PERFORMANCE COMPARISONS BETWEEN US AND EUROPEAN AIRPORTS

Amedeo ODONI, Department of Aeronautics and Astronautics and Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (Room 33-219, MIT, Cambridge, MA 02139; arodoni@mit.edu)

Thomas MORISSET, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology (Room 35-217, MIT, Cambridge, MA 02139; morisset@mit.edu)

ABSTRACT

This paper compares scheduling practices, runway system capacity, air traffic delays, and flight schedule reliability at major airports in Europe and in the United States (US). The comparisons are based on the study of operations at the 34 busiest airports in the United States with those at the 34 busiest airports in Europe using extensive data from 2007 and 2008 flights. Major differences were found in several critical respects with important implications for aviation policy-makers on the two sides of the North Atlantic. In general, US airports achieve higher capacities, in terms of aircraft movements, than their European counterparts by using visual separation procedures, when weather permits, and by not placing slot constraints on the number of movements that can be scheduled at airports. European airports, on the other hand, limit air traffic delays and increase schedule predictability by using slot controls and by determining the number of available slots with reference to airport capacities under instrument meteorological conditions. A few exceptions to the above exist on both sides of the North Atlantic.

Keywords: airports, capacity, delay, schedule reliability, schedule predictability

INTRODUCTION

Europe and the United States are two areas of the world where air transport has developed extensively over the past decades and has established itself as the dominant mode of long-distance transportation. A set of major commercial airports, connected by means of some of the most advanced air traffic management (ATM) systems in the world, provide the infrastructure for the respective air transport networks.

In 2008, the Federal Aviation Administration (FAA), which operates the ATM system of the United States, and EUROCONTROL, the organization responsible for coordinating ATM system planning, development, and operations in the great majority of European countries, undertook a major study aimed at understanding the differences and similarities of the ATM and airport systems in the US and in Europe and at identifying, when possible, best practices. A joint report has been issued (Gulding et al 2009) that presents the first findings of this study. The report deals with comparisons of many aspects of system performance, including flow management, en route, terminal area, and taxiway operations. The work described in this paper supplements the broader joint study. It is focussed on major commercial airports and on the specific questions of how airside airport capacities, airport scheduling practices, and airport air traffic delays on the two sides compare. As will be seen, striking differences do exist in all these respects between airports in Europe and the US. At the root of these differences are the facts that (a) visual separation procedures are broadly utilized by the ATM system at US airports and (b) slot controls limit the number of operations that can be scheduled at the great majority of European airports, with the number of slots determined in most cases with reference to the capacity of these airports under instrument flight rules (IFR). The consequences of (a) and (b) are far-reaching. The resulting airport performance characteristics can also be seen as reflecting “cultural” differences regarding the principal operational objectives of the ATM and airport systems.

Section I provides the background for the paper, identifies some macroscopic characteristics of US and European airports and discusses (a) and (b) in some detail. Section II presents some airside capacity comparisons between airports on the two sides and discusses some of the consequences of prevailing practices. Section III deals with air traffic delays and schedule reliability and with some implications for the airlines. Finally, Section IV summarizes the main conclusions and briefly describes further ongoing work.

I. BACKGROUND

a. The airports

The study concentrated on the 34 commercial airports with the highest number of aircraft movements in 2007 in Europe[1] and their counterpart set of 34 US airports. These airports are listed in Tables 1 and 2, respectively, along with their traffic activity in 2007, which was the worst year in the history of aviation for airport delays (and the best year for the volume of air traffic). The US airports are also known as the "35 leading minus Honolulu" airports and are all in the continental United States.

Rank	City	IATA Code	Aircraft movements	Passengers	Declared capacity
1	Paris CDG	CDG	552,721	59,919,383	112
2	Frankfurt	FRA	492,569	54,161,856	83
3	Madrid	MAD	483,284	52,122,214	90
4	London Heathrow	LHR	481,356	68,068,554	89
5	Amsterdam	AMS	454,357	47,793,602	106
6	Munich	MUC	431,815	33,959,422	90
7	Barcelona	BCN	352,489	32,793,897	61
8	Rome	FCO	334,848	32,855,542	88
9	Vienna	VIE	280,915	18,768,468	66
10	Zurich	ZRH	268,537	20,686,986	68
11	Milan	MXP	267,825	23,885,305	69
12	London Gatwick	LGW	266,495	35,218,399	50
13	Brussels	BRU	264,366	17,838,689	74
14	Istanbul	IST	262,248	25,561,357	n/a
15	Copenhagen	CPH	257,591	21,356,134	83
16	Paris Orly	ORY	236,926	26,440,736	72
17	Dusseldorf	DUS	227,897	17,831,248	47
18	Oslo	OSL	226,232	19,044,011	n/a
19	Manchester	MAN	222,669	22,362,050	n/a
20	Stockholm	ARN	218,549	17,968,023	n/a
21	Dublin	DUB	211,803	23,289,417	46
22	London Stansted	STN	208,601	23,777,194	n/a
23	Athens	ATH	205,294	16,522,680	n/a
24	Palma de Mallorca	PMI	197,354	23,223,963	60
25	Nice	NCE	190,076	10,399,570	n/a
26	Geneva	GVA	190,008	10,807,060	n/a
27	Helsinki	HEL	184,052	12,956,754	80
28	Prague	PRG	174,662	12,478,078	44
29	Hamburg	HAM	173,513	12,780,504	53
30	Stuttgart	STR	167,264	10,321,431	42
31	Warsaw	WAW	153,476	9,268,476	n/a
32	Berlin Tegel	TXL	151,396	13,357,741	48
33	Cologne	CGN	151,020	10,471,657	52
34	Lisbon	LIS	144,797	13,392,131	36

Table 1 – 34 busiest European airports (aircraft movements, 2007) (Airport Council International, 2007)

The column "Declared capacity" in Table 1 indicates the number of slots available at each airport per hour. In the cases of several airports (for example LHR and FRA), the declared capacity varies slightly by time-of-day to take into account the changing mix of arrivals and

[1] "European airports" in the above includes only airports in Member States of EUROCONTROL. Because Russia is not a member of EUROCONTROL, Moscow's Domodedovo Airport is not included in Table 1.

departures in the schedule. For these two airports, Table 1 indicates the largest number of slots available at any hour during the day.

Rank	City	IATA Code	Aircraft movements	Passengers	Optimal capacity	IFR capacity
1	Atlanta	ATL	980,386	85,907,423	180-188	158-162
2	Chicago O'Hare	ORD	927,834	76,159,324	190-200	136-144
3	Dallas	DFW	684,779	59,784,876	270-279	186-193
4	Los Angeles	LAX	681,445	61,895,548	137-148	117-124
5	Denver	DEN	614,169	49,863,389	210-219	159-162
6	Las Vegas	LAS	609,472	47,595,140	102-113	70-70
7	Houston	IAH	603,836	42,978,617	120-143	108-112
8	Phoenix	PHX	538,063	42,197,080	128-150	108-118
9	Charlotte	CLT	522,541	33,383,812	130-131	102-110
10	Philadelphia	PHL	498,963	32,207,709	104-116	96-96
11	Detroit	DTW	467,230	36,126,555	184-189	136-145
12	Minneapolis - St Paul	MSP	450,337	35,160,505	114-120	112-114
13	Newark	EWR	443,952	36,391,911	84-92	61-66
14	New York JFK	JFK	443,004	47,810,630	75-87	64-67
15	Salt Lake City	SLC	414,395	22,029,488	130-131	110-113
16	Boston	BOS	399,537	28,088,855	123-131	90-93
17	New York La Guardia	LGA	389,492	24,940,818	78-85	69-74
18	Miami	MIA	386,981	33,740,416	116-121	92-96
19	Washington Dulles	IAD	382,907	24,494,999	135-135	105-113
20	San Francisco	SFO	379,500	35,793,117	105-110	68-72
21	Memphis	MEM	374,989	10,853,698	148-181	120-132
22	Orlando	MCO	359,101	36,385,300	144-164	104-117
23	Seattle	SEA	346,073	31,303,220	80-84	57-60
24	Cincinnati	CVG	320,449	15,734,322	120-125	102-120
25	Fort Lauderdale	FLL	307,975	22,681,903	60-62	52-56
26	Chicago Midway	MDW	304,657	19,378,546	64-65	61-64
27	Baltimore-Washington	BWI	296,870	21,497,555	106-120	60-71
28	Washington Reagan	DCA	275,433	18,670,924	72-87	48-70
29	Portland	PDX	264,518	14,654,222	116-120	77-80
30	Cleveland	CLE	259,471	11,447,011	80-80	64-64
31	Tampa	TPA	258,349	19,154,957	102-105	74-75
32	St Louis	STL	254,302	15,366,198	104-113	64-70
33	San Diego	SAN	228,902	18,326,761	56-58	48-50
34	Pittsburgh	PIT	209,303	9,821,980	152-160	119-150

Table 2 – 34 busiest American airports (aircraft movements, 2007) (Airport Council International, 2007)

In Table 2, the two rightmost columns indicate the estimated capacity of each airport (number of runway movements per hour) under good weather conditions and in instrument meteorological conditions. The source of these estimates is the FAA Airport Capacity Benchmark Report (2004).

b. VFR and IFR procedures

Although the ATM system, facilities, and equipment at most of the European airports in Table 1 is as advanced as at airports in the United States – and in some cases, more advanced – there is a major difference in the way air traffic management is conducted. The FAA, weather permitting, allows the use of visual flight rules (VFR) for airport operations in the US. Under VFR, pilots of landing aircraft are instructed by air traffic control to maintain visually a safe separation from the aircraft landing ahead of their own aircraft on the same runway and from other traffic in their immediate vicinity. When weather conditions do not permit the use of VFR, instrument flight rules (IFR) are in use, under which air traffic controllers are responsible for maintaining separations between aircraft. In contrast, IFR apply all the time[2] at European airports, independent of weather conditions. VFR results in distances between successive landing aircraft which are, on average, smaller than the prescribed separation standards that are imposed in IFR. The net result is that the capacities of US airports under VFR are always higher than the capacities under IFR.

Equally important, the separation requirements between aircraft operating on different runways may be relaxed when an airport is operating under VFR. For example, in good weather, simultaneous approaches (two aircraft landing side-by-side) are routinely conducted at SFO on parallel runways whose centerlines are only 740 feet (roughly 230 meters) apart – a procedure that requires that the approaches of the two aircraft be coordinated so that neither is affected by the wake vortex generated by the other. In IMC, only one of these runways can be used for landings at a time. As a result, the capacity of SFO for landings is 54 – 60 per hour in VMC and only 32 – 34 in IMC (almost one half, as we might expect when reducing two streams of landing aircraft to a single file).

c. Slot control

Flight scheduling at airports in Europe is “slot-controlled”. This means that the number of movements scheduled at an airport is limited by the so-called “declared capacity” of the airport which specifies the number of slots available per unit of time – typically per hour, although limits for other time units (10 minutes, 30 minutes) are occasionally used at airports with more elaborate slot controls. Airlines must acquire a slot for the right to land and depart at an airport at some particular time. The declared capacity is typically set by the responsible Slot Coordinator, typically in consultation with other entities such as the national ATM system operator, the airport operator, the airlines using the airport, air traffic controller representatives, etc. – for details see de Neufville and Odoni (2003) and Czerny et al (2008). For the European airports considered here, the procedures used to determine the declared capacity range from technically advanced – including the use of simulations and extensive consultation, as in the cases of LHR and FRA – to essentially *ad hoc* approaches with limited documentation. Although theoretically the limit on the number of available slots is the capacity of the most constraining element of the airport (which may be the terminal building or the number of aircraft stands in the case of some smaller airports), for major airports, such as those considered here, the capacity of the runway system is typically the one that determines the declared capacity. The allocation of the slots takes place at Slot Coordination Conferences organized by the International Air Transport Association (IATA) every six months. The Slot Coordinator assigns slots at each airport to airlines on the basis of a set of rules which vary somewhat according to the location of the airport. For example, airports in European Union nations must follow a set of rules promulgated by the European Commission.

---

[2] It is generally understood, however, that these requirements are relaxed, in practice, at a few of the busiest airports in Europe where air traffic control occasionally authorizes “experienced pilots” of some airlines that use the airport heavily to maintain visual separations on final approach in good weather.

In the United States, the concept of “declared capacity” is not used, in contrast to practically everywhere else in the world. The scheduling of flights at airports is not constrained: an airline may schedule a landing or takeoff at any time it wishes, as long as it can obtain access to a terminal building and aircraft stand. The exception is four airports, JFK, LGA, EWR and DCA, known – along with ORD – as the High Density Rule (HDR) airports, where the FAA has historically imposed limits on the number of movements that may be scheduled. This, however, is not, strictly speaking, slot coordination in the same sense it is applied at European airports. The High Density Rule expired in 2007, but the FAA has continued to impose scheduling limits at the three New York metropolitan area airports in order to mitigate congestion[3].

d. Data sources

While numerous data sources have been used in this study, by far the two most important ones were the ASPM database of the FAA and the CODA database of EUROCONTROL. The Aviation System Performance Metric (ASPM) database is comprehensive, as it combines data from a variety of sources. For example, for every scheduled flight at every airport with commercial service, ASPM provides a record of scheduled and actual times of arrival at the gate, departure from the gate, and take-off times, along with information about weather conditions at the time, runway configuration in use, etc. Passenger and cargo flights are included, as well as information on general aviation flights. Coverage is nearly complete, with typically only a very small percentage of flights missing. Access to ASPM can be authorized upon request from the FAA and its use is subject to certain mild restrictions.

The CODA database (EUROCONTROL, 2008) is less complete, but very useful nonetheless. CODA relies on the airlines to provide data on a voluntary basis, subject to certain confidentiality clauses. A majority, but not all, of the major European carriers are participating in the program. Low cost carriers, such as Ryanair, EasyJet, Wizz Air or GermanWings, generally do not. For the 34 European airports in our sample, 69% of all commercial flights are covered, with a range of a low of 31% of all flights at London Stansted to a high of 89% at Oslo Gardermoen. There are also a few notable gaps in the data. For example, no or little information exists about movements taking place in the late evening and night hours[4], especially after 9 pm, at many of the airports. Data coverage is generally better during daytime.

[3] In fact, beginning in mid-2008, the FAA has initiated an effort to apply slot coordination procedures similar to those of IATA at JFK and EWR, with slot limits of 81 movements per hour at each of these airports.
[4] This is because when too few flights are on the record, the confidentiality agreement between the airlines and CODA does not allow for the sharing of data, so that airlines’ identities can be protected.

II. CAPACITY COMPARISONS

a. Impact of the use of VFR procedures in the US

We turn next to an examination of some of the implications of the differences between European and US airports noted in Sections I.b and I.c, beginning with airport capacities. Inspection of the two rightmost columns of Table 2 suggests that the difference between the VFR ("Optimal") capacity and IFR capacity of many US airports is large. Moreover, airport weather information retrieved from the ASPM database indicates that VFR procedures may be used in the large majority of time at the 34 US airports of interest. These 34 airports experienced VFR weather conditions in 2007 for 83% of the hours of the year, on average. The airport with the lowest percent of VFR conditions was Seattle (64% in 2007), and Las Vegas the one with the highest (almost 100%). In general, airports on or near the East and West Coasts had a relatively lower incidence of VFR weather.

If one uses the mid-points of the capacity ranges given in the two rightmost columns of Table 2 as the proxy values for the Optimal and IFR capacities of the US airports (e.g., 184 for Optimal and 160 for IFR in the case of ATL), then, on average, the VFR (Optimal) capacities are 29% higher than the IFR capacities. This gain, as noted in Section 1.b is the result of the combination of higher capacities under VFR for individual runways and of the ability to use combinations of two or more runways more efficiently under VFR. Taking the weighted average of the Optimal and IFR capacities of each airport for local weather conditions, results in an average capacity gain which is 26% higher than the capacity that would be available at US airports had these airports been operating under IFR 100% of the time. Thus, the use of VFR procedures, weather permitting, results in very large capacity increases at US airports, with the precise amount varying from airport to airport, depending on the layout of the runway system and the local weather conditions. In general, gains are smaller at airports with one runway (e.g., SAN has an 18% gain) or with widely separated runways (e.g., CVG where runway separations permit the independent operation of individual runways even under IFR and the gain is only about 10%); and they are larger at airports with complex geometries and closely-spaced parallel runways.

b. Conjecture about the respective capacities of European and US airports

As noted already, European airports operate with Instrument Flight Rules in all weather conditions – at least officially. This means that the declared capacities of European airports are generally in line with the IFR capacities of these airports. Indeed, the declared capacities shown in the rightmost column of Table 1 are generally consistent with the IFR capacities that one would estimate for the corresponding airports using IFR separation standards[5]. In

[5] This has been confirmed through the use of the MACAD theoretical capacity model (Stamatopoulos et al 2003) to estimate capacities for some of the single-runway and close-parallel runway airports.

most cases, the declared capacities are either almost equal to or somewhat lower than these IFR capacities. There are however a few exceptions: LHR, LGW, FRA and MUC, where the declared capacities are a little *higher* than the computed IFR capacities. This is because at these airports, which are among the most highly utilized in Europe, the authorities setting the declared capacities take into consideration the fact that, in good weather conditions, the airports may achieve higher capacities than in IMC. In summary, it can be said that the declared capacities of European airports are largely determined by their IFR capacities, with some differences that may depend on the country in which the airport is located and how intensively the airport is being utilized.

Going one step further, the IFR capacities of US airports would be expected to be generally somewhat higher than those of European airports with similar runway layouts. The reason is that the FAA's IFR separation requirements (FAA, 2008) are similar to but, in some cases, smaller[6] than the requirements of the International Civil Aviation Organization (ICAO, 2005) which are widely used in Europe. It was also just shown in Section a above that the weighted capacity of US airports is significantly higher than the IFR capacity of these airports, because of the use of VFR procedures when weather permits. (A 26% average gain was indicated in **Section a.)** From this line of reasoning, the following conjecture can be posited:

"The weighted capacity of US airports can be expected to be significantly higher than the capacity of European airports with similar runway system layouts."[7]

c. A simple test of the conjecture

To test the above conjecture we have tried to identify airports in the US and in Europe with similar runway layouts. This has proven surprisingly difficult. The reason is that runway layouts tend to be much more complex in the US than in Europe: for our 34 top airports on each side, it turns out that the average number of runways per airport in the US is 4.12, while in Europe it is only 2.47[8]. Most of the 34 airports in Europe have relatively simple runway layouts with a single runway or with sets of two or more parallel runways with a single orientation. By contrast, the great majority of the 34 airports in the US have at least one crosswind runway, i.e., at least one runway pointing to a direction different from that of one or more other runways at the same airport.

For an initial test, we have examined three "families"[9] with distinctive runway layouts and identified airports in Europe and the US that belong to these families[10]; a few of them are

[6] For example, whereas the minimum IFR separation between two B737s (or similar narrow-body commercial jets) on final approach is 3 nautical miles according to ICAO's requirements, it is only 2.5 miles per the FAA at most major airports in the US.
[7] "Weighted capacity" in this instance refers to a long-term average capacity that takes into consideration the amounts of time that the airport operates under VFR or IFR.
[8] It should be emphasized that not all runways at an airport are active all the time. Because of their complex layouts, many US airports, especially, utilize only subsets of their runways, depending on weather conditions.
[9] We have actually identified 5 families and plan to do more extensive and detailed testing with additional airports.

shown in Figures 3 and 4. Table 5 lists the airports of this comparison: it provides the declared capacity of the European airports and the IFR, Optimal and Weighted capacities of the US airports.

Figure 3 – Runway layouts for airports of family A (Google, 2008)

Figure 4 – Runway layouts for airports of family E (Google, 2008)

	US Weighted and benchmark capacities				**Europe Declared capacities**	
Family	**Airport**	**Weighted**	**Optimal**	**IFR**	**Airport**	**DC**
A Single runway	San Diego	**55**	57	49	Gatwick[11] Dublin Berlin Tegel Stuttgart	**50** **46** **48** **42**
B 2 closely spaced, parallel runways	Seattle[12]	**76**	82	59	Düsseldorf	**47**
E 2 pairs of closely spaced parallel runways	Los Angeles	**137**	143	121	Paris	**112**

Table 5 – Comparison of US and European airport capacities for similar airports

It appears from Table 5 that:

[10] We do not claim that the layouts are identical, as the runway systems in each "family" may differ with respect to the length of the runways or the separations between parallel runways or other characteristics. However, the dominant characteristic (single runway, or close parallel runways, or independent pair of close parallel runways) is the same for airports belonging to each of the families.
[11] Gatwick, Dublin and Berlin Tegel all have a secondary runway, but it is only used when the main runway is unavailable, e.g., for repairs.
[12] A third runway opened in Seattle in 2008; the capacities shown in Table 5 apply to the two-runway airport.

(1) The IFR capacities of the US airports are reasonably close to the declared capacities of the European airports, with the possible exception of Seattle, whose IFR capacity is about 25% greater than Düsseldorf's declared capacity. Note that London Gatwick's declared capacity is greater than the IFR capacity of San Diego.
(2) The Optimal capacities of the US airports are much higher than the declared capacities of their European counterparts.
(3) Because VFR procedures are generally used for a very high fraction of the time at US airports, the overall weighted capacities of the US airports are also much higher than the declared capacities of their European counterparts.

Thus, this simple test confirms strongly the conjecture of the previous section, as well as the observations made earlier regarding the comparative magnitude of US IFR capacities and European declared capacities.

d. Implications for airport utilization and scheduling

These findings have important implications for airport utilization and scheduling practices on the two sides. In the case of European airports, there is little ambiguity: by definition, the declared capacities impose an upper limit on the number of movements that airlines may schedule during peak traffic periods or during any specified time interval. This limit, as we saw, is generally dictated by the IFR capacity of each airport – with the possible exception of such airports as LHR, FRA, LGW and MUC.

In the United States, on the other hand – and with the exception of the former High Density Rule airports – there is only a *perceived limit* on how many movements can be scheduled in peak hours. Due to the high percentage of time when VFR procedures are in use, this perceived limit tends to be associated with the VFR (Optimal) capacity of busy airports. Thus, at the busiest airports, airlines schedule movements with reference to this Optimal capacity, while implicitly recognizing that, when weather conditions are less than good, the airport's capacity will fall below demand and long delays will result. In fact, the airlines are under no legal obligation to adhere to any limit and the number of scheduled movements often exceeds even the Optimal capacity at the most popular airports, especially for short intervals of time, such as 15 or 30 minutes[13]. The use of Optimal capacity as the reference point for scheduling purposes is also suggested by practices at the five US airports (ORD, JFK, LGA, EWR, and DCA) where scheduling limits still exist. The scheduling limits set at these airports by the FAA are closer to the capacity that these airports can attain under VFR than to their IFR capacity. Thus, the Optimal capacity is treated even by the regulatory authority as the best guideline for how many movements should be scheduled.

The net result is that US airports are generally called upon to handle a much heavier volume of aircraft traffic[14] and are more heavily utilized on airside than European airports. In a sense,

[13] Ongoing research we are conducting shows that schedule frequently exceeds 23 movements per 15 minute intervals at Newark airport, where the optimal capacity is 84 to 92 per hour.
[14] Note that more aircraft traffic does not necessarily translate into more passenger traffic, as the latter also depends on the size of the aircraft utilized.

it can be stated that available airport capacity is utilized more efficiently in the US than in Europe. What these divergent practices and operating philosophies mean for the performance of European and US airports in terms of delays and schedule reliability will be examined in the next section.

III. DELAY COMPARISONS

a. Measures of level of service on airside

Section II suggests that the available airport infrastructure is, for the most part, utilized more efficiently in the United States than in Europe, in the sense that airports in the United States are able, on average, to serve more aircraft movements per unit of time than those in Europe. A central question, however, is whether the more efficient utilization of the infrastructure in the US also results in significant differences in the level of service provided to the airports' users. The most commonly used measures of airside level of service at airports are related to delays and other characteristics associated with delays, such as schedule predictability. We have therefore elected to focus the analysis in this section on such measures. To this purpose we have:

(a) computed for every flight arriving at one of the selected 34 airports in the US or in Europe during the year 2007, the difference between the actual time of arrival at the gate and the scheduled time at the gate; and

(b) used this information to examine a number of derivative measures of performance.

It should also be noted that, when calculating average delays in a particular hour, we compute the average of the individual delays (including negative values) across all the flights that were *scheduled* to arrive at the gate within this particular hour. For example, the average hourly delay at 3pm is the average of the delays suffered by all the flights whose scheduled (not actual) time of arrival was between 3:00:00 and 3:59:59.

b. Impact of VFR procedures on delays in the US

Before comparing US airports against European airports, we study the impact of using two different sets of air traffic control procedures, depending on weather conditions, at US airports alone. We compare arrival delays (relative to schedule) at US airports in VFR vs. IFR weather. Specifically all flights at the 34 US airports were separated into two categories:

- "IFR flights", defined as those flights whose scheduled arrival time fell during a period of Instrument Meteorological Conditions (IMC) at the destination airport[15]; and
- "VFR flights", all the other flights.

We chose to distinguish flights according to the weather conditions at their *scheduled* time of arrival rather than their *actual* time of arrival, because it is the conditions at the scheduled time of arrival that largely determine how much delay an aircraft will suffer. The actual time of

[15] Meteorological conditions (IMC or VMC) were determined on the basis of hourly weather data for each airport. The implicit assumption is that VFR procedures are used in VMC and IFR procedures in IMC. In reality, however, this may not be strictly true as air traffic may choose to utilize a single set of procedures over several hours, especially when the weather is variable, changing between IMC and VMC.

arrival is simply a consequence of that delay. In any case, had we classified flights according to their actual time of arrival the results would have been very similar.

Figure 6 shows that average delays, computed for all flights in 2007, are strikingly higher in IFR conditions than in VFR conditions. The average delay relative to schedule over the interval of time considered (7am – 10pm) is 9 minutes for VFR flights and 23 minutes for IFR flights, i.e., the impact of reduced capacity in IFR conditions is an increase of the average delay by 150% (while the average decrease in capacity is 29%, as estimated in Section II). This is a good example of the (well-known from queuing theory) non-linear relationship between capacity and delay, when demand is close to capacity.

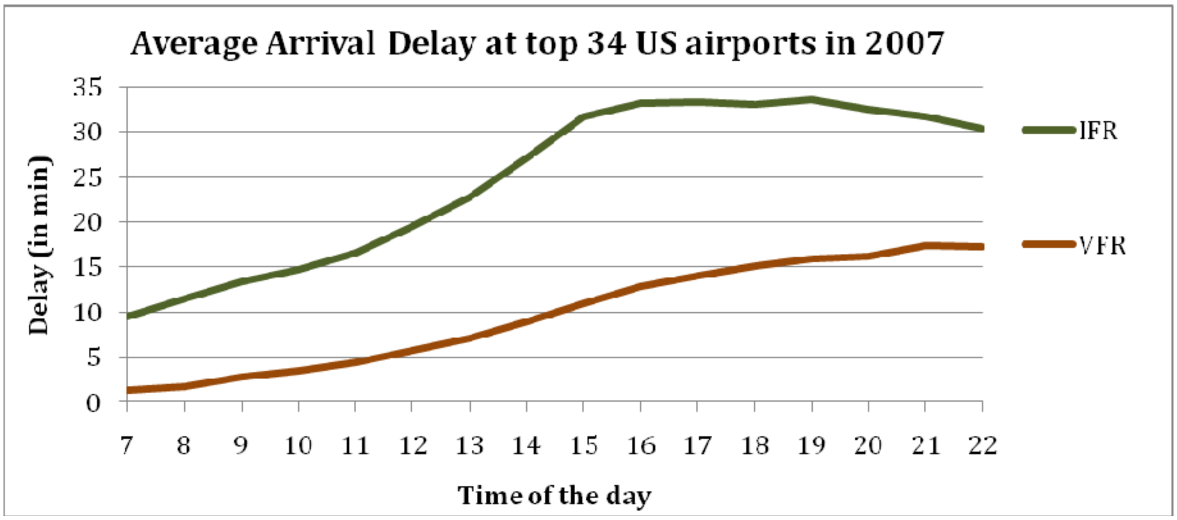

Figure 6 – Average arrival delay relative to schedule at 34 busiest US airports in 2007

Figure 6 can also be seen as confirmation of the observation that, when airlines schedule flights at US airports, they use the VFR capacity of the airports as their notional point of reference. Thus, when weather conditions do not permit VFR procedures and capacity falls short of this notional expectation, delays become very large. Note that in IFR conditions the average delay relative to schedule over all 34 airports (not just the most congested ones) was close to 35 minutes(!) from 3 pm to 9 pm local time in 2007 – a situation that many would consider unacceptable. The reader will recall that such conditions prevail for about 17% of the time (Section I.a) or, roughly, for one out of every six days.

Although weather conditions obviously vary at European airports[16] as well, the use of IFR procedures all the time means that they experience far less capacity variability than US airports. One can therefore speculate that the performance of European airports with respect to air traffic delay is less dependent on weather than at US airports. Unfortunately, the absence of relevant information in the CODA database makes it difficult to classify European

[16] For example, ongoing work using detailed data from Frankfurt Airport shows that conditions corresponding to the definition of IFR weather in the US prevail at FRA for about 12% of the time, in line with what is typical of US airports.

airport operations into "VFR" and "IFR" as was done for US airports in Figure 6 and thus to confirm or refute this conjecture[17].

c. Average delay comparisons

We next compare average delays by time-of-day for all weather conditions at the airports of interest. We have selected for this purpose the principal time window for airport operations, i.e., the period between 7 am and 10 pm (local time for each airport). The results are shown in Figure 7. For example, flights at the 34 US airports which were scheduled to arrive at their destination airport between 4 and 5 pm local time were, on average, about 15 minutes late in 2007.

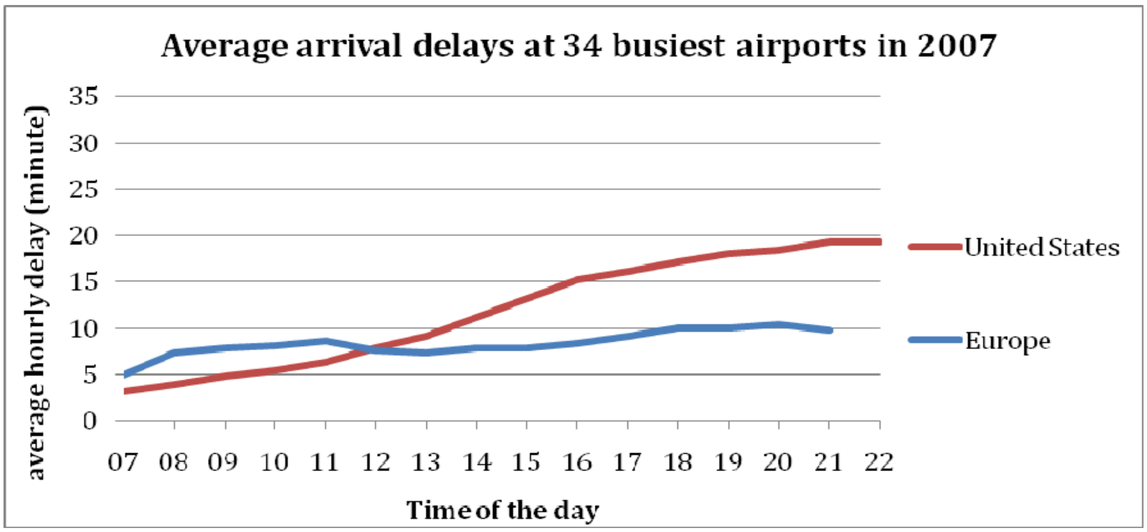

Figure 7 – Average hourly daily for 34 busiest American and European airports

As noted in Section I.d, European data for the time from 8 pm on are sparse and possibly non-representative, as they are unevenly distributed among airports and airlines. We shall therefore limit any comparisons to the time until 8 pm.

Figure 7 shows two different patterns for the evolution of average delays over a day in the US and in Europe. For the former, delays increase steadily through the course of the day, until they reach their maximum level at about 9 pm – declining subsequently during the late night and early morning hours. For Europe, by contrast, average delay relative to schedule remains remarkably constant during the greater part of the day and, after 3 pm, at about one-half the level of delay at US airports. One can conclude that the scheduling of flights at European airports is performed at a more sustainable level than at US airports, resulting in delays that are significantly more predictable and reasonable. Slot-control policies in Europe are clearly instrumental in preventing the build-up of queues as the "typical day" progresses.

It is also worth noting that the absence of data for low-cost carriers in Europe (Section I.d) should not affect the validity of these conclusions. The major low-cost carriers place

[17] We are in the process of testing the conjecture through the analysis of the detailed data from FRA.

particular emphasis on on-time performance: they do not offer tickets that involve connecting flights and will not typically delay a flight to accommodate late passengers (e.g., RyanAir requires that passengers be at their gate at least 30 minutes prior to departure time to allow adequate time to load the aircraft). Low-cost carriers also avoid as much as possible the most congested European airports. It is therefore reasonable to expect that the on-time performance of the flights that are missing from the CODA database is on average better than that of the typical European flight, a hypothesis also supported by a review of the flightstats.com database.

d. Delay distribution analysis

Schedule reliability at an airport can be thought of as the extent to which the actual time of arrival or departure of a flight adheres to the scheduled time. Schedule reliability captures a different dimension of airport performance than average delay. Low schedule reliability, i.e., a large amount of uncertainty about actual flight times, has a significant negative impact on passengers, when it comes to the amount of time they spend at airports and to their ability to plan for and make their flight connections on multi-segment flights. Ultimately, this also has a negative effect on airline profitability.

We have quantified schedule reliability at our subject airports by examining the probability distribution of flight delays and the standard deviations of these delays. For every airport and for every hour of the day for which adequate data were available, we have computed the distribution of the arrival delay of flights, relative to scheduled time at the gate, for the entire year 2007. To this purpose we have calculated the fraction of aircraft that arrived between 0 and 1 minute later than scheduled, 1 and 2 minutes, and so on, for every 1-minute interval, in the range from 60 minutes ahead of schedule (“negative delay”) to 180 minutes later than the scheduled time. Figure 8 presents an example of these distributions for three different hours of the day for EWR and FRA. These distributions give a good indication of the reliability of the schedule at these two airports: if they are very “concentrated” (typically around the value of 0, but possibly around another positive or negative value) then the schedule is strongly adhered to and flights arrive near their “scheduled time plus the average delay”. On the other hand, if there is a large dispersion around the expected value – and the standard deviation of the delay is large – the flight schedule cannot be relied on, and the delays vary within a wide range of values.

In the case of Figure 8, one can immediately observe a striking difference between FRA and EWR. In the case of the former, both the average delay and the standard deviation of the delay remain relatively stable between 8 am and 5 pm; while for EWR, we observe a rapidly deteriorating situation as the day progresses. This example is typical of what we observed at other European and American airports. Specifically, our analysis of the delay distributions by time-of-day leads to the following quasi-general observations:

- in the US, over the course of the day, not only does the average delay increase steadily, but also the distribution of the delay becomes increasingly dispersed – approaching a nearly “flat” shape – suggesting a low reliability of schedule;

- in Europe, even at the busiest airports, the average delay remains relatively constant, and the distribution of the delay reasonably concentrated over the course of the day, suggesting far greater reliability of the schedule. This is true even of London Heathrow, the airport experiencing the highest delays in Europe.

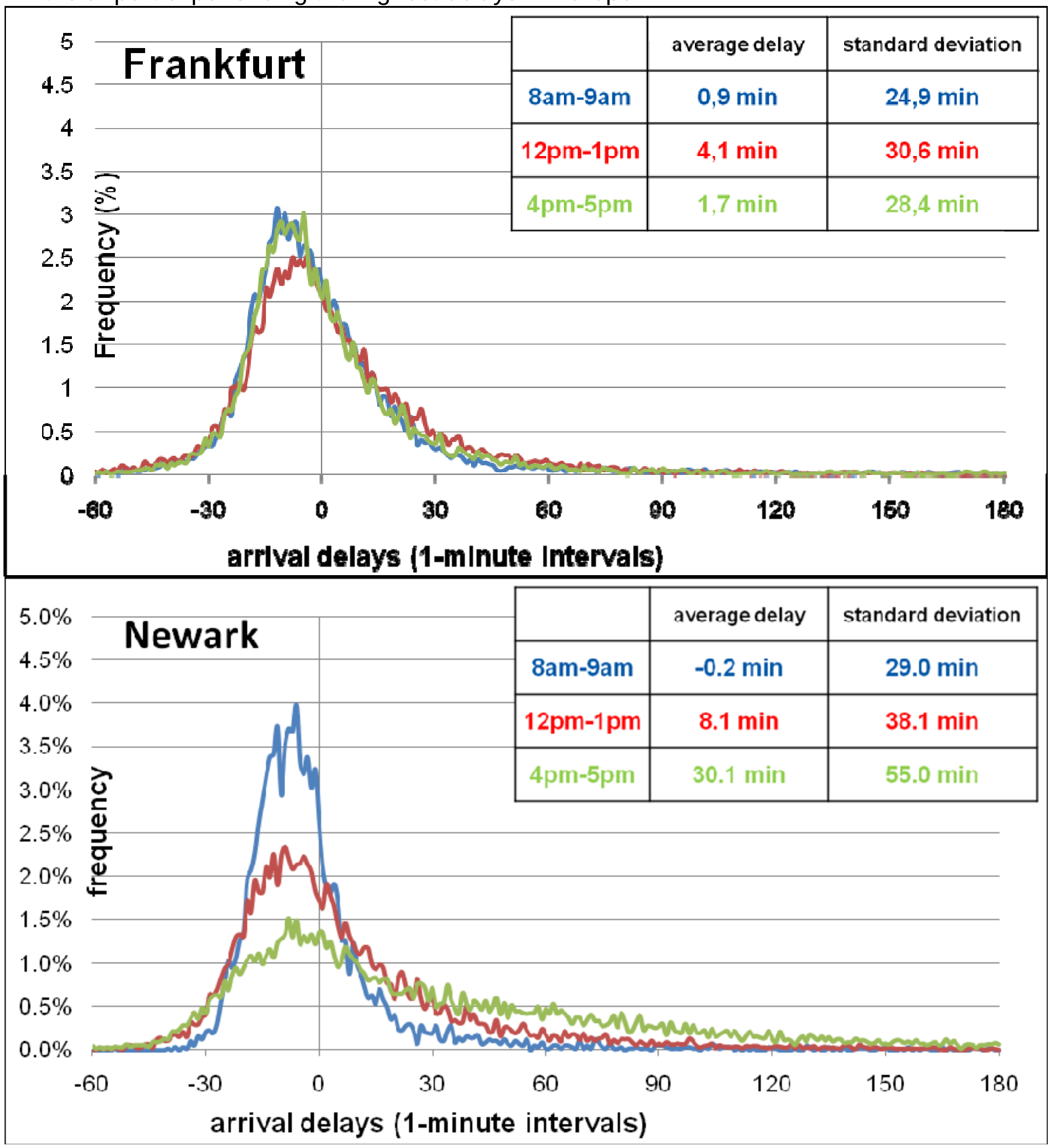

Figure 8 – Average delay distribution curves for FRA v. EWR (8am, 12pm, 4pm)

These observations support the hypothesis that a considerable number of US airports may be "over-scheduled", notably those of the New York area and some of the most important hub airports. Moreover, they suggest that the airport system as a whole cannot sustain the current[18] overall schedule of flights, building up delays and losing scheduling reliability over the course of an average day. In Europe, schedule reliability remains virtually constant over the course of the day supporting the hypothesis that most airports operate at demand levels that are sustainable vis-a-vis their capacities. This is corroborated by the fact that average hourly delay is roughly the same from the morning hours to late evening at these airports.

e. Padding of block times

Schedule padding refers to the practice whereby airlines increase the scheduled gate-to-gate duration, or “block times”, of flights in order to improve on-time performance. This practice is a natural consequence of the increase in flight delays and the loss of schedule reliability noted above. We look briefly at schedule padding in the US and Europe, as it provides additional evidence of differences in performance between the airport systems of the two sides with respect to delays and schedule reliability.

The recent FAA/EUROCONTROL report (Enaud, Gulding, et al 2009) estimated that between 2000 and 2007, the average block time of intra-European flights remained constant, fluctuating only slightly according to season to account for winds and seasonal traffic levels. By contrast, it was found that the average block time for a domestic flight in the United States increased by about 3 minutes during the same period, with larger seasonal fluctuations. Equally important, an MIT study (El Alj, 2003) found that, between 1993 and 2000, the average block time for a large sample of flights increased by an average of 7 minutes. This indicates that block times in the US have increased by a total of roughly 10 minutes between 1993 and 2007. The average domestic flight in the US is about 110 minutes long, therefore “padding” accounts for close to 10% of the average block time! Stated differently, to improve schedule reliability in the face of long and highly variable airport delays, US airlines have been forced to increase certain very important planned flight costs (such as the amount of crew time and of aircraft time allocated to a typical flight) by about 10% over a period of 15 years.

Schedule padding is attracting considerable public interest in the US (McCartney 2010) as airline passengers are increasingly noticing that they frequently arrive ahead of schedule, i.e., that airlines systematically allocate considerably more time to flights than would be needed in the absence of congestion. Our data analysis provides relevant evidence. In Figure 8, for example, a significant fraction (in fact, a majority) of all flights at both EWR and FRA. However, EWR has a higher proportion of very early flights (e.g., more than 30 minutes ahead of schedule) than FRA at all the times of the day shown, despite the fact that average delay is much higher at EWR than at FRA at noon and 4pm. A major reason for this is the larger amount of padding typically assigned to US flights, particularly those to and from New York area airports. Moreover, we also observe an increased percentage of very early flights at EWR in the later hours of the day, as the average delay increases and the schedule gets

[18] The number of flights scheduled in 2009 has declined by a total of about 8% since its 2007 peak, but this may be only temporary relief.

less predictable. This does not happen at FRA. It would seem that schedule padding intensifies in the later parts of the day at EWR.

Schedule padding may distort the apparent on-time performance of airports: the more padding airlines use, the smaller the average delay relative to schedule will seem to be. However, the distribution of delays around the average (and the associated) standard deviations will be less sensitive to schedule padding. This is another good reason why analyses of the type shown in Figure 8 can be informative.

IV. CONCLUSIONS

The US and European networks of major airports and the respective ATM systems utilize about equally advanced technologies, as well as facilities and equipment of similar quality. Important differences exist, however, in terms of operating procedures and scheduling practices that may ultimately reflect different philosophical approaches to the management of air transport's infrastructure. As a result, the performance of major commercial airports on the two sides differs in important ways, with neither system outperforming the other in every respect. Our summary finding is that the comparison between the US and European systems of major airports illustrates a case of a classical trade-off between intensive utilization and high throughput in the use of the available infrastructure on the US side, and predictability and schedule reliability on the European.

Efficient utilization of airport infrastructure

The principal conclusion from the comparisons of airport capacity is that, due to its use of VFR procedures, weather permitting, US airports achieve much higher average capacities than European airports with similar runway layouts. The use of VFR procedures at US airports was possible 83% of the time in 2007. This led to an average increase of 26% in airport capacity over the capacity that would have been achieved if IFR procedures were used all the time, as in Europe. Being able to operate such a large number of additional movements gives airlines in the US greater flexibility in selecting aircraft types and times when flights operate and contributes to an air transportation environment which is more open to competition with large volumes of traffic and attendant economic benefits.

Furthermore, it was observed that, although IFR procedures and separation requirements are comparable between the US and most European countries, IFR capacities at US airports are generally higher than the declared capacities of European airports with similar runway layouts. At the same time, it should be noted that some of the busiest European airports declare and achieve high capacities compared to what might be expected given their runway layouts. UK NATS has some of the best practices in the world, in this respect, at LHR and LGW and has managed to increase gradually the available capacities there. Declared capacities have also been increasing gradually at several other major European airports in recent years, as growing demand has required higher throughput rates.

Level of service: delays and schedule reliability

When it comes to level of service, the use of VFR procedures at US airports may lead to over-scheduling of flights. This, in turn, creates a situation in which performance deteriorates sharply when weather conditions are less than good. In 2007, flights arrived on average 9 minutes behind schedule in VFR weather and 23 minutes in IFR weather. The heavy reliance of US airports on the use of VFR procedures thus makes them more vulnerable to excessive delays (and increased numbers of flight cancellations). By using IFR procedures all the time, the performance of European airports is less variable with weather conditions.

It was also shown that, even in VFR weather, delays increase steadily over the course of the day in the US, from 7am until 9pm. This suggests that flight schedules may have reached an unsustainable level in 2007, not only at individual airports, but also system-wide. In contrast, European airports manage to maintain a roughly constant level of delay for most of the day, meaning that the use of slot controls is effective in protecting airports from excessive demand and sharp deterioration of schedule reliability. This hypothesis is corroborated by the study of delay distribution. In the US, uncertainty about arrival times increases throughout the course of a typical day. European airports are characterized by much more robust behaviour: even when delays increase, the distribution of delays around their average value tends to remain “concentrated”, resulting in more reliable schedules.

Our study of delay performance therefore suggests that declared capacities, if determined carefully, may make it possible for airports to maintain demand at reasonable levels (and thus provide a satisfactory level of service under most weather conditions) while, at the same time, achieving reasonably high levels of infrastructure utilization. However, another school of thought may argue that European slot control policies may be hampering the quest for additional airport capacity by creating the false impression that all existing demand is being served at a reasonable level of service.

Further research

As much as our research has tried to perform fair comparisons between US and European airports, its macroscopic nature and the limitations of the data on the European side make it difficult to arrive at more definitive quantitative conclusions. There is much we do not know about local operating conditions at individual airports. The delays database we used for the European airports also had significant gaps as noted in Section I.d. Additional ongoing research is trying to address some of these deficiencies by looking at two specific airports, EWR and FRA, in much greater depth, taking advantage of a highly detailed set of data concerning Frankfurt Airport. The dependence of performance on weather conditions is an issue of particular interest in this ongoing work, as is the utilization and throughput rates achieved with various runway configurations.

A more detailed account of our research can be found in the Master’s thesis at MIT of one of the authors (Morisset, 2010). For example, in the case of airport capacity, we have performed comparisons involving additional “families” of runway layouts, such as those with two independent parallel runways. A closer look was also taken at the impact on capacities of different separation requirements in IFR. With respect to delays, we have looked at the delay increases over the course of the day at each of the individual airports in the US and Europe and analyzed them as additional evidence of over-scheduling in the US.

BIBLIOGRAPHY

ACI – Airports Council International. (2007). *Worldwide Airport Traffic Statistics.* ACI Statistics, Airport Council International.

Civil Aviation Authority. (2009). *CAP 493 Manual of Air Traffic Services Part 1.* London: CAA.

Czerny, A., P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) (2008). *Airport Slots: International Experiences and Options for Reform.* Ashgate Publishing, Hampshire UK.

de Neufville, R. and A. Odoni (2003). *Airport Systems : Planning, Design, and Management.* McGraw-Hill, New York.

El Alj, Yasmine (2003). *Estimating the true extent of air traffic delays*. MIT libraries, MIT press, Cambridge.

EUROCONTROL. (2008). *CODA OneSky online*. Retrieved from https://extranet.eurocontrol.int/http://onesky2.eurocontrol.int/amserver/UI/Login?gx_charset=UTF-8

Federal Aviation Administration. (2008). *Aviation System Performance Metric home page*. Retrieved from http://aspm.faa.gov

Federal Aviation Administration. (2008). *Order JO 7110.65S Air Traffic Control.* Washington: Federal Aviation Administration.

Federal Aviation Authority. (2004). *Airport Capacity Benchmark Report.* Washington: Federal Aviation Authority.

Google. (2008, September 10). *maps.google.com.* Retrieved from www.google.com.

Gulding, J., Knorr, D., Rose, M., Bonn, J., Enaud, P., & Hegendoerfer, H. (2009). US/ Europe comparison of ATM-related operational performance. *Eighth USA/Europe Air Traffic Management Research and Development*, (p. 2 & p.4). Napa.

International Civil Aviation Organization. (2005). *Rules of the air and air traffic services (DOC 4444).* Montreal: ICAO.

Larson, R.C. and A. Odoni. (2007). *Urban Operation Research.* Dynamic Ideas, Belmont, MA.

McCartney, Scott. (2010). *Why a Six-Hour Flight Now Takes Seven*. The Wall Street Journal, February 04, 2010, New York.

Morisset, Thomas. (2010). *Comparing Capacities and Delays at Major European and American Airports*. MIT libraries, MIT press, Cambridge.

Performance Review Unit. (2008). *Performance Review Report.* Brussels: Eurocontrol.

Stamatopoulos, M.A., K.G. Zografos, and A. Odoni (2004). “A Decision Support System for Airport Strategic Planning”, *Transportation Research C*, *12*, pp. 91-118.