Kim, Amy, Hansen, Mark.

Deconstructing delay: A non-parametric approach to analyzing delay changes in single server queuing systems.

AUTHOR POST PRINT VERSION

Kim, A., & Hansen, M. (2013). Deconstructing delay: A non-parametric approach to analyzing delay changes in single server queuing systems. Transportation Research Part B: Methodological, 58, 119-133. https://doi.org/10.1016/j.trb.2013.10.001

*2. Manuscript
Click here to view linked References

Deconstructing Delay: A Non-Parametric Approach to Analyzing Delay Changes in Single Server Queuing Systems

Amy Kim*
Department of Civil and Environmental Engineering
University of Alberta

Mark Hansen
Department of Civil and Environmental Engineering
University of California, Berkeley

* Corresponding author. Address: 3-007 CNRL/Markin Natural Resources Engineering Facility, University of Alberta, Edmonton, AB, T6G 2W2, Canada. Tel.: +17804929203, Fax.: +17804920249. Part of this work was performed while the corresponding author was at the University of California, Berkeley.

E-mail addresses: amy.kim@ualberta.ca (A. Kim), mhansen@ce.berkeley.edu (M. Hansen).

Abstract

This paper introduces an empirically driven, non-parametric method to isolate and estimate the effects that changes in demand and changes in throughput have on delay – in particular, arrival and departure flight delay at airport runways. Classical queuing model concepts were used to develop a method by which an intermediate, or counterfactual, queuing scenario could be constructed, to isolate the delay effects due to shifts in demand and throughput. This method includes the development of a stochastic throughput function that is based entirely on data and has three key features. Firstly, the function relies on non-parametric, empirically-based probability distributions of throughput counts. Secondly, facility capacity needs not be explicitly defined, as it is implicitly included in the probability distributions of throughput. Thirdly, the throughput performance function preserves the effect of factors that cause capacity (and, therefore, throughput) to fluctuate over a given period. Temporal sequences of high, moderate, and low capacity are maintained between the observed and counterfactual scenarios. The method was applied to a case study of the three major New York area airports of LaGuardia (LGA), Newark Liberty (EWR), and John F. Kennedy (JFK), using operational data extracted from the Federal Aviation Administration's (FAA's) Aviation System Performance Metrics (ASPM) database. The focus was on the peak summer travel seasons of 2006 and 2007, as these airports experienced record levels of delay in 2007. The results indicate that decreases in both demand and throughput were experienced at LGA and EWR, although the decreases in throughput had more significant effects on operational delays as they increased overall at these airports. At JFK, the increase in departure throughput was not sufficient to offset the increase in departure demands. For arrivals, demand increased and throughput decreased. These trends caused a significant growth in delay at JFK between 2006 and 2007.

Keywords

Counterfactual delay scenario; empirical probabilistic simulation; airport runway operations, Aviation Systems Performance Metrics (ASPM) database.

1. Introduction

This paper introduces a method for decomposing the change in delay at an airport – or potentially some other server in a queuing system – between two time periods of roughly equal duration. This duration could be a day, week, month, or longer. In this method, the delay change is divided into a component that results from a change in demand and another component that reflects the ability of the airport to handle demand. The latter would traditionally be termed capacity, but here we will, for reasons discussed below, use the term "throughput performance". When both demand and throughput performance shift simultaneously, it can be difficult to quantitatively estimate how much the resulting change in delay should be attributed to either. In turn, it can be difficult to isolate the role of the air navigation service provider's performance from the role of flight demand in effecting delay changes. As a result, extensive efforts have been made to better identify and understand the factors contributing to delay at airports.

To address this issue, we present a classical queuing analysis that relies on an empirically driven, time-sequential simulation procedure. The aim of this procedure is to predict the resulting flight delay at an airport over a period of time, if demand had remained constant but the throughput performance came from a different time period, or vice versa. The procedure probabilistically simulates throughput, in order to build "counterfactual" flight queuing scenarios. These counterfactual scenarios consist of demands from one time period being served at rates based on throughput performance from another time period. The resulting counterfactual scenarios can be used to identify the change in delay resulting from a shift in either throughput alone or demand alone. The main engine of this procedure is a stochastic throughput function, which was developed to exhibit three key features. Firstly, this throughput function relies on non-parametric, conditional distributions of throughput counts, which are constructed from readily available historical data. Secondly, facility capacity needs not be explicitly defined, as it is implicitly included in the probability distributions of throughput. This is advantageous in that operational capacity is subject to a wide variety of influencing factors, and can often be difficult to accurately estimate or even precisely define. Thirdly, the throughput performance function preserves the effects of factors that cause capacity (and therefore, throughput) to fluctuate over a given period. Temporal sequences of high, moderate, and low capacity are maintained between the observed and counterfactual scenarios. Details of this simulation procedure are presented in Section 3.

An application of this simulation procedure is presented in Section 4. The procedure is applied to a case study of flight delay at the three major New York area airports: LaGuardia (LGA), Newark Liberty International (EWR), and John F. Kennedy International (JFK). Specifically, the arrival and departure operations at these airports were studied to determine the influences of demand and throughput performance in the significant delay increases observed between the 2006 and 2007 summer travel months.

The main objective of the methodology presented in this paper is to decompose delay changes. The ability to isolate the individual contributions of demand and throughput performance to changes in delay is a critical step in systematically diagnosing degraded performance or assigning credit for improved performance. This is particularly important in a setting where both airport demand management and capacity enhancement are considered as candidate approaches in mitigating delay.

2. Background

Delay is measured, estimated, and analyzed for many transportation systems, as it is a fundamental economic and operational indicator of performance. There is a significant body of literature and knowledge on methods used to quantify delay analytically, empirically, and using simulation techniques. In a traditional depiction of queuing at a single stationary point, delay can be characterized using a simple cumulative plot of vehicles arriving to, and passing, that point (Daganzo, 1997). This plot consists of a cumulative demand curve and cumulative throughput curve taken over a time period of interest, T. An example of this simplified queuing scenario is shown in Figure 1. The cumulative demand and throughput curves are actually step functions because "customers" (passengers, vehicles, flights, repairs requests, etc.) are discrete entities. However, the cumulative demand and throughput curves can be approximated as continuous functions (smoothed curves) over sufficiently long periods of time, which simplifies calculations. Assuming first-in first-out (FIFO) conditions, the delay experienced by some customer n_0 is the difference between n_0's desired service time ($t_{desired}$) and the actual time n_0 is served (t_{actual}). In Figure 1, this is the horizontal distance between the two curves at n_0. The number of customers queued for service at, for instance, $t_{desired}$ is the vertical distance between the curves at $t_{desired}$ (depicted by the vertical arrow). Where the demand and throughput curves meet, customers are being served without any delay and, as a result, there are no standing queues for service. In the simplified scenario of Figure 1, this only occurs twice: when the cumulative demand is zero, and N after the queue has cleared. When the two curves are apart, customers wait for service. The throughput curve cannot cross (to exceed) the demand curve because customers cannot be served until they request service.

Place Figure 1 here.

Note that in Figure 1, the cumulative demand and throughput functions are, for simple illustrative purposes, shown to be piecewise linear functions of t. However, in reality, this is rarely the case due to the fact that they are dependent on highly stochastic and variable phenomena (which is discussed in more detail in the following paragraphs). Assuming an empty queue at $t = 0$ and $t = T$, the area between the demand and throughput curves is the total delay experienced by all customers over the observation time $0 \leq t \leq T$, or W. The average delay per customer ($\bar{w}$) can be determined by dividing the total delay W by the total number of customers N that requested service over the observation time $[0, T]$, $\bar{w} = W/N$. In a queuing analysis, it is entirely the interaction of demand and throughput that impacts delay. In the queuing diagrams shown, the time period $[0, T]$ might represent an hour, a day, a period of months, etc.

Queuing analysis is an appropriate and effective tool to assess demand and throughput operations – and the resulting delay – for many types of transportation facilities and transportation systems (Daganzo, 1997). A key performance indicator, delay has been assessed over all the facilities and subsystems of the National Airspace System (NAS) (de Neufville and Odoni, 2002). In this context, the customers shown in Figure 1 could represent flights queuing for takeoff at a runway (Newell, 1979), flights queued for final approach to an airport (Nikoleris and Hansen, 2012; Hansen et al., 2002; Lovell et al., 2013), inbound flights traversing terminal airspace (Mukherjee and Hansen, 2009), flights subject to miles-in-trail separation en route, flights on an origin-destination airport pair, and others. In addition to queuing analysis, statistical models, discrete event simulation, and combinations of these and other techniques (Pyrgiotis et al., 2013) have also been employed to measure as well as estimate flight delay and its causes. Airport delays have also been studied at aggregate levels; Santos and Robin (2010) investigated European flight delays per origin-destination pair by airline and season, and used regression to attribute these delays to market concentration, slot coordination, hub airports and hub airlines. Air navigation service providers, airport authorities, other jurisdictional bodies, and airlines use delay metrics to make investment decisions. Also, air travel delay statistics are disseminated to the public, through agencies such as the U.S. Department of Transportation's Bureau of Transportation Statistics (www.rita.dot.gov/bts).

It is well documented that the dominant proportion of flight delays throughout the NAS originates at airports, and airport-related flight delays contribute heavily to total flight delays in the NAS (Xu et al., 2008). For instance, in the early 2000s, Hansen and Zhang (2005) found that a one-minute increase in total arrival delay at LGA resulted in a three-minute increase in total arrival delay at other major airports. Flight delays that originate at overscheduled airports can propagate rapidly through a highly connected airspace system, causing significant downstream delays (Beatty et al., 1998; Churchill et al., 2010; Nayak and Zhang, 2011). A major study estimated that the total cost of all U.S. air transportation delays in 2007 was $31.2 billion (Ball et al., 2010).

Airport delays have great impact within the NAS. To inform strategic planning, considerable efforts have been made to understand how and what factors contribute to airport delays and changes in delay over time. Airport operational performance has also been studied extensively. Both runway demand and capacity (or, throughput performance) are influenced by many operational, environmental, and human factors that can be highly stochastic. As a result, runway demand and throughput performance can be highly variable and difficult to predict, also contributing to demand-capacity imbalances and delays. Runway demands are set by airline scheduling practices as well as upstream disturbances – including delays, internal airline problems, and traffic management initiatives – that perturb the original schedules. Most major U.S. airports operate under dense schedules in order to accommodate heavy demands; however, four operate under slot controls to reduce queuing delays. Swaroop et al. (2012) investigated the levels of slot control needed to balance schedule delay and queuing delay.

Researchers have also extensively studied runway capacity, to better predict and ultimately reduce the quantity and cost of future airport delays. Airport capacity is affected by runway configurations, weather, airline fleet mix, use of new technologies (Hansen et al., 2002), controller workload and cognition, as well as equipment outages and other factors. In cases where multiple airports share an airspace, airport capacity can be affected by the capacity of the airspace (Donaldson and Hansman, 2011). Many efforts have been made to characterize and estimate airport runway capacity analytically (Newell, 1979), empirically (Hansen, 2004; Kim and Hansen, 2010), and using simulation (Odoni et al., 1997). Researchers have also studied the optimization of airport capacity through aircraft sequencing (Psaraftis, 1978; Venkatakrishnan et al., 1993; Gupta and Malik, 2009) and arrival/departure balancing (Gilbo, 1993), as well as estimating the evolution of capacity at an airport (Liu et al., 2008). For delay analysis, researchers have also focused on relating delay directly to the operational characteristics that might otherwise be used to define the capacity values themselves (Peterson et al., 1995).

The above literature review demonstrates that much effort has been dedicated to understanding airport demand and capacity, to assess their contributions to airport delays. However, when we observe shifts in delay levels from, for example, one year to the next, it can be difficult to quantify how much of the delay shift should be attributed to changes in demand or capacity (throughput performance), not to mention the factors that influence each, when demand-side phenomena and capacity-side phenomena are changing simultaneously. Previous research efforts have attempted to identify these underlying causes of delay changes, but have required explicit definitions and/or estimates of capacity (Hansen and Hsiao, 2005).

In this paper, we present a new queuing analysis-based methodology that decomposes the change in delay from one time period to another, into a part caused by changes in demand and another part caused by changes in throughput performance. This methodology involves the construction of counterfactual scenarios, using demands and throughput performance that are each taken from differing time periods. Because these scenarios never actually existed, we term them “counterfactual”. In a counterfactual scenario, we simulate serving demand from one time period (period 1), assuming throughput performance from another period (period 2). The difference in delay between what is observed in period 1 and what is simulated in the counterfactual can be interpreted as the delay difference caused by the inter-period difference in throughput performance. Using an analogous procedure, we can measure the delay impact of a change in demand as well. We implement this method using historical airport runway operations data from the Federal Aviation Administration’s (FAA’s) Aviation Systems Performance Metrics (ASPM) database, which includes observed demand and throughput counts by both the hour and quarter-hour for the major U.S. airports. The throughput counts implicitly contain runway capacity information, thereby eliminating the need to explicitly (and accurately) define and/or estimate capacities.

To demonstrate the methodology, we analyzed delay changes at LGA, JFK, and EWR airports between the time periods of May-September 2006 and May-September 2007. Delay at all three airports increased markedly from 2006 to 2007, with 2007 recording record-high delays. Our methodology determines the contributions of changes in demand and supply-side factors to this increase. In developing this methodology, we found it necessary to extend the concept of capacity – a single scalar parameter – to a broader concept, which we term “throughput performance”. This concept more systematically captures the relationship between demand and throughput observed in the data. The methodology allows us to quantitatively attribute changes in overall delay to changes in demand and throughput performance. To our knowledge, there are no previous efforts described in the transportation literature to build a methodology similar to the one presented in this paper.

3. Methodology

3.1 Deterministic Counterfactual Delay

Figures 1 and 2 illustrate our methodology in a stylized manner. They show fictional cumulative curve diagrams for two time periods, p_1 (Figure 1) and p_2 (Figure 2), of equal duration (T) that are to be compared[1]. In these figures, for illustrative purposes we assume that the total cumulative demand N remains unchanged from p_1 to p_2. The total area between the demand and throughput curves represents the delay for each period. In Figure 2, the total delay is represented by the sum of all shaded areas A, B, and C. The slope of the throughput curve (when it is greater than zero) is the capacity of the system. The change in total delay from p_1 to p_2 – the difference of the two shaded areas shown in Figures 1 and 2, represented by areas B and C in Figure 2 – could be caused by changes in the demand rate, changes in capacity, or both. In the particular example shown in these figures, note that the demand rate increased and the capacity decreased between periods p_1 and p_2, contributing to an overall increase in delay. The figures also clearly show that the delay increase caused by the reduction in capacity from p_1 to p_2 corresponds to area B, while that caused by the increase in the demand rate corresponds to area C.

[1] The time periods referred to here consist of two time blocks from differing time periods; for instance, in this analysis we are interested in comparing the throughput performance, demand, and delay changes between the time period of May 1 – September 30, 2006 and that of May 1 – September 30, 2007.

Place Figure 2 here.

To decompose the delay changes resulting from demand and throughput changes in a more formal and systematic manner, we construct a "counterfactual" scenario where the demands from p_1 are served by the capacity from p_2. The counterfactual scenario is represented in Figure 3, and the total area between the demand and throughput curves (including both solid (A) and striped (B) areas) represents the resulting counterfactual delay. The striped area B represents the difference between the resulting counterfactual delay and the delay in period p_1; this is the change in total delay due to the change in capacity from p_1 to p_2. The change in total delay due to a demand increase from p_1 to p_2 is represented by the p_2 delay of Figure 2 minus the counterfactual delay, and is depicted by the crosshatched area (area C) in Figure 2. The entire process is summarized in Table 1.

Place Figure 3 here.

Place Table 1 here.

We now extend the approach depicted in Figures 1–3 so that we can employ it using the real-world operational data briefly introduced in the previous section (and to be discussed in greater detail in Section 4). We first discretize the two time periods being compared (each of duration T) into 15-minute intervals. Next, we define demand in a particular interval as the cumulative number of flights that wish to arrive (or depart) by that time interval, minus the number that have already landed or departed by the end of the previous time interval. In other words, demand consists of flights that were not served in previous interval(s) plus "new" flights desiring service in the current interval. Demand in a given interval is thus the upper bound of the throughput for that interval, which is the number of flights that actually do arrive or depart.

3.2 Stochastic Counterfactual Delay

We next introduce the concept of throughput performance. This is an extension of the concept of capacity motivated by empirical observations of throughput and demand at airports. Figure 4 is based on such observations. Figure 4 shows the average observed arrival throughput counts at LGA, JFK, and EWR airports with respect to arrival demand over a 15-minute time interval (depicted on the x-axis) and visibility condition (Visual Meteorological Conditions, or VMC, in Figure 4), for the May-September period of 2006 (solid dots) and 2007 (hollow dots), as collected from the ASPM database. The visibility condition affects runway throughput because it determines whether aircraft operate under visual or instrument flight rules. This designation is included as a factor in describing throughput performance because of its significant impact on runway operational capacity. An airfield operating under Visual Meteorological Conditions (VMC) can process up to two times as many flights than if it were operating under Instrument Meteorological Conditions (IMC). In Figure 4, the average arrival count per 15-minute demand level was calculated by averaging all counts recorded at each 15-minute demand level from 0 to 60+ flights. The sizes of the plotted points reflect the number of counts observed at a given demand level. Figure 4 shows fairly wide ranges of arrival throughput counts observed at each airport and is roughly consistent with what would be expected of a server with finite capacity. The average throughput counts track demands up to a certain point, after which the trend stops as the facility cannot serve flights at the rate demanded any longer. After this peak throughput count level, counts decrease and then stabilize around a lower count. Capacity is expected to lie in the region where the average throughput counts peak (Hansen, 2004). In the cases where the throughput count attains the demand in a time interval, the demand is cleared and there is no residual queue that must be served in a later period. In cases where the throughput count is less than the demand, the result is a residual (or unserved) demand that queues and continues to contribute to demand in the next period. The details of the data on which Figure 4 is based, and the ASPM database in general, are discussed in greater detail in Section 4.2.

Place Figure 4 here.

The queuing diagrams of Figures 1-3 suggest that throughput is a deterministic quantity – the minimum of the capacity and total (new + queued) demand. Figure 4 is inconsistent with this idea, because a deterministic throughput function that depends on capacity and demand is an idealized representation of real-world operations. The alternative to such a deterministic throughput function is a stochastic throughput function that incorporates variability in the relationship between throughput and demand. We use the term throughput performance to represent the set of throughput count probability distributions conditional on demand and visibility condition. These distributions can be readily constructed for the airports included in the ASPM database, at quarter-hour time intervals. As these probability distributions are constructed without any assumptions about their shapes and parameters, they are non-parametric.

Figures 1-3 demonstrate how to construct counterfactual scenarios in an idealized deterministic queuing setting; we now seek a method for doing so using throughput performance. The demand is represented using a cumulative curve similar to those shown in Figures 1-3, constructed using empirical data available from ASPM. Given the cumulative demand curve from p_1, we wish to simulate a cumulative throughput curve based on the throughput performance in another period p_2, to ascertain how the change in throughput from p_1 to p_2 contributed to a change in delay. We achieve this by developing a stochastic throughput performance simulation method. Two unsatisfactory methods were constructed before a satisfactory third approach was developed. Each subsequent approach is a refinement of the last.

- In the first approach, we simulated throughput as the conditional mean based on demand and visibility condition. With demand and the simulated throughput for time interval j, we can simulate the total demand in time interval $j+1$. In turn, using demand in interval $j+1$ combined with the visibility condition that prevailed in $j+1$, we are able to select the appropriate conditional mean throughput for that time interval.
- In the second approach, we simulated throughput by sampling randomly from the appropriate conditional throughput distribution. In time interval j, we made a random draw from the period 2 throughput distribution at the period 1 demand and visibility condition in j. The throughput value obtained was used to calculate the demand in the next period in a manner similar to the first approach.
- The final method, which was the only satisfactory one, is discussed in greater detail in the following section (3.3). This method uses a probabilistic approach to simulate the time series of throughput values in the counterfactual scenario. It matches the counterfactual throughput distribution quantiles to those of the period p_2 baseline scenario, when demand and visibility condition are taken from period p_1. As a result, if observed throughput in a given time interval in period p_2 is unusually high or low for the given demand and visibility condition, the simulated counterfactual throughput will likewise be unusually high or low. We refer to this approach as the "quantile equivalence" method, because it uses period p_2 throughput quantiles to build the throughput function of the counterfactual scenario.

The performance of each candidate method was first tested by observing how successful it was in reproducing the delays in observed datasets. As mentioned previously, the datasets used in this analysis are from the ASPM database for LGA, JFK, and EWR airports from May-September of 2006 and 2007. The first two methods did not accurately reproduce delays in these periods. This is likely because delay is highly dependent on the preservation of underlying phenomena embedded in the time sequence of throughput levels, and neither of the first two methods were able to retain these time sequences accurately. The quantile equivalence method, however, is guaranteed to reproduce the delay in observed baseline scenarios. When the method is applied to an observed period, it precisely replicates the time series of throughputs in that period. Therefore the quantile equivalence method was used in the subsequent analysis, and we now present this method in more detail.

3.3 Quantile Equivalence Method

For the quantile equivalence approach, we require the conditional cumulative distribution functions (CDFs) of throughput for a given time period and operation type (arrivals or departures) at an airport. More specifically, define $F_{o,p}(q|d,w)$ as the CDF of throughput q for operation type $o \in \{arrivals, departures\}$ in time period p, conditional on demand d (for operation o) and visibility condition $w \in \{VFR, IFR\}$. It is calculated as follows:

$$F_{o,p}(q|d,w) = \sum_{k \le q} n_{o,p}(k,d,w) \,/\, \sum_{all\ k} n_{o,p}(k,d,w) \quad (1)$$

where $n_{o,p}(k,d,w)$ is the number of time intervals in p when the throughput count for operation type o is k, the demand is d, and the visibility condition is w.

Let j be a time interval within p, where $j = 1, \dots, J$. Let $g_{o,p}(j)$ be a time series of $F_{o,p}$ values in p, such that $g_{o,p}(j) = F_{o,p}\big(q(j)\big|d(j), w(j)\big)$. The time series $g_{o,p}(j)$ represents the throughput performance of the airport in j given the demand and visibility condition in j, relative to the distribution of all counts observed in time period p. Since $F_{o,p}(q|d,w)$ are proportions, all $g_{o,p}$ values are in the interval (0,1].

Because there are fewer instances of throughput count data recorded at very high demand levels (i.e., airport runways do not operate under extremely heavy queuing conditions as frequently as they do under lighter queuing conditions), the cumulative probability distributions of throughput counts conditional on higher demand values are often based on small numbers of observations. To avoid relying on probabilities constructed using sparse data, all throughput counts recorded in time intervals with demands beyond a predetermined threshold were combined into a single probability distribution. This threshold was chosen based on empirical information about demand-throughput relationships (as shown in Figure 4).

Suppose we now want to construct a counterfactual scenario that predicts the delay resulting from serving demand from period p_1 using the throughput performance from p_2. In this counterfactual scenario, the cumulative demand is based on the flight demand in p_1. We define $a_{o,p_1}(j)$ as the increment of demand added to the cumulative demand (for operation class o) in time interval j; $a_{o,p_1}(j)$ is taken directly from the data. We also develop a cumulative throughput curve, through simulation, that replicates the series $g_{o,p_1}(j)$ based on the CDFs from time period p_2. Assuming that the queue is empty at the beginning of p_1, we proceed as follows:

1. Initialize the simulated counterfactual demand, $\hat{d}_o(j)$, for the first time interval as $\hat{d}_o(1) = a_{o,p_1}(1)$.
2. For $j = 1, \dots, J$:

a. Find the conditional cumulative distribution function $F_{o,p_2}\left(q\middle|\hat{d}_o(j), w_{p_1}(j)\right)$ for all possible throughput levels $q = 0, \ldots, q_{max}$ in p_2, given demand $\hat{d}_o(j)$ and visibility $w_{p_1}(j)$. q_{max} is the highest recorded throughput count in p_2 for demand $\hat{d}_o(j)$ and visibility $w_{p_1}(j)$. Denote all the values of $F_{o,p_2}\left(q\middle|\hat{d}_o(j), w_{p_1}(j)\right)$ in ascending order as $P^0(j), P^1(j), \ldots, P^q(j), \ldots, P^{q_{max}}(j)$. For instance, when $q = 0$, $P^0(j) = F_{o,p_2}\left(0\middle|\hat{d}_o(j), w_{p_1}(j)\right)$. Also define q_{max} intervals, where interval $i(j)$ is $\left(P^{i-1}(j), P^i(j)\right]$. Note that $P^{q_{max}}(j) = 1, \forall j$.

b. Determine interval $i(j)$ to which $g_{o,p_1}(j)$ belongs. Recall that $g_{o,p_1}(j)$ is a time series of F_{o,p_1} values, representing the relative throughput performance of the airport in period p_1 and time interval j given prevailing demand and visibility conditions. Set the "lower" and "upper" counts for interval $i(j)$ as $q^L(j) = F^{-1}_{o,p_2}\left(P^{i-1}(j)|\hat{d}_o(j), w_{p_1}(j)\right)$ and $q^U(j) = F^{-1}_{o,p_2}\left(P^{i}(j)|\hat{d}_o(j), w_{p_1}(j)\right)$, respectively.

c. Randomly draw a value for n, where n is distributed uniformly on [0,1]. Assign $\hat{Q}_o(j)$ – the simulated counterfactual throughput in j – to be either $q^L(j)$ or $q^U(j)$, based on the condition:

$$\hat{Q}_o(j) = \begin{cases} q^L(j) & if\ n \leq h(j) \\ q^U(j) & otherwise \end{cases} \tag{2}$$

where $h(j)$ is the probability that $\hat{Q}_o(j)$ is assigned the lower throughput value q^L, and takes the form

$$h(j) = \left(P^i(j) - g_{o,p_1}(j)\right) / \left(P^i(j) - P^{i-1}(j)\right) \tag{3}$$

Equation 3 indicates that $h(j)$ is based on the position of $g_{o,p_1}(j)$ on the interval $(P^{i-1}(j), P^i(j)]$.

d. If $j < J$, calculate the simulated counterfactual demand $\hat{d}_o(j+1)$. It is comprised of the "new" demand to the queue in time interval $(j+1)$, $a_{o,p_1}(j+1)$, plus the simulated queued (unserved) aircraft from the previous interval j, $\left[\hat{d}_o(j) - \hat{Q}_o(j)\right]$, such that $\hat{d}_o(j+1) = a_{o,p_1}(j+1) + \left[\hat{d}_o(j) - \hat{Q}_o(j)\right]$.

3. If $j = J$ and all demand has been served such that $\hat{d}_o(j) = \hat{Q}_o(j)$, then stop. If, on the other hand, there is a residual queue in time period J such that $\hat{d}_o(j) > \hat{Q}_o(j)$, continue the simulation by setting $\hat{d}_o(j+1) = \left[\hat{d}_o(j) - \hat{Q}_o(j)\right]$ until there is no queue. Use $g_{o,p_1}(J)$ in step 2b to simulate the count. Define J' as the time period when the queue becomes empty, or $\hat{d}_o(J') \leq \hat{Q}_o(J')$.

4. Calculate the average delay per flight for operation class o by summing the resulting flight queues over all time intervals $j = 1, \ldots, J'$, and dividing by the total demand:

$$\widehat{w}_o = \frac{\Delta t \cdot \sum_{j=1}^{J'}\left[\hat{d}_o(j) - \hat{Q}_o(j)\right]}{\sum_{j=1}^{J} a_{o,p_1}(j)} \tag{4}$$

where $\hat{\bar{w}}_o$ is the average simulated counterfactual delay per flight for operation class o, Δt is the length of one time interval, and $\sum_j^J a_{o,p_1}(j)$ is the total demand for operation o in p_1.

The above procedure can be summarized as follows. Steps 2.a-2.c describe how the stochastic throughput function $\hat{Q}_o(j)$ is obtained; the empirically-derived conditional cumulative probabilities of throughput counts in p_1 – that is, $g_{o,p_1}(j)$ – are compared to the cumulative probabilities of throughput counts from p_2. In Step 2.a, we first determine the family of CDF values from which we will obtain our counterfactual throughput. Recall that the CDFs are conditional on the visibility condition from p_1; this means that we are using the throughput performance CDFs from p_2, but choosing the appropriate CDF according to the visibility condition of p_1. In 2.b, we determine the interval of the CDF family to which $g_{o,p_1}(j)$ belongs. In 2.c, we randomly assign $\hat{Q}_o(j)$ to the throughput associated with the lower bound $\left(q^L(j)\right)$ or upper bound $\left(q^U(j)\right)$ of the interval $i(j)$. In Step 2.d, we calculate the counterfactual demand queue in the proceeding time interval $j+1$, based on how much of the demand in j was served by the throughput simulated in 2.c. After Step 2 has been performed for all J sequentially, in Step 3 we either stop the simulation or continue assigning throughput until the demand queue is cleared. The average counterfactual delay per flight is then calculated in Step 4.

With the counterfactual scenario generated above, we can determine how much the changes in delay from period p_1 to the counterfactual scenario, and from the counterfactual scenario to p_2, can be attributed to changes in throughput and demand. If $\hat{\bar{w}}_o - \bar{w}_{o,p_1} > 0$, where $\hat{\bar{w}}_o$ is the average counterfactual delay per flight and $\bar{w}_{o,p_1}$ is the average delay per flight recorded in p_1, it is an indication that throughput performance has degraded between p_1 and p_2, contributing to a delay increase from p_1 to p_2. If $\hat{\bar{w}}_o - \bar{w}_{o,p_1} = 0$, we know that throughput performance has, on the whole, remained unchanged. If $\hat{\bar{w}}_o - \bar{w}_{o,p_1} < 0$, this indicates that the throughput performance has improved between p_1 and p_2. We can also compare the average delay of the counterfactual scenario against the average recorded delay in p_2, or $\bar{w}_{o,p_2}$, to determine the effects that changes in demand may have had on changes in delay. However, recall that the counterfactual scenario is constructed based on the visibility condition of p_1; therefore, any difference between $\bar{w}_{o,p_2}$ and $\hat{\bar{w}}$ is not only due to a shift in demand levels from p_1 to p_2 but also to overall differences in visibility conditions. Suppose that visibility conditions did not change significantly between p_1 and p_2. In this case, if $\bar{w}_{o,p_2} - \hat{\bar{w}}_o > 0$, we can say that a shift from demand levels in period p_1 to those in period p_2 contributed to an increase in delay, and therefore demand increased between p_1 and p_2. A similar logic can be applied for the situations where $\hat{\bar{w}}_o - \bar{w}_{o,p_2} < 0$ and $= 0$[2].

4. Case Study

4.1 Background

To demonstrate the methodology presented in Section 3, we analyze demand and throughput operations at the New York area airports of LaGuardia (LGA), Newark Liberty (EWR), and John F. Kennedy (JFK) between May-September 2006 and 2007. During the summer of 2007, flight delays reached record high levels throughout the National Airspace System (NAS). National and international news agencies reported headline after headline describing the extreme wait times resulting from flight departure delays and missed connections that plagued

[2] One could extend our methodology to explicitly account for visibility conditions, and decompose a delay change into three parts – one caused by a change in throughput performance, one by a change in visibility conditions, and one by a change in demand. We combine the visibility and demand components because they are both outside the control of the air navigation service provider.

airlines and air travelers during this peak travel season. LGA, EWR, and JFK experienced some of the highest delays in the NAS, with travelers spending an additional 3.9 million hours waiting for their aircraft to take off after leaving their gates in 2007 as compared to a decade earlier (Belson, 2007). Knorr (2007) reported that the growth in total aircraft arrival and departure operations from 2006 to 2007 at these airports was approximately 3-4%, but the increase in delay was in the order of about 28%. Using the presented methodology, we investigate the change in delay between the summer months of 2006 and the summer months of 2007, in order to assess the contributions of demand and throughput performance to this remarkable reported increase.

We note here that LGA, EWR, and JFK share the most congested airspace in the NAS. Flights that arrive and depart from these airports share many of the same routes through this airspace, and as a result, operations at these three airports are not always independent of one another. In other words, constraints in the airspace can impact individual airport capacities and not always in the same manner or magnitude (Donaldson and Hansman, 2011). However, operational interaction with another airport (or airports) is simply another reason why the capacity of an airport can fluctuate. As a result, we argue that with or without interactions, it is appropriate to perform this analysis at the individual airport level.

We first looked at the FAA's Operations Network (OPSNET) database, a publicly available source of historical NAS operational data, to provide some general metrics, trends, and magnitudes regarding operational changes at these airports between 2006 and 2007. OPSNET defines a delay event as an occasion when an aircraft is detained 15 minutes or more, and attributes each event to a particular NAS facility. Table 2 contains the results of OPSNET airport delay data extracted for LGA, EWR, and JFK for May-September of 2006 and 2007. The first half of the table indicates that the total number of operations (throughput) decreased at both LGA and EWR, but significantly increased at JFK. The second half of the table shows the number of flights that were delayed more than 15 minutes from their flight plans. It can be observed that the number of delayed flights increased at LGA and JFK by 13% and 76% respectively, and decreased at EWR by 9% from 2006 to 2007.

Place Table 2 here.

4.2 Description of Data

The Aviation System Performance Metrics (ASPM) database is part of the FAA Operations and Performance Data system, and is available at https://aspm.faa.gov/. Data from the "Download/Airport" section of the ASPM database, which can be accessed online with permission from the FAA, was used for this analysis. The data includes hourly as well as quarter-hourly arrival and departure counts, demands, and visibility conditions (either VMC or IMC). The data is available for 77 major airports in the United States.

ASPM count data is based on individual aircraft landing (wheels-on) and take-off (wheels-off) times, as supplied through Airline Service Quality Performance (ASQP) data or Enhanced Traffic Management System (ETMS) radar surveillance data. Demand in ASPM is not based on airline schedules, but rather, updated flight plans filed approximately 15 minutes before a flight is due to push back from its gate at the origin airport. However, the demand is calculated such that the effects of a Ground Delay Program (GDP) are accounted for and attributed to the airport at which the GDP was called. Arrival demands at an airport are calculated as follows: the start of arrival demand for a flight is calculated by adding its filed en-route time to its wheels-off time at the departure airport, and subtracting ground delay time imposed on this flight at the departure airport due to a GDP or ground stop at the (arrival) airport. Therefore, the arrival demand at an airport is adjusted to account for ground delays at that airport. The start of departure demand for a flight at an airport is calculated by adding the flight's unimpeded taxi-out time and any ground delay time to its filed gate-out time. The departure demand is adjusted such that any ground delay imposed on a departing flight (due to operational problems at the destination airport) is not attributed to the departure airport. As a result, for a flight departing an airport, the departure delay calculated in this analysis includes the delay incurred between the time the flight was filed to depart and the time it actually does, adjusting for any ground delay. To summarize, a flight will contribute towards demand starting in the time interval it was first filed to arrive/depart, until the time interval it actually does arrive/depart, adjusting for ground delay appropriately. Demand $d_o(j)$ includes the "new" demand $a_o(j)$ in time interval j, plus the queued (unserved) aircraft from the previous time interval $j-1$. For our counterfactual simulations, we require $a_o(j)$ for each 15-minute interval $j=1,\dots,J$. Although it is not directly available in the ASPM dataset, it is easily calculated from quantities that are directly provided according to Equation 5 below:

$$a_o(j) = d_o(j) - [d_o(j-1) - q_o(j-1)] \quad (5)$$

where $a_o(j)$ is the "new" demand for operation type o in time interval j, $d_o(j)$ is the total demand for operation o in time interval j, and $q_o(j-1)$ is the throughput for operation o in time interval $j-1$.

Note that $[d_o(j-1) - q_o(j-1)]$ is the number of unserved aircraft from the previous time interval, $j-1$, that wait for service in j.

In practice, an arriving or departing flight may require service earlier than planned. When this occurs, the time interval in which the flight is served, and therefore included in the operational throughput count, is recorded as the first and only time interval in which it is counted towards demand. Thus, any time interval in which a flight is served is also an interval when it is included in the demand. However, a delayed flight may count towards demand in time intervals when it does not contribute to the operational count, because it must wait until a future time interval to be served. As a result, the throughput count can never exceed demand, just as the cumulative count curve can never cross the cumulative demand curve in a deterministic queuing diagram.

ASPM data from LGA, EWR, and JFK was obtained for May 1 through September 30 of 2006, and May 1 through September 30 of 2007. Period 1 (p_1) as identified in Section 3 will correspond to May-September 2006, while Period 2 (p_2) corresponds to May-September 2007. Based on the data, Figure 5 displays cumulative hourly arrival demands at each airport by year, averaged over all days from May 1 through September 30. It can be observed that the average cumulative daily demand at LGA and EWR decreased (between 2-3%) from 2006 to 2007 while it increased significantly (by 18%) at JFK. The departure demands exhibit very similar trends and as a result are not displayed here in the interest of length.

Place Figure 5 here.

Figure 6 displays the average departure counts recorded during VMC conditions plotted against departure demand; Figure 6 is identical to Figure 4 except that it is constructed based on flight departures at each airport, while Figure 4 is for arrivals. The average departure throughput count per demand was calculated by averaging all counts recorded at each demand level from 0 to 60+. In Figure 6, like Figure 4, the average throughput counts track demands up to a certain point, after which this trend stops as the facility cannot serve flights at the rate demanded any longer. Again, capacity is expected to lie in the region where the average throughput counts peak. Figure 4 suggested that the arrival capacities of all three airports decreased from 2006 to 2007, since the peak average counts are smaller in the latter year. Although not easily observable in Figure 4, the highest arrival demands reported at LGA and JFK in 2007 were higher than those of 2006, suggesting that queues were longer in 2007. This in turn suggests that aircraft waited longer for service, and therefore experienced greater delay in 2007. This phenomenon, however, was not observed at EWR. A similar interpretation applies for the departure throughput counts in Figure 6. The data suggests that departure capacities may also have decreased at LGA and EWR between 2006 and 2007, but increased at JFK.

Place Figure 6 here.

Recall that delay is calculated against flight plan demand, and the data is recorded in quarter-hour intervals ($\Delta t = 15$ min). Each dataset from May 1 through September 30 contains $J = 14{,}688$ quarter-hour intervals. The average delay per flight (in minutes) was calculated for arrival and departure operations at each airport from May through September of 2006 and 2007, as per Equation 6:

$$\bar{w}_{o,Y} = \frac{15 \cdot \sum_{j}^{14{,}688} \left[d_{o,Y}(j) - q_{o,Y}(j) \right]}{\sum_{j}^{14{,}688} q_{o,Y}(j)} \tag{6}$$

where $\bar{w}_{o,Y}$ is the average delay per flight of operation o from May through September of year Y (2006 or 2007). Demand (new + queued) in time interval j is represented by $d_{o,Y}(j)$, and $q_{o,Y}(j)$ is the throughput in interval j. The results are summarized in Table 3.

Place Table 3 here.

According to Table 3, the average delay per flight increased at both LGA and JFK between 2006 and 2007, and significantly so for JFK arrivals (4.88 minutes per flight). At EWR, average delay actually decreased by about 1.6 minutes per departing flight, while for arrival flights it increased 0.6 minutes. These results from ASPM are consistent with the OPSNET data previously discussed.

4.3 Counterfactual Delay Results

The results of the counterfactual delay simulation and analysis for the three major New York airports – where 2006 demand is served using a throughput performance function from 2007 – are shown in Table 4. Each of the six counterfactual delays ($\hat{\bar{w}}_o^1$) shown (arrival and departure operations for each of the three airports) is the average result of 1,000 simulation runs. The coefficient of variation of each set of 1,000 runs of the counterfactual scenario range from 0.5-2.1%; they are also reported in Table 4.

Place Table 4 here.

The results of Table 4 are consistent with the trends observed in the OPSNET data summarized in Table 2. At LGA and EWR, when we compare the 2006 average flight delays for both arrivals and departures against the counterfactual scenario delays, we can observe that changes in throughput performance from 2006 to 2007 have resulted in increases in average delays for both departure and arrival operations at both airports. This implies that throughput performance deteriorated at these airports from 2006 to 2007.

Recall that the counterfactual scenario is based on 2006 demands and visibility conditions. An analysis of the data indicates that overall, visibility conditions in 2006 and 2007 were similar. In addition, weather in the New York area was not reported to be significantly different in the summer of 2007 compared to previous summers. Therefore, when we compare the counterfactual delay results against the 2007 average flight delays, we observe that changes in demand have contributed to the lessening of delays, and thus departure and arrival demands decreased at both LGA and EWR from 2006 to 2007.

The above results suggest that the effects of both departure and arrival demand reductions at LGA were not sufficient to offset the larger effects of throughput performance declines, resulting in the overall increase in delay experienced at this airport from 2006 to 2007. In other words, a substantial decline in throughput performance at LGA was partially masked by an offsetting reduction in demand. At EWR, the average departure delay fell between the 2006 and 2007 summer seasons because decreases in demand more than offset the concurrent (and large) decline in throughput performance. On the other hand, arrival demand did not decrease sufficiently to completely offset the decline in arrival throughput performance at EWR, resulting in a relatively small net increase in average delay per flight (0.6 minutes, according to Table 3) from 2006 to 2007.

Other sources (including OPSNET; see Table 2) have documented that throughput at LGA and EWR declined from 2006 to 2007, although scheduled demands in the summer of 2007 continued to exceed published airport capacities. A study by the MITRE Corporation's Center for Advanced Aviation System Development (CAASD) found that the average capacity at EWR was 83 operations per hour for the period of September 2006-August 2007, down almost 5% from the previous year (FAA, 2007). Also, GDPs and ground stops were used nearly every weekday at LGA and EWR in 2007, which limit throughput in addition to impacting demands. In addition, decreases in overall throughput performance might also be attributed to causes that are not controlled for in our methodology. For instance, the methodology does not control for runway configurations used or aircraft fleet mixes, both of which can impact runway capacity (Newell, 1979). Our methodology also does not control for congestion in the shared airspace above the airport, which may also affect capacities. From the demand side, published in December 2006 was the FAA Order Operating Limitations at New York LaGuardia Airport, which imposed a reservation system for unscheduled operations that permitted a maximum of six unscheduled operations per hour (FAA, 2008).

In comparing the 2006 average flight delays against the counterfactual delays at JFK, we can observe that departure throughput performance improved in 2007, but not enough to offset the larger growth in departure demand. Overall, this has resulted in an approximately 19% increase in average flight delay from 2006 to 2007. JFK arrival delays, which were relatively low in 2007, increased 151% in 2007. From the counterfactual scenario results of Table 4 ($\widehat{w}_o^1 = 6.26$), it appears that the majority of this increase (62%) can be attributed to the decrease in arrival throughput performance. Schedule increases at JFK likely led to a growth in demand, given the growth of JetBlue's operations and the establishment of a Delta Airlines hub at JFK in late 2006 (FAA, 2010). The arrival throughput performance may have decreased for the same reasons posited above for LGA and EWR.

The counterfactual scenario can also be constructed by serving 2007 demands with a throughput function based on the 2006 throughput performance; in other words, we can use the same procedure described above and in Section 3, but with the demand and throughput years swapped. The average delay per flight of this version of the counterfactual scenario is denoted by $\widehat{w}_o^2$. In this case, the difference between the counterfactual average flight delay $\widehat{w}_o^2$ and the 2006 average flight delay can be attributed to changes in demand (rather than changes in throughput), given that visibility conditions on the whole were not significantly different between the two years. Also, the differences between the 2007 delays and the counterfactual delays can be attributed to changes in throughput performance. Table 5 contains the results of this simulation. Again, the counterfactual scenario delay is the average of 1,000 runs.

Place Table 5 here.

The overall directional trends and insights gained from Table 5 are consistent with those of Table 4. However, the magnitudes of the changes in delay differ. The greatest discrepancies are observed in the arrival results for LGA and JFK, with the latter difference by far the largest. It is not surprising that the results vary with the "direction" of the simulation. Firstly, delay depends on both the magnitudes and interactions of throughput performance and demand, and their effects are not strictly additive. As demand queues grow, delay also increases at increasing rates (Hansen, 2004). The effect of a change in throughput performance on delay also depends on demand. It is this dependence that accounts for the major differences in the results of the two counterfactual simulations. Secondly, recall that the counterfactual results of Table 4 are based on visibility conditions experienced in 2006 while the results of Table 5 are based on those in 2007. Although visibility conditions were generally not significantly different between the two years, this may have contributed at least in a small way to the differences in the results. Nonetheless, there is a strong correlation between the two sets of results. For example, the correlation between the changes in delay due to throughput performance for the two sets of results is 0.91. Moreover, the signs and relative magnitudes of the delay changes are consistent. Only the arrival delay results at JFK are significantly different in magnitude; the Table 5 results indicate that the change in delay due to changes in demand versus change in delay due to changes in throughput performance can be calculated at 11%, while in Table 4 it is 61%.

We can make several inferences based on the results of Tables 4 and 5. Firstly, of the three airports, JFK experienced the largest overall increase in average flight delay between 2006 and 2007, with the growth in departure demand the largest single contributor to delay (4.97 and 6.14 minutes per flight, as shown in Tables 4 and 5 respectively). This finding is supported by the fact that JFK experienced large schedule increases from 2006 to 2007. Overall throughput performance at JFK has not been a major driver of delay changes, with improvements in departure throughput performance largely offsetting declines in arrival throughput performance. In Table 4, the ratio of the decrease in JFK departure delay due to changes in throughput performance (-2.65) to the increase in arrival delay due to changes in throughput performance (3.03) is -0.87; in Table 5, this ratio is also -0.87. Moderate increases in delay at LGA reflect the net difference between a substantial decline in throughput performance and a more modest decrease in demand. At EWR, departure delay decreased while the arrival operations experienced a small increase in delay. Similar to LGA, these trends at EWR reflect declines in throughput performance offset by reductions in demand, which were discussed earlier. We again mention here that aircraft interactions in the airspace shared by these three airports may be partially responsible for the throughput performance results summarized above. However, we reiterate that these interactions are simply another cause of airport throughput performance fluctuations, and the analysis and results are valid at the individual airport level.

5. Conclusions & Future Work

This paper introduced a methodology to isolate and quantify the effects that changes in demand and throughput have on delay. When both demand and throughput performance are evolving simultaneously, it can be difficult to ascertain how much the resulting changes in delay should be attributed to either of these causes. We have proposed an empirical, stochastic simulation methodology that is based on classic queuing concepts. The main engine of this methodology is a simulation procedure that implements the "quantile equivalence" method, which is used to create a counterfactual queuing scenario that consists of demand from one time period being served using throughput performance from another. The key advantages of this procedure are twofold. Firstly, it uses non-parametric probability distributions of throughput counts that are constructed from readily available data. Secondly, the method instantiates the observed stochastic and dynamic properties of the relationship between throughput and demand at airports, one that is only roughly approximated by a model based on the conventional notion of capacity. The resulting counterfactual scenarios were then used to attribute changes in delays to changes in throughput or demand individually.

The methodology was applied to a case study of flight delays at the three major New York airports of LaGuardia (LGA), Newark Liberty (EWR), and John F. Kennedy (JFK). These airports experienced significant increases in delays during the summer of 2007 compared to the previous summer. We applied the methodology to determine why this occurred. Data from the FAA's ASPM database was used for the analysis. The counterfactual scenario was first constructed with 2006 demands served using a throughput function based on 2007 throughput performance. At both LGA and EWR, the decreases in demand were not generally significant enough to offset the declines in throughput performance, and, as a result, delays increased at both (except for EWR departures where a larger decrease in demand was experienced). At JFK, both departure demands and throughput performance increased, with the net result being an overall increase in delays. On the other hand, arrival demands increased and throughputs decreased, significantly increasing what were relatively low arrival delays in 2006.

This procedure is a step towards further identifying and assessing the causes of operational delay at airports. It would be beneficial to identify and control for factors other than designated visibility conditions that influence demand and throughput performance. These factors might include runway configuration used, fleet mix changes, and arrival/departure interaction effects. They could be used to re-specify the throughput performance function in order to control for additional factors. A statistical capacity model using the above factors as explanatory variables could also be specified and estimated to compare results (Kim and Hansen, 2010). Also, the application of the methodology in this study is highly aggregate; it would be helpful to hone in on specific time periods within the larger time period simulated (May-September) to help isolate "hotspots" where more careful analysis is required. More widespread application of the methodology across airports and time periods would lead to a more comprehensive picture of how throughput performance is evolving.

Finally, while airport delay is the focus of this study, the methodology has wider potential applications, including to other types of transportation facilities and queuing systems in which throughput performance cannot be easily characterized using a simple capacity parameter.

Acknowledgments

This work was sponsored by the Dwight D. Eisenhower Transportation Fellowship Program, as well as the Air Traffic Organization at the FAA. An early version of this paper was presented at the 3rd International Conference on Research in Air Transportation in Fairfax, Virginia. The authors would like to thank Joe Post, Michael Wells, and Dan Murphy at the FAA for providing invaluable insights and suggestions. The authors would also like to acknowledge Dave Knorr at the FAA, Ken Wright at MITRE, and Marc Rose at MCR for their assistance. Finally, the authors also thank the editor and two anonymous reviewers for their comments and critique.

References

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Zou, B. 2010. Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States. NEXTOR Technical Report: October 2010.

Beatty, R., Hsu, R., Berry, L. Rome, J., 1998. Preliminary Evaluation of Flight Delay Propagation Through an Airline Schedule. 2nd U.S.–Europe Air Traffic Management Research and Development Seminar. Orlando, FL.

Belson, K., 2007. Study puts price tag on delays at airports. The New York Times. <http://www.nytimes.com/2007/12/02/nyregion/02airport.html>.

Churchill, A. M., Lovell, D. J., Ball, M. O., 2010. Flight Delay Propagation Impact on Strategic Air Traffic Flow Management. Transportation Research Record 2177, 105-113.

Daganzo, C., 1997. Fundamentals of Transportation and Traffic Operations. New York: Emerald.

de Neufville, R. & Odoni, A., 2002. Airport Systems: Planning, Design and Management. McGraw Hill.

Donaldson, A. D. & Hansman, R. J., 2011. Improvement of Terminal Area Capacity in the New York Airspace. MIT International Center for Air Transportation (ICAT) Technical Report: February 2011.

FAA, 2007. Notice of Airport Level Designation for Newark Liberty International Airport for the Summer 2008 Scheduling Season. Federal Register 72(247), 73418-73419. <http://www.gpo.gov/fdsys/pkg/FR-2007-12-27>.

FAA, 2008. Congestion Management Rule for LaGuardia Airport. Federal Register 73(75), 20846-20868. <http://www.gpo.gov/fdsys/pkg/FR-2008-04-17/html/E8-8308.htm>.

FAA, 2010. New York Flight Delays Have Three Main Causes, but More Work Is Needed To Understand Their Nationwide Effect. Federal Aviation Administration Report Number AV-2011-007: October 2010. <http://www.oig.dot.gov/sites/dot/files/NY%20Delays%20Final.pdf>.

Gilbo, E. P., 1993. Airport Capacity: Representation, Estimation, Optimization. IEEE Transactions on Control Systems Technology 1(3), 144-154.

Gupta, G. & Malik, W., 2009. A Mixed Integer Linear Program for Airport Departure Scheduling. 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO). Hilton Head: AIAA.

Hansen, M., 2004. Post-deployment analysis of capacity and delay impacts of an airport enhancement: case of a new runway at Detroit. Air Traffic Control Quarterly 12(4), 339-365.

Hansen, M. & Hsiao, C. Y., 2005. Going South? Econometric Analysis of U.S. Airline Flight Delays from 2000 to 2004. Transportation Research Record 1915, 85-94.

Hansen, M., Mukherjee, A., Knorr, D. & Howell, D., 2002. Effect of T-TMA on Capacity and Delay at Los Angeles International Airport. Transportation Research Record 1788, 43-48.

Hansen, M. & Zhang, Y., 2005. Operational Consequences of Alternative Airport Demand Management Policies: Case of LaGuardia Airport, New York. Transportation Research Record 1915, 95-104.

Kim, A. & Hansen, M., 2010. Validation of Runway Capacity Models. Transportation Research Record 2177, 69-77.

Knorr, D., 2007. Delay Metrics: Why are Delays Increasing?. Pacific Grove, CA.

Liu, P.-c. B., Hansen, M. & Mukherjee, A., 2008. Scenario-based air traffic flow management: From theory to practice. Transportation Research Part B 42(7-8), 685-702.

Lovell, D. J., Vlachou, K., Rabbani, T. & Bayen, A., 2013. A diffusion approximation to a single airport queue. Transportation Research Part C 33, 227-237.

Mukherjee, A. & Hansen, M., 2009. A dynamic rerouting model for air traffic flow management. Transportation Research Part B 43(1), 159–171.

Nayak, N. & Zhang, Y., 2011. Estimation and Comparison of Impact of Single Airport Delay on National Airspace System with Multivariate Simultaneous Models. Transportation Research Record 2206, 52-60.

Newell, G., 1979. Airport capacity and delays. Transportation Science 13(3), 201-241.

Nikoleris, T. & Hansen, M., 2012. Queueing Models for Trajectory-Based Aircraft Operations. Transportation Science 46(4), 501-511.

Odoni, A. R. et al., 1997. Existing and Required Modeling Capabilities for Evaluating ATM Systems and Concepts. NASA Final Report: March 1997.

Peterson, M. D., Bertsimas, D. J. & Odoni, A. R., 1995. Models and Algorithms for Transient Queuing Congestion at Airports. Management Science 41(8), 1279-1295.

Psaraftis, H., 1978. A Dynamic Programming Approach to the Aircraft Sequencing Problem. Cambridge.

Pyrgiotis, N., Malone, K. M. & Odoni, A., 2013. Modelling delay propagation within an airport network. Transportation Research Part C 27, 60-75.

Santos, G. & Robin, M., 2010. Determinants of delays at European airports. Transportation Research Part B 44(3), 392–403.

Swaroop, P., Zou, B., Ball, M. O. & Hansen, M., 2012. Do more US airports need slot controls? A welfare based approach to determine slot levels. Transportation Research Part B 46(9), 1239–1259.

Venkatakrishnan, C., Barnett, A. & Odoni, A. R., 1993. Landings at Logan Airport: Describing and Increasing Airport Capacity. Transportation Science 27(3), 211-227.

Xu, N., Sherry, L. & Laskey, K. B., 2008. Multifactor Model for Predicting Delays at U.S. Airports. Transportation Research Record 2052, 62-71.

List of Figures

List of Tables

3. Table 1

Table 1. Demand and throughput scenarios

Demand	Throughput	Total Delay	Δ in Total Delay
Period p_1	Period p_1	(A) Period p_1	n/a
Period p_1	Period p_2	(A)+(B): Counterfactual	(B); due to capacity change
Period p_2	Period p_2	(A)+(B)+(C) Period p_2	(C); due to demand change

3. Table 2

Table 2. OPSNET data for New York airports, May-September 2006 & 2007

Total Aircraft Arrival and Departure Operations (May-Sept)

	2006	2007	*% change, 2006-2007*
LGA	172,142	168,616	*-2.0%*
EWR	191,531	188,211	*-1.7%*
JFK	169,957	197,626	*16.3%*

Total Number of Flights Delayed >15 Minutes (May-Sept)

	2006	2007	*% change, 2006-2007*
LGA	12,992	14,647	*12.7%*
EWR	20,051	18,251	*-9.0%*
JFK	7,855	13,799	*75.7%*

3. Table 3

Table 3. Average delay per flight, May-September 2006 and 2007

		Average delay per flight, $\overline{w}_{o,Y}$, (min)		Change	
		2006	**2007**	**Δ06-07**	**% Δ06-07**
LGA	Departure	8.56	10.72	2.16	25%
	Arrival	8.85	10.70	1.85	21%
EWR	Departure	11.53	9.95	-1.58	-14%
	Arrival	11.46	12.06	0.60	5%
JFK	Departure	12.06	14.38	2.32	19%
	Arrival	3.23	8.11	4.88	151%

3. Table 4

Table 4. Counterfactual delay simulation results (demand p, throughput p')

	Average Delay per Flight, $\bar{w}_o$ (min)			Δ Delay due to Δ Throughput	Δ Delay due to Δ Demand	σ**
	2006	$\hat{\bar{w}}_o^1$*	2007			
LGA						
Departure	8.56	13.99	10.72	5.43 [a]	-3.27 [b]	0.125
Arrival	8.85	19.73	10.70	10.88 [a]	-9.03 [b]	0.388
EWR						
Departure	11.53	15.03	9.95	3.50 [a]	-5.08 [b]	0.095
Arrival	11.46	18.94	12.06	7.48 [a]	-6.88 [b]	0.177
JFK						
Departure	12.06	9.41	14.38	-2.65 [c]	4.97 [d]	0.060
Arrival	3.23	6.26	8.11	3.03 [a]	1.85 [d]	0.129

* Simulated counterfactual delay with 2006 demand & 2007-based throughput performance function

** Standard deviation of counterfactual delay, for 100 simulation runs made

a Throughput has decreased

b Demand has decreased

c Throughput has increased

d Demand has increased

3. Table 5

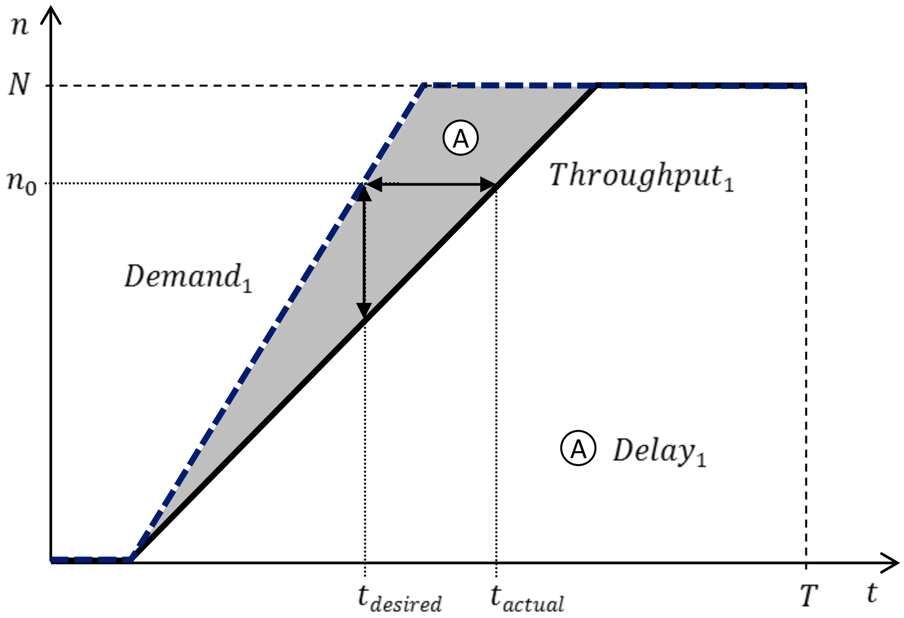
Table 5. Counterfactual delay simulation results (demand p', throughput p)

	Average Delay per Flight, $\bar{w}_o$ (min) 2006	$\hat{\bar{w}}_o^2$*	2007	Δ Delay due to Δ Demand	Δ Delay due to Δ Throughput	σ**
LGA						
Departure	8.56	5.96	10.72	-2.60	4.76	0.102
Arrival	8.85	4.43	10.70	-4.42	6.27	0.100
EWR						
Departure	11.53	7.46	9.95	-4.07	2.49	0.064
Arrival	11.46	7.13	12.06	-4.33	4.93	0.123
JFK						
Departure	12.06	18.20	14.38	6.14	-3.82	0.081
Arrival	3.23	3.72	8.11	0.49	4.39	0.116

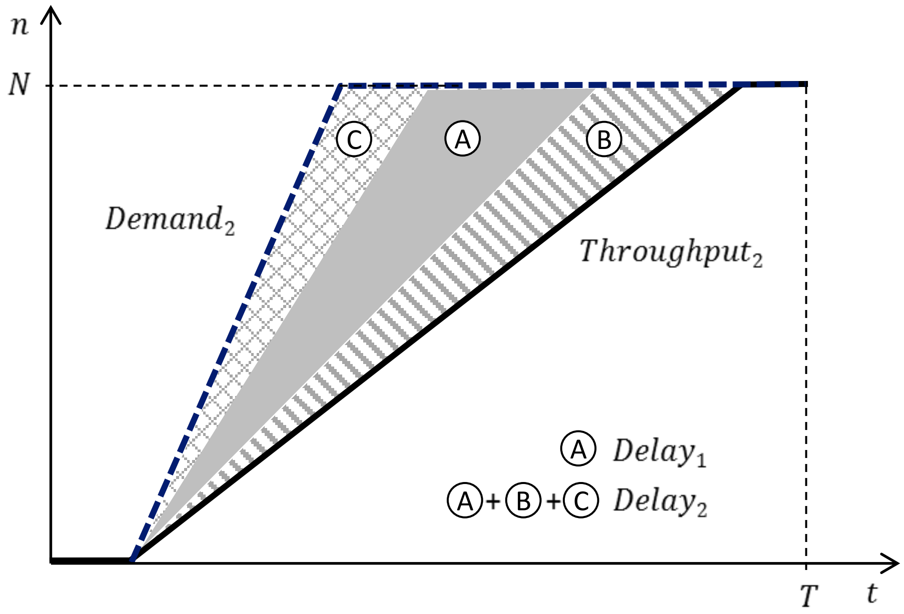
* Simulated counterfactual delay with 2007 demand & 2006-based throughput performance function

** Standard deviation of counterfactual delay, for 100 simulation runs made

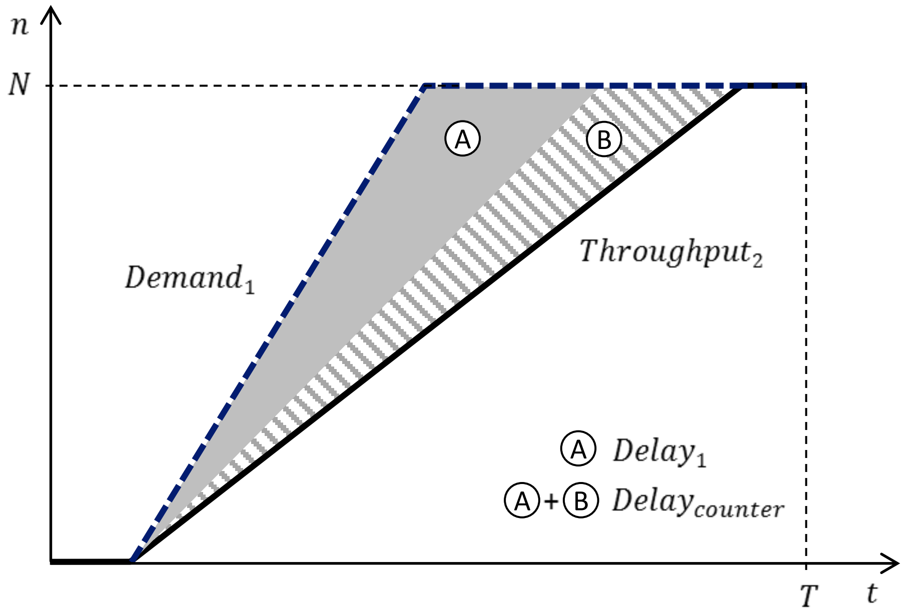
4. Figure 1



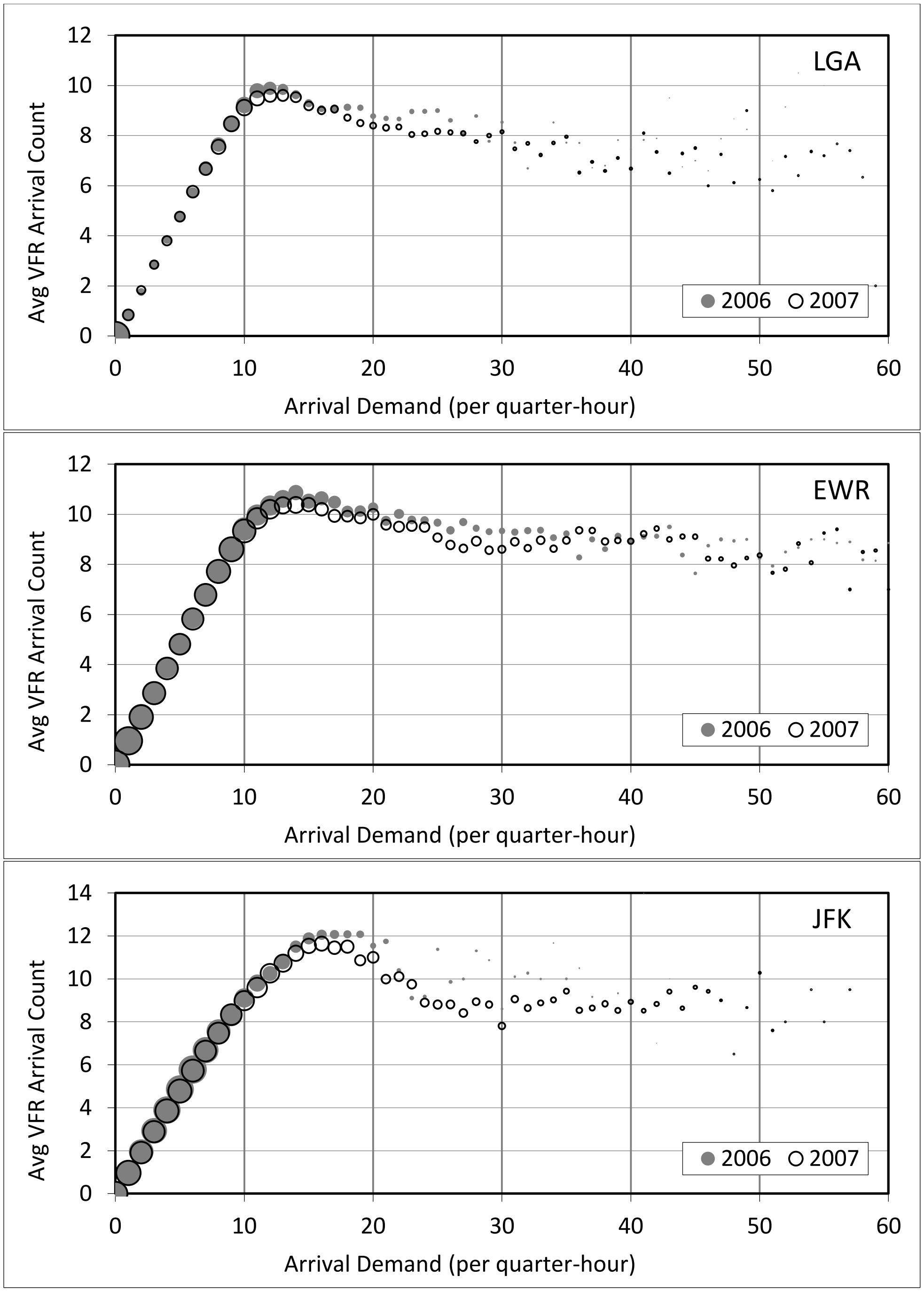
4. Figure 2



4. Figure 3

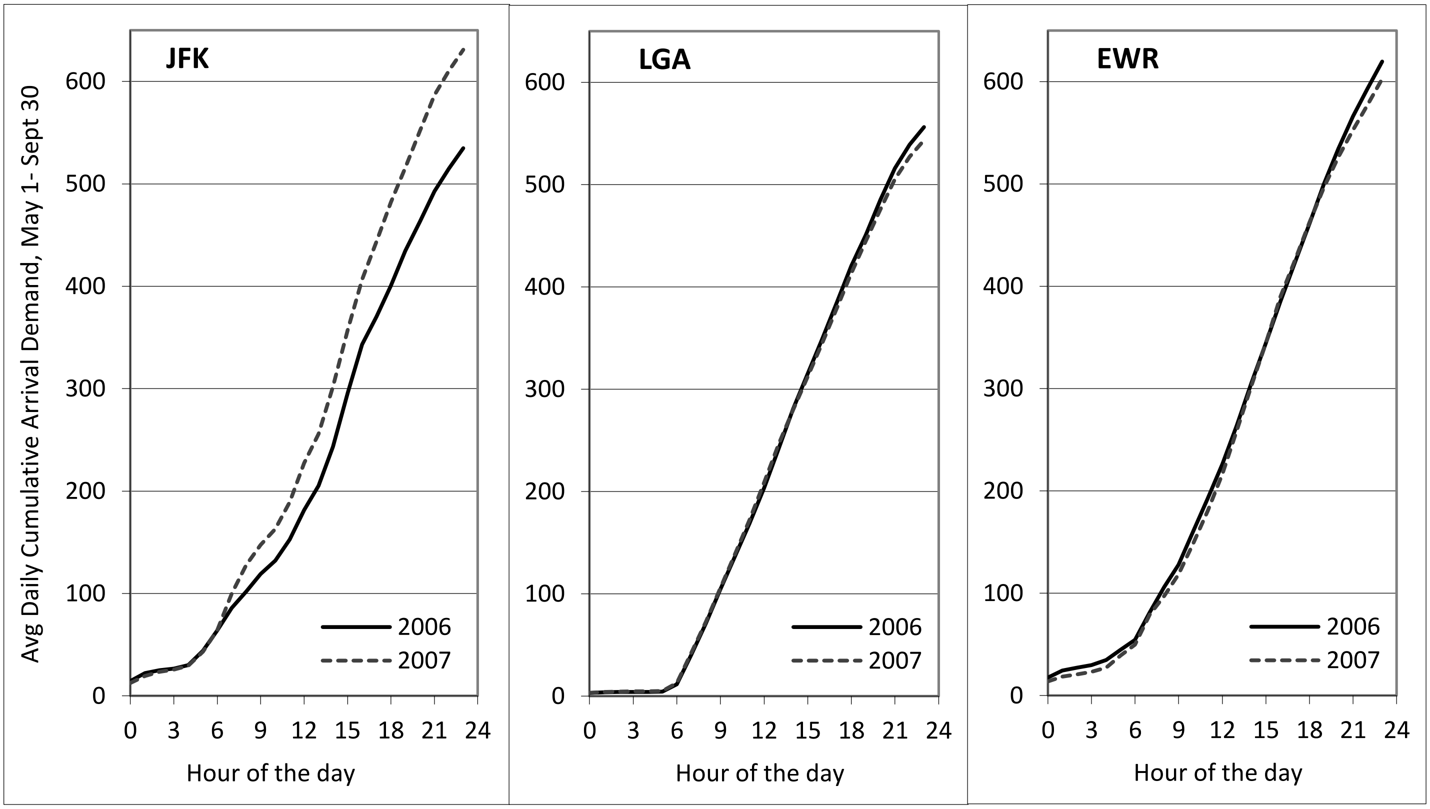


4. Figure 4

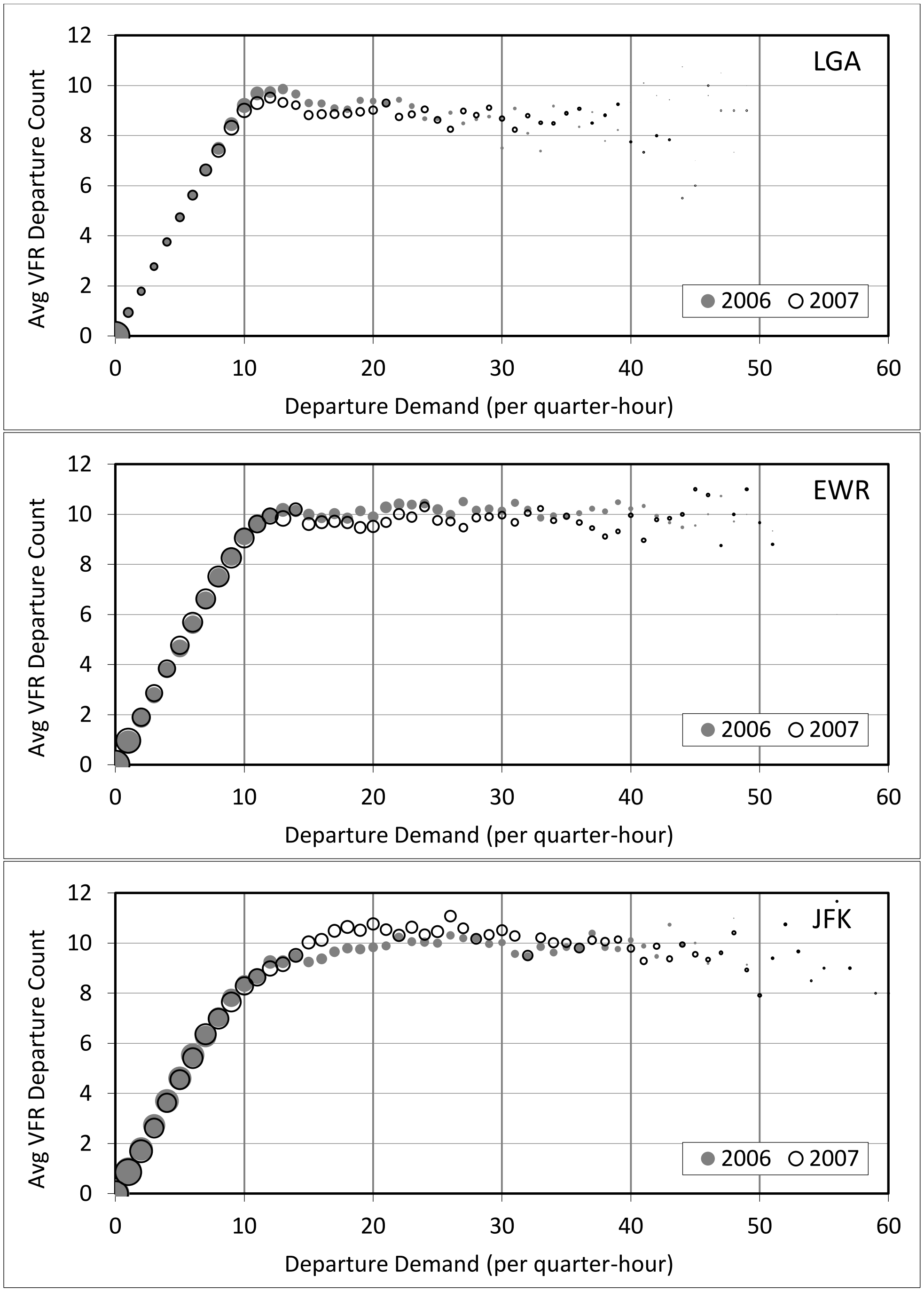


LGA
2006
2007
Avg VFR Arrival Count
Arrival Demand (per quarter-hour)
0
2
4
6
8
10
12
0
10
20
30
40
50
60
EWR
2006
2007
Avg VFR Arrival Count
Arrival Demand (per quarter-hour)
0
2
4
6
8
10
12
0
10
20
30
40
50
60
JFK
2006
2007
Avg VFR Arrival Count
Arrival Demand (per quarter-hour)
0
2
4
6
8
10
12
14
0
10
20
30
40
50
60

4. Figure 5


JFK
LGA
EWR
Avg Daily Cumulative Arrival Demand, May 1- Sept 30
0
100
200
300
400
500
600
0 3 6 9 12 15 18 21 24
Hour of the day
2006
2007

4. Figure 6



LGA
EWR
JFK
Avg VFR Departure Count
Departure Demand (per quarter-hour)
2006
2007
0
2
4
6
8
10
12
0
10
20
30
40
50
60