Analysis of Air Transportation for the New York Metroplex: Summer 2007

Liya Wang; George Donohue;
Karla Hoffman; Lance Sherry
Center of Air Transportation and Systems Research
George Mason University
4400 University Dr.
Fairfax, VA 22030

Rosa Oseguera-Lohr
NASA Langley Research Center
Aeronautics Systems Analysis Branch M. S. 442
Hampton, VA 23681

***Abstract*—The New York metroplex airports (JFK, LGA, EWR) provide air transportation service to this critical international economic hub. In the summer of 2007 the flights servicing the NYC metroplex airports experienced excessive delays and cancellations that added significant costs to doing business in New York. These delays can be attributed to changes in daily airport capacity (due to weather) and to airline practices, in accordance with regulations, of scheduling in excess of airport capacity. Previous research has demonstrated that maintaining airline seat capacity by increasing aircraft size and reducing frequency is an economically efficient and feasible solution. This paper analyzes the characteristics of the air transportation service to the New York metroplex airports. The metroplex has service to 109 domestic airports. Thirty six percent of airports serve all three New York airports, while 35% serve two of the airports. For the airports that serve all three NY airports, the average number of flights per day is 24 with a maximum of 60. These flights have an average seat size ranging from 41 to 178 with an average of 91 seats per flight. During the summer of 2007, these flights had passenger load factors ranging from 53% to 91% with an average of 71%. This yields an average of 633 unused seats per day on these routes. Additional statistics and discussion of these results on the implications for consolidation of service with larger aircraft and reduced frequency is discussed.**

Keywords- JFK; LGA; EWR; metroplex; air transportation

I. INTRODUCTION

In 2007, domestic U.S. airline travelers experienced the lowest on-time performance on record. Approximately 30 percent of all flights were either cancelled or delayed more than 15 minutes [1].

Airline service to the international economic hub of New York City (NYC) was particularly hard hit. The on-time percentage for the three New York airports (EWR, LGA, and JFK) was 71.5% departures, 62% arrivals, and 3.46% cancellations [2]. The national average was 76.5%, 73.4% and 2.16% respectively [1]. Also, airline service experienced the worst cancellation rate in the nation. Since approximately a third of the nation's air traffic passes through NY airports, delays in NYC ripple through the system causing delays at other airports [3].

Analysis of these delays identified that two functional causes of delays; (i) changes in daily airport capacity (due to weather) as high as 20% reductions from good weather capacity, and (ii) airline practices, in accordance with regulations, of scheduling in excess of airport capacity.

Previous research [4] demonstrated that maintaining airline seat capacity by increasing aircraft size and reducing frequency is an economically efficient and feasible solution. Airlines flying larger aircraft, with higher load factors increase revenue. Air Traffic Control has reduced operations leading to marginal delays. The airport increases passenger throughput and provides reliable service to its customers.

This paper describes the results of an analysis of the air transportation characteristics of the NYC metroplex airports:

- The metroplex has service to 109 domestic airports.
- Thirty six percent of airports serve all three New York airports, while 35% serve two of the airports.
- For the airports that serve all three NY airports, the average number of flights per day is 24 with a maximum of 60. These flights have an average seat size ranging from 41 to 178 with an average of 91 seats per flight.
- During the summer of 2007, the flights from the same airport to all three NYC airports had passenger load factors ranging from 53% to 91% with an average of 71%.
- This yields an average of 633 unused seats per day on these routes the equivalent of 7 flights per day.

Additional statistics are also provided: number of markets served, redundant service at NYC airports, flight number per day of NYC markets, mumber of competing airlines of NYC markets, load factor of NYC markets, aircraft seat size of NYC markets, seat size vs. load factor for NYC Markets, unite revenue ($/mile) vs. load factor for NYC Markets, flight frequency vs. seat size classified by load factor for NYC Markets, and flight frequency vs. seat size classified by unit revenue for NYC Markets.

The paper is organized as follows: Section 2 provides an overview of the scheduled/actual flights and available capacity at NYC airports. Section 3 describes the methodology and algorithms for analysis the air transportation data. Section 4 describes the results of the analysis. The conclusions and future work are discussed in Section 5.

II. BACKGROUND: DEMAND VS. CAPACITY

Competition coupled with high demand force the airlines to schedule multiple flights during peak hours. To cut costs, flights are served by smaller planes thereby allowing each airline to provide frequency during the most popular times while maintaining reasonable costs. The result for the airport, is that more flights are scheduled than the runway can handle. Figure 1 through 3 show that 2007 summer (06/01-08/30) departure capacity of mean ADR (Airport Departure Rate), average number of scheduled departures, and wheels-off delays in 15 minutes bin from 6:00 am to 12:00 am for NYC metroplex three airports. Obviously, during peak hours of 8:00am-9:00 am and 4:00pm-6:00pm, the airports are over-scheduling. Figure 1 to 3 also illustrate how over-scheduling in one period without sufficient under-scheduling to allow the queues to dissipate, forces delays to become longer as the day progresses. Examining Figure 3, one sees that although there is only a small amount of over-scheduling in any given period, the delays continues to get larger and larger as the day progresses. In addition, JFK has the worst wheels-off delays from volume standpoint. Table 1 summarized the over-scheduling time percentage from 6:00am-10:00pm in 2007 summer for scheduled and actual demands respectively. Table 1 shows that departures are more over-scheduled than arrivals. From time standpoint, LGA is the most over-scheduled airport.

In addition, the study also shows the airport capacity is not fully, efficiently, and properly used. Figure 4 and 5 show the actual and scheduled demands distribution for 2007 summer respectively for JFK. In each cell, the probability of demands is calculated, and then colors are used to distinguish them. Figure 4 shows some quarterly scheduled demands are exceeding the airport capacity very much. For example, (23, 7) cell shows a period of time where there are 23 scheduled arrival demand when the airport is capable of handling only 14. At the same time, Figure 5 also shows that system is under used at most of the time in that high probability of demands is distributed far away from the capacity line.

Therefore, to reduce the delays at NYC metroplex airports, first the transportation characteristics should be achieved. Based on that, then the possible improvement space can be identified. In the next part, algorithms to processing data for achieving air transportation characteristics are in detail presented.

Table 1: NYC airports quarterly overscheduled percentage (6:00am-10:00pm)

Airport	Actual Arr.(%)	Actual Dep.(%)	Scheduled Arr. (%)	Scheduled Dept. (%)
EWR	1.8%	4.8%	6.5%	6.6%
JFK	2.9%	8.9%	5.4%	13.8%
LGA	5.1%	10.3%	9.3%	12.2%

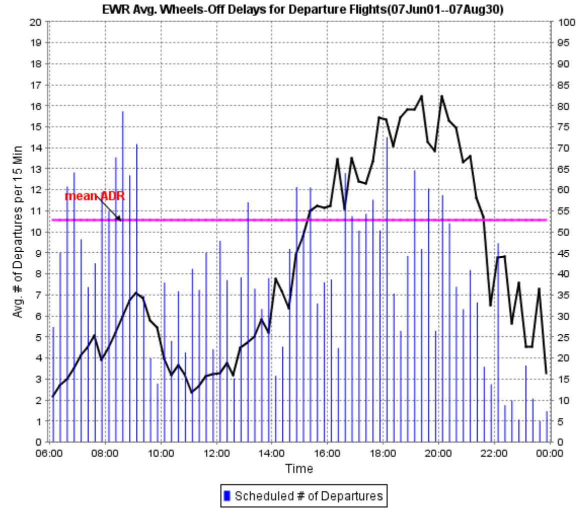

Figure 1. 2007 summer EWR mean capacity, departures and wheels-off delays per 15 min

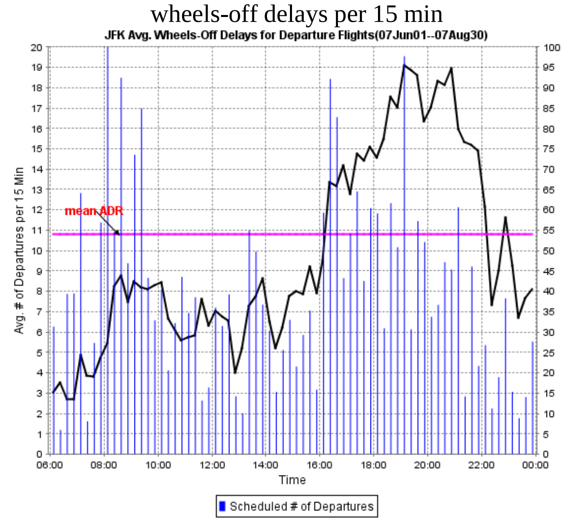

Figure 2. 2007 summer JFK mean capacity, departures and wheels-off delays per 15 min

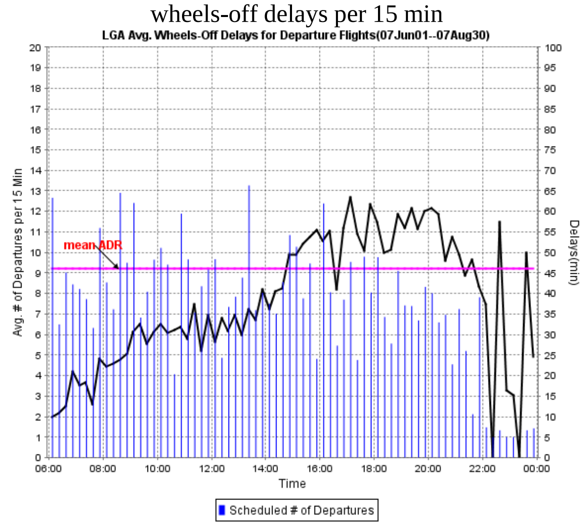

Figure 3. 2007 summer LGA mean capacity, departures and wheels-off delays per 15 min

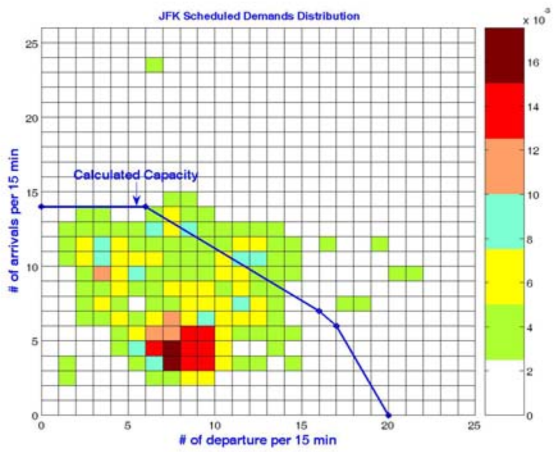

Figure 4. 2007 summer JFK scheduled demands distribution

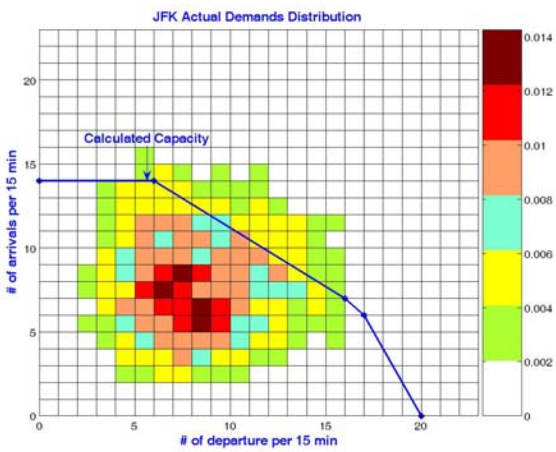

Figure 5. 2007 summer JFK actual demands distribution

III. METHODOLOGY

To get NYC metroplex airports aviation transportation characteristics, a number of aviation databases from Bureau of Transportation Statistics (BTS) [5] and Federal Aviation Association (FAA) Aviation System Performance Metrics (ASPM) [6] databases are used for the study. This section mainly describes the data processing needed to perform our analysis. Figure 6 shows the data processing flow chart. Next, the details of data processing algorithms will be discussed.

A. Data Processing Algorithms

1) Determining seat size, load factor, frequency using T-100 Domestic Segment (U.S. Carriers) Dataset

T-100 Domestic Segment (U.S. Carriers) database contains domestic non-stop segment data reported by U.S. air carriers, including carrier, origin, destination, aircraft type and service class for transported passengers, freight and mail, available capacity, scheduled departures, departures performed, aircraft hours, and load factor when both origin and destination airports are located within the boundaries of the United States and its territories.

Therefore, from the T100 database, we can extract market information for each airport such as the number of carriers, average plane seat size, load factor, and average flight number per day. Average airplane seat size can be calculated in Eq. (1). The calculation of load factor and average flight number per day is expressed in Eq. (2) and (3) respectively.

Seat Size= total # of seats / total # of departures (1)
Load factor= total # of passengers / total # of seats (2)
Flight number= total # of departures / total # of days (3)

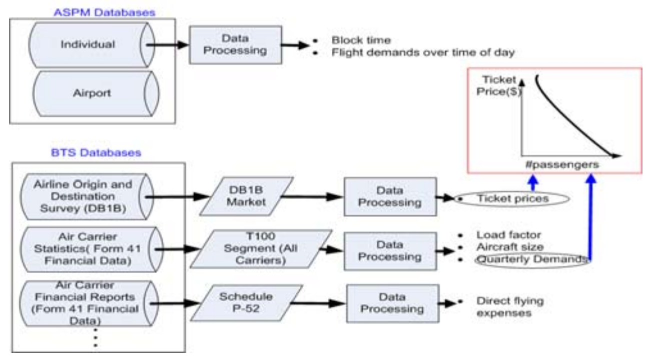

Figure 6. Data processing flow chart

2) Calculations of segment prices using DB1B Market Dataset

DB1B market data table contains directional market characteristics of each domestic itinerary of the Origin and Destination Survey, such as the reporting carrier, origin and destination airport, prorated market fare, number of market coupons, market miles flown, and carrier change indicators. Therefore, DB1B can be used to extract ticket price for market A to B.

The ticket prices in DB1B market table are for itinerary tickets. To get ticket price from airport A to airport B, two situations should be considered. In the first situation, A is origin and B is destination. Therefore, A to B ticket price can be fetched directly from DB1B. In the second situation, A is origin; however B is a transfer airport for flying A to C. Under this situation, the ticket price from A to B cannot be directly got from DB1B. Segment fares are traditionally prorated from itinerary fares. However, there is a fixed cost in any flight leg. This port of fixed cost is large in flight legs of short distance, and decreases in legs of longer distance. We compute segment fares proportionally to the squared root of distances of segments in the itinerary [4]. Therefore, Eq. (4) will be used to get segment fare from A to B.

$$T_{AB} = T_{AC} \times \frac{\sqrt{d_{AB}}}{\sqrt{d_{AB}} + \sqrt{d_{BC}}} \tag{4}$$

Where T_{AB} is the ticket price form A to B, and d_{AB} is the distance form A to B.

Specifically, if a flight has two legs of 100(=10^2) miles and 225(=15^2) miles, and has the one-way ticket price of $100, then leg one is allocated $40 (=$100\times\frac{\sqrt{100}}{\sqrt{100}+\sqrt{225}}$) and leg two 60$ (=100-40).

3) Extracting airport capacity from ASPM Airport Dataset

ASPM airport database can provide detail information by quarter hour or hour on the airport, which includes AAR, ADR, wind speed, visibility, runway configuration, scheduled departures, scheduled arrivals, efficiency of departures, and ETMS departures etc.

Therefore, from this database, we can directly fetch airport capacity of AAR and ADR.

4) Determing demands and delays data from from ASPM Individual Dataset

ASPM individual database can provide detail schedule information on a flight including carrier, origin, destination, aircraft type, departure date, arrival date, scheduled in time, actual in time, scheduled out time, actual out time, scheduled taxi-out time, actual taxi-out time, scheduled taxi-in time, actual taxi-in time, delays, and block time etc.

Hence, from ASPM individual database, schedule information such as departure and arrival demands, delays in each quarter or hourly can be extracted. Eq. (5) calculates the demands per quarter. ASPM individual database can also provide information scheduled block time. Eq. (6) shows how to calculate delays per quarter.

$$\bar{d}_i = (\sum d_i)/(\sum I_{d_i>0}) \tag{5}$$

Where $I_{d_i>0} = \begin{cases} 1 & if \quad d_i > 0 \\ 0 & othwerwise \end{cases}$

$$\bar{L}_i = \sum_j L_{ij} / d_i \tag{6}$$

Where d_i is the demands in the quarter i , $I_{d_i>0}$ is to determine whether the there is demand in quarter i or not, and L_{ij} is the delays of flight *j* at quarter *i*.

IV. DATA ANALYSIS RESULTS

This part will discuss the data analysis results for 2007 summer at NYC airports. Next, we will discuss the results for number of markets NYC served, average flight number per day, number of competing airlines, airfares, the load factor, and the average aircraft seat size for the identified markets.

A. Number of Markets to NYC

First, a market is defined as the airport has more than 10 flights in 2007 summer flying from NYC airports. After processing the data, we got that the NYC severs 109 domestic markets in 2007 summer (Figure 7). And EWR serves 88 of them, JFK severs 64, and LGA serves 73 (Table 2).

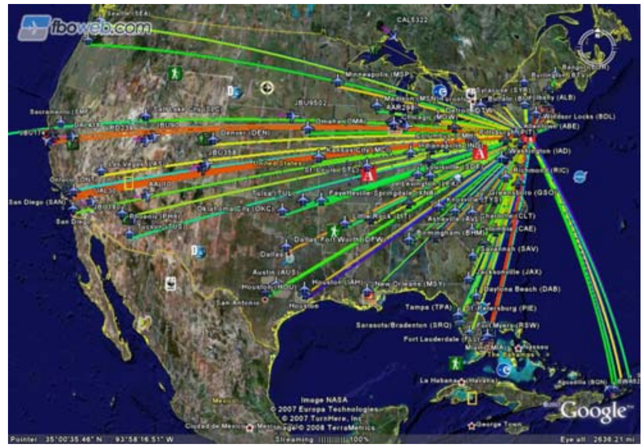

Figure 7. NYC metroplex served markets route map

Table 2. Number of markets served by NYC three airports

Airport Code	# of Markets Served
EWR	88
JFK	64
LGA	73

Table 3 also lists the details about how NYC airports sever the 109 markets. 39 (35.8%) of the identified 109 markets are served by all 3 airports; 38 (34.9%) are served by two of three airports; and only 32 (29.4%) are served by only one of three airports

Table 3. Redundancy of service in NYC airports

# of NYC serving airports	1	2	3
# of markets served	32	38	39
% of markets served	29.4%	34.8%	35.8%

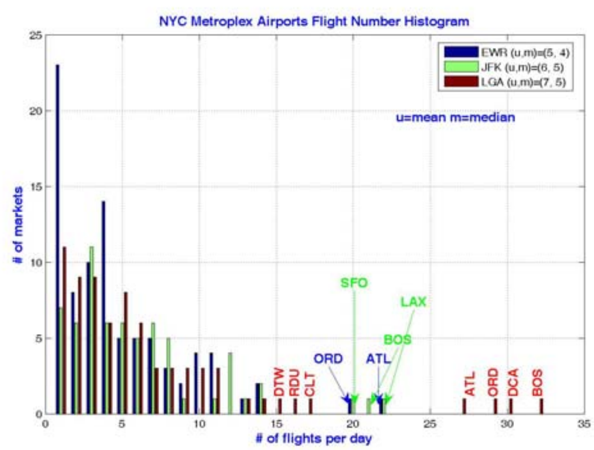

Figure 8. NYC metroplex airports flight number/day histogram

B. Flight Number per day for NYC Markets

Figure 8 shows the average flight number per day histogram for NYC metroplex airports. Figure 8 shows that the average flight number per day per market are 5 to EWR, 6 to JFK, and 7 to LGA. And the medians are 4 to EWR, 5 to JFK and LGA. Therefore, approximate (5x88+6x64+7x73)=1335 flights each day are flying to/from NYC metroplex airports. In addition, most of markets have less than 8 flights per day. Table 4 also lists the top 20 markets to NYC, and 51% of flights are from these top 20 markets.

Table 4. Top 20 markets to NYC airports

Airport	Code	# of flights/day			
		EWR	JFK	LGA	NYC
Boston Logan MA	BOS	10	21	32	63
Chicago O'Hare IL	ORD	20	11	29	60
Atlanta Hartsfield GA	ATL	22	5	27	55
Ronal Reagan National DC	DCA	7	8	30	45
Raleigh Durham NC	RDU	11	13	17	40
Fort Lauderdale FL	FLL	13	12	14	39
Charlotte NC	CLT	14	7	17	38
Orlando FL	MCO	14	14	9	36
Detroit MI	DTW	11	5	15	32
Los Angeles CA	LAX	10	22		31
Dallas Ft. Worth TX	DFW	11	5	13	29
San Francisco CA	SFO	9	20		29
Dulles VA	IAD	10	12	6	28
Buffalo NY	BUF	5	14	7	26
Pittsburg PA	PIT	8	8	8	24
Miami FL	MIA	8	6	10	24
Houston Bush Int. TX	IAH	11	4	9	23
Port Columbus OH	CMH	5	7	11	23
Cincinnati OH	CVG	8	4	10	22
Cleveland[Hopkins Intl] OH	CLE	6	4	11	21

C. Number of Airlines for NYC Markets

Figure 9 shows the number of airlines histogram for the NYC metroplex airports where most markets are served by 11 airlines, when considering both major carriers and their regional associated airlines as separate entities.

D. Airfares for NYC Markets

Figure 10 is the airfare histogram for NYC metroplex airports. The average segment airfare to/from EWR is $179, to/from JFK is $154, and to/from LGA is $159. In addition, virtually all markets have segment airfares below $250. Table 5 also lists the segment prices to top 20 markets. This table illustrates that the price can vary substantially among the three NYC airports. For example, BOS to EWR has an average segment fare of $201, for JFK to BOS, the average segment fare is $87, and for LGA is $149. Since the prices are often dependent on the prices charged by the competitor airlines in that market and whether these flights segments are part of longer non-direct flights and reflect only part of total airfare to the passenger.

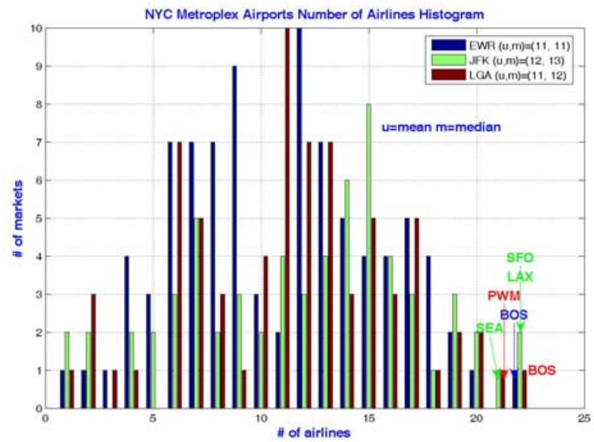

Figure 9. NYC metroplex airports number of reporting airlines histogram

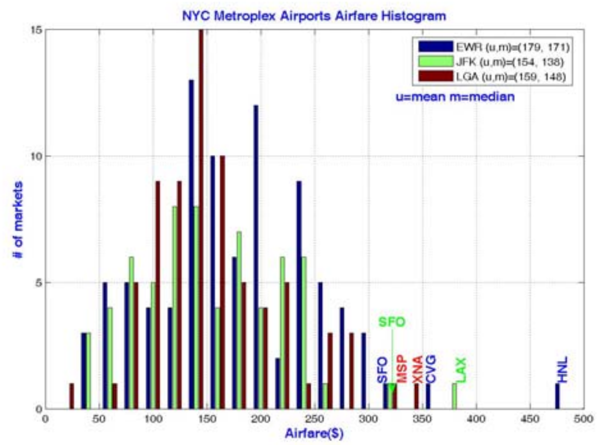

Figure 10. NYC metroplex airports airfares histogram

Table 5. Airfares of top 20 markets to NYC airports

Airport	Code	Airfares($)			
		EWR	JFK	LGA	NYC
Boston Logan MA	BOS	201	87	149	141
Chicago O'Hare IL	ORD	170	130	169	164
Atlanta Hartsfield GA	ATL	180	197	171	175
Ronal Reagan National DC	DCA	210	112	165	165
Raleigh Durham NC	RDU	136	113	130	126
Fort Lauderdale FL	FLL	130	133	133	132
Charlotte NC	CLT	155	122	149	143
Orlando FL	MCO	135	136	143	137

Detroit MI	DTW	269	151	141	159
Los Angeles CA	LAX	294	374		344
Dallas Ft. Worth TX	DFW	312	329		323
San Francisco CA	SFO	283	221	286	279
Dulles VA	IAD	112	96	182	112
Buffalo NY	BUF	109	95	98	97
Pittsburg PA	PIT	133	88	129	113
Miami FL	MIA	153	138	146	147
Houston Bush Int. TX	IAH	300	203	264	267
Port Columbus OH	CMH	147	96	116	116
Cincinnati OH	CVG	352	265	279	291
Cleveland[Hopkins Intl] OH	CLE	211	145	165	182

E. Aircraft Seat Size for NYC Markets

Figure 11 shows the mean aircraft size histogram for the NYC metroplex airports. The average aircraft seat size to LGA is the smallest (75) of the three airports, while JFK is the largest (114). In addition, at EWR and LGA, over one third of the markets are serving with aircrafts having 50 seats or less.

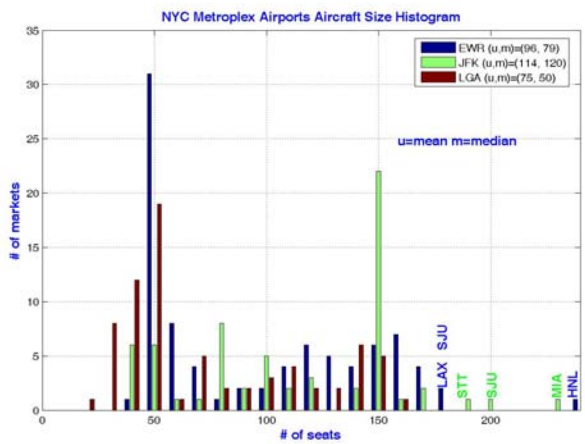

Figure 11. NYC metroplex airports aircraft seat size histogram

F. Load Factor for NYC Markets

Figure 12 presents a load factor histogram for NYC metroplex airports. The graph illustrates that EWR and LGA has better load factors than LGA, with the smallest load factor being the Lewisburg West Virginia market whose load 0.26. At EWR and JFK, over 90% of markets have load factors greater than 0.6, assuring the profitability of most flights during the summer of 2007.

G. Seat Size vs. Load Factor for NYC Markets

Figure 13 shows the relationship between aircraft size and load factor for NYC metroplex. For most markets, aircraft size and load factor are positively correlated. However, for high frequent markets such as DCA, BOS, CLE, the load factors are not so high (less than 0.7). Thus, for these highly-competitive markets, if each of the airlines serving these regions wishes to maintain frequency, they must choose smaller aircrafts to maintain profitability. Alternatively, passengers could be equally well-served with larger aircraft but fewer airlines serving these markets. It is on these high-demand routes where upguaging would help rationalize aircraft and runway capacity.

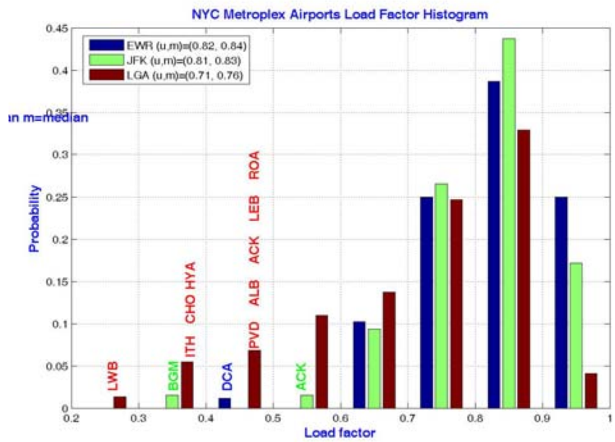

Figure 12. NYC metroplex airports load factor histogram

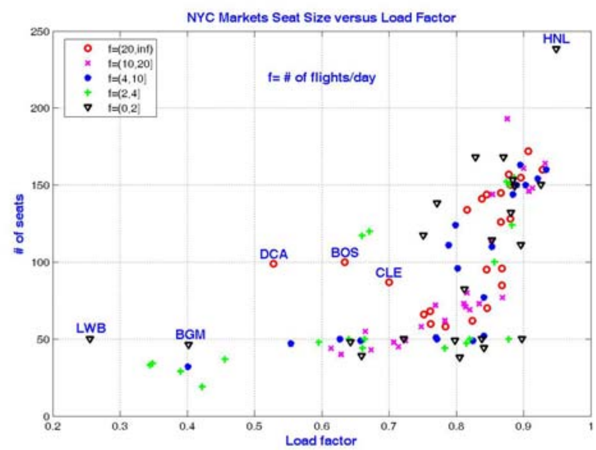

Figure 13. NYC metroplex seat size vs. load factor classified by flight frequency

H. Unit Revenue vs. Load Factor for NYC Markets

Figure 14 presents the relationship between unite revenue(=airfare/distance) and load factor. It shows that there are negative correlations between the unite revenue and load factor.

I. Flight Frequency vs. Seat Size for NYC Markets

Figure 15 shows the relationship between frequency (# of flights per day) and seat size, and it is classified by load factor. It shows that most of low load-factor aircrafts are small and providing services infrequency. Not surprisingly, flights to

Hawaii (HNL) are serviced by large airplanes and have very high load factors.

Moreover, Figure 16 shows the relationship between frequency (# of flights per day) and seat size, and it is classified by unit revenue. It shows that long distance flights usually have small unit revenue.

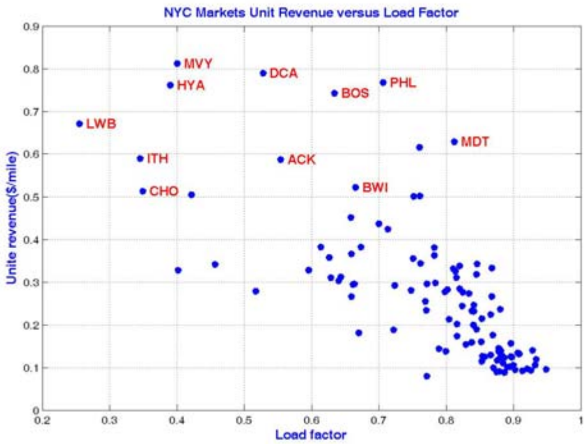

Figure 14. NYC metroplex unit reveune vs. load factor

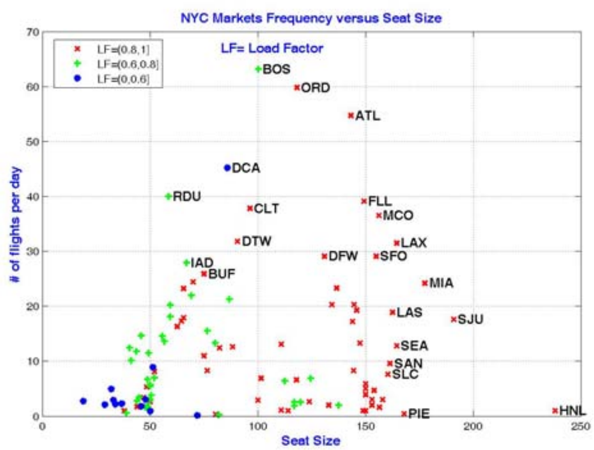

Figure 15. NYC markets frequency vs. seat size and classified by load factor

V. CONCLUSION

The paper is the first in a series of papers that studies the NYC Metroplex of airports. This paper took a first look at the status of each airport in terms of the markets served, seat capacity, delays, locad factors, and other similar characteristics. The results identify the NYC metroplex serving 109 domestic markets (EWR:88, JFK:64, LGA:73). The results also indicate that NYC metroplex exhibits redundant service. 39 (35.8%) of the identified 109 markets are served by all 3 airports; 38 (34.9%) are served by two; and only 32 (29.4%) are served by one. Demands such as the number of flights per day per market (EWR: mean=5 and median=4; JFK: mean=6 and median=5; LGA: mean=7 and median=5) and load factors (EWR: mean=0.82 and median=0.84; JFK: mean=0.81 and median=0.83; LGA: mean=0.71 and median=0.75) imply the heavy demands to NYC airports. The number of airlines serving a market (EWR: mean=11 and median=11; JFK: mean=12 and median=13; LGA: mean=11 and median=12) presents that NYC markets are highly competitive, which also forces the airfares (EWR: mean=$179 and median=$171; JFK: mean=$154 and median=$138; LGA: mean=$159 and median=$148) down. The average aircraft sizes (EWR: mean=96 and median=79; JFK: mean=114 and median=120; LGA: mean=75 and median=50) presents us the opportunity of upguaging in order to reduce demands for the crowded NYC airports. The relationship between seat size and load factor discloses a possible way to help reduce airport demands by proper upguaging small airplanes to big ones. The data will also be also used in our future research that will examine the impacts of both regulation and cost on these markets with the goal of determining how best to allocate the scare runway capacity that exists within this region.

ACKNOWLEDGMENTS

Work was supported by NASA Grants NRA NNH06ZEA001N-AS. Thank you for technical comments and suggestions to David Smith, Bengi Manley, John Shortle, Rajesh Ganesan, and Vivek Kumar (GMU).

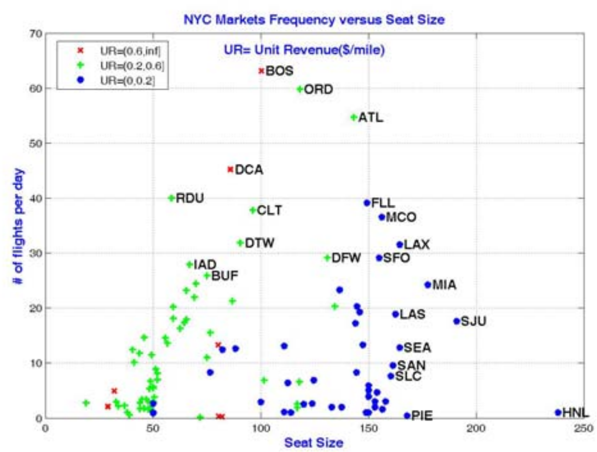

Figure 16. NYC markets frequency versus seat size and classified by unit revenue

REFERENCES

[1] Summary of airline on-time performance year-to-date through december 2007. http://www.bts.gov/programs/airline_information/airline_ontime_tables/2007_12/html/table_01.html. Accessed December 2007.

[2] Airport snapshort. http://www.transtats.bts.gov/airports.asp?pn=1/. Accessed December 2007.

[3] X. Ning, "Method for deriving multi-factor models for predicting airport delays," Ph.D. dissertation, George Mason University, December 2007.

[4] L. Le, "Demand management at congested airports: how far are way from utopia?" Ph.D. dissertation, George Mason University, August 2006.

[5] Bureau of transportation statistics (BTS) databases and statistics. http://www.transtats.bts.gov/. Accessed December 2007.

[6] Aviation system performance metrics (ASPM)–complete. FAA. http://www.apo.data.faa.gov/aspm/entryASPM.asp. Accessed December 2007.