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 104 domestic airports. 36.5% of airports serve all three New York airports, while 35.6% serve two of the airports. For all the routes to NYC, the average number of flights per day is 6 with a maximum of 32. These routes have an average aircraft seat size ranging from 19 to 238 with an average of 94 seats per flight. These routes had passenger load factors ranging from 0.26 to 0.95 with an average of 0.78. This yields an average of 281 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 104 domestic airports.
- 36.5% of airports serve all three New York airports, while 35.6% serve two of the airports.
- For all the routes to NYC, the average number of flights per day is 6 with a maximum of 32.
- These routes have an average aircraft seat size ranging from 19 to 238 with an average of 94 seats per flight.
- These routes had passenger load factors ranging from 0.26 to 0.95 with an average of 0.78.
- This yields an average of 26,197 unused seats over 119, 004 provided seats on all the routes each day.

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

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.

ISBN: 978-0-615-20720-9

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%

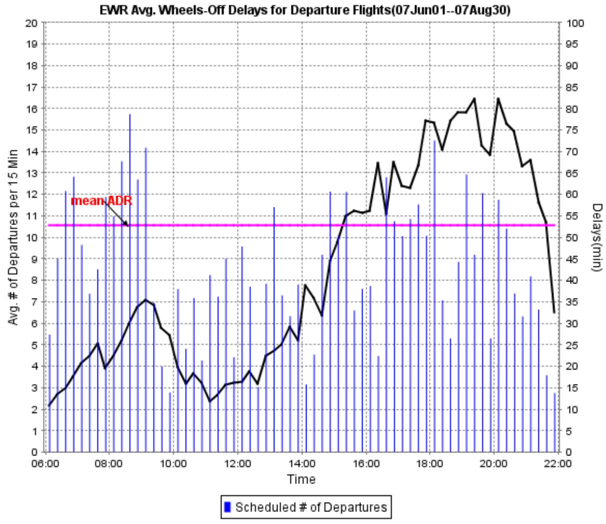

Figure 1. 2007 summer EWR mean capacity, departures and wheels-off delays per 15 min

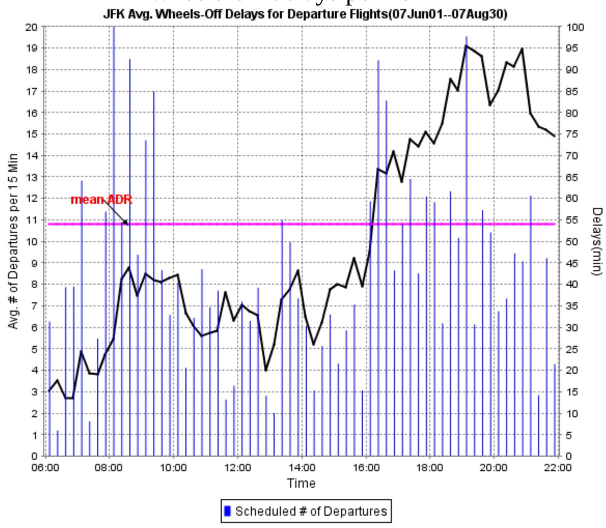

Figure 2. 2007 summer JFK mean capacity, departures and wheels-off delays per 15 min

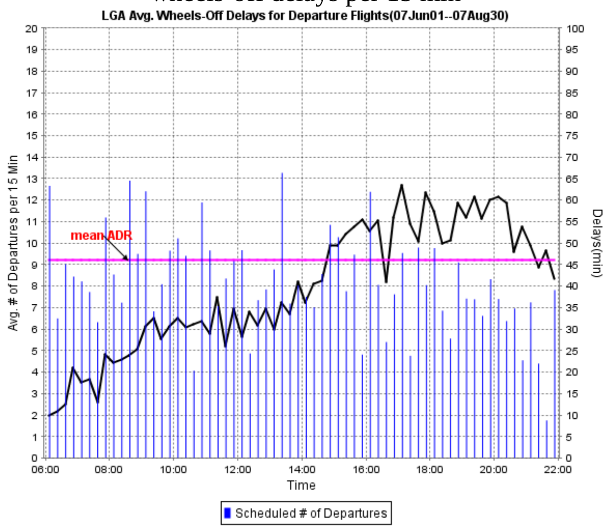

Figure 3. 2007 summer LGA mean capacity, departures and wheels-off delays per 15 min

 ISBN: 978-0-615-20720-9

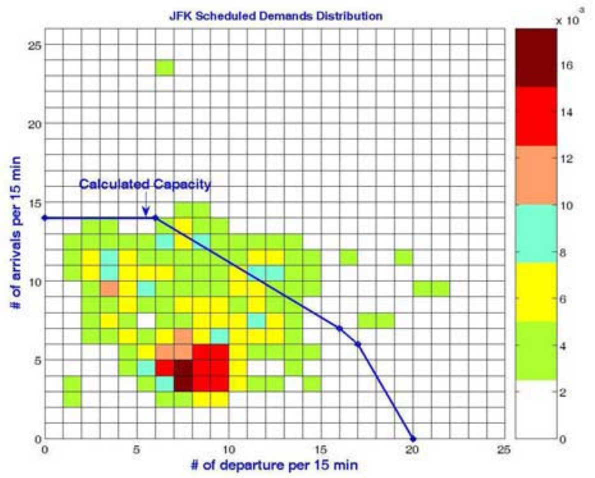

Figure 4. 2007 summer JFK scheduled demands distribution

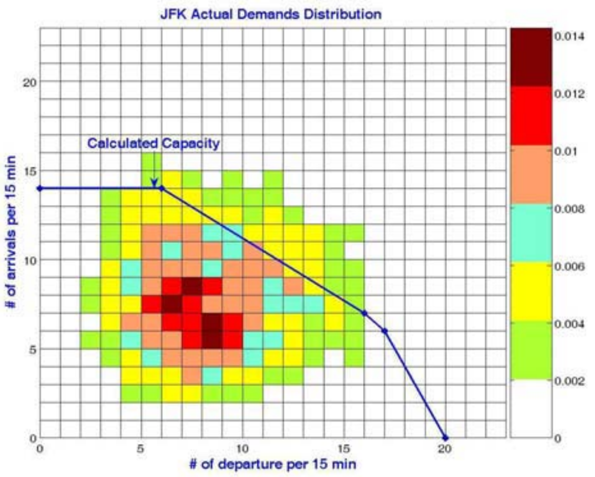

Figure 5. 2007 summer JFK actual demands distribution

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)

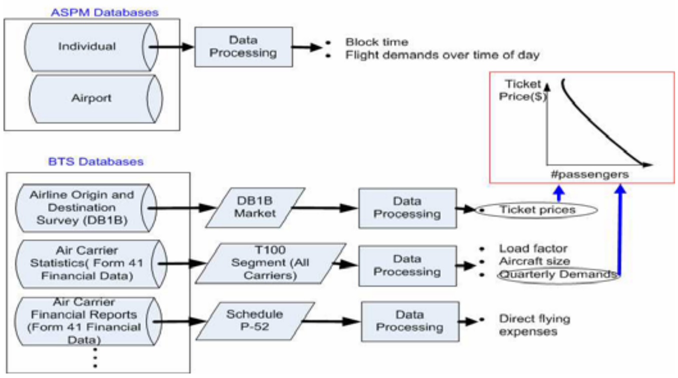

Figure 6. Data processing flow chart

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.

$$T_{AB} = T_{AC} \times \frac{\sqrt{d_{AB}}}{\sqrt{d_{AB}} + \sqrt{d_{BC}}} \quad (4)$$

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

ISBN: 978-0-615-20720-9

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.

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

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

$$\bar{L}_i = \sum_j L_{ij} / d_i \quad (6)$$

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. The data from major airlines are only counted in T100. A major airline is defined as the airline which has more than 60 arrival flights to the studied airport in 2007 (Jun 01-Aug 30). Next, we will discuss the detail statistics results such as number of airports NYC served, average flight number per day, number of competing airlines, airfares, the load factor, and the average aircraft seat size for the identified airports etc.

A. Number of Airports to NYC

After processing the data, we got that the NYC severs 104 domestic airports in 2007 summer (Figure 7). And EWR serves 81 of them, JFK severs 62, and LGA serves 68 (Table 2).

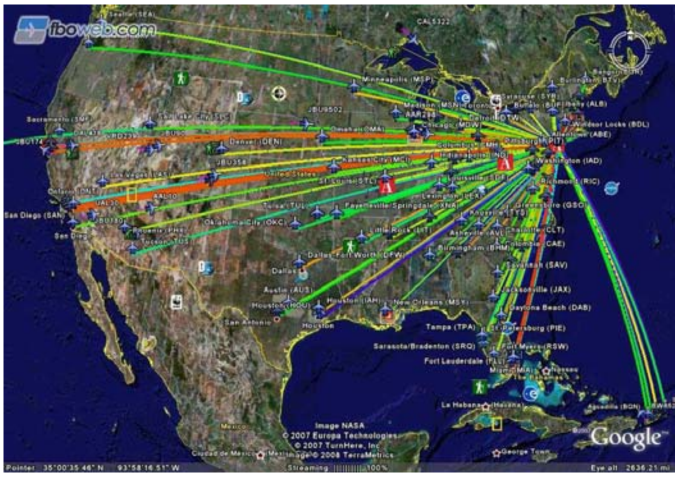

Figure 7. NYC metroplex served airports route map

Table 2. Number of airports served by NYC three airports

Airport Code	# of Airports Served
EWR	81
JFK	62
LGA	68

Table 3 also lists the details about how NYC airports sever the 104 airports. 38 (36.5%) of the identified 104 airports are served by all 3 airports; 37 (35.6%) are served by two of three airports; and only 29 (27.9%) are served by only one of three airports

Table 3. Redundancy of service in NYC airports

# of NYC serving airports	1	2	3
# of airports served	29	37	38
% of airports served	27.9%	35.6%	36.5%

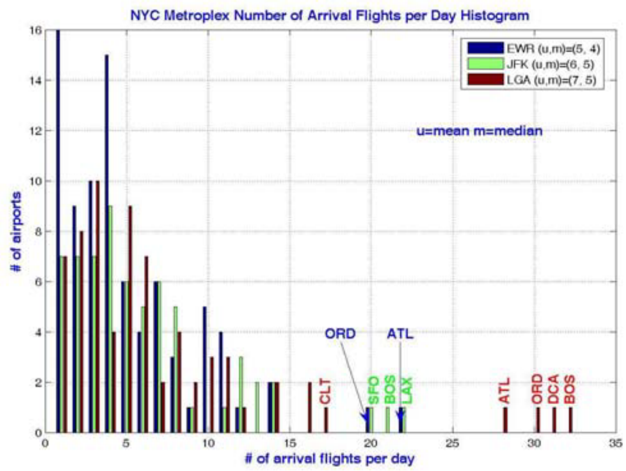

Figure 8. NYC metroplex arrival flight number/day histogram

B. Flight Number per day for NYC Airports

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 (5x81+6x62+7x68)=1253

ISBN: 978-0-615-20720-9

flights each day are flying to/from NYC metroplex airports. In addition, most of airports have less than 15 flights per day. Table 4 also lists the top 20 airports to NYC, and 53% of flights are from these top 20 airports.

Table 4. Top 20 airports to NYC airports

Airport	Code	# of arrival flights/day			
		EWR	JFK	LGA	NYC
Boston Logan MA	BOS	10	21	32	63
Chicago O'Hare IL	ORD	20	11	30	61
Atlanta Hartsfield GA	ATL	22	5	28	55
Ronal Reagan National DC	DCA	7	8	31	46
Raleigh Durham NC	RDU	11	13	16	40
Fort Lauderdale FL	FLL	12	13	14	39
Charlotte NC	CLT	14	7	17	38
Orlando FL	MCO	14	14	9	37
Detroit MI	DTW	10	22		32
Los Angeles CA	LAX	11	5	16	32
Dallas Ft. Worth TX	DFW	10	20		30
San Francisco CA	SFO	11	5	14	30
Dulles VA	IAD	10	12	6	28
Buffalo NY	BUF	5	14	7	26
Pittsburg PA	PIT	8	6	11	25
Miami FL	MIA	11	4	9	24
Houston Bush Int. TX	IAH	8	8	8	24
Port Columbus OH	CMH	5	7	11	23
Cincinnati OH	CVG	8	4	10	22
Cleveland[Hopkins Intl] OH	CLE	6	4	11	21

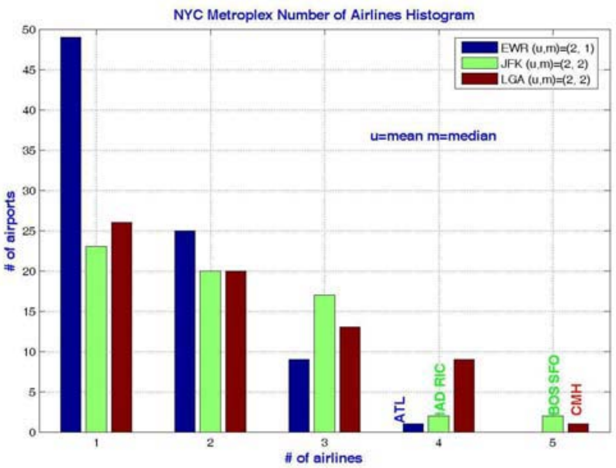

Figure 9. NYC metroplex airports number of airlines histogram

C. Number of Airlines for NYC Airports

The airline number is got from T100, and only the major airline is counted. Figure 9 shows the number of airlines histogram for the NYC metroplex airports where each route are served by 2 airlines in average. Route CMH-LGA, BOS-JFK, SFO-JFK are most competitive routes, with 5 airlines severing on these three routes.

D. Airfares for NYC Airports

Table 5 also lists the segment prices to top 20 airports. 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 $200, for BOS to JFK, the average segment fare is $86, and for LGA is $148. Therefore, BOS-EWR is most expensive way to NYC. This can be explained by airline competition which drives the price down. From BOS to EWR, only one airline flies on this route, however five airlines fly from BOS to JFK, and 3 airlines are flying from BOS to LGA. In addition, Figure 10 is the airfare histogram for NYC metroplex airports. The average segment airfare to EWR is $175, to JFK is $149, and to LGA is $151. In addition, virtually all airports have segment airfares below $250.

Table 5. Airfares of top 20 airports to NYC airports

Airport	Code	Airfares($)			
		EWR	JFK	LGA	NYC
Boston Logan MA	BOS	200	86	148	140
Chicago O'Hare IL	ORD	164	128	157	156
Atlanta Hartsfield GA	ATL	152	161	147	150
Ronal Reagan National DC	DCA	192	102	142	143
Raleigh Durham NC	RDU	134	112	127	125
Fort Lauderdale FL	FLL	131	132	133	132
Charlotte NC	CLT	127	117	127	125
Orlando FL	MCO	134	135	142	136
Detroit MI	DTW	290	361		335
Los Angeles CA	LAX	193	127	136	146
Dallas Ft. Worth TX	DFW	309	326		320
San Francisco CA	SFO	249	189	239	237
Dulles VA	IAD	81	90	106	93
Buffalo NY	BUF	105	95	97	97
Pittsburg PA	PIT	152	137	146	146
Miami FL	MIA	264	188	212	228
Houston Bush Int. TX	IAH	122	88	125	111
Port Columbus OH	CMH	146	95	115	115
Cincinnati OH	CVG	173	141	181	174
Cleveland[Hopkins Intl] OH	CLE	198	137	150	168

ISBN: 978-0-615-20720-9

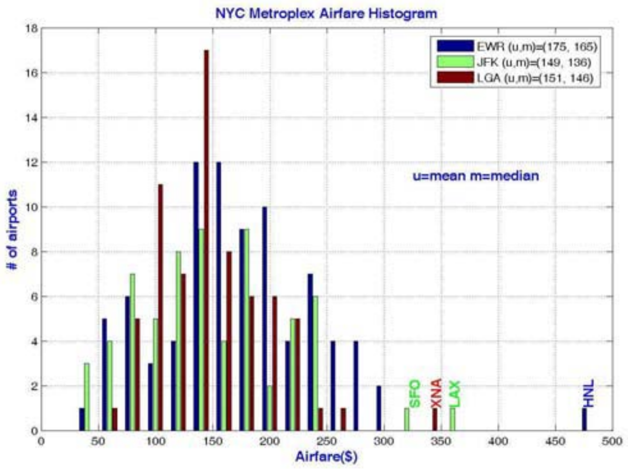

Figure 10. NYC metroplex airports airfares histogram

E. *Aircraft Seat Size for NYC Airports*

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 (113). In addition, at EWR and LGA, over one third of the airports are serving with aircrafts having 50 seats or less.

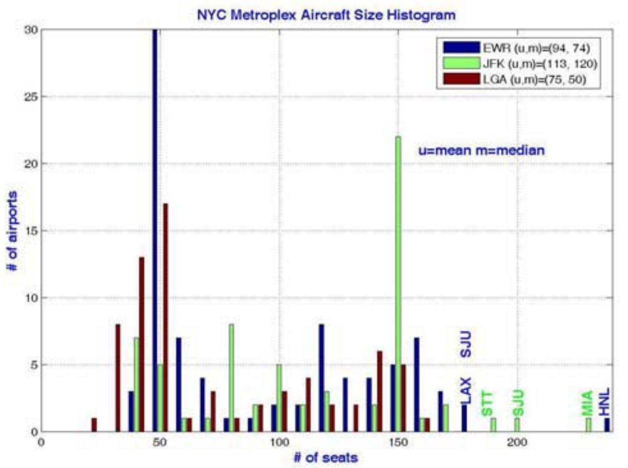

Figure 11. NYC metroplex airports aircraft seat size histogram

F. *Load Factor for NYC Airports*

Figure 12 presents a load factor histogram for NYC metroplex airports. The graph illustrates that EWR and LGA has better load factors than LGA. LGA-LWB (Lewisburg, West Virginia) has the smallest load factor (0.26). At EWR and JFK, over 90% of airports 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 Metroplex*

Figure 13 shows the relationship between aircraft size and load factor for NYC metroplex. For most airports, aircraft size and load factor are positively correlated. However, for high frequent airports such as DCA, BOS, CLE, the load factors are not so high (less than or equal to 0.7). Thus, for these highly-competitive airports, 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 airports. It is on these high-demand routes where upguaging would help rationalize aircraft and runway capacity.

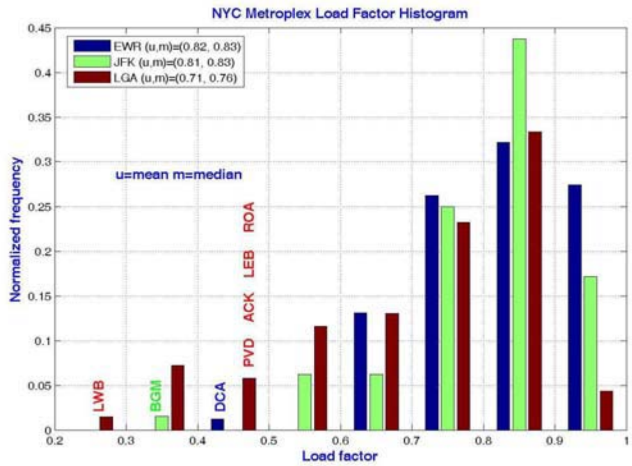

Figure 12. NYC metroplex airports load factor histogram

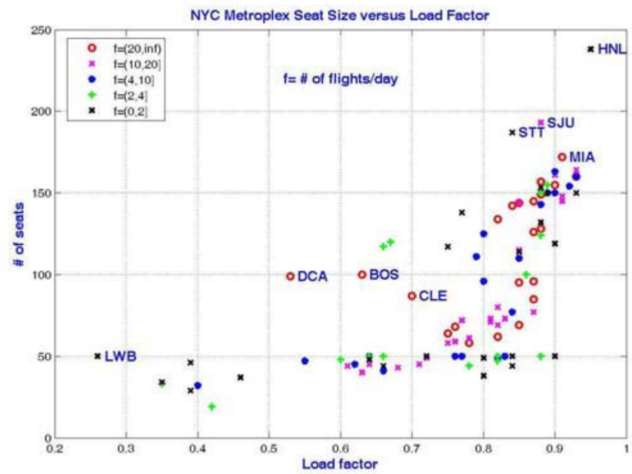

Figure 13. NYC metroplex seat size vs. load factor classified by flight frequency

H. *Unit Revenue vs. Load Factor for NYC Metroplex*

Figure 14 presents the relationship between unit revenue (=airfare/distance) and load factor. It shows that there are negative correlations between the unite revenue and load factor. MVY (Martha's Vineyard Airport, MA) to NYC has the largest unite revenue (0.83$/mile), however, TUS (Tucson, AZ) to NYC has the smallest unite revenue (0.08$/mile).

I. *Flight Frequency vs. Seat Size for NYC Metroplex*

Figure 15 shows the relationship between frequency (# of arrival flights per day) and seat size, and it is classified by load factor. It shows that most low load-factor aircraft are small and

ISBN: 978-0-615-20720-9

providing service infrequently. 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 arrival flights per day) and seat size, and it is classified by unit revenue. It shows that long distance routes usually have lower unit revenue than some short, low load factor routes.

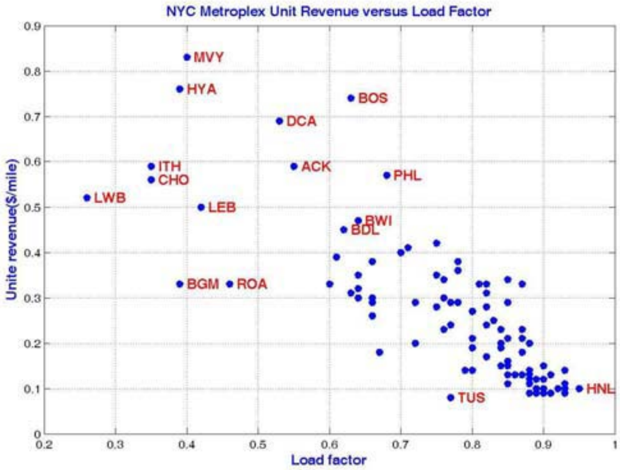

Figure 14. NYC metroplex unit reveune vs. load factor

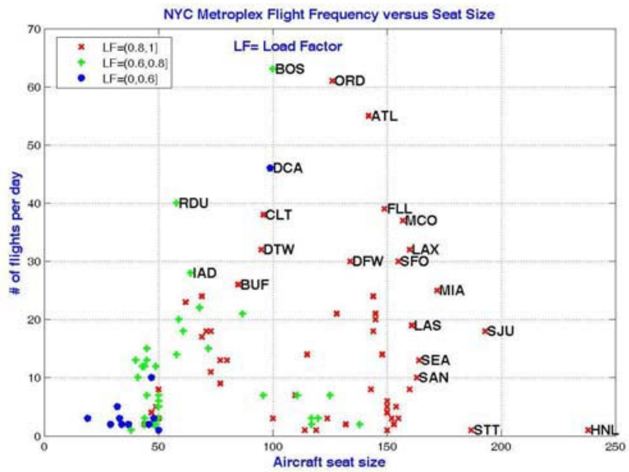

Figure 15. NYC airports frequency vs. seat size and classified by load factor

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 NYC airports in terms of the airports served, seat capacity, airfares, load factors etc. The results identify the NYC metroplex serving 104 domestic airports (EWR:81, JFK:62, LGA:68). The results also indicate that NYC metroplex exhibits redundant service. 38 (36.5%) of the identified 104 airports are served by all 3 NYC airports; 37 (35.6%) are served by two; and only 29 (27.9%) are served by one. Demands such as the number of arrival 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=0.83; JFK: mean=0.81 and median=0.83; LGA: mean=0.71 and median=0.76) imply the heavy passenger demands to NYC airports. The number of airlines serving a market (EWR: mean=2 and median=1; JFK: mean=2 and median=2; LGA: mean=2 and median=2) presents that NYC airports are competitive, which also forces the airfares (EWR: mean=$175 and median=$165; JFK: mean=$149 and median=$136; LGA: mean=$151 and median=$146) down. The average aircraft sizes (EWR: mean=94 and median=74; JFK: mean=113 and median=120; LGA: mean=75 and median=50) presents us the opportunity of upguaging in order to reduce congestion for the crowded NYC airports. The relationship between seat size and load factor discloses a possible way to help reduce congestion by proper upguaging small airplanes to big ones. The data will also be used in our future research that will examine the impacts of both regulation and cost on these airports 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).

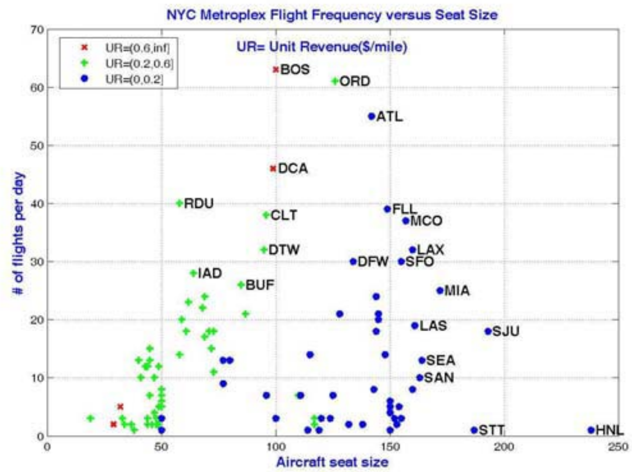

Figure 16. NYC airports frequency versus seat size and classified by unit revenue

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.

ISBN: 978-0-615-20720-9

[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.

ISBN: 978-0-615-20720-9