Estimation of Arrival-Departure Capacity Tradeoffs in Multi-Airport Systems

Varun Ramanujam and Hamsa Balakrishnan

Abstract—The accurate estimation of airport capacity is critical for the efficient planning of landing and takeoff operations, and the mitigation of congestion-induced delays. The analysis of tradeoffs between arrival and departure capacity at an airport, represented by the airport capacity envelope, has been the main focus of prior research. The increasing demand for air traffic operations has resulted in the growth of multi-airport systems, in which several major airports that are in close proximity of each other serve the same geographical region. The arrival and departure flows into these airports interact with each other, and it is necessary to consider inter-airport arrival-departure capacity tradeoffs while scheduling operations.

This paper proposes a statistical technique based on quantile regression, for systematically analyzing arrival-departure capacity tradeoffs in multi-airport systems using observations of flight operations. The proposed technique enables the identification of key factors (such as, runway configuration geometry, weather conditions, etc.) that influence both the capacity envelopes of individual airports, and the capacity envelope of the multi-airport system as a whole. The approach is demonstrated through an analysis of the capacity envelopes of the New York area multi-airport system (comprising Newark (EWR), John F. Kennedy (JFK) and LaGuardia (LGA) airports).

I. INTRODUCTION

The safe and efficient planning of airport and terminal-area operations is essential for meeting the predicted increase in demand in the Next Generation Air Transportation System (NextGen) without incurring unacceptably large delays. The effects of congestion are beginning to be seen even in today’s system: in the United States, between 2006 and 2007, there was a 30% rise in delays due to terminal-area volume, while there was only a 1% increase in traffic [4,6]. A critical step to meeting the expected increase in demand is through the improved utilization of airport capacity, especially in congested terminal-areas. With the growth of secondary and even tertiary airports in the most congested regions of the country (for example, the New York area, the San Francisco Bay area, the Los Angeles area, Boston and Chicago [3]), the problem of coordinating operations in multi-airport systems to use terminal-area capacity more efficiently has also become increasingly important.

The operating capacity of an airport is given by its arrival capacity (the number of aircraft landings per hour) and its departure capacity (the number of departures per hour). Due to the shared nature of ground resources such as runways and taxiways, there is a tradeoff between the arrival and departure capacity at an airport [2,14]. An airport capacity envelope is the boundary (generally approximated as a convex polygon on the plane with the arrival and departure rates as axes) that defines the envelope of the maximum capacities that can be achieved under specified operating conditions. Fig. 1 illustrates a representative capacity envelope for an airport, for a given runway configuration. The capacity envelope describes the capacity available for a runway configuration under a specific set of conditions, and captures the tradeoff between the maximum arrival and departure rates [5]. The extent of the tradeoff depends on factors such as the relative alignment of the active runways, meteorological factors like wind and visibility, as well as the aircraft fleet mix.

The emergence of several core, secondary and regional airports (known collectively as a metaplex [15]) in an already congested airspace results in the complex interaction of traffic flows. The interactions between arrival and departure operations at proximate airports could potentially make it infeasible to simultaneously operate all the airports at their individual optimal runway configurations. This implies that the resource allocation and schedule optimization at an airport would have to not only take into consideration the arrival-departure tradeoffs for that airport, but also the tradeoffs with arrival and departure capacities at the other airports within the multi-airport system. While there is anecdotal and descriptive evidence of such interactions between air traffic flows into and out of neighboring airports, there have been no attempts to quantify these interactions and their impact on capacity using operational data.

A detailed understanding of airport capacity tradeoffs and their dependence on external factors such as operating conditions and airport layout, both in the single- and multi-
airport settings, is necessary for the efficient utilization of airport capacity. While these tradeoffs have not been studied so far in the multi-airport setting, the estimation of single-airport capacity envelopes has traditionally been restricted to theoretical approaches [2,11,12]. Empirical methods that use operational data have only been employed in recent times, but have generally been ad hoc approaches [10]. This paper proposes a systematic statistical inference technique, based on quantile regression, to estimate piecewise-linear, convex capacity envelopes using observed throughput data, for both individual airports as well as the corresponding inter-airport dependencies in a multi-airport system. The problem of estimating the capacity envelopes as quantiles (or percentiles) of the reported data is reduced to the solution of a linear program. The proposed technique is illustrated using reported data on arrivals and departures at the three major New York area airports, namely John F. Kennedy (JFK), Newark (EWR) and LaGuardia (LGA), for the years 2006 and 2007. The analysis identifies the key factors that influence the intra- and inter-airport capacity tradeoffs, and determines the associated capacity envelopes.

II. RELATED WORK

Airport capacity is affected by various external factors such as, pilot and air traffic controller procedures, approach and departure speeds, runway and taxiway occupancy times, weather, etc. Theoretical approaches to capacity estimation have traditionally modeled these factors through simplified stochastic models of aircraft behavior, and estimated the capacity using the expected separation time between successive aircraft operations [2,11,12,19,22,23]. These models construct the capacity envelope through linear interpolation between capacity estimates estimated at specific arrival/departure mix ratios. Newell [20] and Odoni et al. [21] provide comprehensive overviews of analytical and simulation methods that adopt the above approach.

Empirical estimation approaches have the potential to reflect the practical operating capacity envelopes at airports, as opposed to idealized models of capacity tradeoffs. Recognizing this, Gilbo [10] proposed a quasi-statistical procedure for estimating the capacity envelope of a single airport, for a given runway configuration. 15-min arrival and departure counts were used to estimate the capacity envelope as the convex hull (polytope) of the planar scatter of the counts. Frequency-based filtering was employed on the outer perimeter of the data to eliminate outliers, and alternative outlier rejection criteria such as proximity to neighboring observations and rank order statistics were also mentioned. In contrast, this paper proposes a systematic statistical approach to estimating intra- and inter-airport capacity envelopes from observed data.

On the statistical estimation front, quantile regression techniques attempt to determine statistics such as the median or a general percentile of the conditional distribution of a response variable which is a function of observed variables [17]. For example, Bernini et al. modeled the production frontier in classical economics as a higher-order quantile (90–100%ile), and studied the interaction between underlying determinants at intermediate quantiles [1]. We extend these techniques to the case of airport capacity envelope estimation; in particular, our approach is suitable for determining piecewise-linear, concave quantile functions.

III. PROBLEM STATEMENT

Given the counts of arrival and departure throughput at an airport per time interval (say, 15 minutes), we would like to determine the capacity envelope by suitably excluding outliers; we would also like to identify key factors that affect the capacity tradeoffs; finally, we would like to develop a technique that is capable of analyzing both intra- and inter-airport capacity envelopes.

Our approach to solving this problem draws from the field of quantile regression. A data point is said to be at the \(\tau^{th}\) quantile \((\text{or} \ (100 - \tau^{th}) \text{percentile})\) if it is larger than a proportion \(\tau/100\) of the data points, and less than a proportion \((1-\tau/100)\) [17]. Similar to least-square regression techniques that estimate the mean of a response variable given values of the predictor variables, quantile regression techniques estimate other statistics, such as the median or a quantile [17]. Since the airport capacity envelopes represent the upper limits of operating capacity, quantile regression techniques, with \(\tau\) sufficiently large, are suitable mechanisms for estimating them. In other words, if \(\tau = 99.5\), we would like to draw the capacity envelope such that 99.5% of all reported operating points fall within the feasible region.

The quantile (value of \(\tau\)) chosen to represent the airport capacity is conceptually similar to the frequency-based filter adopted by Gilbo [10] to discard spurious data (for example, reporting errors). While quantile regression has traditionally been used to determine linear quantile functions, in the case of airport capacity envelopes we would like to determine a piecewise-linear, concave, continuous function that represents the quantile. We show that the estimation of parameters of such a quantile function can be conducted by solving a linear program. In addition to airport-specific capacity envelopes, we use the proposed approach to study the presence and magnitude of tradeoffs between operations at neighboring airports.

IV. MODEL DESCRIPTION AND SOLUTION FRAMEWORK

In this section, we describe the modeling of capacity envelopes, and the formulation of a linear optimization problem to estimate them.

A. DATA SOURCES

The FAA’s Aviation System Performance Metrics (ASPM) database provides records of flight activity at 77 of the major airports in the United States [7]. For each airport, for every 15-min interval, the database includes reports of the number of arrivals, the number of departures, prevailing weather conditions (Visual or Instrument Flight Rules, wind speed and direction), and the runway configuration that was used. This paper focuses on the three major New York area airports. Overnight periods of operations (midnight-6AM) and...
15 minutes before and after runway configuration changes are filtered out from the analysis, due to the increased tendency for reporting errors during these periods.

B. Capacity envelopes

The capacity envelope representation is decomposed into two parts: an arrival rate threshold and a departure capacity function. The arrival rate is treated as the independent variable, and the departure capacity for any time interval is estimated as a function of the arrival rate using quantile regression. This decision is based on the observation that arrivals are generally given priority at airports over departures. It is therefore reasonable to assume that while the arrival rate is not impeded by departures, it corresponds to the unhindered arrival capacity, which represents the capacity envelope, \( y \) at a given airport, the process of estimating the piecewise coefficients \((\alpha_k, \beta_k)\) for a chosen quantile \( \tau \) involves solving the following linear optimization problem:

\[
\text{Minimize } \sum_{n=1}^{N} Z_n \quad (3)
\]

subject to:

\[
Z_n \geq y_n - \left[ \sum_{i} \alpha_i^k \theta_i^k + \left( \sum_{i} \beta_i^k \theta_i^k \right) x_n \right] \quad \text{if } k - 1 \leq x_n \leq k, \forall n \quad (4)
\]

\[
Z_n \geq \omega_k \left[ \sum_{i} \alpha_i^k \theta_i^k + \left( \sum_{i} \beta_i^k \theta_i^k \right) x_n - y_n \right] \quad \text{if } k - 1 \leq x_n \leq k, \forall n \quad (5)
\]

\[
\sum_{i} \beta_{i+1}^k \theta_{i+1}^{i,m} \leq \sum_{i} \beta_i^k \theta_{i}^{i,m}, \forall k \in \{1, \ldots, x_{\text{max}} - 1\}, \forall m \in \{\text{min}, \text{max}\}, \forall i \quad (6)
\]

\[
\sum_{i} \beta_i^k \theta_{i}^{i,m, \text{m}} \leq 0, \forall m \in \{\text{min}, \text{max}\}, \forall i \quad (7)
\]

\[
\sum_{i} \alpha_i^k \theta_{i}^{i,m} + \left( \sum_{i} \beta_i^k \theta_{i}^{i,m} \right) k = \sum_{i} \alpha_{i+1}^k \theta_{i+1}^{i,m} + \left( \sum_{i} \beta_{i+1}^k \theta_{i+1}^{i,m} \right) (k) \quad \forall k \in \{1, 2, \ldots, x_{\text{max}} - 1\}, m \in \{\text{min}, \text{max}\} \quad (8)
\]

In the above formulation, constraints (4)-(5) define the absolute deviation of the estimated quantile function from the observed value for \( y \). Constraints (6)-(7) help ensure concavity and non-positivity of the piecewise slope estimates for all values of \( \theta^i \in \Theta \) in the range \([\theta_{\text{min}}, \theta_{\text{max}}]\). This feature is motivated by the observation that the magnitude of tradeoffs monotonically increases with arrival rate. Constraints (8) ensure continuity of adjacent segments. \( \omega_k = \left(100 - \tau\right)/\tau \) is the asymmetric weight applied upon the negative deviations, where \( \tau \) denotes the order of the quantile estimated [18]. The size of the formulation depends on the number of observations in the data set \( N \), the maximum arrival count in a 15-minute interval \( x_{\text{max}} \) and the number of factors considered (that is, the size of \( \Theta \)).

The unhindered arrival (or arrivals only) capacity can be estimated using a simplified version of the above formulation. The data set is restricted to observations where the arrival rate is not impeded by departures. The unhindered arrival capacity is estimated by solving the following linear program:

\[
\text{Minimize } \sum_{n=1}^{N} Z_n \quad (9)
\]

subject to:

\[
y_{n,\text{cap}} = \sum_{i} \alpha_{i,\text{cap}}^k \theta_i^k + \left( \sum_{i} \beta_{i,\text{cap}}^k \theta_i^k \right) x_{\text{max}} \quad (10)
\]

\[
Z_n \geq y_n - \sum_{i} \theta_i^k \quad \text{if } y_n \leq y_{n,\text{cap}} \quad (11)
\]

\[
Z_n \geq \omega_k \left[ \sum_{i} \gamma_i^k \theta_i^k \right] - x_n \quad \text{if } y_n \leq y_{n,\text{cap}} \quad (12)
\]
where $\gamma_i$ and $\phi_i$ are the counterparts of $\alpha_i^c$ and $\theta_i^c$ in Equation (2). The lowest value of the departure capacity ($V_{\text{cap}}$) is realized at the highest observed arrival count ($x_{\text{max}}$) owing to the concavity assumption.

The choice of quantile $\tau$ for the unhindered arrival capacity estimate is independent of its counterpart for the departure capacity function. The principle governing this choice in both cases is effective outlier elimination [10]. The choice is determined by iterating over a progression of quantiles descending from 100%ile at chosen step sizes (set at 0.25%ile in this study), until stable functional parameter estimates are obtained. In other words, the choice of $\tau$ was such that the capacity envelope did not change significantly for a small change in $\tau$. The statistical significance of each incremental vector of features ($\theta$ and $\phi$) is ascertained through quantile likelihood ratio tests [18].

D. Comparison of proposed method with frequency-based filtering

As pointed out earlier, the chosen quantile ($\tau$) for representing the capacity envelope in the proposed approach is notionally similar to frequency-based filtering. However, in contrast to frequency-based filtering, a regression-based approach has the ability to quantify underlying factors that influence the capacity curve, through hypothesis testing over a range of specifications. For example, in addition to estimating capacity envelopes for each runway configuration, the proposed approach can identify characteristics of the configuration (such as, the angle between the active arrival runways) that impact the capacity envelope to a statistically significant extent. Also, the proposed LP-based estimation framework will consider factors such as the proximity of a point to the quantile function, in addition to its observation frequency, when discarding points, thereby simultaneously accounting for two of the potential outlier rejection criteria suggested by Gilbo [10]. This property is enabled by the concavity constraints (6)-(7) of the LP formulation for estimating the capacity quantile function, and is illustrated in Figure 2.

The quantile regression methodology could retain point B due to its conformance with the concave shape of the capacity envelope, as determined by the neighboring high-frequency observations (C and D).

V. CASE STUDY: THE NEW YORK MULTI-AIRPORT SYSTEM

The proposed formulations were applied to obtain capacity envelope estimates for the three major airports in the New York region: JFK, LGA and EWR. In addition, the interaction between operations at different airports within the system was also investigated using pairwise inter-airport capacity envelopes (for example, JFK arrivals vs. LGA departures, JFK departures vs. EWR departures, etc.). This decomposition of inter-airport relationships into pairs is motivated by the observation that at any given time, there is a dominant inter-airport capacity interaction. Under this framework, the relationships among operational capacities at the three NY airports can be represented through 15 capacity envelopes (3 intra-airport, and 3 × 4 inter-airport pairs). The capacity for an operation at any time would then be determined by the most restrictive of its pairwise envelopes for that time interval.

The choice of the dependent and independent variables in the pairwise capacity envelopes is determined by the relative congestion experienced at the airports, combined with the earlier observation that arrivals have greater priority than departures. In particular, from analyzing the data for the years 2005 and 2006, we arrive on the following precedence order: JFK arrivals $\geq$ LGA arrivals $\geq$ EWR arrivals $\geq$ JFK departures $\geq$ LGA departures $\geq$ EWR departures.

The following sections describe the estimation results for the pairwise capacity profiles at the New York area airports, and discuss the implications.

VI. RESULTS AND DISCUSSION

As described in Section IV, quantile function specifications were developed and estimated for the 15 pairwise capacity relationships (3 intra- and 12 inter-airport pairs) involving the three New York airports (JFK, EWR and LGA). The unhindered arrival capacities of the three airports were also estimated. The three airports are equipped with 4, 3 and 2 runways respectively, and their relative alignments are shown in Figure 3. We note that the principal alignments at the three airports are aligned with each other.

Data for capacity estimation was extracted from the ASPM archives, which were described in Section IV-A. Apart from arrival and departure counts for every 15-min interval, the database also provides the corresponding information on the prevailing weather conditions (wind and visibility), the active runway configuration and the operating conditions (VFR/IFR). The estimation data set covered the years 2005 and 2006 for JFK and EWR, but was restricted to the year 2006 for LGA due to inconsistencies in the throughput reports during 2005. The linear programs for estimating the capacity envelopes were coded in AMPL [8], and solved

A. Intra-airport capacity tradeoffs

The influencing factors on the capacity tradeoffs that were considered included visibility (VFR or IFR), alignment of the arrival and departure runways (parallel or crossing), and the number of additional runways for arrival or departure operations (beyond the primary runway). These factors were statistically tested for their influence on the intra-airport capacity envelopes.

The orders of quantile (\(\tau\)) that yield robust estimates of the capacity envelope were found to be 99%ile, 99%ile and 99.5%ile for JFK, EWR and LGA respectively. The \(\tau\) values for the corresponding unhindered arrival capacities were found to be relatively higher, at 99.75%ile, 99.5%ile and 99.75%ile respectively. Figures 4-6 illustrate the effect of each of the factors found to have a statistically significant influence on the capacity envelopes, under VFR and IFR conditions, for the three airports.

Figures 4-6 show that visibility has a significant influence on the unhindered arrival capacities, but does not appear to have a tangible effect on departure capacities. The relative alignment of the primary runways used for arrivals and departures plays a critical role in determining the capacity envelopes. Figures 4-5 show that the area under the capacity envelope progressively increases as we go from mixed arrival/departure operations on a single runway, through separate arrival and departure runways that intersect or converge, to additional parallel runways, demonstrating the benefit of independent operations on runways. The use of an additional departure runway at JFK and EWR provides the most benefit in terms of departure capacity at low values of arrival rate. Figure 4 also shows that the use of an additional arrival runway at JFK flattens the slope of the tradeoff curve, indicating the effective redistribution of operations across the two runways, while also increasing the unhindered arrival capacity. It is also observed that the use of the parallel runway configurations (22R, 22L) or (4R, 4L) at JFK results in a lower capacity as compared to their perpendicular counterparts (31R, 31L) or (13R, 13L). This is possibly explained by the smaller separation between the runway centerlines of the former pairs (resulting in a greater coupling of operations).

B. Inter-airport capacity tradeoffs

Since inter-airport interactions are expected to involve the airspace rather than the airport surface, the overlap between approach or departure paths is considered instead of the
runway alignment attribute used for intra-airport capacity envelopes in Section VI-A. The approach and departure paths were approximated by two-dimensional conics extrapolated from the runway in the direction of operation, and binary terms were used to signify the intersection of these 2D conics. Representative inter-airport capacity envelopes for pairs of airports and arrival-departure operations, under different flight conditions are shown in Figures 7-9.

From the inter-airport capacity envelopes, it is observed that capacity tradeoffs are prominent at higher throughput values than those seen in the intra-airport envelopes. This observation suggests that airport (ground) capacity is a more dominant operational bottleneck than the capacity of the surrounding airspace. Another consequence of this observation is the impact of IFR conditions and single runway configurations on the inter-airport capacity curves, mainly due to the reduced throughput. Figures 7-9 illustrate this phenomenon for selected airport operational pairs that exhibited tradeoffs close to the limits of their respective operational capacities. The approach path overlap attribute was not found to be statistically significant for any of these pairwise envelopes, possibly due to the limited operational range over which the tradeoff effects were found to be prominent.

VII. CONCLUSION

This paper proposed a statistical framework for identifying and quantifying arrival-departure tradeoffs in a multi-airport system, and for estimating airport capacity envelopes using observations of flight operations. Quantile regression, the technique adopted within this framework, is shown to be well-suited for systematic outlier elimination. The proposed approach generates credible capacity estimates, and also assesses the influence of underlying factors that impact the capacity tradeoffs.

The estimation of capacity envelopes was conducted in two parts: the first part estimated a piecewise-linear function
that related the departure capacity to the arrival rates, and the second part required the estimation of the unhindered arrival capacity. It was shown that both estimation problems can be solved using linear programming. The proposed approach was demonstrated through the estimation of capacity envelopes for the three New York area airports: JFK, EWR and LGA. This paper also extended the estimation methodology to analyze interactions between operations at different (and nearby) airports, by considering the New York multi-airport system.

REFERENCES