

AIRPORT THROUGHPUT CAPACITY LIMITS FOR DEMAND MANAGEMENT PLANNING

Vivek Kumar, Lance Sherry

Center for Air Transportation Systems Research, George Mason University, Fairfax, VA

Abstract

The success of demand management approaches to curbing flight delays is predicated on an accurate and reliable *a-priori* determination of airport arrival and departure capacity. The assignment of declared capacity, completed more than one year in advance, must consider the inherent variability in the actual throughput capacity for each 15 minute period during the day. For airports with large variations in airport capacity, the capacity limit must be set to high enough avoid long periods of under utilization of the airport resources and low enough to avoid extended periods of excessive delays.

This paper examines: (i) the variability of throughput capacity at the OEP-35 airports during the convective weather season in 2008, (ii) established the average costs of delays due to reduced capacity and the average profits per flight at each airport, (iii) given the cost of delays and lost profits, an optimum capacity is determined at each airport to minimize both the under-utilization of airport capacity and flight delays due to variability in the available throughput capacity

Keywords – Airport Capacity, arrival rate, departure rate, demand management, delays, over-scheduling, underutilization.

I. Introduction

The IATA Schedule Coordination (IATA, 2000) and High Density Rule (FAA, 1968), curtail scheduled arrivals and departures within the capacity limits of the airport. For example, the High Density Rule established limits on the number of takeoffs and landings allowed by the incumbent airlines at five U.S. airports (Chicago O'Hare International, Newark, JFK, LaGuardia, and Washington Reagan National). JFK and Newark are capped at 81 arrivals and departures slots per hour (i.e. 1458 slots between 6am and 10pm or 20.25 slots per 15 minutes), and La Guardia is capped at

75 arrival and departure slots per hour (i.e. 1350 slots between 6am and 10pm or 18.75 slots per 15 minutes).

The success of these programs in curtailing flight delays is predicated on the viability of the assigned capacity limits. There are two symptoms of problems with the way capacity limits are determined. First, the presence of excessive delays at these airports and the need for Ground Delay programs at these airports indicates that the assigned capacity limits (and the associated rules and exemptions) that are used for scheduling do not reflect the actual available capacity.

The success of demand management approaches to curbing flight delays is predicated on an accurate and reliable *a-priori* determination of airport arrival and departure capacity. The assignment of declared capacity, completed more than one year in advance, must consider the inherent variability in the actual throughput capacity for each 15 minute period during the day. Sources of variation include the reduced capacity runway configurations due to wind, reduced throughput due to: weather, arrival-demand profiles, fleet-mix, unscheduled operations, and environmental (noise) considerations. For airports with large variations in airport capacity, the capacity limit must be set to high enough avoid long periods of under utilization of the airport resources and low enough to avoid extended periods of excessive delays.

This paper examines: (i) the variability of throughput capacity at the OEP-35 airports during the convective weather season in 2008, (ii) established the average costs of delays due to reduced capacity and the average profits per flight at each airport, (iii) given the cost and lost profits, an optimum capacity was determined at each airport to minimize both the under-utilization of airport capacity and flight delays due to variability in the available throughput capacity (see Figure 1).

The results of this analysis are as follows:

- 1 13 of the OEP-35 airports exhibit variations in runway configurations that result in a reduction of more than 20% in capacity more than 10% of the time. Included in this category are Cleveland, Boston, Cincinnati, San Francisco, Chicago O’Hare and JFK.
- 2 24 of the OEP-35 airports exhibit an average cost of delays per flight in excess of the average profit generated by a flight. Included in this category are JFK, Newark and La Guardia
- 3 Optimum airport capacity, computed by trading-off flights delays and underutilization ranged from 81% to 100% of the maximum airport capacity. The average optimum airport capacity was 93% of the maximum airport capacity. Twelve airports, including Boston, JFK, Newark, and San Francisco, exhibited an optimum airport capacity below the average 93%.
- 4 These results have several implications. First, the variability in the available airport capacity throughput is critical in the *a priori* establishment of airport capacity limits for demand management. Second, the range of the ratio of costs of delays due

to reductions with the potential profits per flight indicates the non-homogeneous nature of the service providers in the National Airspace System and the need to take into account the airline revenue and cost structures in determining resource availability. Third, “economically optimal” capacity limits established to trade-off delays and underutilization, provide a rational approach to assigning capacity limits that take into account airline preferences.

This paper is organized as follows: Section II provides background on the factors that contribute to throughput capacity variability. Section III describes the sources of data and methods for the analysis. Section IV describes the results of the analysis. Section V discusses the implications of these results and limitations of this study.

II. Determining Airport Capacity

The a-priori computation of airport capacity limits for the purpose of regulating scheduled arrivals and departures is an inherently complex process. First, the runway capacity of an airport is a random variable that is determined by probabilistic phenomena, such as weather, that by definition are difficult to predict. Second, the capacity of the airport is also dependent on the fleet mix of aircraft. Fleet-mixes with larger aircraft will require greater separation distances between arrivals/departures to account for wake vortex and runway occupancy times. Third, the capacity of the airport is also dependent on the orderliness of the arrival flow. Variances in inter-arrival times can result in “bunching” of arriving aircraft that result in excessive queuing delays. Fourth, the capacity limit is also determined by the proficiency of the Air Traffic Controllers on a given shift.

There are several alternative measures of runway capacity that have been proposed and used for policy-making and operational control. All of the measures provide an estimate of the number of aircraft movements (arrivals and departures) that can be performed on the runways of the airport during a specified time frame.

For the purpose of a-priori slot coordination, Declared Capacity is used to set a limit on the number of movements that can be scheduled at an airport. Declared Capacity is a measure defined as

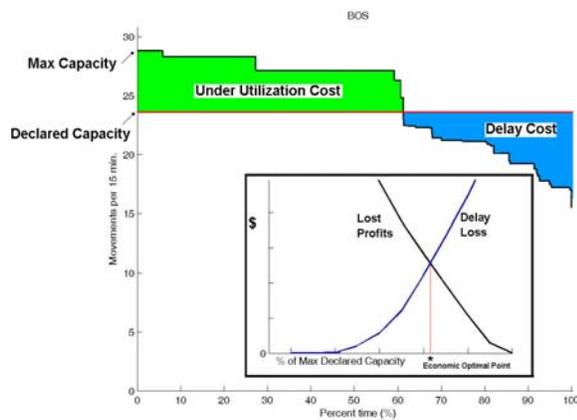


Figure 1: Probability distribution of airport throughput capacity. Shaded regions represent the under-utilized capacity (resulting in lost revenues) and the over-utilized regions (resulting in additional costs due to delays). The inset chart shows the tradeoff between lost revenues and additional costs for airport throughput capacity limits set below the maximum airport capacity. The intersection of the curves represents the economic optimum for the throughput capacity for the airport.

the number of aircraft movements per hour that an airport can accommodate at a “reasonable level of service” defined in terms of delays. To ensure reliability in the estimate, Declared Capacity is generally set in the range of 85% to 90% of the Maximum Throughput Capacity.

The fundamental measure of airport capacity is known as the Maximum Throughput Capacity (or Saturation Capacity) of the airport (Neufville, Odoni, 2004). This measure defines the expected number of movements that can be performed in one hour. This measure assumes operations that do not violate Air Traffic Control rules (e.g. separation distances, and runway occupancy) and continuous, evenly spaced arrival flow. The Maximum Throughput Capacity does account for the features of the airport and the expected fleet mix. The Maximum Throughput Capacity is not set with consideration for the number of minutes of delay acceptable.

Two other measures used include: (1) Practical Hourly Capacity (PHCAP) is a measure of the capacity of the airport in a one hour period when the arrivals and departures experience an average delay of 4 minutes (Neufville, Odoni, 2004). Due to queuing phenomena of stochastic arrivals, PHCAP is generally between 80% and 90% of Maximum Throughput Capacity. (2) Sustained Capacity is a measure defined as the number of movements per hour that can be sustained over a period of several hours. This measure of capacity is designed to take into account the effects of ATC workload. Although desirable to operate at maximum capacity, studies have demonstrated that this level of performance cannot be maintained for more than two hours at a time. Sustained Capacity is generally set at 90% of Maximum Throughput Capacity in favorable weather conditions to manage variations in inter-arrival times. Sustained capacity is set at 100% of the low Maximum Throughput Capacity in unfavorable weather conditions to ensure a high level of throughput.

Reduction in airport capacity is predominantly due to change in Runway configuration. For certain airports, reduction in airport capacity due to change in runway configuration can be considerably large and may result in excessive delays. For example: CLE airport when operating under the configuration 6L,6R|6R (Arrival|Departure) has a capacity of 28

operations per 15 min. If however, the runway configuration changes to 6L|6C (Arrival|Departure), the airport capacity is reduced by 29% to 18 operations per minute. These changes in Runway configuration are mainly influenced by wind direction and strength, which by definition is difficult to predict.

All airports have unique runway layout and weather/wind conditions because of their geographical locations. Airports like Boston experience a lot of fluctuation in capacity due to runway configuration changes primarily due to wind effects; whereas airports like MCO(Orlando International) hardly see any change in capacity. Capacity Coverage Chart (CCC) is a convenient way to summarize the range of capacities at an airport and the frequency with which various levels of capacity are available (Neufville, Odoni, 2004). A ‘step-like’ CCC represents an airport with ‘many’ levels of capacity. For example, Figure 3 shows the CCC for BOS for the period of June-August 2007. BOS operates at or above its mean capacity of approximately 23 operations per 15 min, only about 60% of the time. However, MCO(Figure 4) operates close to its maximum capacity almost the entire time.

Given the different distribution of airport capacity for different airports, to achieve identical ‘level of service’, it is not mathematically prudent to set their declared capacities to a fixed percentage (80 or 90% for example) of maximum capacity. A traditional approach to quantifying the measure of reliability in such cases as used by electrical engineers is Signal-to-Noise Ratio(SNR), which is as a way of comparing desired signal(runway capacity) to the level of background noise(variance in capacity). High values of SNR imply lesser relative variance and therefore higher reliability in capacity. In terms of CCC, high SNR implies a flatter curve and lower SNR ratio implies a more step-like shape of CCC. SNR is therefore a measure of reliability in airport capacity and can be thought of as ‘Reliability Index’.

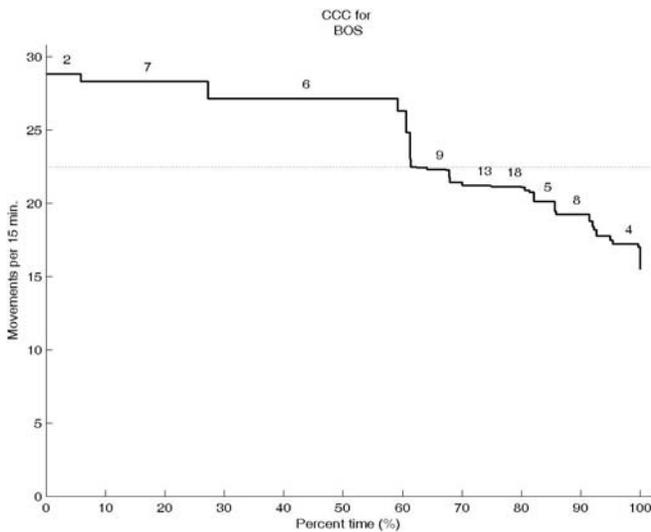


Figure 3: CCC for BOS (Step-like)

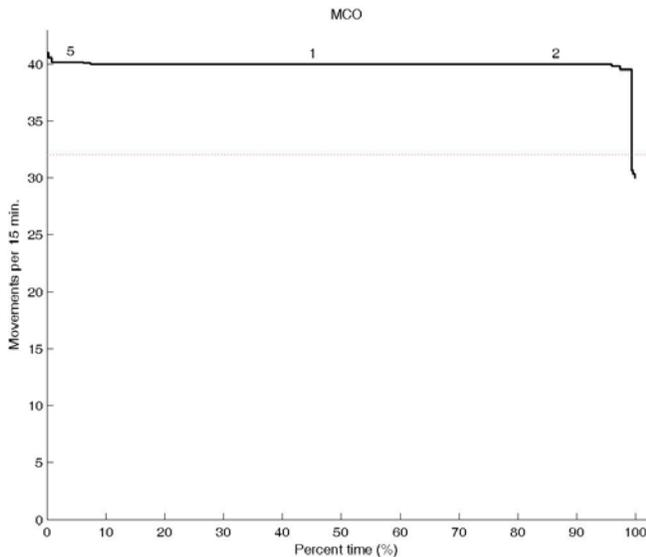


Figure 4: CCC for MCO (Flat)

Figure 5 shows the OEP-35 airports ranked according to their Signal-to-Noise ratio. MCO, which has a flat CCC, ranks highest and CLT which operates at less than 50% of its maximum capacity ranks lowest about 25% of the time ranks lowest.

III. Technical Approach and Data Sources

This section describes the database used, data collection method and method of airport capacity analysis in this paper.

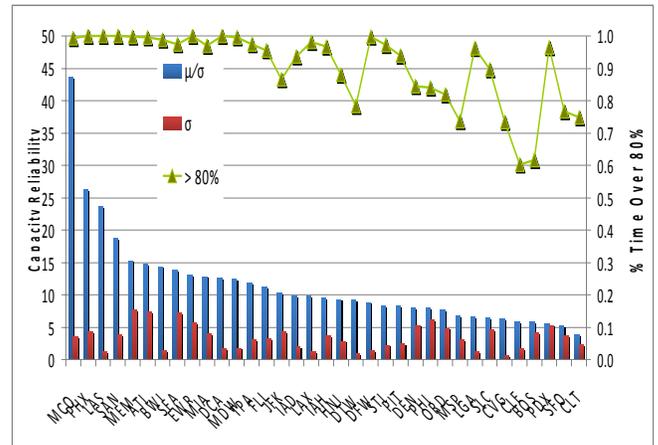


Figure 5: Reliability Index of OEP-35 airports

Source of Airport Capacity Data

The data used in this study was derived from the Federal Aviation Administration (FAA) Aviation System Performance Metrics (ASPM) airport database. The data used was for the OEP-35 airports (ref) for the period from June 2008 to August 2008. The following fields were used:

- 5 Airport ID – this field identifies the airport with the FAA 3 letter code
- 6 AAR (Airport Arrival Rate) – declared arrival capacity of airport for each 15 minute period between 6am and 10pm local time
- 7 ADR” (Airport Departure Rate) – declared departure capacity of the airport for each 15 minute period between 6am and 10pm local time
- 8 RwyConf (Runway configuration) – the configuration of the runways during the 15 minute period

To calculate the average delay “Airline On-Time Performance data” from BTS was used. The following fields were used: origin airport, destination airport, departure delay and arrival delay. For the “number of passengers” information from the “T100 data for domestic segment (All Carriers)” from BTS was used. The following fields were used: origin airport, destination airport and number of passengers. For the airfare (revenue) information, the DB1BMarket database from BTS was used. The following fields were used: origin airport, destination airport, number of passengers and market fare.

Computational Methods

The following calculations were performed.

Average Capacity for each Runway Configuration

Average number of movements for each runway configuration for an airport is computed by summing the AAR and ADR for every 15-min bin when that particular runway configuration was used. This sum is then divided by the corresponding instances of the runway configuration to obtain an average. For example, a given runway configuration was used for 3 15-min periods during the entire duration of summer. During those periods the airport's declared capacity (AAR, ADR) was (10,9),(8,10) and (12,6). The average capacity for all movements for the runway configuration would be:

$$((10+9)+(8+10)+(12+6)) / 3 = 18.33 \text{ movements per 15-min.}$$

Average Excess Costs of Delays due to Over-Utilization of Airport

When the scheduled arrivals and/or departures exceed the Declared Capacity of the airport, it results in queuing delays. The average delay per departure and arrival operation is computed by the sum of the delays for all arrivals and departures in the 15 minute period, divided by the number of arrivals and departures.

For every airport, arrival and departure revenue is calculated by summing across all the O-D pairs for that given airport.

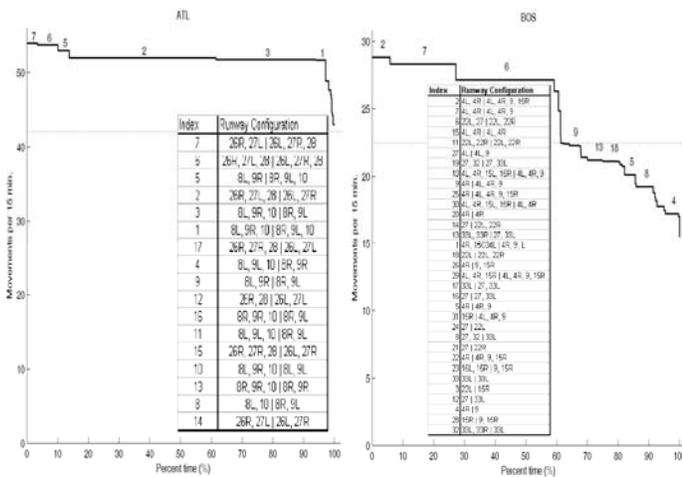


Figure 6: CCC for ATL and BOS for Summer 2007

INDICES:

Airport index $i = \{1, 2, \dots, 35\}$ – All the OEP-35 airports

arr(i) = All the airports which are origin for flights ending at airport i. Note: arr(i) is not a subset of i.

dep(i) = All the airports which are destination for flights out of airport i. Note: dep(i) is not a subset of i.

VARIABLES:

delArr_i = average arrival delay(in minutes) for flights coming into airport i

delDep_i = average departure delay(in minutes) for flights leaving airport i

numArr_i = total number of arrivals into airport i

numDep_i = total number of departures from airport i

avgDel_i = average delay(in minutes) per operation(arrival or departure) at airport i

$$= \frac{[(delArr_i * numArr_i) + (delDep_i * numDep_i)]}{(numArr_i + numDep_i)} \quad \text{---- 1}$$

Average delay per flight for a given airport is obtained by calculating a weighted sum of the average arrival delay and average departure delay with the number of arrivals and departures as weights and then dividing the sum-product with the total number of flights, i.e. number of arrivals and departures.

This simplified method of calculating average delay does not take into account the inherent non-linear (exponential) nature of delays. It assumes that delays are linear and the delays from one 15 minute bin do not cascade into subsequent 15 minute bins.

Using an estimated direct operating cost of about \$2000 per aircraft hour (Neufville, Odoni, 2004) the average Cost of Delays is computed as follows:

avgLoss_i = average dollar loss(opportunity cost) per operation at airport i

$$= avgDel_i * 2000/60$$

$$= (\$ 33.33 * avgDel_i) \text{ per operation} \quad \text{---- 2}$$

Average Lost Revenue due to Under-Utilization of Airport Capacity

When the Declared Capacity is set below the available capacity of airport, arrival and departure slots go unused, resulting lost revenue and lost profit.

$ticketArr(j,k)$ = average ticket price for j-k O-D pair, where $j \in arr(i), \forall k \in i$

$ticketDep(j,k)$ = average ticket price for j-k O-D pair, where $k \in dep(i), \forall j \in i$

$paxArr(j,k)$ = average # of passengers on j-k O-D pair, where $j \in arr(i), \forall k \in i$

$paxDep(j,k)$ = average # of passengers on j-k O-D pair, where $k \in dep(i), \forall j \in i$

$avgRev_i$ = average revenue per operation(arrival or departure) at airport i

$$= \frac{1}{2} * \left(\sum_{j \in arr(i)} paxArr(j,k) * ticketArr(j,k) + \sum_{k \in dep(i)} paxDep(j,k) * ticketDep(j,k) \right)$$

Average revenue per flight for a given airport is obtained by calculating a weighted sum of the arrival revenue and departure revenue with the number of arrivals and departures as weights. To calculate average revenue per operation, the weighted sum is multiplied by 1/2. This method of calculating average revenue per flight assumes that there is no economy of scale.

Average Lost Profit for Under-Utilization of Airport Capacity

For the purpose of this study, a profit margin of 2.5% of revenue was used (ref).

$$avgProfiti = \$ 2.5\% * avgRevi \text{ per operation, } \forall i \text{ ---- } 1$$

Relative Value of Loss due to Delays and Loss due to Unused Capacity

Airports exhibit different profiles of airfares (e.g. competing modes of transportation, inter-airline competition), different load factors (i.e. balance between seat capacity and passenger demand), and queuing delays due to over-scheduled peak periods and/or periods of reduced capacity.

To understand the relative value between the lost profit due to setting the declared capacity below the actual available capacity, and the loss due to delays when the schedule is in excess of the capacity, the Delay/Under Utilization Ratio is computed. For example, a ratio of two implies that average cost of delay per flight costs twice as much as the profit derived by a possible additional flight. As a result it is more economically efficient to set the Declared Capacity lower than max capacity to avoid excessive delays.

In contrast, a ratio of 0.5 implies that the average cost of delays by a flight is valued half as much as the value of profit derived from an additional flight. For this scenario it makes more economical sense to set the Declared Capacity closer to the maximum capacity to avoid excessive under-utilization.

Computing the Economic Optimum Seasonal Declared Capacity

The declared capacity (the horizontal line in Figure 7) is varied from 50% of the Max Capacity to 100% of Maximum Capacity in steps of 5%. For each value of Declared Capacity, the total number of flights that would be delayed is determined by the area between the Seasonal Declared Capacity(red) and Actual Declared Capacity (blue). For example, in Figure 7 this area is represented by the sum of area of the rectangles represented by:

$D_i = (d1 + d2 + d3 + d4)$ flights. (where i is the airport index) . Similarly, when Seasonal Declared Capacity is set lower than the actual Declared Capacity, the total number of extra flights that could have been serviced is represented by:

$U_i = (u1 + u2)$ flights. (where i is the airport index)

Thereafter, the loss due to the delays(LD) is determined by multiplying the number flights delayed, i.e. D_i by the $avgLoss_i$ (average loss per flight for airport i due to delays). The result of this computation is the Total Loss due to Delays for that airport, i.e.

$$LD_i = D_i * avgLoss_i$$

Similarly, the Loss of Profit due to under utilization(LP) is computed by multiplying the number of extra operations that could have been

serviced, i.e. U_i by the avgProfiti (average profit per operation for airport i), i.e.

$$LP_i = U_i * \text{avgProfiti}$$

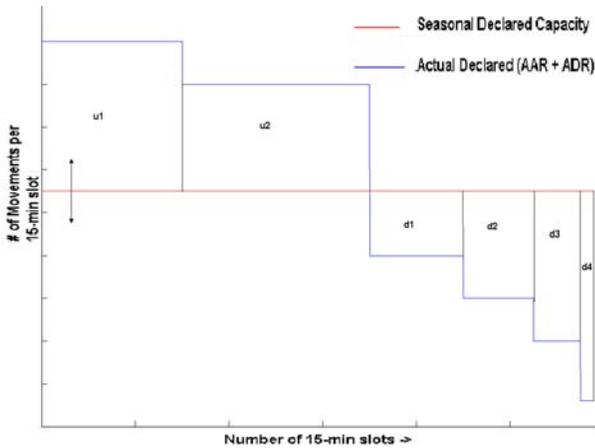


Figure 7: Calculation of Delayed and Underutilized Slots

When the loss due to delays and lost revenue due to under-utilization are plotted together, their intersection gives the optimum declared capacity for the airport. Figure 8 shows the result for Boston-Logan airport. The optimum Annual Declared Capacity for the airport is 86.2% of the Maximum Capacity or 25 operations per 15 minutes.

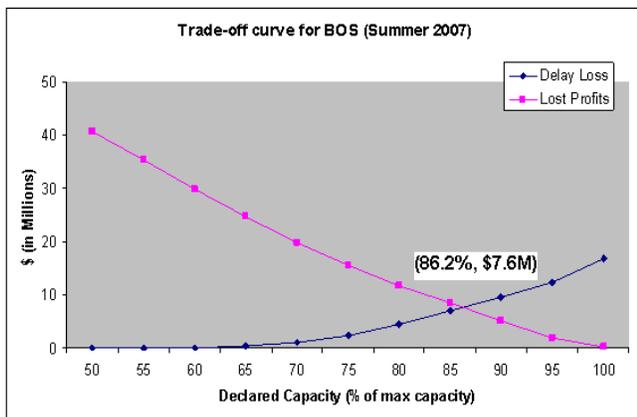


Figure 8: Trade-off curve showing economic optimal operating point for BOS.

IV. Results

Degree of Asymmetry in Actual Capacity at U.S. Airports

All airports have unique runway layout and weather/wind conditions primarily due to their geographical location. As a result, the amount of variation in their capacity is not similar. For example, airports like ATL which have more or less flat CCC operate at or near capacity mostly. And airports like BOS and CLT have higher variation in their capacity and therefore a more step-like CCC. Reduction in airport capacity is attributed to a combination of factors including environmental effects like wind and inclement weather, environmental (noise and emission) constraints. Capacity limits for airports with step-like CCC (high variation in capacity) should be set more conservatively since the likelihood of an event under which their capacity may be reduced is higher. For airports with flat CCC

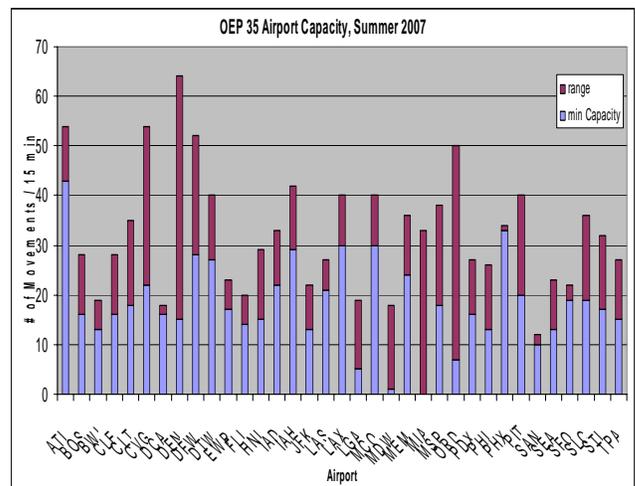


Figure 9: Range of Capacity of OEP-35 Airports

Seven airports experience declared capacity below 80% of their maximum capacity more than 20% of the time. Thirteen airports experienced declared capacity below 80% maximum capacity more than 10% of the time (Figure 10).

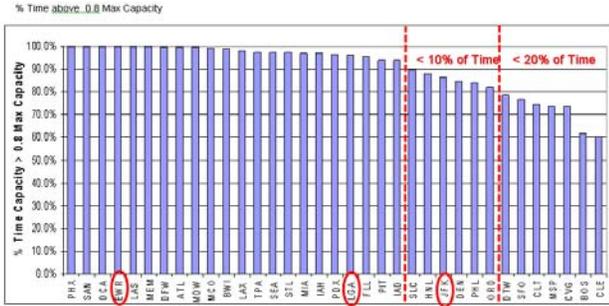


Figure 10: Percentage of time OEP-35 airports experienced capacity greater than 80% of the maximum throughput capacity for the airport.

The cost of delay per operation is not identical for different airports. Under our model assumptions (section III), the cost of delay at an airport is directly proportional to the corresponding average delay per operation. As a result New York multiplex area airports like JFK, EWR and LGA have higher cost of delays compared to airports like SAN, MEM or HNL. (Figure 11)

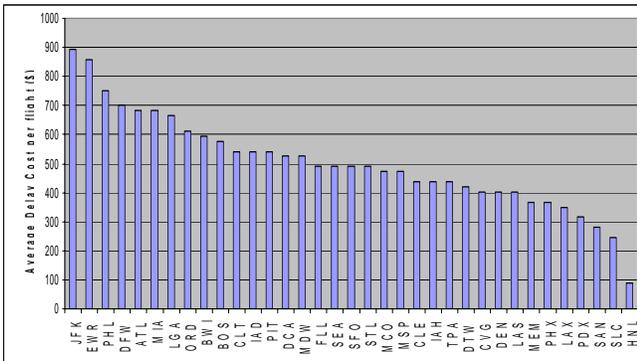


Figure 11: Estimated Average Cost of Delays at OEP-35 Airports

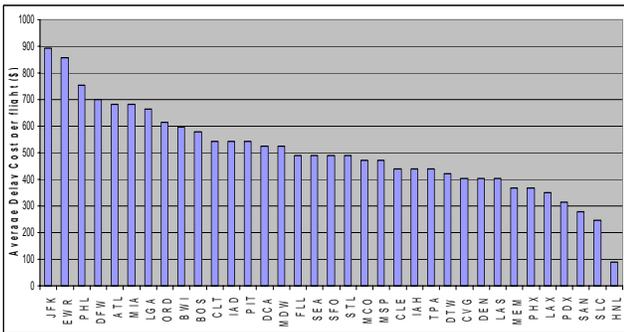


Figure 12: Estimated Loss of Profit due to Unused Capacity at OEP-35 Airports

The estimated loss of profit is the opportunity cost foregone as a result of not scheduling an extra operation at the airport. Under the model

assumptions of no ‘economies of scale’ and fixed profit margin (2.5%), this value is directly proportional to airline revenue for the given airport. Airports like SEA, HNL, MIA and SFO have high estimate of lost profit (Figure 12). These airports are high revenue generators for the airlines because they mostly cater to long distance flights. It must be noted here that the results might be different if actual aggregated profit margins for airlines for specific O-D pairs is used instead of a common fixed 2.5% profit margin. This is because few smaller distance routes are more profitable for airlines than the long distance routes due to factors like higher demand, higher ticket price, low crew cost etc.

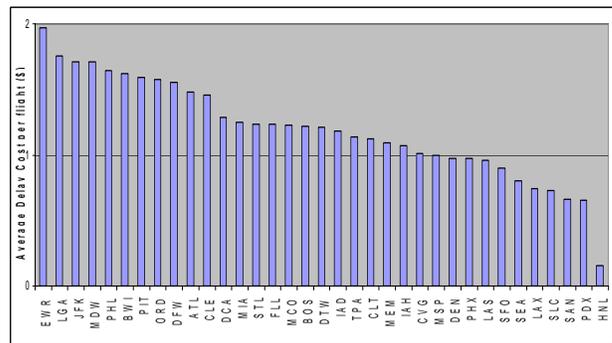


Figure 13: Ratio of Estimated Cost of Delays and Estimated Loss of Profit due to Unused Capacity at OEP-35 Airports

Statistics	Loss of Profit	Average Delay Cost	RDU
mean	427	501	1.197023
median	412	490	1.212209
Min	300	87	0.156302
max	612	892	1.969849

Table 1: Summary Statistics of RDU (Ratio of Delay and Underutilization cost)

The ratio of unit delay cost to unit opportunity (under-utilization) loss yields what we call the RDU(Ratio of Delay and Underutilization) . RDU varies from 0.15 to 1.97 for the OEP-35 airports. RDU value greater than 1(EWR, JFK, LGA) means that delay costs more than underutilization and vice versa for airports with RDU less than 1(HNL, SFO, SEA). For example, cost of one extra operation at LGA in terms of delay is less when compared to an extra operation at EWR [19 (min) *33.3 (\$/min) < 24.5 * 33.3(\$/min) => 632<808]. Opportunity loss

for one extra operation at LGA and EWR is \$380 and \$435 respectively. Clearly, we see that there is higher incentive to have an extra operation at EWR when compared to LGA in terms of dollar profit, but also on the other end the expected loss of this extra operation at EWR is higher than LGA.

Economic Optimum for Seasonal Declared Capacity at OEP-35 Airports

The theoretical economic optimum (within the underlying assumptions of our analysis) of an airport is a combination of factors mentioned in bullets 1 and 2 above, i.e. the combined effect of stochasticity in capacity(natural uncontrollable factor) and the unit profit/loss per extra operation(decided by market supply/demand characteristic). And this economic optimum is unique for each airport because the factors that influence it, i.e. capacity and profit/loss are unique for each airport.

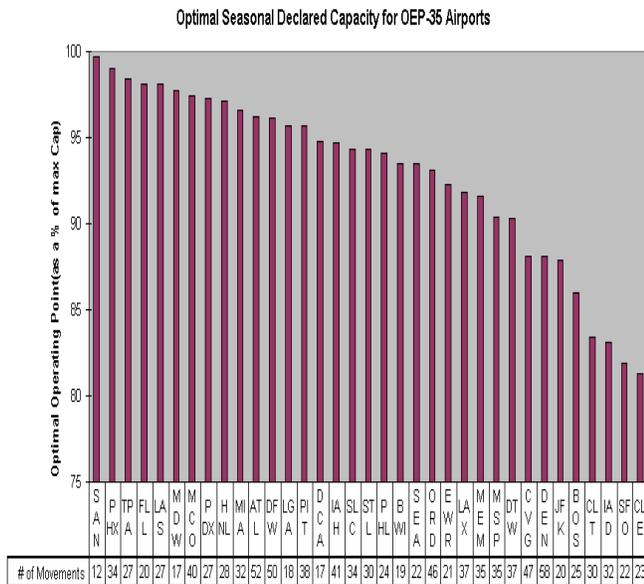


Figure 14: Optimal Seasonal Declared Capacity for OEP-35 airports

Min	81.3
Max	99.7
Mean	92.9
Median	94.3

Table 2: Summary Statistics of Optimal Operating Points of OEP-35 airports

It is expected that airports which have flatter CCC (eg: ATL), i.e. lesser variation in their

operating capacity have higher ‘optimal operating point’ than the airports which have more step-like shape (eg: BOS), and higher variation/uncertainty in their operating capacity. ATL has an optimal operating point of 96% of max capacity whereas BOS has an optimal operating capacity of 86%.

CONCLUSIONS

This paper examines: (i) the variability of throughput capacity at the OEP-35 airports during the convective weather season in 2008, (ii) established the average costs of delays due to reduced capacity and the average profits per flight at each airport, (iii) given the cost and lost profits, an optimum capacity was determined at each airport to minimize both the under-utilization of airport capacity and flight delays due to variability in the available throughput capacity.

The results of this analysis are as follows:

- 13 of the OEP-35 airports exhibit variations in runway configurations that result in a reduction of more than 20% in capacity more than 10% of the time. Included in this category are Cleveland, Boston, Cincinnati, San Francisco, Chicago O’Hare and JFK.
- 24 of the OEP-35 airports exhibit an average cost of delays per flight in excess of the average profit generated by a flight. Included in this category are JFK, Newark and La Guardia
- Optimum airport capacity, computed by trading-off flights delays and underutilization ranged from 81% to 100% of the maximum airport capacity. The average optimum airport capacity was 93% of the maximum airport capacity. Twelve airports, including Boston, JFK, Newark, and San Francisco, exhibited an optimum airport capacity below the average 93%.

These results have several implications. First, the variability in the available airport capacity throughput is critical in the *a priori* establishment of airport capacity limits for demand management. Second, the range of the ratio of costs of delays due to reductions with the potential profits per flight indicates the non-homogeneous nature of the service providers in the National Airspace System and the need to take into account the airline revenue

and cost structures in determining resource availability. Third, “economically optimal” capacity limits established to trade-off delays and underutilization, provide a rational approach to assigning capacity limits that take into account airline preferences.

Future Work

Future work is planned to improve the accuracy of these results by (1) increasing the fidelity by taking into account opportunity costs in the formulation of the profit, (2) increasing the fidelity in the cost of delays.

Future work is also planned to use an alternate approach to this analysis is to use marginal costs. This will include the effects of economies-of-scale.

Instead of using average delay for an airport as an estimate of delay for an extra operation at the

airport, a simulation which takes into account the non-linear nature of delay and predicts delays would improve the results.

References

1. deNeufville & Odoni, Airport Systems : Planning, Design and Management, 2003.
2. Churchill Andrew et al, “Determining the Number of Airport Arrival Slots”.
3. Bratu, S. and Barnhart, C. (2005). An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data, *Air Traffic Control Quarterly*, Volume 13, Number 1, 2005
4. Ahuja, R.K., Magnanti, T.L., Orlin, J.B., “Network Flows, Theory, Algorithms, and, Applications”, Prentice Hall, 1993.