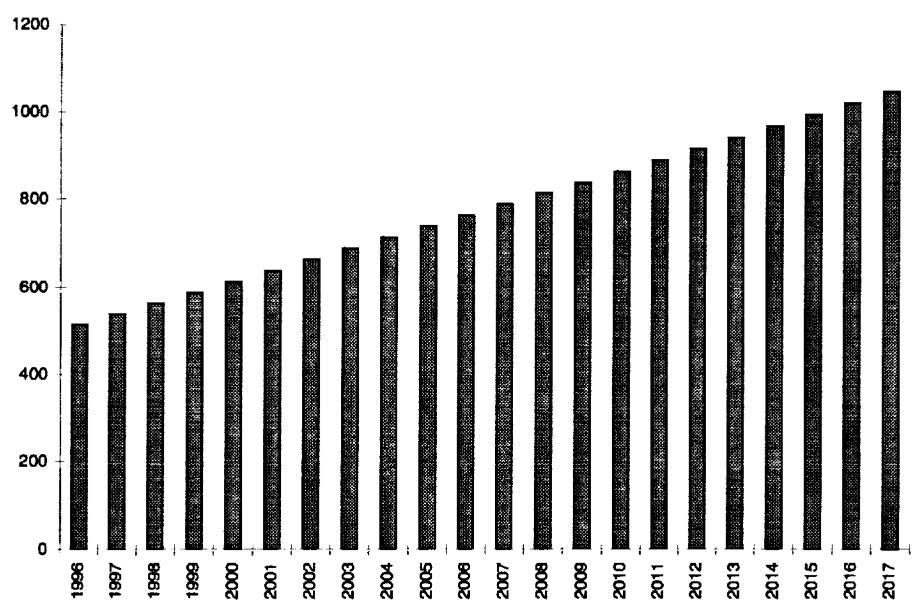


Figure 3-2. Total LMINET Airport Annual Enplanements (Millions)



Current Air Traffic Schedule

We considered both scheduled air transport service and itinerant GA traffic. We based demand for scheduled air transport service on the schedule published by the

OAG. We constructed the time variation of GA demands from data recorded in the FAA's Enhanced Traffic Management System (ETMS). The schedules on April 8, 1996—whether the commercial operations represented by the OAG or the GA extracted from ETMS—are selected to run our model. Everything that we do to construct the traffic schedule in the future will be set in early April on a weekday.

Since the OAG schedule is the planned rather than observed air traffic schedule, and only the GA flights filing instrument flight rules (IFR) flight plans will be recorded in ETMS, there will be differences between the traffic reported by the OAG and ETMS and the FAA's TAF. Since the TAF is recorded by traffic control towers, which are believed reliable, both the OAG and GA schedules have to be scaled to conform to the data in the TAF.

Table 3-4 lists the traffic adjustment factors for each LMINET airport. We first compute the total annual commercial operations, per airport, based on the entire 1996 OAG. The Commercial Adjustment Factor, α, is given by the commercial operations in the TAF (air carrier and air taxi) divided by the operations given by the OAG. The Total Adjustment Factor, γ, is given by the total airport operations in the TAF (air carrier, air taxi, and itinerant GA) divided by the operations given by the OAG.

Table 3-4. Commercial and Total Traffic Adjustment Factors

Airport	Commercial	Total	Airport	Commercial	Total
BOS	1.035	0.974	MKE	1.457	1.15
BDL	1.585	1.203	ORD	1.019	0.975
HPN	3.314	1.067	MDW	1.893	1.251
ISP	3.591	1.189	STL	1.075	1.005
TEB	1	1	IAH	1.014	0.956
LGA	1.037	0.979	HOU	1.897	1.048
JFK	1.061	1.017	AUS	2.246	1.119
EWR	1.033	0.99	SAT	2.5	1.321
PHL	1.086	0.966	DAL	2.23	1.268
BWI	1.235	1.134	DFW	1.102	1.068
DCA	1.206	1.016	MSP	1.147	1.025
IAD	1.225	1.023	MCI	1.046	0.975
GSO	2.757	1.836	DEN	0.966	0.917
RDU	1.599	1.11	ABQ	1.557	1.107
CLT	1.233	1.077	ELP	1.981	1.216
ATL	1.053	1.023	PHX	1.295	1.083
MCO	1.111	1.016	SLC	1.419	1.113
PBI	2.442	1.208	LAS	1.553	1.24
FLL	1.639	1.126	SAN	1.138	1.052
MIA	1.247	1.099	SNA	3.667	1.016

Airport	Commercial	Total	Airport	Commercial	Total
TPA	1.463	1.231	LGB	40.116	2.034
MSY	1.307	1.119	LAX	1.014	0.982
MEM	1.738	1.452	BUR	2.699	1.468
BNA	1.77	1.303	ONT	1.419	1.204
SDF	1.53	1.24	RNO	1.634	1.137
CVG	1.033	0.997	SMF	1.409	1.165
DAY	1.68	1.161	OAK	3.155	1.859
CMH	1.679	1.331	SFO	1.048	0.984
IND	1.948	1.543	SJC	1.618	0.979
CLE	1.116	1.006	PDX	1.294	1.101
DTW	1.239	1.058	SEA	1.094	1.073
PIT	1.083	1.022	OTR	1.299	1.07
SYR	1.75	1.247			

Obviously from the factors' definitions, if we scale the OAG operation by the adjustment factors α and γ, we will get the *actual* commercial operations and total operations, respectively. This method is fine so long as we are not concerned with the *schedule* of the operations. Since commercial and GA operate on different schedules and adjust at different rates, a GA adjustment factor, β, is needed.

Let T_{Total}, $T_{Commercial}$, T_{OAG}, T_{GA}, T_{ETMS} be the traffic indicated by the subscripts. By the definitions,

$$T_{Total} = T_{Commercial} + T_{GA}, \qquad \text{[Eq. 3-1]}$$

$$T_{Total} = \gamma T_{OAG}, \qquad \text{[Eq. 3-2]}$$

$$T_{Commercial} = \alpha T_{OAG}, \qquad \text{[Eq. 3-3]}$$

$$T_{GA} = \beta T_{ETMS}, \qquad \text{[Eq. 3-4]}$$

we have

$$\beta = (\gamma - \alpha) \times T_{OAG}/T_{ETMS}. \qquad \text{[Eq. 3-5]}$$

Since we do not have any particular knowledge about the missing flights from OAG to commercial, and from ETMS to GA, we have to assume they are random or that the missed flights are proportional to the ones in the current schedule. The scaling-up of OAG for commercial traffic takes an application of the Fratar algorithm, and the scaling-up from ETMS for GA traffic takes the simple form of multiplying all the flights by the adjustment factor, β, of the departure airport, based on the same arguments that will be presented in the next section.

AIR TRAFFIC SCHEDULE IN THE FUTURE

The future air traffic demand, expressed in terms of the schedule, s_{ijkl}, must be constructed, although the only thing that we know is the total airport operations. In fact, generating demand schedules for the entire network is a challenging task. Although the academic literature is rife with models and algorithms, they are geared to providing the forecast of single variable systems or a non-networked multivariable system.

There are two major intellectual challenges:

- The interaction of the NAS network's nodes and arcs and the possibility of achieving the goal of a specific traffic level via different means
- The prediction of air carriers' behavior, even at some high aggregate level.

This section presents our modeling considerations and the algorithm that we used to forecast air traffic schedules in the future.

Modeling Assumption

Our modeling is based on available data and models; on their integration; and, more importantly, on the desired properties of our forecast. We require our approach for forecasting the unconstrained air traffic demand to satisfy the following:

1. The schedule provided by the air carriers is the variable of interest, which reveals everything about air carriers' operations.

2. We will construct an industry-wide model instead of one that integrates carrier-specific models. The air transport industry in the United States is an oligopoly, consisting of 10 major carriers with about 90 percent of total domestic operation and three dozen or so affiliated and unaffiliated commuter, cargo, and chartered passenger and cargo carriers. If we just concentrate on having individual models for each of the 10 major passenger carriers—if we could accomplish the tremendous amount of work involved—it is still impossible to predict the industry configuration, or market share in the future, in this dynamic environment. The recently announced virtual merger between Northwest Airlines and Continental Airlines, and the marketing alliance between the two former foes, American Airlines and US Airways, are good examples of these difficulties.

   Taking the whole industry together, while still assuming the existence of competition among the carriers, we avoid attempting to predict winners and losers in the competition. A representative of one major U.S. carrier told us that his airline aggressively seeks opportunities to grow, since if they do not someone else will. This means that the air carriers put their resources where the demand is on the aggregate level. On the other hand, we do not really need an air-carrier-specific model if our model will be used by other models to quantify the benefits of new air traffic management (ATM) procedures or decision support tools. Individual air carriers will benefit indirectly from our industry-wide model, in that it is up to them to make up the market share to best utilize their resources.

3. The FAA's TAF will be used as an input, so the future schedule we derive must meet the TAF at the airport level. Due to the way the TAF is produced, delineated previously in this chapter, we assume that it already reflects air carriers' choices to use less congested airports or new hubs, and the use of larger equipment in the short haul market and the replacement of turboprops by regional jets. Other terminal area forecasts will suffice.

4. The traffic growth rate between two cities must be proportional to the traffic growth rates in both cities, respectively.

5. Air carriers' operation practices will be unchanged. Specifically, we assume the current air carriers' operations are rational and will continue to be rational in the future. By *rational* we mean that the air carriers, being commercial companies, will try to maximize their profits by putting their resources or schedules where the demand is. Battling for market share, just for the sake of market share, by providing more schedule than demand, is not a rational behavior. This seems a good assumption, since the air transport industry appears finally to have reached maturity after two decades of deregulation. Evidence of this comes from comparing the record profits and relative stability enjoyed by the industry in the past few years to the record losses, massive traffic growth, labor disputes, and industry instability (with a plethora of low-cost start-up carriers and merger and acquisition activities) seen right after the deregulation in 1980s and early 1990s.

   The assumption of rationality of air carriers can be decomposed into the following:

   a. The current OAG schedule is the best schedule to meet the air travel demand. One example is Continental Airlines' decision in the past few years to cancel their hubs at Denver, Cleveland, and Greensboro/High Point and redeploy the flights to Houston and Newark to get better yields.

   b. The air carriers will continue to conduct bank operations in hub airports. Since airline deregulation in 1978, the carriers have had the freedom to design their schedules as they see fit except for a few slot-controlled airports.[1] Since then, air carriers have consolidated their op-

[1] The slot-controlled airports are ORD, JFK, LGA, and DCA.

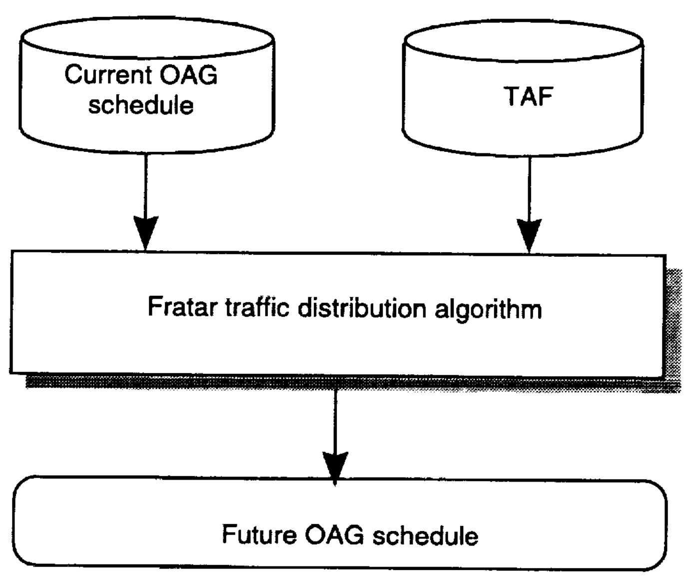
erations to concentrate on a few hub airports, which are characterized by alternating banks of arrivals and departures. There are two major advantages of bank operations: first, the number of markets, through connection at the hub, is massively expanded—offering travelers choices that cannot be made through point-to-point operations; second, the airline that has the dominant market share at the hub cities commands premium fares.

6. The time-of-day demand pattern will not change. Given the total number of people willing to travel from A to B in a day, research by airlines and Boeing has shown that the distribution of that demand across the day depends on the local departure and arrival times and the journey time, where business travelers and leisure travelers may have different demand patterns, and, of course, different demand elasticities. Thus, unless there are new technologies that will drastically reduce the journey time, the travelers' time-of-day demand patterns will not change.

Fratar Algorithm

This algorithm is the most widely used method of generating trip distributions based on the terminal area forecast. It has been used by both the DOT and FAA in their transportation planning models, such as NASPAC, an event simulation model of NAS. A schematic diagram of the algorithm is shown in Figure 3-3.

Figure 3-3. The Fratar Traffic Growth Distribution Algorithm



The daily traffic, t_{ij}, from airport i to airport j, total daily departures, d_i, from airport i, and total daily arrivals, a_j, to airport j are related to the schedule, s_{ijkl}, as follows:

$$t_{ij} = \sum_{kl} s_{ijkl}, \qquad \text{[Eq. 3-6]}$$

$$d_i = \sum_j t_{ij}, \qquad \text{[Eq. 3-7]}$$

$$a_j = \sum_i t_{ij}. \qquad \text{[Eq. 3-8]}$$

If the schedule is balanced, or the network does not have any sinks, then $d_i = a_i$, $\forall\, i \in I$.

Let D_i, $i \in I$ be the total number of departures in the target year taken from the forecast. The Fratar method is an iterative algorithm that takes the following steps:

Step 0: Assign t_{ij}, d_i, a_j, $\forall i, j \in I$, based on the current year schedule.

Step 1:

$$g_i = \frac{D_i}{d_i}, \forall i \in I, \qquad \text{[Eq. 3-9]}$$

Step 2:

$$T_{ij} = t_{ij} \times g_i \times g_j \times \frac{1}{2}\left[\frac{d_i}{\sum_m t_{im} \cdot g_m} + \frac{a_j}{\sum_m t_{mj} \cdot g_m}\right], \forall i, j \in I \qquad \text{[Eq. 3-10]},$$

Step 3:

If $\sum_m T_{im} = D_i, \forall i \in I$, then go to Step 4;

else

$t_{ij} = T_{ij}$, $\forall i, j \in I$, and update d_i, a_j, $\forall i, j \in I$ accordingly, go to Step 1.

Step 4: Compute the traffic growth factor, r_{ij}, $\forall i, j \in I$, by dividing the traffic, T_{ij}, in the target year by the one in the current year; compute the schedule, S_{ijkl}, in the target year by multiplying the schedule in the current year by the traffic growth factor, r_{ij}. Stop.

Now let us check that the schedule in the target year made by the Fratar algorithm has the desired properties. First, the schedule will always meet the terminal departure totals predicted in the TAF, which is embedded in the algorithm.

Second, $r_{ij} = r_{ji}$, which means the traffic growth is undirectional. This is an implicit desired property in a travel network, although not explicitly stated in the previous subsection.

Third, the growth factor is uniform across the entire day, which is a desired property if we assume that the current schedule is rational and the travelers' time-of-day demand pattern will not change.

The fact that the growth factor is uniform across the day implies another property of the schedule in the target year: the airport traffic is dynamically balanced, and the bank operations in hub airports are preserved. Let d_{ik}, a_{ik}, $\forall i \in I$, $\forall k \in K$, be the total departures and arrivals in time k at airport i.

$$d_{ik} = \textstyle\sum_{jl} s_{ijkl},$$

$$a_{ik} = \textstyle\sum_{jl} s_{jilk}.$$

An airport i is said to be dynamically balanced if $d_{ik} = a_{ik}$, $\forall k \in K$, which means there are no idle aircraft sitting on the ground. In reality, a flight has to spend some time in the terminal before taking off, but we will keep this simple definition, and real operations can be modeled by shifting the time index. Let D_{ik}, A_{ik}, $\forall i \in I$, $\forall k \in K$, be the total departures and arrivals at airport i at time k in the target year. By the Fratar algorithm,

$$D_{ik} = \sum_{jl} s_{jkl} = \sum_{jl} r_{ij} s_{ijkl} = G_i \sum_{jl} u_j s_{ijkl}, \qquad \text{[Eq. 3-11]}$$

$$A_{ik} = \sum_{jl} S_{jilk} = \sum_{jl} r_{ji} s_{jilk} = \sum_{jl} r_{ij} s_{jilk} = G_i \sum_{jl} u_j s_{jilk}, \qquad \text{[Eq. 3-12]}$$

where

$$G_i u_j = r_{ij},\ \forall i, j \in I,\ and \sum_{jl} u_j = 1. \qquad \text{[Eq. 3-13]}$$

The right hand sides of D_{ik} and A_{ik} resemble the expectations of the product of two discrete random variables. If two random variables are independent, then the expectation of their product is equal to the product of their expectations. If we assume that the traffic growth rate is independent of the current schedule, which is a reasonable assumption, then

$$D_{ik} \cong G_i \left(\sum_{jl} u_j\right)\left(\sum_{jl} s_{ijkl}\right) = G_i d_{ik},$$

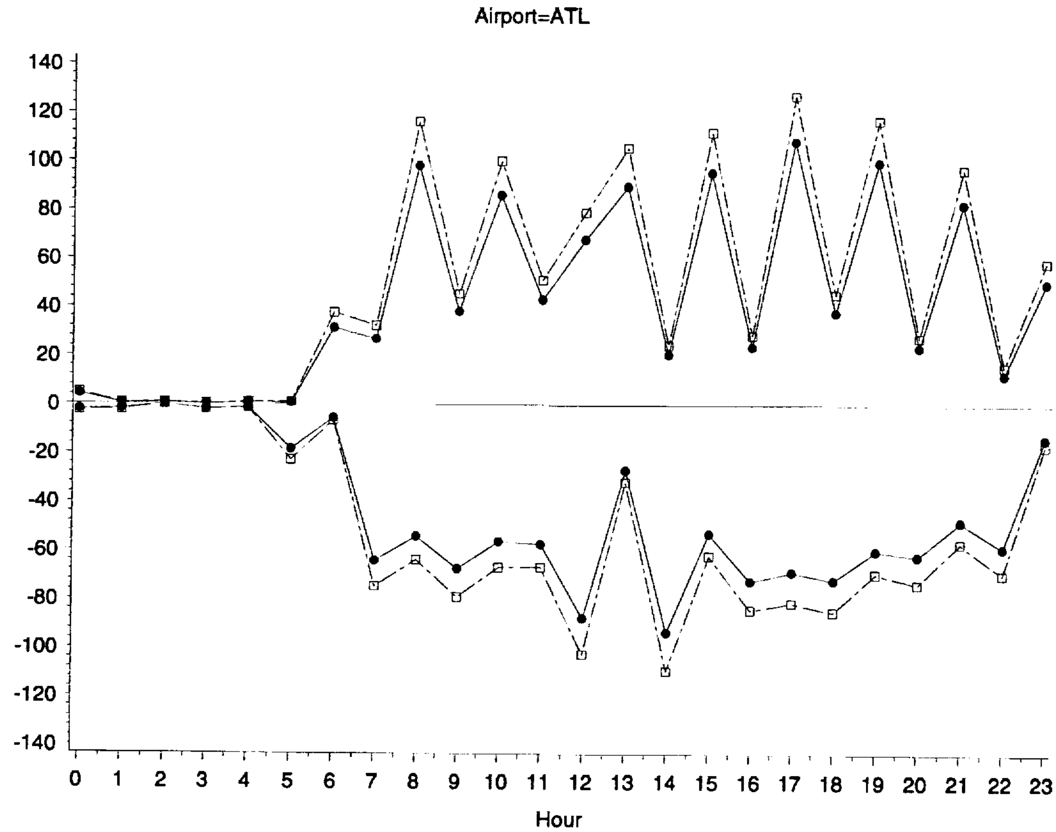
similarly,

$$A_{ik} \cong G_i a_{ik}.$$

Since $d_{ik} = a_{ik}$, $\forall i \in I$, $\forall k \in K$, then $D_{ik} \cong A_{ik}$. And, interestingly, G_i must be the growth factor implied by the TAF in order to satisfy the binding terminal total departure constraint.

Figure 3-4 shows an example of applying our method to ATL.

Figure 3-4. 1996 and Forecast Unconstrained 2007 Hourly Operations at ATL



Note: Positive curves = departure; negative curves = arrival;
dot = 1996; square = 2007.

Chapter 4

Modeling Constrained Air Traffic Demand in the Future

This chapter describes a model for forecasting future air traffic schedules when insufficient capacities or excessive flight delays must be considered.

GENERAL MODELING APPROACH

Generic Air Carriers' Operations Practices

There are three flows in airline operations: passenger, equipment, and crew. They are subject to the following constraints: fleet, maintenance location, crew location, and regulations governing equipment and crew. As a commercial entity, an airline seeks to maximize its revenue under the given constraints. Although the details of their practices may be different, air carriers all seem to take the following steps to conduct their operations:

1. *Market planning*. This process estimates the passenger demand between two points, including connecting passengers, the passenger time-of-day demand pattern, and the yield. The airlines' flight schedule is designed on the basis of this information. Market assessments tend to be carried out by a combination of models and detailed, market-by-market analysis. Statistical time series forecasting models, the Quality Service Index (QSI), and Decision Window Path Methodology are among the techniques that carriers have used to design their schedules, but their implementations vary. Yield information is an airline's most closely guarded proprietary data.

2. *Fleet assignment*. Given the fleet composition, the maintenance location, and the regulations governing the equipment, the fleet schedule is assigned to maximize the expected revenue. This is a vehicle routing problem that connects the schedule given from the market planning with the desired equipment with the most efficiency. The solution to the fleet assignment problem is highly automated among the carriers by software run on powerful computers. The software packages used by the carriers are almost identical; differences among the packages lie mainly in their speed of execution and flexibility.

3. *Crew assignment*. Given the fleet assignment and crew and its base, crew trips are generated with minimum crew cost while not violating regulations and the pilot labor contract. Like the solution to the fleet assignment

problem, the solution to crew assignment is also accomplished via the use of highly sophisticated, mixed integer programming mathematical models run on powerful computers. Like fleet assignment, crew assignment software is essentially a commodity, i.e., the models are similar across the carriers, and the differences lie mainly in computation speed and their flexibility.

4. *Operation control.* This involves rescheduling equipment and rematching appropriately qualified crews with equipment in response to disturbances in the schedule, such as those caused by inclement weather. Operation control demands real-time solutions, which are typically obtained by essentially heuristic methods, although the methods usually are implemented with computers. Different carriers may have different policies regarding the restoration of normal operations. For example, one major U.S. carrier told us that they emphasize the preservation of revenue, while another told us their goal is to return to normal operations as soon as possible.

Some iterations between market planning and fleet assignment are needed mainly to satisfy station managers. We are not concerned with operation control since it deals with day-to-day disturbances and we are interested in planning. Actually, we need to model only the market analysis, since given a reasonable schedule there are always feasible equipment routing and crew assignment solutions. Market planning is the least understood of the four schedule planning steps, and the carriers do it using a hodgepodge of techniques.

One can see that delay is not explicit in the schedule planning process. However, this does not mean the management of a carrier does not pay attention to delay statistics. The recent examples of Continental Airlines' paying employees bonuses for achieving certain delay targets, and of US Airways' forming special employee committees to reduce delays, testify to management's serious resolve to combat them. Management's intention is to improve their service quality and, thus, draw more revenue.

Most delays experienced by the airlines are caused by poor coordination, insufficient staffing, equipment breakdown, or inadequate procedures (such as those waiting for connecting passengers), apart from weather delays over which management does not have control. Once these elements were well addressed, Continental and US Airways went from bottom to top in on-time statistics.

Another reason that the carriers have not explicitly considered delays due to inadequate NAS capacity is simply that these delays have not yet become excessive. One major U.S. air carrier told us they tend to lengthen the block time for the flight at the end of the departure push to accommodate the ATC-induced delays. Pleas of carriers' operations staffs to reduce the number of departure flights at rush hours have been overruled by their market staffs because they can make money even with some delays.

Model Survey

While the literature and applications are abundant on the topics of estimating the underlying variable distribution with constrained observations, we have not identified any similar work—either in academic publications or in airlines' practices—in developing air carrier operations models that incorporates delays. Some NAS software simulation packages, like NASPAC and AIRNET, have the capability to generate the future OAG schedule based on the current schedule and on the specified terminal growth rate, but they lack the capability to adjust the schedule due to excessive delay, which is the core of our present model. The NAS simulation packages may claim to have the capability to construct a new schedule for new hub cities, but they are essentially itinerary builders, and their users have to supply the hub cities' schedules and equipment types.

Modeling Premise

If we assume the carriers are rational (in that they are in business to make profit for their shareholders whatever operating conditions they may have to face), then it is most straightforward to think of solving the problem of modeling delay effects on air traffic operations by estimating delay elasticities of air traffic demand over the entire system. If this could be done, it would be the ultimate model for capturing the future air traffic demand reduction due to delays.

However, it is impossible for many reasons. First, we have not seen any data that show substantial air travel delays, as predicted for the future by the planning models [1,2] It is impossible to estimate without data.

Second, it is hard to separate the delay factor in the overall demand function. Actually, the airports with more delays have more traffic due to high demand from socioeconomic factors. Other ways that indirectly solve the problem must be sought.

Here, let us explore what the carriers will do in the future, if they are rational. Suppose there are no excessive delays. Then they will conduct their operations as predicted by the unconstrained model in the last chapter. The reason that they will have such schedules is that they are the most economical ones, absent excessive delay.

Now suppose that some flights at some airports at some times during the day have to be canceled due to excessive delays induced by excessive demand. Do those cancellations affect schedules in other parts of NAS where flights do not have excessive delays? We believe not, since flight schedule is specified to meet the traveling demand. Of course there will be some network effects, but here we concentrate on the main effects.

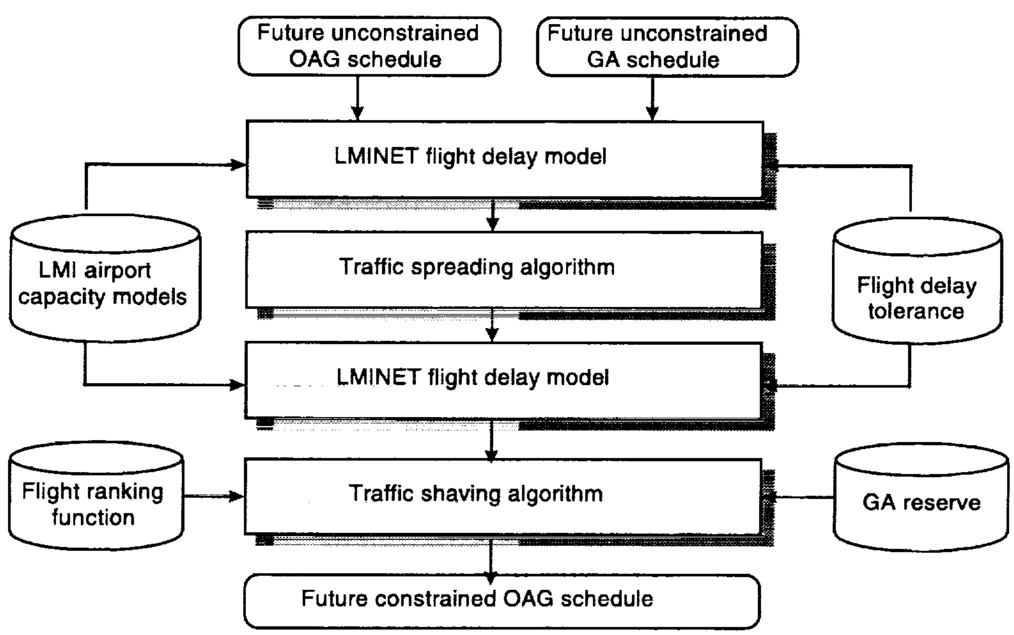
The following assumptions seem to be reasonable for modeling an air carrier's operations due to flight delays:

- As demand increases, the airlines will keep growing their schedules to meet it according to the unconstrained model, until the resulting operations produce excessive delays.
- The delays chiefly important in airlines' schedule planning are those that occur in terminal areas due to insufficient runway and taxiway capacities in universally good weather conditions. Capacities of TRACON and en route sectors have second-order effects on limiting air traffic growth [1,2,8]
- Any canceled flights must have the least economic value to the carriers. Since the variable cost to carry an additional passenger is so negligible in the flight operation, and most other costs must be distributed to the entire network, we are not going to consider cost explicitly in modifying the schedules. Instead, we focus on revenues.

MODEL FOR FORECASTING CONSTRAINED SCHEDULE DUE TO INSUFFICIENT CAPACITIES

Figure 4-1 is a schematic of the model that we propose for forecasting air carriers' future flight schedules as they respond to increasing delays.

Figure 4-1. A Model for Forecasting Constrained Flight Schedule in the Future



The following subsections detail the input data sources, individual components of the model, and their working and usage. A user is free to modify or use any other comparable data input for the model. For the sake of completeness, the model for forecasting the unconstrained flight schedule, which is part of the overall model, is included.

Current Flight Schedule for Air Carriers

This is a data input; we use the OAG schedule.

Current Flight Schedule for General Aviation

This is a data input; we use the schedule extracted from ETMS based on April 8, 1996.

Terminal Area Forecast

This is a data input that provides airport-specific traffic (operations) growth rates. We use the growth rates given in the FAA's TAF.

Unconstrained Air Traffic Demand Forecast Model

For the air carriers, we expand the current OAG schedule using the TAF's airport-specific growth rates and the Fratar algorithm. For GA, we expand the current representative GA schedule, determined from ETMS data, using the TAF's airport-specific growth rates for GA and the simple algorithm of multiplying the departures by the GA growth rate of the departure airport. The reason that we do not use any sophisticated algorithm here is that GA flights do not follow any particular schedule, and they tend to leave for a limited number of airports from a given departure airport, thus making the Fratar algorithm unstable.

Airport Capacity Models

This is a data input. For each of 64 LMINET airports, we need two sets of models: the runway capacity model and the taxiway capacity model. The runway capacity model is characterized by a piecewise linear Pareto frontier characterized by four parameters. The taxiway capacity model is based on the *M*/*M*/1 delay model and by matching the mean delays at each airport to those in the FAA's PMAC. (References [1] and [2] give details of these models.)

Airport Delay Tolerance

This is a data input. This parameter controls the maximum allowable demand. The user has the flexibility to specify the tolerance individually for each airport, and by departure or arrival. The tolerance can be imposed either as a maximum average delay or as a minimum probability that the delay will not exceed 7.5 minutes. (Air carriers and DOT usually use delays of more than 15 minutes to calculate the delay statistics, which splits to 7.5 minutes for both departure and arrival). The relationship between the maximum tolerable average delay minutes and the minimum tolerable on-time probability depends on the initial condition of the airport (whether the airport is busy or not during the last time period) and on the capacities of the runway and taxiways. Numerical calculations for the 64 LMINET airports show that the two tolerances share a fairly stable relationship, with an error of 1 to 2 percent of the on-time probability, which is shown in Table 4-1.

Table 4-1. Maximum Tolerable Average Delay Minutes Versus Minimum Tolerable On-Time Probability

Maximum Tolerable Average Delay Minute	5.0	4.0	3.0	2.0
Minimum Tolerable On-Time Probability	78%	85%	92%	98%

The overall on-time percentage for 1996 as reported by ASQP is 78.9. This statistic includes delays caused by many sources. What is desired here is the tolerance just for the delay caused by the insufficient airport capacities, which certainly must be higher than the statistics from ASQP. Based on Table 4-1, the maximum tolerable delay, system-wide, must be less than 5 minutes. In Chapter 2, we derived the average padding in the schedule, based on the flights among the 29 FAA large airports, to be 3.83 and 2.71 minutes for departure and arrival, respectively. Since these paddings are estimated for all the flights, whether the airport is busy or not, the system-wide tolerance must be higher, because that tolerance must be based on the padding when the airports are busiest. On the other hand, the 29 FAA large hub airports, being the busiest in the United States, experience most of their operations while operating close to their capacities. The tolerances inferred from these airports should not be far from the right one for the overall network.

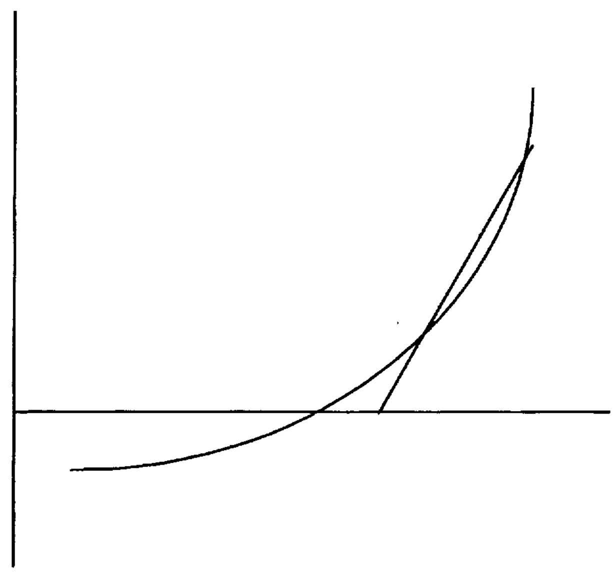
Accordingly, throughout this report when we present the illustration of this model, we use 3.83 and 2.71 minutes as the maximum tolerable average delays for departures and for arrivals, respectively. (Very likely, the levels of tolerance are different at different airports. Lacking data on which to base good estimates of airport-specific tolerances, however, we used common values for all airports.)

Airport Delay Model

This model computes the delay distributions for both departures and arrivals, based on demands and capacities. Appendix A details the development of our model, which is based on accurate numerical solutions of the exact equations for two-member tandem queues representing arrival and departure processes.

Here is how our model determines the maximum tolerable demand, given the maximum tolerable delay, using the airport delay model. We do this with an application of the secant method (Figure 4-2).

Figure 4-2. The Secant Method



For our problem of finding the maximum tolerable demand such that the maximum average delay is T, our algorithm works as follows. Let x be the unconstrained demand; x_1, x_2, x_3 be successive approximations of the maximum tolerable demand to meet the target; and y_1, y_2, y_3 be the average delays associated with x_1, x_2, x_3, respectively.

Step 0: *Let* $x_1 = min(x, 0.9 \times min(runway\ capacity, taxiway\ capacity))$.

If $y_1 \leq T$ and $x_1 = x$, stop.

Step 1: If $y_1 > T$, then let $x_2 = 0.95 \times x_1$.

If $y_1 < T$ and $x_1 < x$, then let $x_2 = min(x, 1.05 \times x_1)$.

Step 2: If $y_2 = T$ then stop; else apply the secant method to generate x_3,

$$x_3 = x_2 - \frac{x_2 - x_1}{y_2 - y_1}\left(y_2 - T\right), \quad \text{[Eq. 4-1]}$$

If $x_3 > x$, then let $x_3 = x$. If $x_3 < 0$ then let $x_3 = 0$.

$x_1 \leftarrow x_2, \quad y_1 \leftarrow y_2;$

$x_2 \leftarrow x_3, \quad y_2 \leftarrow y_3.$

The solution of the secant method is the maximum allowable demand. It turns out that the secant method converges fairly quickly for our problem. We have at most needed three iterations in step 2 with 0.1 convergence bound, which corresponds to at most an uncertainty of 0.5 operation/hour in demand.

Traffic Spreading Model

When demand exceeds maximum tolerable demand, we adjust the schedule to "spread" excess demand into neighboring epochs. This procedure is based upon the result proved in Appendix C, that balancing arrival and departure operations across all epochs in a day will maximize the total number of operations.

If the demand in 1 hour exceeds the maximum tolerable demand, the excess demand—i.e., the difference between demand and maximum tolerable demand—is accommodated by spreading it evenly into the neighboring hours (1 hour before and 1 hour after), so that in each epoch all the flights have the same proportion of retained flights from the original demand and the same proportion of spread flights from the neighboring hours. Due to the nature of our problem, the number of flights in an epoch is not necessarily an integer.

Flight Ranking Function

This is a data input that will rank the desirability of flights. It will be used by the traffic shaving module (described later) to cut flights when necessary to meet the maximum tolerable delay.

Ideally, the desirability of a flight should be directly related to its revenue contribution, or fares and load factor, computed based on the historical data, but those data are not available to us. As stated before, they are the most closely guarded of all the air carriers' data. Instead, we use a surrogate, defined as the number of connection possibilities of a flight, which is the sum of the total arrivals to the departure airport 1 hour before the flight departure and the total departures to the arrival airport 1 hour after the arrival.

The implication of this policy is that flights to or from small airports will be cut first. Flight connection is widely used by the market planning department of the carriers in forecasting passenger demand.

Airport GA Reserve

This is a data input per airport and per epoch. It is the number of reserved GA slots, which the GA demand cannot exceed. The unused GA slots cannot be used by commercial flights even if the GA demand is below the assigned number of slots. This is currently airport operating policy under slot control. For the model runs in the report, we have set the GA reserves to 0 throughout.

Traffic Shaving Model

Some flights have to be cut if the combined commercial and GA demands exceed the airport capacity. The flight with least desirability and the flight that is mutually least desirable from both departure and arrival airports will be cut first. While the least desirable flight is well-defined by the ranking function, finding it involves a search throughout the entire network on both the departure and arrival airports' lists.

If the departure demand exceeds the capacity at one airport and we cut all the excessive departure flights in order to meet the airport capacity, then it is quite possible that some additional departures from this airport will also have to be cut due to the cutting of arrival flights at some airports, which will cause under-utilization of airport capacity even with excessive demand. Another problem with cutting all the excessive demand from airport to airport is that the flight-ranking function will be significantly modified; the sequence in which the airports' excessive demand is cut makes a big difference.

We solved this technical problem by incrementally shaving off excessive demand throughout the entire network. First, we start off with a padding added to the airport capacities such that there are no excessive flights. Then the padding is reduced by an increment, the lowest-ranking excessive flights are cut off, and the fight-ranking function is updated. This process continues until the padding is 0.

This algorithm reduces the aforementioned problem to the minimum that the demand may be over-cut by a delta at the worst, and the flight-ranking function is current.

Measures of NAS Operation Concentration

One can directly obtain a picture of network concentration of NAS operations from the TAF for the current time and for the future. But the TAF misses one important element of airport operations: time distribution of flights. We also need to study the intensity or the concentration of the air carriers operations at the airports, which is characterized by the bank operations at hub airports. First, measures to quantify the airport operations concentration must be determined.

It seems that the Lorenz curve and Gini index, widely used in welfare economics for measuring wealth inequality, can be adopted for measuring the concentration of airport operations. Our basic data are a set of $64 \times 24 = 1{,}536$ cases of operations counts, one for each epoch at each airport.

To build the Lorenz curve, we first sort the operations counts in increasing order. We then index the cases in increasing order from 1 to 1,536. Cases with identical counts are indexed arbitrarily, subject to the constraint that the cases' counts are a monotone nondecreasing function of their indices. Let the *i*th indexed operations count be O_i. Then the Lorenz curve is the set of points (i/1536, O_i/C), where C denotes the total number of operations. If every airport and hour had the same number of operations, the Lorenz curve would be the straight line from (0,0) to (1,1). In the other extreme case, that of total concentration of operations, in which all operations take place at just one airport in just one epoch, the Lorenz curve would be 0 for abscissas larger than 0 and less than 1, and 1 for abscissa 1.

The Gini index is defined as twice the area between the Lorenz curve and the line (0,0) to (1,1) that represents the uniform distribution. One can see that the Gini index ranges from 0 to 1, with 0 corresponding to the uniformly distributed case and 1 corresponding to the completely concentrated cases.

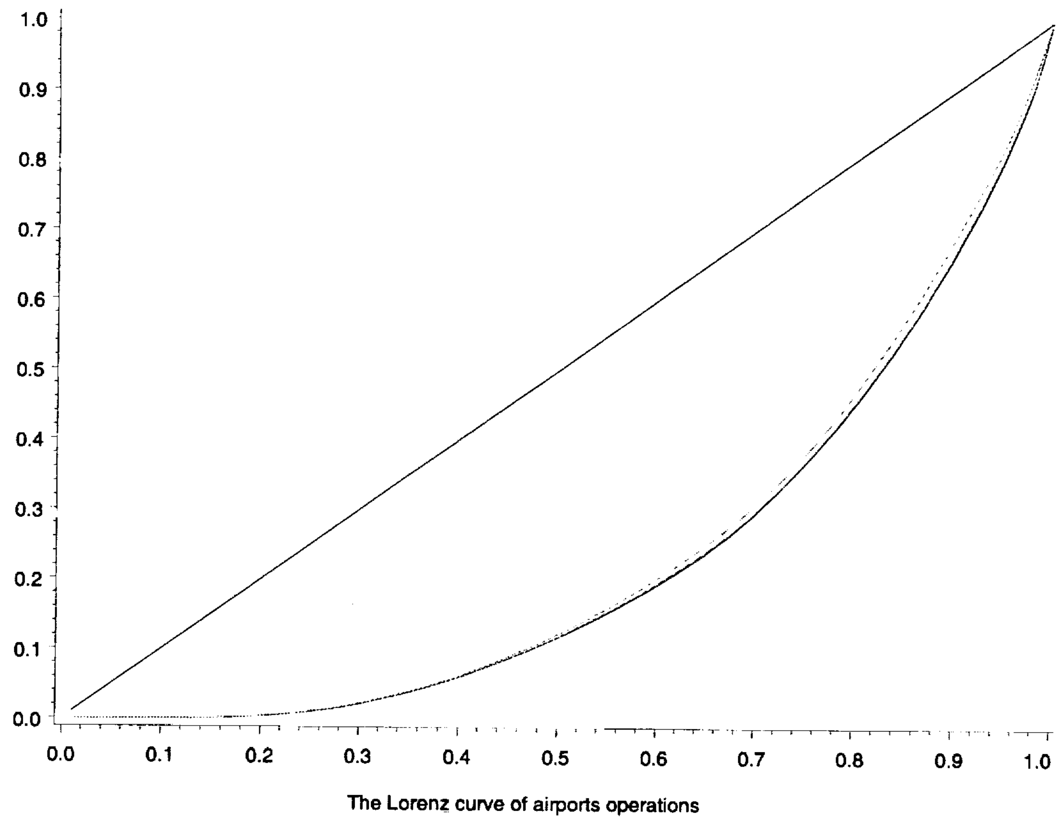
If one schedule dominates another with more concentrated operations, then its Lorenz curve is below the other's and its Gini index is larger.

Figure 4-3 shows the Lorenz curve for the 64 LMINET airports, based on the April 1996 OAG. The Gini index of this case is 0.549, which means that there is still a fair amount of airport operations concentration even among the heavily traveled airports. The Gini index, again for the 64 LMINET airports, but based on the unconstrained future schedule in 2007 from the Fratar algorithm, will be 0.547, an indication of essentially unchanged concentration of airport operations if the TAF forecasts are met and the current air carriers' bank operations at hub cities are kept. The Gini index based on constrained 2007 schedule is 0.534, a reduction that makes sense because some flights at peak hours in hub airports have either been spread to neighboring hours or been cut. The Lorenz curve of this case also is depicted in Figure 4-3.

IMPACT OF ATM TECHNOLOGIES ON TRAFFIC GROWTH

Clearly, by the examples shown in the previous sections, either some traffic growth has to be curtailed because of insufficient NAS capacity, or the capacity must be increased. Since it becomes harder and harder to construct new airports or augment existing airports by paving new runways or taxiways—due to the available land, financing, environmental regulations, etc.—the opportunities to improve capacity by physical improvements to the airports' capacities is quite limited. (We do include "certain-to-happen" projects from the airport improvement plan published by the FAA in our airport capacity models for 2007.)

Figure 4-3. The Lorenz Curves of 64 LMINET Airport Operations



Another way to improve the NAS capacity is by improving airports' efficiencies. To meet the national requirement for sustained air traffic growth, the NASA Aeronautics Enterprise has recently asked LMI to undertake a study to evaluate the contribution of the existing programs [2] Two NASA programs, the Terminal Area Productivity (TAP) and the Advanced Air Transport Technologies (AATT), comprise tools that will essentially either reduce the uncertainties of the aircraft position or enhance the airport utilization by better sequencing of flights and balancing runway usage. Both types of technologies, in addition to some other benefits in terms of safety or flight operating cost, will enlarge airport capacities.

To assess the tools' effectiveness, we mapped the operating characteristics of each set of tools into the parameters of the LMINET airport capacity models [1,2] This yields another set of runway capacity frontiers and another set of taxiway service rates. Figure 4-4 shows an example of a changed airport capacity Pareto frontier. The capacity Pareto frontier with the technology dominates the one without.

Nationwide for the 64 LMINET airports the daily operations in 2007 are 61,668 with ATM technologies versus 60,120 daily operations without ATM technologies, achieving 92.7 percent versus 81.3 percent potential traffic growth from 1996–2007. Table 4-2 summarizes the total daily operations for each of the 64 LMINET airports and for the overall 64 LMINET airports.

Figure 4-4. Airport Runway Capacity Comparisons

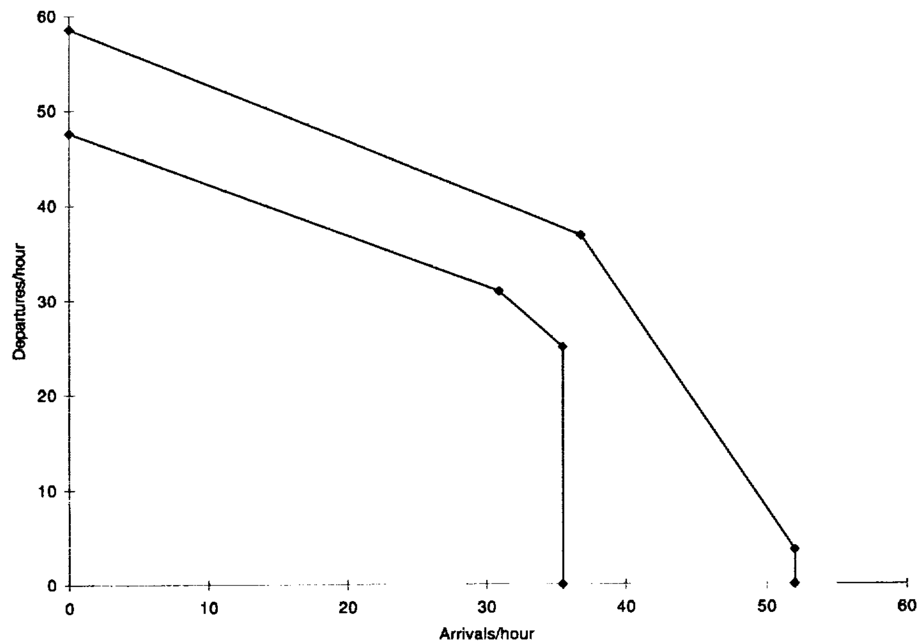


Table 4-2. Daily Carriers Operations by Airport

Airport	1996	Unconstrained 2007	Constrained 2007	Growth attained (%)	Constrained 2007 with tech	Growth attained (%)
BOS	1,262	1,374	1,371	97.4	1,374	99.6
BDL	347	431	422	89.9	428	96.8
HPN	165	209	204	88.8	208	96.9
ISP	119	150	149	97.4	149	99.4
TEB	0	0	0	100.0	0	100.0
LGA	960	1,065	963	2.5	1,056	90.9
JFK	907	1,005	987	81.9	1,005	100.0
EWR	1,306	1,605	1,603	99.6	1,604	99.7
PHL	1,036	1,314	1,311	98.9	1,314	100.0
BWI	702	913	883	85.7	911	99.0
DCA	758	800	774	38.8	790	76.7
IAD	777	963	961	99.0	960	98.5
GSO	275	357	353	94.8	357	100.0
RDU	445	552	547	96.0	551	99.0
CLT	1,184	1,467	1,445	92.3	1,468	100.1
ATL	2,093	2,464	2,350	69.4	2,463	99.9
MCO	890	1,341	1,322	95.7	1,340	99.7
PBI	322	375	364	79.9	372	95.4
FLL	512	696	686	94.7	689	96.2

Table 4-2. Daily Carriers Operations by Airport (Continued)

Airport	1996	Unconstrained 2007	Constrained 2007	Growth attained (%)	Constrained 2007 with tech	Growth attained (%)
MIA	1,348	1,801	1,617	59.5	1,762	91.4
TPA	696	874	873	99.8	874	100.0
MSY	402	481	477	94.3	479	96.6
MEM	916	1,172	1,087	67.1	1,168	98.4
BNA	492	588	586	97.4	587	98.6
SDF	445	564	544	83.5	558	95.2
CVG	1,135	1,711	1,600	80.7	1,660	91.2
DAY	312	361	348	73.3	359	96.4
CMH	446	575	572	97.2	573	98.1
IND	567	788	786	99.0	787	99.4
CLE	798	1,033	1,030	98.7	1,033	100.0
DTW	1,351	1,815	1,742	84.3	1,814	99.9
PIT	1,220	1,470	1,468	99.4	1,467	98.8
SYR	285	353	347	90.7	353	100.0
MKE	460	584	563	83.3	582	98.8
ORD	2,367	2,793	2,698	77.7	2,793	100.0
MDW	479	608	599	93.3	607	99.2
STL	1,414	1,753	1,618	60.1	1,744	97.5
IAH	1,053	1,495	1,415	82.0	1,488	98.5
HOU	397	487	466	77.1	449	57.6
AUS	303	406	392	86.1	403	97.1
SAT	363	487	479	93.2	479	93.3
DAL	377	523	523	100.0	515	95.1
DFW	2,443	3,372	2,575	14.2	2,778	36.0
MSP	1,257	1,611	1,524	75.3	1,585	92.5
MCI	548	669	667	98.4	669	99.9
DEN	1,148	1,391	1,390	99.7	1,391	100.0
ABQ	366	490	487	97.4	486	96.5
ELP	206	233	225	73.0	230	90.3
PHX	1,312	1,779	1,632	68.6	1,736	90.8
SLC	830	1,138	1,129	97.1	1,136	99.4
LAS	1,024	1,510	1,506	99.2	1,510	99.9
SAN	611	798	750	74.3	786	93.4
SNA	303	418	402	86.4	414	96.9
LGB	38	54	40	13.6	44	35.8
LAX	2,042	2,517	2,457	87.3	2,487	93.8
BUR	289	391	386	95.3	384	92.9
ONT	362	442	434	89.4	437	93.3
RNO	295	414	410	96.7	411	98.0
SMF	354	489	480	93.3	484	96.4

Table 4-2. Daily Carriers Operations by Airport (Continued)

Airport	1996	Unconstrained 2007	Constrained 2007	Growth attained (%)	Constrained 2007 with tech	Growth attained (%)
OAK	698	887	877	94.7	884	98.7
SFO	1,111	1,443	1,422	93.8	1,442	99.6
SJC	363	506	499	94.8	501	96.2
PDX	731	987	986	99.3	987	99.9
SEA	1,062	1,319	1,318	99.6	1,318	99.6
Total	49,073	62,656	60,120	81.3	61,668	92.7

SENSITIVITY OF THE TOLERANCE PARAMETERS

The tolerance to the delays caused by insufficient airport capacities seems to be the variable that the user will most likely modify. Indeed, the attained traffic growth, relative to the unconstrained forecast, is sensitive to the tolerance levels selected. The tolerance levels presented in Table 4-3 give the likely upper and lower bounds based on the corresponding on-time probabilities. These results lead us to conclude that NAS can accommodate about 75 to 85 percent of potential traffic growth in a decade, absent improvements such as the NASA ATM technologies.

Table 4-3. Delay Tolerances and Total System Operations in 2007

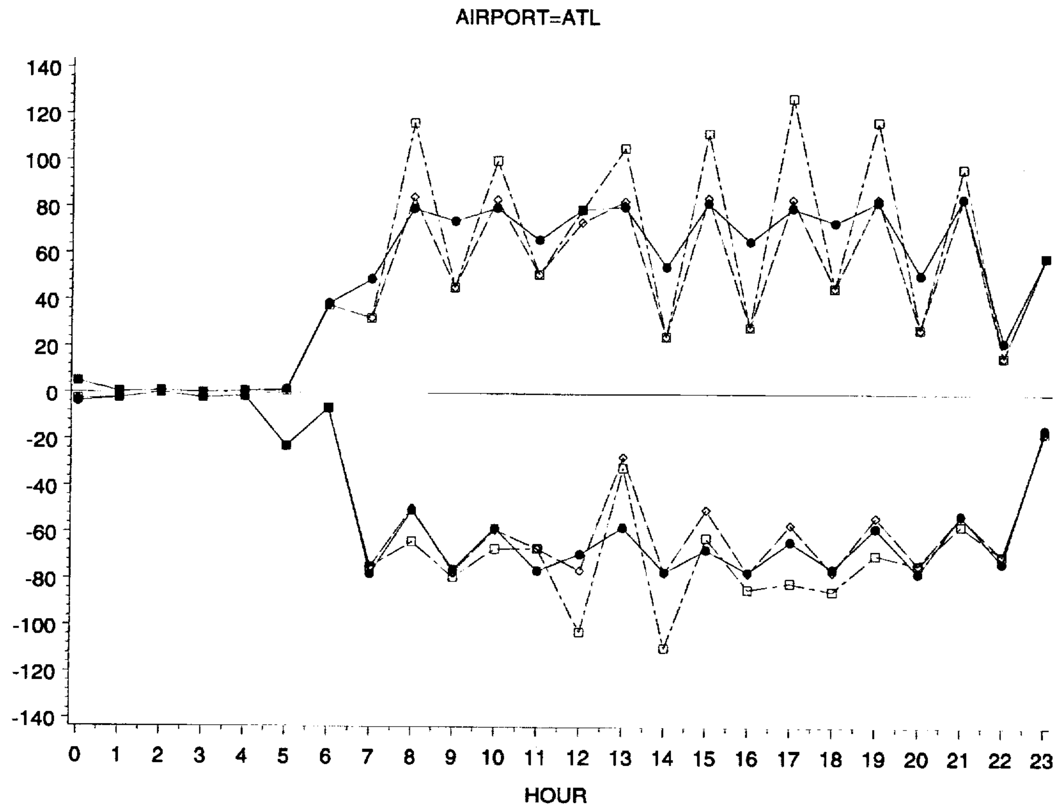
Departure tolerance (minutes)	Arrival tolerance (minutes)	Total daily operations	Attained traffic growth (%)
3.0	2.0	59,106	73.9
3.83	2.71	60,120	81.3
5.0	4.0	60,849	86.7

EXAMPLES OF MODEL RUNS

Smooth Operations

Figures 4-5 and 4-6 compare airport operations of different operating policies at ATL, where the airport capacities are assigned by the rule that the maximum tolerable average departure delay is 3.83 minutes and the maximum average arrival delay is 2.71 minutes. Three curves are presented: *unconstrained*; *no spreading*, in which demands causing excess delays are eliminated; and *spreading*, where excess demands, as computed in the no-spread case, are first spread to the neighboring times, before applying the tolerance again to reduce the excess demand.

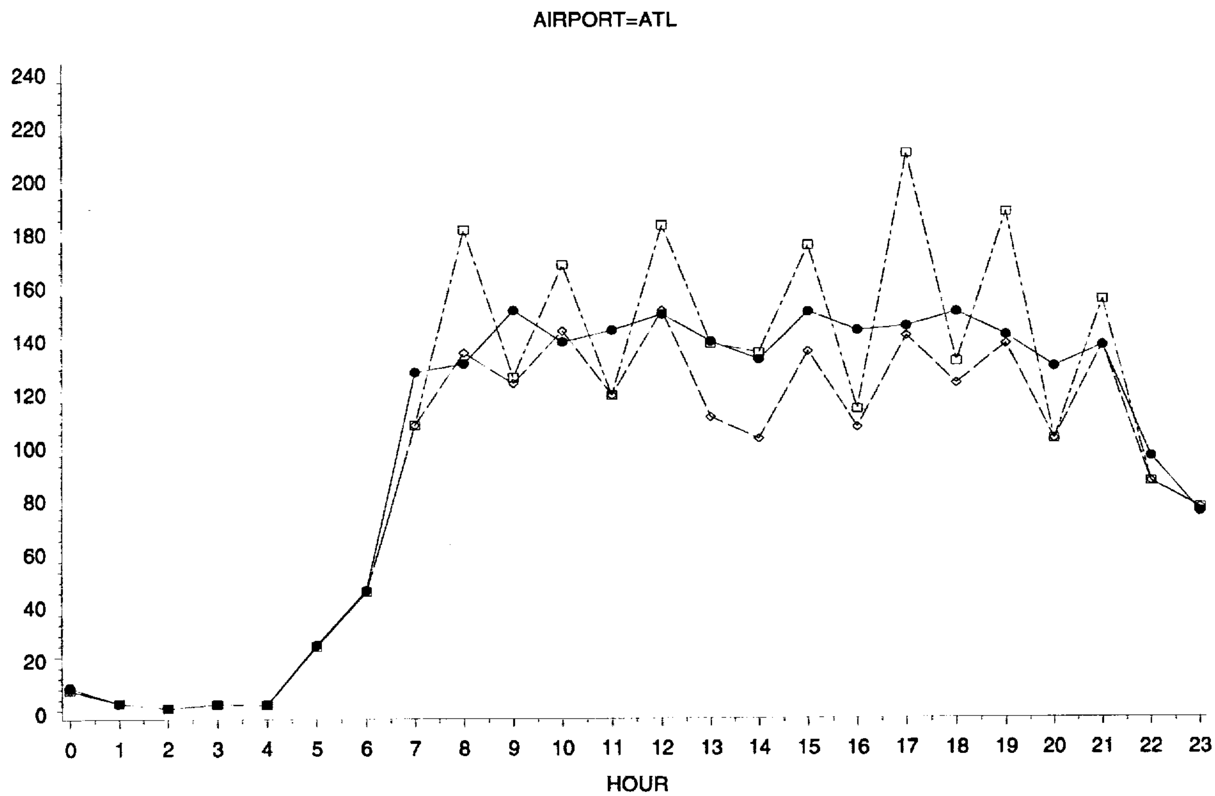
Figure 4-5. Operations at ATL



Note: Square = unconstrained; diamond = no spreading;
dot = spreading; positive curve = departure; negative curve = arrival.

One can see that if the carriers continue to follow their current bank operation practices, the operation in the future will mimic the current one—albeit with less intensity because the airport cannot sustain the traffic. Just by spreading the excess demand, departure, or arrival, to the neighboring time—which does not much change the point-to-point passenger demand based on the time-of-day demand pattern—ATL will have a smoother operation, with the total daily operations increasing from 2,109 to 2,350 out of the unconstrained demand of 2,464, rising from accommodating just 4.4 percent of the potential traffic increase to accommodating 69.4 percent of potential traffic increase. For the 64 LMINET airports, spreading will increase the total number of daily operations from 58,501 to 60,120 out of the unconstrained demand of 62,656, attaining 69.4 and 81.3 percent of potential operations growth.

Figure 4-6. Total Airport Operations (Departure + Arrival) at ATL



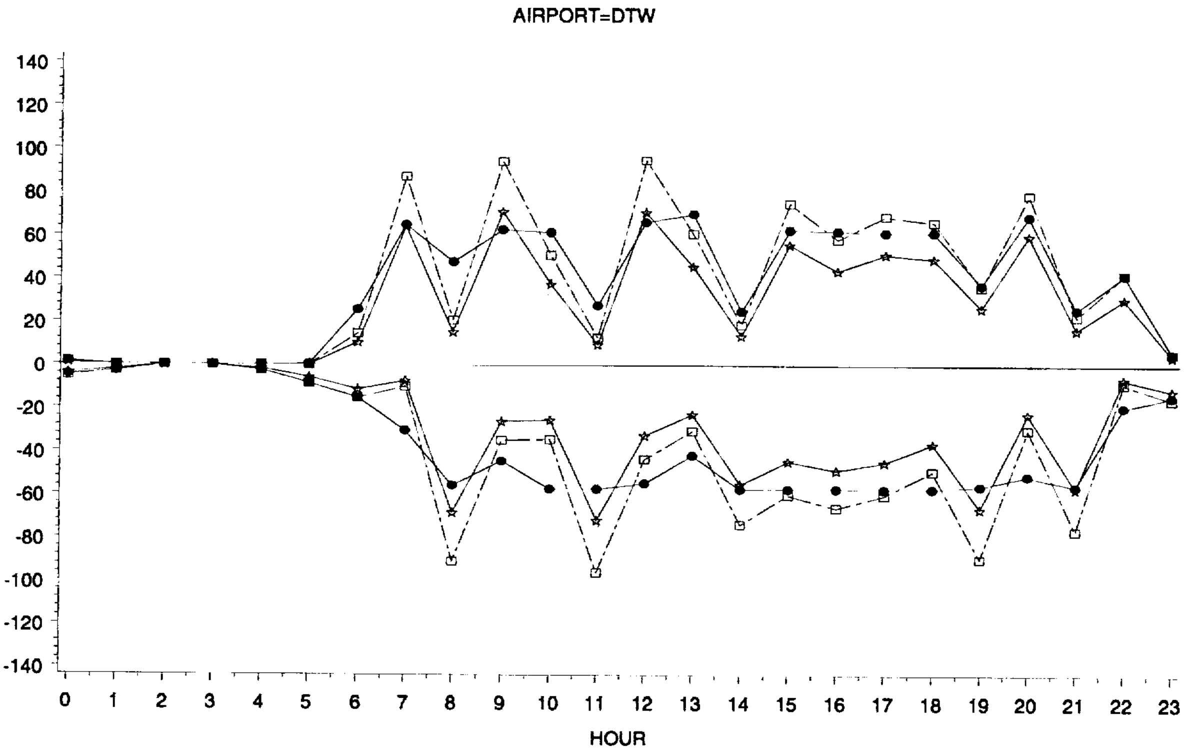
Note: Square = unconstrained; diamond = no spreading; dot = spreading.

Spreading works magically to improve the total number of operations without costing too much to the local travel, but it does have negative impact on the connecting traffic. Suppose an airport has a capacity of 100 operations per hour for combined departures and arrivals, which can trade with each other one for one. If the carriers opt to conduct their bank operation so that there will be 100 arrivals at one hour and 100 departures in the following hour, then each arrival will have 100 connections and the total connections for the bank is 10,000. If the carriers opt to even out their operations at 50 arrivals and departures at both hours, then each arrival will have 50 connections and the total connections for arrivals within an hour are 2,500. Thus, a traveler's options when using the hub as a connection point are severely reduced, from 100 connections to 50 connections, a 50 percent reduction, unless the traveler waits another hour to have the same range of connection options as in the bank operation case.

For the carriers, for the 2-hour period over which a bank operation lasts, there will be 50,000 connections, which implies 50 percent connection reduction. Of course, this is the limiting case. In the current network, about 60 percent of passengers are direct. Also, not all the possible connections have profitable passenger demand, and the carriers can still couple the most likely connections.

Another example of smooth operations by carriers, where the excessive flights are either spread to the neighboring hours or cut, is depicted by the operations at DTW (Figure 4-7).

Figure 4-7. Operations at DTW



Note: Star = 1996; square = unconstrained 2007; dot = constrained 2007.

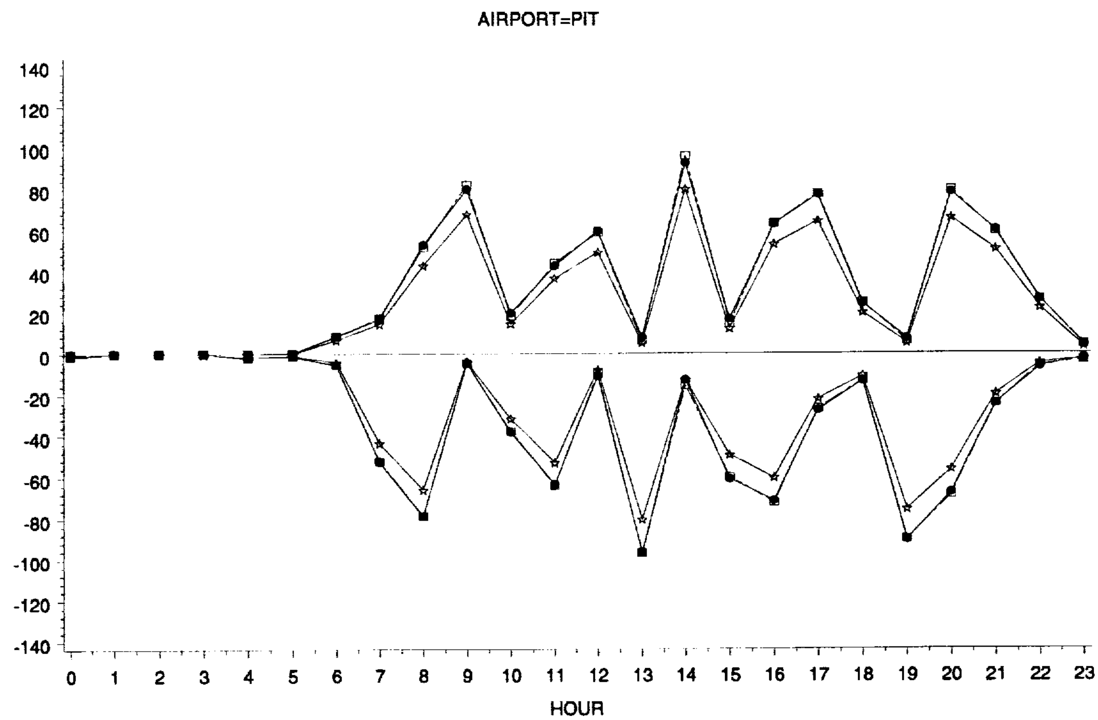
However, not all the airports will see constraints imposed by the airport capacities. An example of such an airport is Pittsburgh International Airport (PIT), a major hub for US Airways. PIT's existing capacities will meet the potential demand in 2007 (Figure 4-8).

Effect of ATM Technologies on Operations

With NASA-proposed ATM technologies, the improvement to carriers' operations at ATL is shown in Figures 4-9 and 4-10.

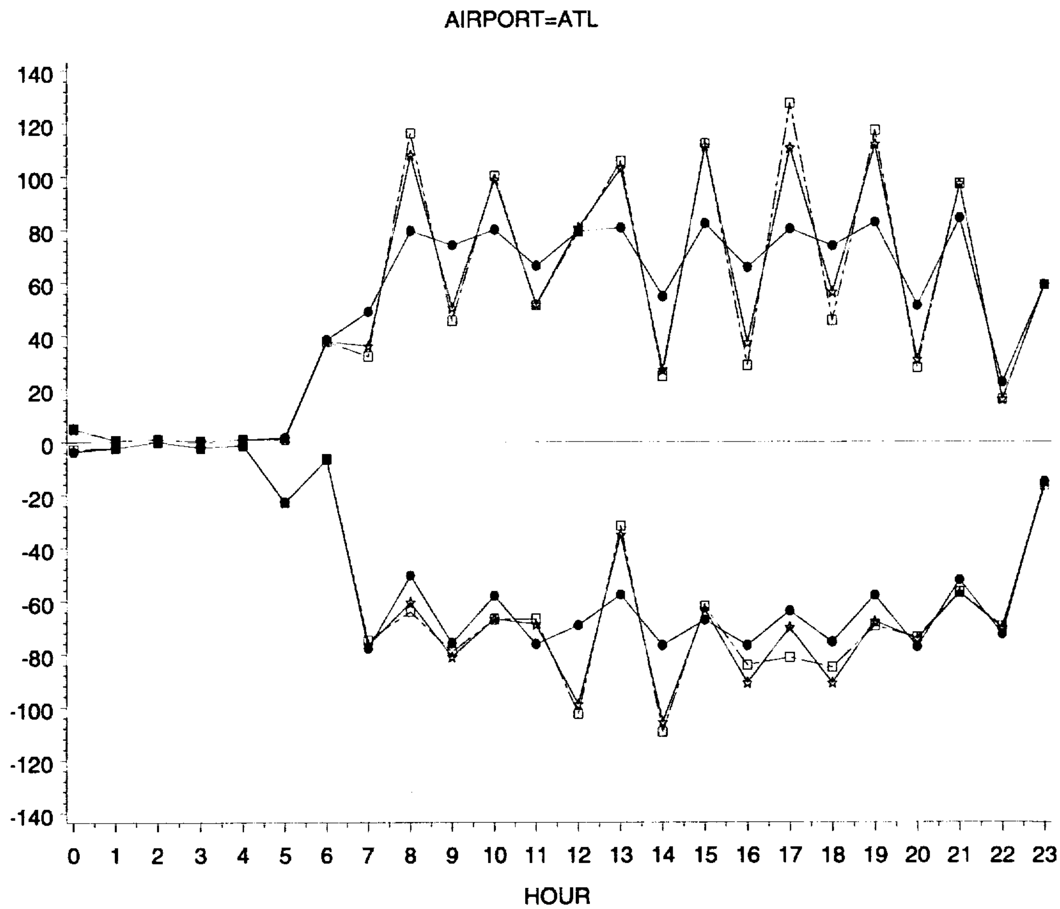
One can see that with the ATM technologies, operations at ATL are very close to the unconstrained case, conducting 2,463 daily operations with ATM technologies versus 2,350 daily operations without ATM technologies, achieving 99.9 percent of potential traffic growth from 1996 to 2007 with ATM technologies versus 69.4 percent without ATM technologies. With the ATM technologies, the bank operations are preserved.

Figure 4-8. Operations at PIT



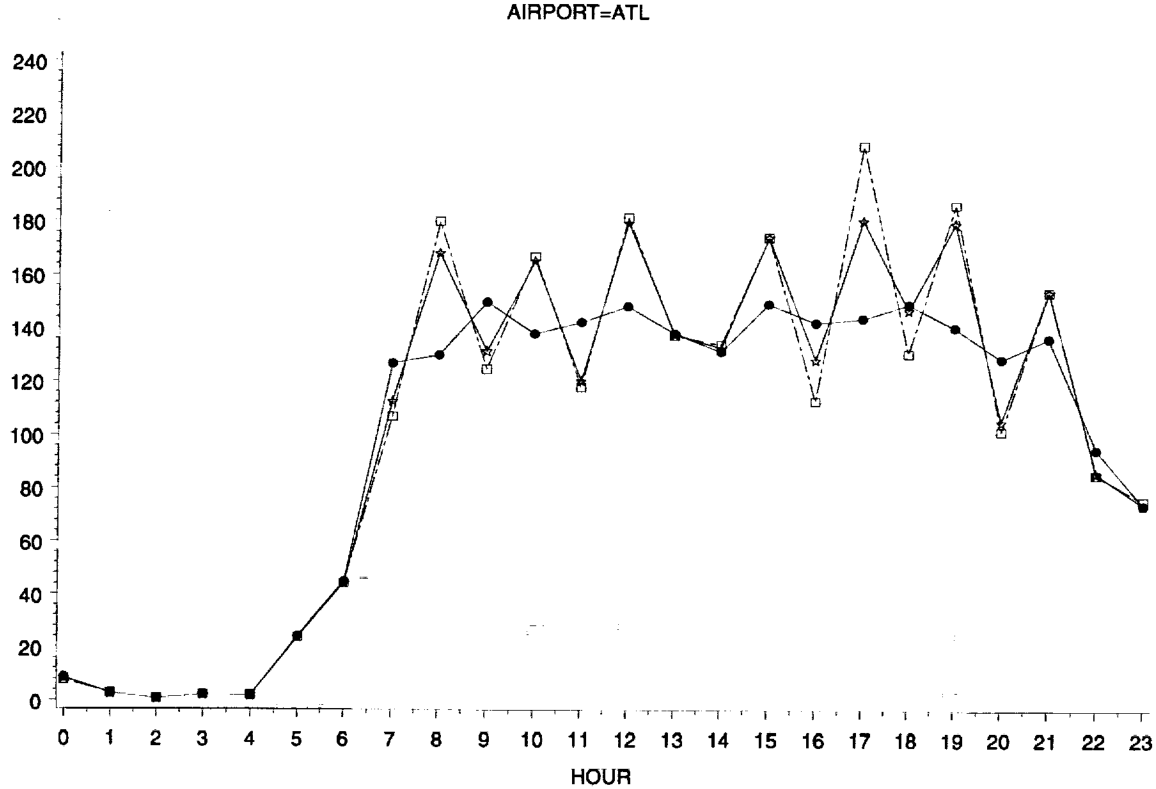
Note: Star = 1996; square = unconstrained 2007; dot = constrained 2007.

Figure 4-9. Operations at ATL With ATM Technologies



Note: Square = unconstrained operation; dot = constrained operation; star = constrained operation with ATM technologies.

Figure Chapter 4 -10. Total Operations at ATL with ATM Technologies



Note: Square = unconstrained operation; dot = constrained operation; star = constrained operation with ATM technologies.

Chapter 5

An Economic Approach to Estimating the Impact of Congestion on Air Carrier Operations

Introduction

This chapter introduces an economic approach to estimating the impact of congestion on air carrier operations. The basic premise of this approach is that delay imposes additional costs on air carriers, which are passed on to consumers in the form of higher fares. Because air travelers are sensitive to price changes, the higher fares work to mitigate the growth rate of demand. Thus, the economic approach seeks an equilibrium in which the increased costs of delay are balanced against the consumers' willingness to pay for air travel. Implicitly, this approach is rooted in the assumption that air carriers will continue to operate in highly congested environments, but will charge prices that cover the additional costs of delay.

Our approach links models of airspace capacity and delay with models of air carrier operating costs and air travel demand. Specifically, we link delay results from LMINET to inputs of the Aviation System Analysis Capability (ASAC) Air Carrier Investment Model (ACIM) in order to estimate the impact of congestion on industry equilibrium. The remainder of this chapter provides a brief introduction to the ACIM, discusses our methodology in detail, and presents our results. An introduction to LMINET is provided in Appendix A.

Aviation System Analysis Capability Air Carrier Investment Model

To link the technology of flight with its economics, ASAC requires a parametric model of airline costs and air travel demand. As such, the ACIM incorporates air travel demand, airline productivity, input prices, and profit considerations into the airline investment decision. The result is a forecast for air travel throughput, fares, airline employment, and aircraft fleet requirements for a given technology scenario. By comparing the results from a technology scenario with those from a baseline scenario, we, therefore, obtain estimates of throughput benefits.

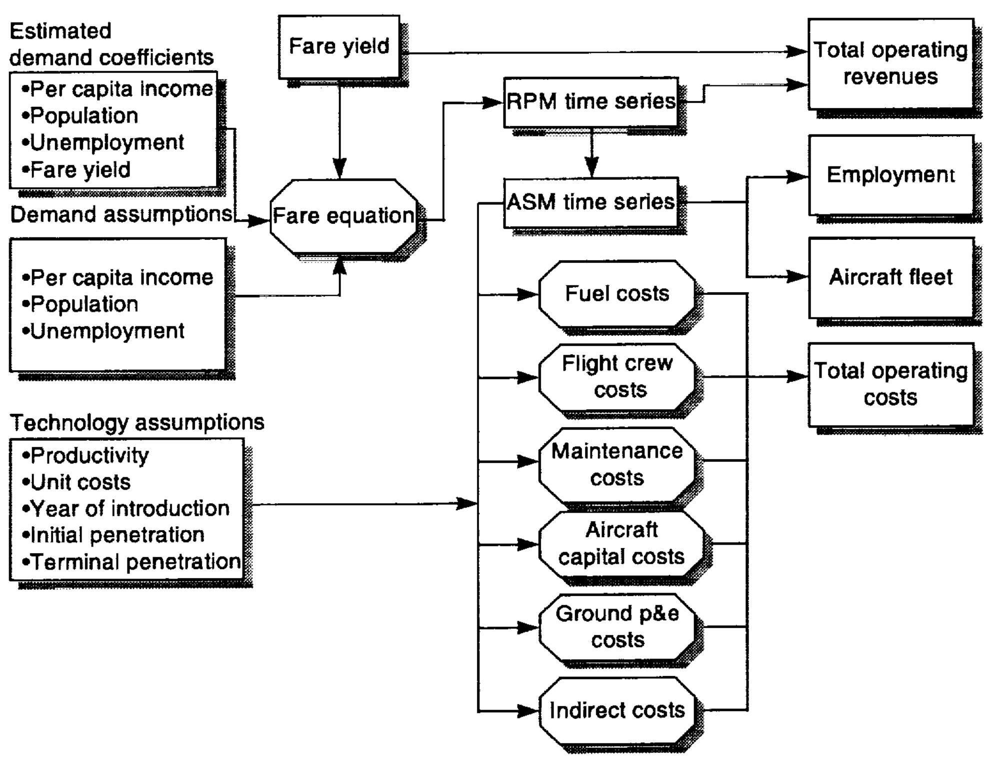
The ACIM consists of four core modules: the U.S. Econometric Module, U.S. Functional Cost Module (FCM), Asian Econometric Module, and European Econometric Module.[1] In addition, extension modules have been developed to

[1] See Reference [6] for more information on the ACIM.

map the aggregate fleet requirements into seat-size categories and estimate the impact of changes in technology on U.S. aircraft manufacturing and related employment. The distinction between the U.S. Econometric Module and the FCM is the approach by which airline operating costs are estimated. The econometric module uses an econometric approach to estimate air carrier costs, while the FCM uses activity-based costing. For this study, we employed the FCM exclusively.

In the FCM, operating costs are calculated for six functional cost categories as functions of input prices, input factor productivities, and total output. As shown in Figure 5-1, the cost categories consist of fuel, flight personnel labor, maintenance, aircraft capital, ground property and equipment capital, and a residual category termed other indirect. The measure of output that drives the cost calculations is available seat miles (ASM). The FCM solves for industry equilibrium by iterating fare yields until the specified profit constraints are satisfied. Thus, any changes in airline operating costs are passed on to the traveling public in the form of changes in fares. Implicitly, such analysis assumes that the commercial air travel industry will remain price competitive.

Figure 5-1. Functional Cost Module Schematic



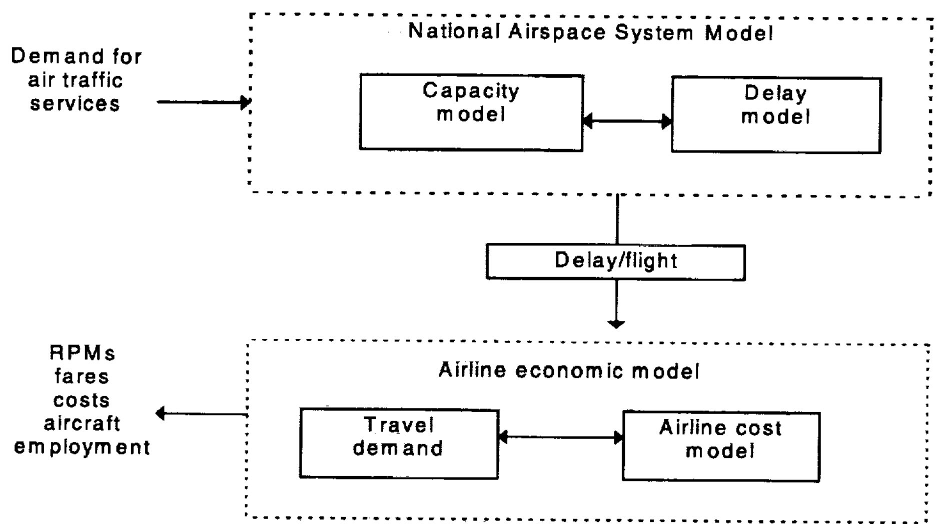
The model is based upon DOT Form 41 reports for 26 of the largest U.S. air carriers. The FCM represents approximately 91 percent of the 1995 U.S. flag scheduled traffic. The default assumptions of the model determine an unconstrained forecast that compares quite favorably with other published forecasts by Boeing and the

FAA [4,5] It is important to note, however, that these unconstrained forecasts are predominately driven by demand and have little or no consideration of possible capacity shortfalls. The following section discusses our modifications to the unconstrained forecast in order to measure the impact of delay. Scenarios for delay-reducing technologies are subsequently compared with this baseline forecast.

USING THE ACIM TO EVALUATE CHANGES IN SYSTEM THROUGHPUT

The basic premise of using the ACIM to evaluate the impact of delay on throughput is that delay imposes additional costs on air carriers, drives up fare yields, and depresses air travel demand growth. The result is a revised forecast for throughput, as measured by revenue passenger miles (RPMs), which takes into account the impact of delay. We then convert RPMs to both operations and enplanements to get a more complete picture of the impact of delay and delay-reducing technologies. Figure 5-2 is a schematic of this approach.

Figure 5-2. Schematic of Air Carrier Investment Model Approach



As shown in Figure 5-2, the primary linkage between the NAS Model—LMINET—and the ACIM is the projected delay per flight. The block time required to complete a flight segment of a given length is a measure of aircraft productivity employed by the ACIM. Generally, with the stage length held constant, shorter block times imply more productive aircraft and consequently more departures per aircraft per day. The ACIM accepts changes in aircraft productivity as changes in aircraft block speed which, by definition, is inversely proportional to changes in block time. Aircraft block speed, in conjunction with the number of seats per aircraft, therefore determines the number of ASM that can be flown by a given aircraft. Since ASM drive costs, the changes in aircraft productivity implied by changes in delay are central to the model's calculations.

As shown in Table 5-1, we begin with the total delay minutes from the NAS model for each scenario. Since the FCM has no representation from GA and air taxi, and only limited representation from commuter carriers, we deduct the proportion of operations and delays attributed to these groups. The remainder is denoted as delays attributed to commercial carriers. The next column contains the projected number of commercial operations from LMINET, with which we calculate the average delay per flight for each scenario in the final column.[2]

Table 5-1. FCM Delay Inputs

Scenario	Total delay minutes (millions)	Total commercial delay minutes (millions)	Commercial operations (millions)	Average delay minutes per flight
1996 baseline	167.7	94.3	11.9	15.90
2007 baseline	525.5	312.8	15.4	40.72
2007 all technologies	229.5	136.6	15.4	17.71

Translating these delay projections into the airline economic model requires several assumptions about the likely airline response to the increase in congestion. We assume that variable operating costs increase proportionally with the increase in delay as scheduled block times get longer and less predictable. However, since the airlines operate a highly coordinated schedule of aircraft and crew movement, delay in one portion of the system can have repercussions system wide. For this reason, the cost of delay can be significantly larger than variable operating costs indicate.

To address this issue we employ a set of cost multipliers from a study by American Airlines [7] Since the value of the cost multiplier varies significantly by time of day and the duration of the initial delay, it was necessary to apply the multipliers within LMINET directly. The cost multipliers, therefore, are implemented as delay multipliers, although the distinction is irrelevant since the impact of congestion is determined by the product of delay and cost, i.e., because multiplication is commutative. Table 5-2 summarizes the impact of the delay multiplier.

[2] In this section, the *delays* are the ones from LMINET, which are measured against the ideal flight times that would be seen if only one aircraft were using each NAS component. These delays are larger than those measured against a predetermined schedule that includes buffers.

Table 5-2. Delay Multiplier Values

Scenario	Delay minutes per flight without multiplier	Delay minutes per flight with multiplier	Implicit value of delay multiplier
1996 Baseline	15.90	33.65	2.12
2007 Baseline	40.72	90.29	2.22
2007 All Technologies	17.71	35.51	2.01

The key variables we modify to estimate the impact of delay on throughput are average aircraft block speed (inversely proportional to block time) and aircraft utilization. Converting from changes in delay, shown in Table 5-2, to changes in block speed requires an assumption regarding the average block time in the absence of delay. To specify this assumption,we researched the current average block time for domestic flights departing from, and arriving to, LMINET airports. For 1996, this average block time was 129 minutes. Subtracting the estimated average delay figure for 1996 of 15.9 yields an average block time of 113.1 minutes in the absence of delay. With the average stage length held constant, therefore, the addition of the estimated delay for each scenario to the 113.1 minute base yields projections for average block time.

The final step was to convert average block times to average block speed and to compute the compound annual rate of change in block speed for each scenario. As shown in Appendix B, this computation is derived from Equation 5-1 in which t denotes time in years, the subscript 0 denotes the current time period, and 1 denotes the future time period. Table 5-3 summarizes these calculations:

$$\textit{Annual Change in Block Speed} = \left(\frac{\textit{Average Block Time}_0}{\textit{Average Block Time}_1} \right)^{\frac{1}{(t1-t0)}} - 1 . \qquad \text{[Eq. 5-1]}$$

Table 5-3. Derivation of ACIM Inputs

Scenario	Initial block time (minutes)	Ending block time (minutes)	Annual rate of change in block speed (percent)
Baseline 1996–2007	146.75	203.39	–2.924
All Technologies 1996–2007	146.75	148.61	–0.114

One issue that arises in modeling aircraft productivity in this way is that in order to fly the same schedule in the face of rising delays, the airline would have to purchase additional aircraft. This may be unreasonable since airline financial departments would resist additional investments for increasingly less productive aircraft. We developed an alternate approach, which assumes that airlines would stretch their schedules out during the day while simultaneously increasing the number of daily aircraft block hours. Operationally, this appears as an increase in aircraft utilization (more aircraft hours per day), although the total output as measured by miles flown or RPMs remains constant because of the delay. Specifically, our approach assumes an equal and opposite impact on aircraft utilization in relation to any block speed productivity changes.

Since the unconstrained ACIM forecast already contains certain assumptions regarding aircraft productivity and utilization, we measured changes in productivity from these default values. Table 5-4 summarizes the ACIM inputs for each scenario.

Table 5-4. ACIM Inputs

Parameter	Unconstrained default value (%)	Baseline value (%)	All technologies value (%)
Average block speed (1996–2007)	0.253	–2.661	0.139
Aircraft utilization (1996–2007)	0.00	2.924	0.114

Note: All parameters denote annual rate of change.

An additional issue that arises in modeling system throughput in this way is that the revised forecast generated by the ACIM will have considerably less traffic than was assumed by the NAS model in generating the initial delay estimates. In this case, the initial delay estimates may be overstated given the corresponding reduction in traffic. Thus, the need arises for a feedback loop between the revised traffic forecasts and the estimated delay.

The most accurate method to implement this feedback effect would be to pass revised traffic forecasts from the ACIM to the NAS model and recompute the delay estimates. This process could then be repeated until the difference between subsequent revisions converged. Unfortunately, the NAS model requires more detailed input than the ACIM can provide. Specifically, LMINET requires a flight schedule with highly detailed information on individual flights and airport operations, while the ACIM functions at the aggregate level only. Thus, to implement a feedback loop involving LMINET would require a complex algorithm to distribute changes in the aggregate traffic forecast to the underlying schedule. At this time, no such algorithm exists.

In light of the difficulties in disaggregating the revised traffic forecast for input to LMINET, we developed an alternate implementation for the feedback effect. The approach is based upon a piecewise log-linear approximation of the delay model applied to the aggregate traffic forecasts from the ACIM. To see the logic of this approach, consider that the core function of the delay model is to calculate delay as a function of traffic throughput and system capacity (as determined by the technology scenario assumptions). Thus, for any fixed capacity, the delay model represents a nonlinear mapping from throughput to delay.

Our approach is to use the actual output from the LMINET delay model to estimate this nonlinear relationship for each technology scenario. A good approximation of the delay model was achieved using a piecewise log-linear specification for each scenario. Thus, the results from the initial run of the ACIM are fed to the approximated delay model to produce revised delay estimates. The revised delay estimates subsequently provide new inputs for an additional run of the ACIM. This process is repeated until the change in system throughput for subsequent runs of the ACIM converges to less than one-half of 1 percent. The final iteration of these inputs is summarized in Table 5-5.

Table 5-5. Final ACIM Inputs

Parameter	Unconstrained default value (%)	Baseline value (%)	All technologies value (%)
Average block speed (1996–2007)	0.253	–2.032	0.141
Aircraft utilization (1996–2007)	0.00	2.285	0.112

Note: All parameters denote annual rate of change.

ACIM THROUGHPUT RESULTS

Implementing the ACIM under the previously discussed methodology yields a time series of system throughput, as measured by RPMs, for each scenario. To convert RPMs to operations and enplanements, we use the projected ratio of RPMs to operations and RPMs to enplanements from the 1998 FAA forecast [5] Implicit in the forecast projections are assumptions about increasing average stage length and increasing average seat size, which cause RPMs to grow considerably faster than enplanements and operations.

Comparing estimates from each scenario with the unconstrained forecast provides estimates of the impact of congestion and the benefits of technologies that mitigate delay. Tables 5-6, 5-7, and 5-8 present our results.

Table 5-6. Commercial RPM Results (Billions)

Year	Unconstrained	Baseline	All technologies
1996	563.2	563.2	563.2
1997	593.7	589.1	593.4
1998	625.8	616.2	625.3
1999	659.7	644.5	659.0
2000	695.4	674.1	694.4
2001	733.0	705.2	731.7
2002	765.1	729.8	763.5
2003	798.6	755.2	796.6
2004	833.6	781.6	831.2
2005	870.1	808.8	867.3
2006	908.2	837.1	905.0
2007	947.9	865.1	937.8

Table 5-7. Commercial Enplanement Results (Millions)

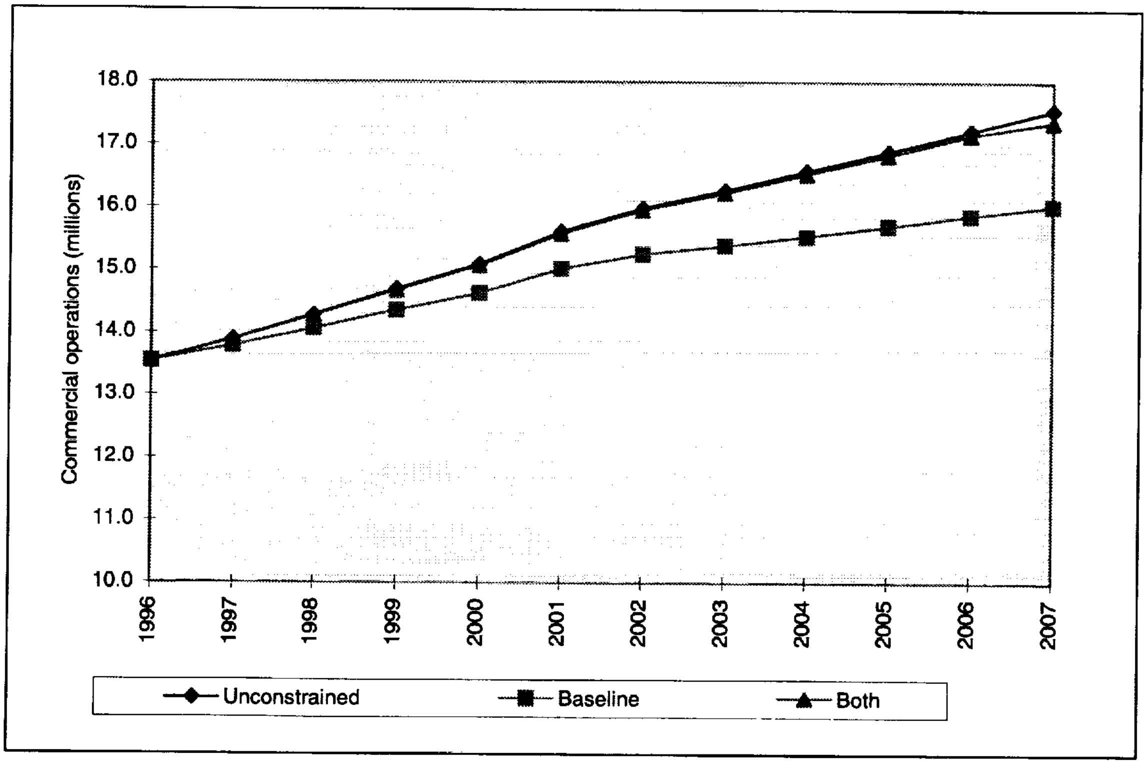
Year	Unconstrained	Baseline	All technologies
1996	593.1	593.1	593.1
1997	625.2	620.4	625.0
1998	659.0	648.9	658.6
1999	694.7	678.8	694.0
2000	732.3	710.0	731.3
2001	772.0	742.6	770.6
2002	805.8	768.5	804.1
2003	841.1	795.4	839.0
2004	877.9	823.1	875.4
2005	916.4	851.8	913.4
2006	956.5	881.6	953.1
2007	998.2	911.1	987.7

Table 5-8. Commercial Operation Results (Millions)

Year	Unconstrained	Baseline	All Technologies
1996	13.54	13.54	13.54
1997	13.89	13.78	13.89
1998	14.28	14.06	14.27
1999	14.69	14.35	14.68
2000	15.09	14.62	15.06
2001	15.61	15.02	15.58
2002	15.99	15.25	15.95
2003	16.27	15.38	16.23
2004	16.58	15.54	16.53
2005	16.89	15.70	16.83
2006	17.22	15.87	17.15
2007	17.55	16.02	17.37

Figure 5-3 depicts the results for commercial operations graphically. As shown, the scenario for all technologies closely approaches the unconstrained forecast through 2007. Basically, the combined technologies allow the projected traffic growth between 1996 and 2007 to occur without increasing delay. Finally, the baseline projections show that growth in operations will be considerably restrained by delay in the absence of any new technologies.

Figure 5-3. Commercial Operations



Chapter 6
Summary and Conclusions

This report presents an integrated set of models that forecasts air carriers' future operations when delays due to limited terminal-area capacity are considered. The suite has two outputs. The more detailed output consists of flight schedules, which convey much useful information about the air carriers' operations, including origins, destinations, and planned block times. The schedules are made by models that restrict traffic growth to levels that the NAS can accommodate with not more than user-specified values of mean arrival delay and departure delay per flight.

The other output is forecasts of commercial RPMs, enplanements, and total operations for the entire U.S. passenger air carrier industry. These results are made by linking econometric models to a NAS model to determine operations levels that will generate user-specified profit levels under delay-induced reductions in productivity.

This report models the industry as a whole, avoiding unnecessary details of competition among the carriers. To develop the schedule outputs, we first present a model to forecast the unconstrained flight schedules in the future, based on the assumption of rational behavior of the carriers.

Then we develop a method to modify the unconstrained schedules, accounting for effects of congestion due to limited NAS capacities. Our underlying assumption is that carriers will modify their operations to keep mean delays within certain limits. We estimate values for those limits from changes in planned block times reflected in the OAG.

Our method for modifying schedules takes many means of reducing the delays into consideration, albeit some of them indirectly. The direct actions include depeaking, operating in off hours, and reducing hub airports' operations. Indirect actions include the using secondary airports, using larger aircraft, and selecting new hub airports, which, we assume, have already been modeled in the FAA's TAF. Users of our suite of models can substitute an alternative forecast for the TAF.

Users can modify other features of our schedule-generating suite. In addition to the TAF forecasts, the parameters users are most likely to want to change are the airport delay tolerances and the airport capacity models. The users may also want to integrate their own individual models into the suite. However, the overall suite's present configuration is sound, and it will give defensible results efficiently.

Turning now to the models that forecast industry-level economic outputs, we note first that these generate remarkably similar forecasts for total system-wide operations to those made by the schedule-oriented models. The forecasts from the economic models lie at the low end of those from the flight-specific models. It is quite possible that the economic models underforecast—while the flight-specific models overforecast—future operations because of their differing treatments of GA and the flight reduction in general.

GA and commercial flights are assumed to have equal relative flight reductions in the economic models, while GA is assumed to be cut first in the flight-specific models. Also, all the flights are assumed to be reduced equally in the economic model, while only flights in the congested airports are cut in the schedule-producing, flight-specific models. This means that more flights are cut in the economic models, since the delays are caused by intense operations in the hub airports.

No matter which models we use, either flight-specific or economic, the methods we have developed have profound implications for the evaluation of ATM technologies and procedures. First, it is likely that we are not going to see the dramatic increases of air traffic delays as predicted by other planning models, because of the airlines' adaptations to the constraining capacities. We see early signs of the validity of our approach and results because the seriously increased delays forecast by some researchers a decade or so ago have not materialized.

The second implication is based on the first, that we need new methods to evaluate the benefits of new ATM technologies and procedures. The model suites presented here—based on carriers' limited tolerance for delays and on their economic incentives to change operations in the face of productivity lost due to delays—are an initial response to this.

The models, accounting rationally for airlines' reactions to delay, are likely to be reasonably robust. This is not necessarily true for this report's specific detailed forecasts. Many uncertainties, some quite small, may change air carriers' operations in the future. We base our conclusions on the assumptions of rational behavior of the carriers and a relatively stable operations environment. Changes in the legal framework in which carriers operate—reflecting concerns for competition, environment, ATC user fees, etc.—could substantially affect air carriers' operations and might call for changes to our models. Changes in the economic forecasts that drive the TAF forecasts will not change our models, although, of course, changing these inputs will change our models' forecasts. Unforeseen technological breakthroughs, or political instabilities like terrorist attacks or war, could also have substantial impacts on air carriers.

References

[1] Logistics Management Institute, *A Method for Making Cross-Comparable Estimates of the Benefits of Decision Support Technologies for Air Traffic Management*, NASA Contractor Report 208455, David A. Lee, Dou Long, Melvin R. Etheridge, Joana R. Plugge, Jesse P. Johnson, and Peter F. Kostiuk, March 1998.

[2] Logistics Management Institute, *Modeling Air Traffic Management Technologies with a Queuing Network Model of the National Airspace System*, Preliminary report, Dou Long, David A. Lee, Jesse P. Johnson, Eric M. Gaier, and Peter F. Kostiuk, June 1998.

[3] American Airlines, *Free Flight: Preserving Airline Opportunity*, Captain Russell G. Chew, September 1997.

[4] Boeing, *1998 Current Market Outlook*, Boeing Commercial Airplane Group Marketing, P.O. Box 3707, Seattle, WA 98124-2207.

[5] Department of Transportation, *FAA Aviation Forecast—Fiscal Years 1998–2009*, Report No. FAA APO-98-1, Federal Aviation Administration, Office of Aviation Policy and Plans, Statistics and Forecast Branch, Washington, DC, March 1998.

[6] NASA, *The ASAC Air Carrier Investment Model (Third Generation)*, NASA Contractor Report 207656, Earl R. Wingrove III, Eric M. Gaier, and Tara E. Santmire, 1998.

[7] American Airlines, *Preliminary Evaluation of Flight Delay Propagation Through an Airline Schedule*, Roger Beatty et al., 1998.

[8] Logistics Management Institute, *The Economic Impacts of Air Traffic Congestion*, Report IR811R1, Peter F., Kostiuk, Eric M. Gaier, and Dou Long, July 1998.

Appendix A
Calculating Total Waiting Times for Two Queues in Tandem

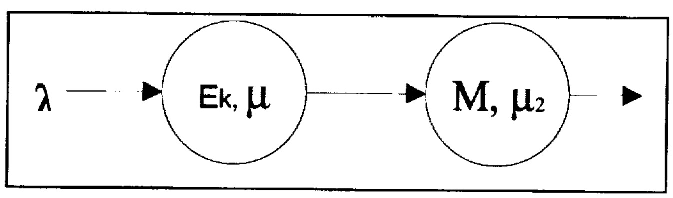
This appendix describes our method for calculating expected delays for traffic arriving at and departing from LMINET airports. In LMINET, a flight arriving at an airport encounters two tandem queues: an *M/Ek/*1 queue for arrival runway service, followed by a queue with Poisson service times modeling arrival surface-movement delays. When there are no delays due to lack of ready-to-depart aircraft, a departing flight encounters an *M/M/*1 queue modeling departure surface-movement delays, followed by a queue for departure runway service. The runway-service queue's service times have the Erlang-k distribution.

Standard LMINET calculations produce mean aircraft-minutes of delay in the model's several queues by integrating the queue lengths with respect to time. For the present study, we are interested in statistics of the waiting times experienced by flights arriving at and departing from an airport, at specific times. We generated these from numerical solutions of the relevant tandem-queue systems. The following subsections treat the arrival and departure cases.

*M/EK/*1 PRECEDING A SINGLE-SERVER QUEUE WITH POISSON SERVICE

Figure A-1 diagrams this tandem-queue network.

Figure A-1. Tandem Queueing Network for Arrivals



Aircraft arrive at the network in a Poisson process with rate λ, before a single-server queue whose service times have the Erlang-k distribution with mean rate μ. After service at this queue, they immediately enter a second single-server queue whose service times have the Poisson distribution with mean rate μ_2.

We characterize the state of this tandem-queue network with the ordered pair (n_1, n_2), where n_1 denotes the number of *phases* in the queue with Erlang-k service times, while n_2 denotes the number of *clients* in the Poisson-service queue.

State-Transition Diagrams

If the system is at rest—i.e., in state (0, 0)—then (potential) service events have no effect. The only possible transition occurs with rate λ, and the system transitions to state (*k*, 0). This is shown in the state-transition diagram of Figure A-2.

Figure A-2. State Transitions from Rest State

$$(0, 0) \xrightarrow{\lambda} (k, 0)$$

States in which the number of phases in the *M/Ek/*1 queue is one more than an integer multiple of *k* are special: a service event in the Erlang queue completes the required set of *k* services and increments the number of clients in the Poisson queue. The diagrams of Figures A-3 through A-7, which complete the set of state-transition diagrams, show transitions from these states separately.

Figure A-3. Transitions from ($n1$, 0), $n1 > 0$ and $n1 \neq mk + 1$

$$(n1, 0) \xrightarrow{\lambda} (n1 + k, 0)$$
$$(n1, 0) \xrightarrow{k\mu} (n1 - 1, 0)$$

Figure A-4. Transitions from (mk + 1, 0), m = *0*, 1, ...

$$(mk + 1, 0) \xrightarrow{\lambda} ((m + 1)k + 1, 0)$$
$$(mk + 1, 0) \xrightarrow{k\mu} (mk, 1)$$

Figure A-5. Transitions from (0, n2), n2 > 0

$$(0, n2) \xrightarrow{\lambda} (k, n2)$$
$$(0, n2) \xrightarrow{\mu 2} (0, n2 - 1)$$

Figure A-6. Transitions from (n1, n2), n*1* > 0, n2 > 0, *and* n1 $\neq$ mk +1

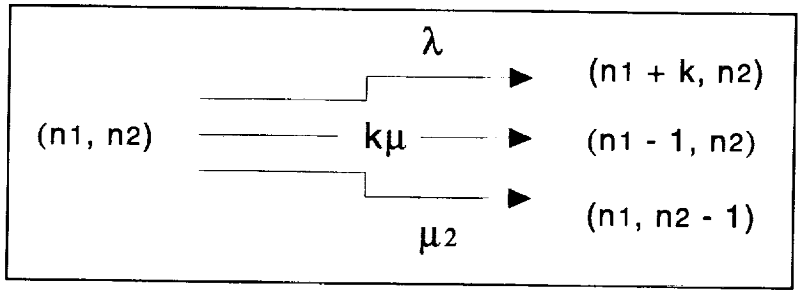
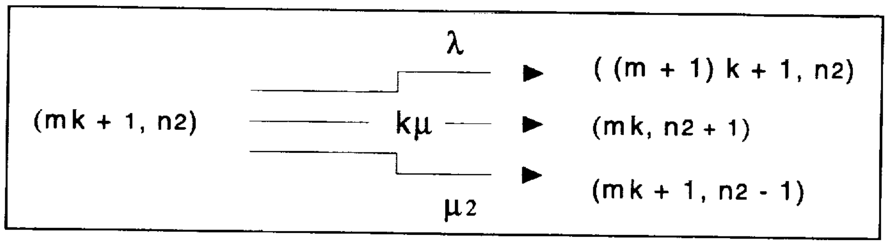


Figure A-7. Transitions from (mk + *1*, n2*)*, m = *0, 1, 2, ...,* n2 > *0*



Evolution Equations

The state transitions diagrammed in the previous section lead directly to the following evolution equations for the tandem queue of Figure A-1. In these equations, $p(n_1, n_2, t)$ is the probability that the system is in state (n_1, n_2) at time t.

$$\dot{p}(0,0,t) = -\lambda p(0,0,t) + \mu_2 p(0,1,t) \qquad \text{[Eq. A-1]}$$

$$\begin{aligned} \dot{p}(0,n_2,t) = &-(\lambda + \mu_2)p(0,n2,t) \\ &+ k\mu p(1,n2-1,t) \\ &+ \mu_2 p(0,n2+1,t), \;\; n_2 > 0 \end{aligned} \qquad \text{[Eq. A-2]}$$

$$\begin{aligned} \dot{p}(n_1,0,t) = &-(\lambda + k\mu)p(n_1,0,t) \\ &+ k\mu p(n_1+1,0,t) \\ &+ \mu_2 p(n_1,1,t), \qquad 0 < n_1 < k \end{aligned} \qquad \text{[Eq. A-3]}$$

$$
\begin{aligned}
\dot{p}(n_1,0,t) = &-(\lambda + k\mu)p(n_1,0,t) \\
&+ k\mu p(n_1+1,0,t) \\
&+ \mu_2 p(n_1,1,t) \\
&+ \lambda p(n_1-k,0,t), \quad n_1 \geq k, n_1 \neq mk, m = 1,2,\ldots,
\end{aligned}
$$

$$
\begin{aligned}
\dot{p}(mk,0,t) = &-(\lambda + k\mu)p(mk,0,t) \\
&+ \mu_2 p(mk,1,t) \\
&+ \lambda p((m-1)k,0,t) \quad m = 1,2,\ldots,
\end{aligned}
\qquad \text{[Eq. A-4]}
$$

$$
\begin{aligned}
\dot{p}(n_1,n_2,t) = &-(\lambda + k\mu + \mu_2)p(n_1,n_2,t) \\
&+ k\mu p(n_1+1,n2,t) \\
&+ \mu_2 p(n_1,n_2+1,t), \qquad 0 < n_1 < k; n_2 > 0
\end{aligned}
\qquad \text{[Eq. A-5]}
$$

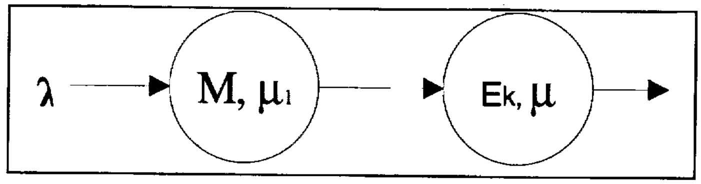
$$
\begin{aligned}
\dot{p}(n_1,n_2,t) = &-(\lambda + k\mu + \mu_2)p(n_1,n_2,t) \\
&+ k\mu p(n_1+1,n2,t) \\
&+ \mu_2 p(n_1,n_2+1,t) \\
&+ \lambda p(n_1-k,n_2,t), \qquad n_1 \geq k; n_1 \neq mk; n_2 > 0
\end{aligned}
\qquad \text{[Eq. A-6]}
$$

$$
\begin{aligned}
\dot{p}(mk,n_2,t) = &-(\lambda + k\mu + \mu_2)p(mk,n_2,t) \\
&+ k\mu p(mk+1,n2-1,t) \\
&+ \mu_2 p(mk,n_2+1,t) \\
&+ \lambda p((m-1)k,n_2,t), \qquad n_2 > 0; m \geq 1
\end{aligned}
\qquad \text{[Eq. A-7]}
$$

*M/M/*1 PRECEDING A SINGLE-SERVER QUEUE WITH ERLANG-*K* SERVICE

When there are no delays for ready-to-depart aircraft, departures in LMINET first enter a *M/M/*1 queue that models surface-movement delays. After service in this queue, departing flights enter a queue with the Erlang-*k* distribution of service times, for departure runway service.

Figure A-8. Tandem Queuing Network for Departures



We now treat this queuing network as we did the one for arrivals, characterizing its state with the ordered pair (n_1, n_2), where now n_1 is the number of *clients* in the first queue, and n_2 is the number of *phases* in the second.

State-Transition Diagrams

The number of different transitions to consider is less for this tandem network than for the network of Figure A-1, because here the more complex workings of the queue with E_k service do not affect another queue. Just four diagrams suffice; they follow:

Figure A-9. Transition from the Rest State

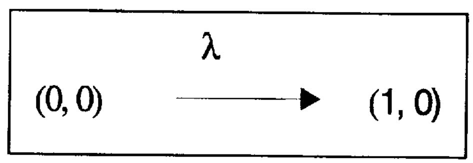


Figure A-10. Transitions from $(n_1, 0)$, $n_1 \geq$ (*1*

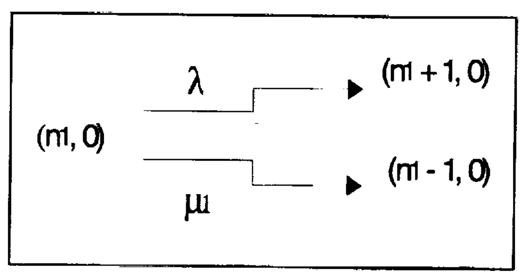


Figure A-11. Transitions from $(0, n_2)$, $n_2 > 0$

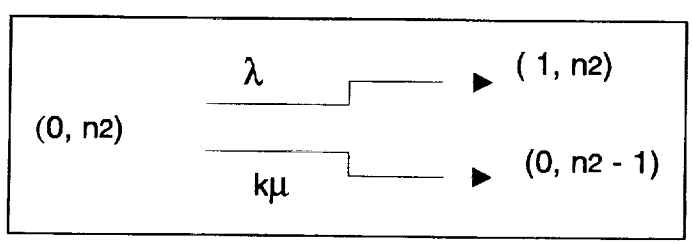
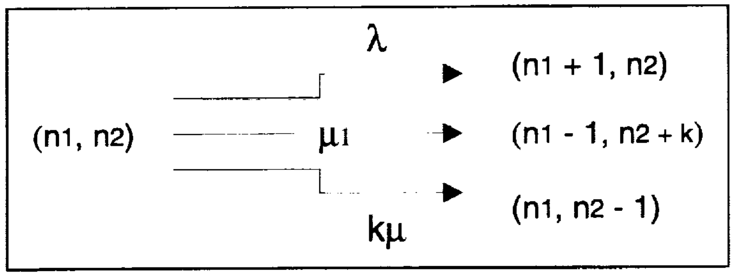


Figure A-12. Transitions from (n_1, n_2), $n_1 > 0$ *and* $n_2 > 0$



Evolution Equations

The evolution equations also are simpler for this case than for the case of the previous subsection. Writing $p(n_1, n_2, t)$ for the probability that the network is in the state (n_1, n_2), and remembering that now n_1 is the number of clients in the first queue while n_2 is the number of phases in the second, we have:

$$\dot{p}(0,0,t) = -\lambda p(0,0,t) + k\mu p(0,1,t) \qquad \text{[Eq. A-8]}$$

$$\dot{p}(0,n_2,t) = -(\lambda + k\mu)p(0,n_2,t) + k\mu p(0,n_2+1,t),\ 0 < n_2 < k \qquad \text{[Eq. A-9]}$$

$$\begin{aligned}\dot{p}(0,n_2,t) = &(\lambda + k\mu)p(0,n_2,t) + k\mu p(0,n_2+1,t)\\ &+ \mu_1 p(1,n_2-k), \qquad n_2 \geq k\end{aligned} \qquad \text{[Eq. A-10]}$$

$$\begin{aligned}\dot{p}(n_1,0,t) = &-(\lambda + \mu_1)p(n_1,0,t) + k\mu p(n_1,1,t)\\ &+ \lambda p(n_1-1,0), \qquad n_1 > 0\end{aligned} \qquad \text{[Eq. A-11]}$$

$$\begin{aligned}\dot{p}(n_1,n_2,t) = &-(\lambda + k\mu + \mu_1)p(n_1,n_2,t)\\ &+ k\mu p(n_1,n_2+1,t)\\ &+ \lambda p(n_1-1,n_2,t), \qquad n_1 > 0; 0 < n_2 < k\end{aligned} \qquad \text{[Eq. A-12]}$$

$$\begin{aligned}\dot{p}(n_1,n_2,t) = &-(\lambda + k\mu + \mu_1)p(n_1,n_2,t)\\ &+ k\mu p(n_1,n_2+1,t)\\ &+ \mu_1 p(n_1+1,n_2-k,t)\\ &+ \lambda p(n_1-1,n_2,t), \qquad n_1 > 0; n_2 \geq k\end{aligned} \qquad \text{[Eq. A-13]}$$

NUMERICAL SOLUTION OF THE TANDEM-QUEUE EQUATIONS

Generating numerical solutions of the system of Equations A-1 through A-8 or Equations A-9 through A-14 is challenging because the dimensions of the systems are rather large for the cases of interest. Some airports have arrival or departure capacities as large as 150 operations per hour. We are interested in cases where delays do not exceed time on the order of 5 minutes. The *k* parameter of our Erlang-*k* distributions is 22. Thus, the number of phases to be considered may be on the order of 10^3. Associated surface-movement queues can be simultaneously around five or ten. Thus, the system of first order, linear-ordinary differential equations that we must consider can have several thousand unknown functions to determine.

Fortunately, the systems are quite sparse, which makes numerical solution feasible. We experimented with several numerical methods, and finally settled on a second-order Runge-Kutta scheme that appears to give a useful balance of storage and execution speed.

We determined the number of phases, and of clients, to be considered at each epoch by computing the steady-state queue lengths for both queues, when utilization ratios were less than one, and by computing the fluid-approximation queue lengths when those ratios exceeded one. We then considered at least three times the number of phases, or of clients, found in that way, for the numerical integration. We renormalized probabilities to sum to one, at each integration step.

WAITING TIME CALCULATION BASED ON THE SYSTEM STATE

For arrivals,

$$\textit{expected runway delay} = \sum_{n_1, n_2} \frac{n_1\, p(n_1, n_2)}{k\mu} \qquad \text{[Eq. A-14]}$$

$$\textit{expected taxiway delay} = \sum_{n_1, n_2} \frac{n_2\, p(n_1, n_2)}{\mu_2}, \qquad \text{[Eq. A-15]}$$

$$\textit{expected arrival delay} = \sum_{n_1, n_2} \frac{n_1\, p(n_1, n_2)}{k\mu} + \sum_{n_1, n_2} \frac{n_2\, p(n_1, n_2)}{\mu_2}. \qquad \text{[Eq. A-16]}$$

Let

$$T_a = \{n_1,n_2: n_1/k\mu + n_2/\mu_2 \leq (7.5/60)\}, \qquad \text{[Eq. A-17]}$$

which is the set of states that the arrival delay is no more than 7.5 minutes, then on-time probability is given by

$$p_{a7.5} = \sum_{n_1,n_2 \in T_a} p(n_1,n_2). \qquad \text{[Eq. A-18]}$$

Similarly, for the departure system

$$\textit{expected departure runway delay} = \sum_{n_1,n_2} \frac{n_1\, p(n_2,n_1)}{k\mu}, \qquad \text{[Eq. A-19]}$$

$$\textit{expected departure taxiway delay} = \sum_{n_1,n_2} \frac{n_2\, p(n_1,n_2)}{\mu_1}, \qquad \text{[Eq. A-20]}$$

$$\textit{expected departure delay} = \sum_{n_1,n_2} \frac{n_1\, p(n_2,n_1)}{k\mu} + \sum_{n_1,n_2} \frac{n_2\, p(n_1,n_2)}{\mu_1}. \qquad \text{[Eq. A-21]}$$

and on-time probability is given by

$$p_{d7.5} = \sum_{n_1,n_2 \in T_d} p(n_1,n_2), \qquad \text{[Eq. A-22]}$$

where

$$T_d = \{n_1,n_2: n_1/k\mu + n_2/\mu_1 \leq 7.5/60\} \qquad \text{[Eq. A-23]}$$

Appendix B
Derivation of Air Carrier Investment Model Inputs

This appendix documents the translation from the delay output of LMINET to the productivity input of ACIM. As described in the main body of this report, the main economic impact of delay is captured with the aircraft block speed parameter of the ACIM. Specifically, this parameter is the ratio of aircraft miles to block hours as shown in Equation B-1:

$$\mathit{Block\ Speed} = \frac{\mathit{Aircraft\ Miles}}{\mathit{Block\ Hour}} \qquad \text{[Eq. B-1]}$$

To begin, we recognize the need to measure the change in aircraft productivity in compound annual percentage terms. This calculation is expressed by equation B-2, in which t denotes time in years, the subscript 0 denotes the year, and the subscript 1 denotes a future year.

$$\mathit{Annual\ Rate\ of\ Change} = \left(\frac{\dfrac{\mathit{Aircraft\ Miles}_1}{\mathit{Block\ Hours}_1}}{\dfrac{\mathit{Aircraft\ Miles}_0}{\mathit{Block\ Hours}_0}} \right)^{\frac{1}{(t_1 - t_0)}} - 1. \qquad \text{[Eq. B-2]}$$

Holding average stage length constant, Equation B-2 reduces to Equation B-3, which is the result reported in the main text:

$$\mathit{Annual\ Rate\ of\ Change} = \left(\frac{\mathit{Average\ Block\ Time}_0}{\mathit{Average\ Block\ Time}_1} \right)^{\frac{1}{(t_1 - t_0)}} - 1. \qquad \text{[Eq. B-3]}$$

Appendix C
Balanced Departures and Arrivals as a Way to Maximize Total Operations

Maximizing the total number of operations at an airport is equivalent to the problem.

$$Max \sum_{i=1}^{n} (x_i + y_i) \qquad \text{[Eq. C-1]}$$

$$\textit{subject to: } \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i , \qquad \text{[Eq. C-2]}$$

$$x_i \geq 0;\ y_i \geq 0,\ i=1, \ldots, n, \qquad \text{[Eq. C-3]}$$

$$f(x_i, y_i) \leq 0. \qquad \text{[Eq. C-4]}$$

where x_i and y_i are the number of arrivals and departures a time i, and $f(\cdot,\cdot)$ is the Pareto frontier of the airport capacity function, which is concave. Equation C-2 means that operations at the airport are balanced, i.e., that the airport is not a sink or source of flights.

Before we solve the problem in general, let us see the solutions of some particular cases.

If $n=1$, then Equation C-2 is reduced to $x_1 = y_1$. The objective function of Equation C-1 becomes $x_1 + y_1 = 2x_1 = 2y_1$. Then in order to maximize the objective function and also to satisfy Equation C-4, the optimal solution is $x_1 = y_1$, and $f(x_1,y_1) = 0$. This means we should operate the airport on the Pareto frontier such that the arrival rate is equal to the departure rate.

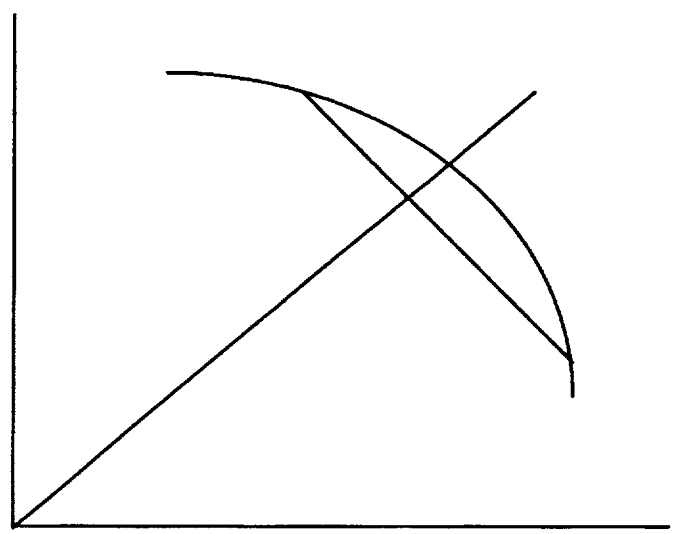
If $n=2$, then Equation C-2 becomes $x_1 + x_2 = y_1 + y_2$, which means (x_1, y_1) and (x_2, y_2) are mirror points along the line $y = x$. Let (x_0, y_0) be the point of intersection of the line connecting (x_1, y_1) and (x_2, y_2) and of the line $y = x$. Since $(x_1\ y_1)$ and (x_2, y_2) are mirror points along the line $y = x$, then the objective function, Equation C-1, becomes

$$x_1 + x_2 + y_1 + y_2 = 2x_0 + 2y_0 = 2(x_0+y_0). \qquad \text{[Eq. C-5]}$$

It is then clear that the objective function (Equation C-5) reaches its maximum value when (x_0, y_0) is on the Pareto frontier. Therefore, the line connecting the points (x_1, y_1) and (x_2, y_2) is degenerate, and $x_1=x_2=x_0$, $y_1=y_2=y_0$.

This means that if we want to maximize the total airport operations in just two time periods, then we have to set the arrival equal to departure, which are also on the Pareto frontier for both periods. In many ways, this problem is solved, since the airport balancing equation (Equation C-2) must be valid for every two periods, evidenced by the bank operations in many hub airports.

Figure C-1. Airport Operations in Two Periods



For any general positive integer n, the conclusion is also true that we have to set arrivals equal to departure on the capacity Pareto frontier for each time period. We have to resort to nonlinear programming techniques to prove that. Suppose the Pareto frontier function is differentiable, then the Kuhn-Tucker condition of optimality can be stated as follows:

$$\frac{\partial}{\partial x_i}\left[\sum_{i=1}^{n}(x_i+y_i)-u\sum_{i=1}^{n}(x_i-y_i)-v_i f(x_i,y_i)\right]=0,$$

$$\frac{\partial}{\partial y_i}\left[\sum_{i=1}^{n}(x_i+y_i)-u\sum_{i=1}^{n}(x_i-y_i)-v_i f(x_i,y_i)\right]=0, i=1,2,\ldots,n, \quad \text{[Eq. C-6]}$$

$$v_i f_i(x_i,y_i)=0, i=1,2,\ldots,n, \quad \text{[Eq. C-7]}$$

$$v_i \geq 0, i=1,2,\ldots,n. \quad \text{[Eq. C-8]}$$

Taking derivatives in Equation C-6, we have

$$1-u-v_i\frac{\partial f}{x}\bigg|_{y=y_i}^{x=x_i}=0,$$

$$1+u-v_i\frac{\partial f}{y}\bigg|_{y=y_i}^{x=x_i}=0, i=1,2,\ldots,n \qquad \text{[Eq. C-9]}$$

Equations C-7 and C-8 force the optimal solution $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ to be all on the Pareto frontier f. Suppose the point (x_j, y_j) is not on the Pareto frontier, i.e., $f(x_j, y_j) < 0$, then by Equations 4-7, $v_j = 0$, which means Equation C-9 becomes

$$1 - u = 0,$$

$$1 + u = 0. \qquad \text{[Eq. C-10]}$$

There is no solution to Equations C-10, which completes the proof that all optimal solutions are on the Pareto frontier. If the Pareto frontier $f(\cdot,\cdot)$ is strictly concave, then $f(\cdot,\cdot)$ will have unique Jacobian derivative at each point, which means, by Equations C-9,

$$x_i = x_j = x_0,$$

$$y_i = y_j = y_0, \; i,j \in \{1,2,\ldots,n\}. \qquad \text{[Eq. C-11]}$$

By Equation C-2 and the fact that all the points of the solution must be on the Pareto frontier, the optimal operation policy that will maximize the total number of operations is given by the point on the Pareto frontier with equal arrival and departure capacity. Since $f(\cdot,\cdot)$ is concave, the sufficient condition of the optimality is also satisfied. *Q.E.D.*

This is a rather interesting theorem. It is well recognized in the research community that bank operations in hub airports have to be reduced in order to avoid excessive delay. However, this calling for reductions of bank of operations does not specify how to reduce the operations, arrival or departure. If we want to maximize the total operations, then based on the theorem, we should spread the arrival and departure evenly throughout epochs, and abandon the current practice of *departure push.*

Some airports, e.g., BWI, LAX, LGA, currently have dedicated departure runways in the airport capacity configuration. This practice makes sense to accommodate the departure push, since departure is more predictable than arrival. However, if the demand constantly exceeds the airport capacity, then it is questionable to conduct the departure push and the dedicated departure runway represents a waste of resources.

Appendix D
Abbreviations

AATT	Advanced Air Transport Technologies
ACI	Airport Council International
ACIM	Air Carrier Investment Model
ASAC	Aviation System Analysis Capability
ASM	Available Seat Mile
ASQP	Airlines Service Quality Performance
ATC	Air Traffic Control
ATM	Air Traffic Management
CRS	Computer Reservation System
DOT	Department of Transportation
ETMS	Enhanced Traffic Management System
FAA	Federal Aviation Administration
FAM	Fleet Assignment Model
FCM	Functional Cost Module
GA	General Aviation
GDP	Gross Domestic Product
IFR	Instrument Flight Rule
IMC	Instrument Meteorological Condition
LMI	Logistics Management Institute
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NASPAC	National Airspace System Performance Analysis Capability
OAG	Official Airlines Guide
O-D	Origin - Destination
QSI	Quality Service Index
PMAC	Performance Monitoring System
RPM	Revenue Passenger Mile

TAF	Terminal Area Forecast
TAP	Terminal Area Productivity
TRACON	Terminal Radar Approach Control
VFR	Visual Flight Rule
VMC	Visual Meteorological Condition

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
January 1999

3. REPORT TYPE AND DATES COVERED
Contractor Report

4. TITLE AND SUBTITLE
A Method for Forecasting the Commercial Air Traffic Schedule in the Future

5. FUNDING NUMBERS
C NAS2-14361
WU 538-16-11-01

6. AUTHOR(S)
Dou Long, David Lee, Eric Gaier, Jesse Johnson, and Peter Kostiuk

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Logistics Management Institute
2000 Corporate Ridge
McLean, VA 22102-7805

8. PERFORMING ORGANIZATION REPORT NUMBER
NS806S1

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA/CR-1999-208987

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Robert E. Yackovetsky
Final Report

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified - Unlimited

Subject Category 01
Availability: NASA CASI (301) 621-0390
Distribution: Nonstandard

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***

This report presents an integrated set of models that forecasts air carriers' future operations when delays due to limited terminal-area capacity are considered. This report models the industry as a whole, avoiding unnecessary details of competition among the carriers.

To develop the schedule outputs, we first present a model to forecast the unconstrained flight schedules in the future, based on the assumption of rational behavior of the carriers. Then we develop a method to modify the unconstrained schedules, accounting for effects of congestion due to limited NAS capacities. Our underlying assumption is that carriers will modify their operations to keep mean delays within certain limits. We estimate values for those limits from changes in planned block times reflected in the OAG. Our method for modifying schedules takes many means of reducing the delays into considerations, albeit some of them indirectly. The direct actions include depeaking, operating in off-hours, and reducing hub airports' operations. Indirect actions include using secondary airports, using larger aircraft, and selecting new hub airports, which, we assume, have already been modeled in the FAA's TAF. Users of our suite of models can substitute an alternative forecast for the TAF.

14. SUBJECT TERMS
Air Carrier Operation, Air Traffic Schedule, Airport Capacity, Flight Delay

15. NUMBER OF PAGES
84

16. PRICE CODE
A05

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited