Transportation Research Part A xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

A roadmap toward airport demand and capacity management

Alexandre Jacquillat[a,*], Amedeo R. Odoni[b,c]

[a] *Carnegie Mellon University, Heinz College, 5000 Forbes Avenue, Building HBH 2118J, Pittsburgh, PA 15213, USA*
[b] *Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, 77 Massachusetts Avenue, Building 33-324, Cambridge, MA 02139, USA*
[c] *Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, 77 Massachusetts Avenue, Building 33-324, Cambridge, MA 02139, USA*

ARTICLE INFO

Keywords:
Airport management
Capacity planning
Airport operations
Demand management

ABSTRACT

This paper synthesizes the major interventions available to manage airport demand and capacity, the analytical tools that may support the underlying policy, managerial and operational decisions, and guidelines for policy and practice obtained from recent research. The resulting insights fall into three broad categories. First, airport throughput exhibits significant variability, and airport capacity depends on the available infrastructure and operating procedures. Second, airport on-time performance is highly non-linear, and thus sensitive to variations in demand and capacity. Third, airport demand management involves a trade-off between mitigating congestion and maximizing capacity utilization, and scheduling mechanisms can support and enhance existing practices. The implications for the development and management of airport systems worldwide are discussed.

1. Introduction

1.1. Problem of demand and capacity management

Airports play a central role in urban development by connecting individuals, businesses and governments, and spurring indirect commercial activities. Over the past decades, airports have accommodated increasing numbers of operations to support regional and national growth and airline business development. Despite declines following 9/11 and during the economic downturn in 2008 and 2009, air traffic has grown significantly in the United States and Europe, and even more rapidly in Asia and Oceania and, more recently, in Africa and Latin America. At the same time, airport throughput is limited by the existing infrastructure and operational capabilities. At many of the world's busiest airports—and despite a number of capacity expansion projects (e.g., the construction of new runways)—demand for airport access has grown to often exceed airport capacity in many metropolitan areas. The impact of the resulting imbalances between demand and capacity depends on access policies (see Section 1.2). At airports with largely unconstrained access (e.g., the overwhelming majority of US airports), the result can be over-capacity scheduling and delays, with significant congestion costs—for instance, the nationwide impact of flight delays in the United States was estimated at over $30 billion in 2007 (Ball et al., 2010). At airports where access is restricted (e.g., most of the busiest European airports), the restrictions can result in demand losses and/or demand displacement (e.g., to less preferred times of the day or to other airports). At the opposite end, airline demand at less busy airports may fall well below available capacity, resulting in the under-utilization of infrastructure resources. For airports under development, this underscores the need for proactive management of demand and

* Corresponding author.
E-mail addresses: ajacquil@andrew.cmu.edu (A. Jacquillat), arodoni@mit.edu (A.R. Odoni).

http://dx.doi.org/10.1016/j.tra.2017.09.027
Received 27 April 2016; Accepted 26 September 2017

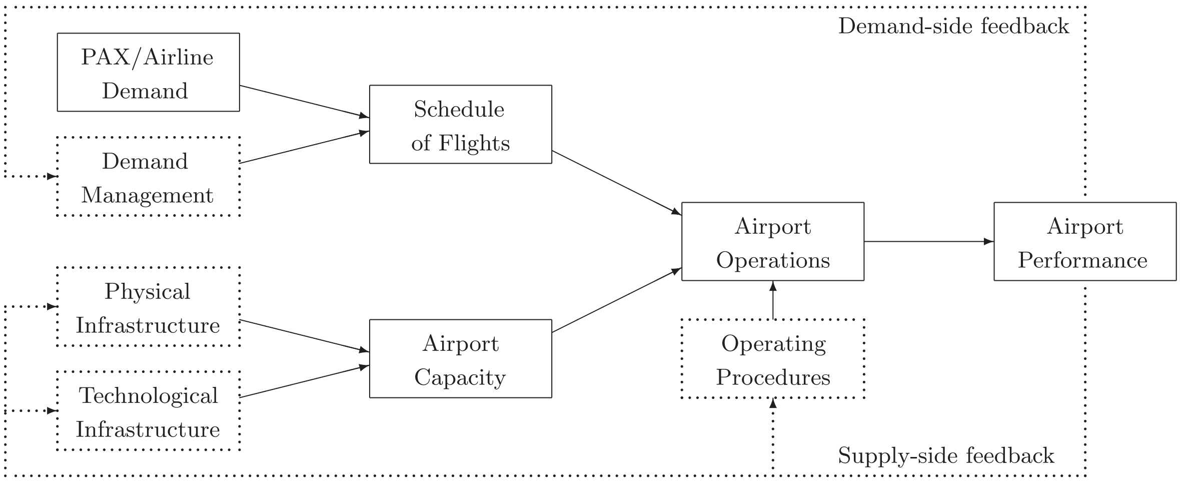


Fig. 1. Schematic representation of airport planning, management and operations (decisions are indicated in dashed lines).

capacity, ranging from long-term infrastructure planning to medium-term infrastructure management and short-term infrastructure operations.

This paper synthesizes the major operational, managerial and policy mechanisms available to manage airport demand and capacity, the analytical tools supporting the underlying decisions, and the implications for the development and management of airports. As indicated by Keeney (1973) in the context of the (existing) Mexico City airport, this requires consideration of multiple criteria, such as ensuring safety, maximizing throughput, minimizing capital expenditures and operations costs, promoting airline competition, mitigating air traffic congestion, and promoting environmental sustainability. These objectives can be aligned (e.g., mitigating congestion has a positive environmental impact), but may also give rise to some trade-offs (e.g., increasing throughput generally requires significant investments). Moreover, airport operations affect a number of stakeholders, including Civil Aviation Authorities, airport operators, commercial airlines and other aircraft operators, passengers, local communities, etc. Therefore, the management of airport demand and capacity creates complex decision-making challenges that require the identification of airport performance objectives and stakeholder incentives over the course of airport lifecycles (Neufville et al., 2013; Zografos et al., 2013).

Airport performance depends on three primary factors: (i) airport capacity, (ii) operations management, and (iii) flight scheduling. Demand and capacity management interventions can thus be classified into three categories. First, infrastructure expansion aims to increase potential capacity through the development of greenfield airports or the expansion of existing airports. Second, operational enhancements aim to improve the efficiency, reliability and sustainability of airport operations, given the available physical and technological infrastructure and fully complying with safety constraints. Third, demand management aims to modify the temporal and/or spatial characteristics of demand through access regulation that controls the number of peak-hour flights scheduled at the busiest airports, or through incentives to spur demand at off-peak hours or at underserved airports. Although interdependent, these decisions are typically made in sequence: airports first plan their capacity based on demand forecasts, then optimize air traffic handling procedures to minimize operating costs, and, last, may need to implement demand management schemes if capacity lies well below airline demand. Fig. 1 provides a schematic representation of these decisions and their impact on airport demand, capacity and, ultimately, performance.

The design and optimization of demand and capacity management interventions requires contributions from the fields of transportation economics, management, and operations. From an economic standpoint, an efficient scheme allocates scarce airport capacity to the users that assign the highest value to it, through quantity-based or price-based mechanisms. From a managerial standpoint, decision-making tools support the planning of airport capacity (a supply-side intervention) and the scheduling of flights (a demand-side intervention). From an operational standpoint, the characterization of airport capabilities makes it possible to predict and, where possible, improve performance. For comprehensive reviews of the underlying modeling advances in these different fields, we refer the reader to (Czerny et al., 2008; Zhang and Czerny, 2012; Zografos et al., 2013, 2016; Gillen et al., 2016). While the economic aspects of the problem have been the subjects of extensive research (some of which is reported in this volume), this paper focuses on the management and operations problems. Specifically, it describes a bottom-up approach that begins with the characterization of airport capacity, operating capabilities and demand patterns, and provides decision-making support for enhancing airport performance through infrastructure expansion, operational enhancements and demand management.

1.2. Overview of current practices

We briefly review the major differences observed worldwide in airport infrastructure, operating procedures and flight scheduling. The range of these practices provides guidelines to address demand and capacity management trade-offs.

First, the physical infrastructure of busy airports varies significantly by size and physical layout. For instance, most European airports exhibit simple layouts, with a single runway or 2–3 parallel runways. In contrast, many US airports have more than three runways, with at least one intersecting runway.[1] The largest airports, such as Dallas/Fort Worth (DFW) and Chicago O'Hare (ORD), have up to 7 and 8 runways, respectively. Other multi-runway airports are currently being constructed or planned in Asia-Pacific, the Middle East and Latin America. Equally important, the technological infrastructure supporting air traffic management systems is being upgraded in several parts of the world. The Next Generation Air Transportation System (NextGen) in the United States and the Single European Sky ATM Research (SESAR) system in Europe aim to supplement, and eventually replace, current radar-based surveillance systems with new satellite-based aircraft technologies to better monitor and manage aircraft operations. In summary, the physical and technological infrastructure of airports impacts strongly their long-term operational capabilities.

From an operating standpoint, air traffic control procedures also exhibit significant differences worldwide. At most airports outside the United States, Instrument Flight Rules (IFR) separations between landing and/or departing aircraft are maintained regardless of weather conditions. In contrast, at US airports (and, recently, at a few airports elsewhere), Visual Flight Rules (VFR) are used under Visual Meteorological Conditions (VMC), i.e., pilots are often requested, weather permitting, to maintain visually a safe separation from preceding aircraft during their final approach to the runway. This practice results in lower separations between consecutive aircraft and more efficient use of multiple runways than with strict adherence to IFR.[2] In addition, operating performance also depends on air traffic flow management procedures, which manage the flow of aircraft in a network of airports through such interventions as ground holding, re-routing, speed control and airborne holding.

On the demand side, the dynamics of flight scheduling also differ from one jurisdiction to another. When unconstrained airline demand exceeds available capacity, busy airports outside the United States are subject to strict schedule coordination (or *slot control*). Under this mechanism, each airport provides a value of its *declared capacity*, and allocates a corresponding number of slots to the airlines through an administrative procedure based on the guidelines from the International Air Transport Association (2015), including *grandfathered rights* and *use-it-or-lose-it rules*. In contrast, scheduling constraints are much weaker in the United States. At the overwhelming majority of US airports, the number of flights that may be scheduled per time period is not capped by any pre-specified limits. Historically, 4–5 airports were subject to access restrictions under the High Density Rule (HDR), but, in 2000, the Wendell H. Ford Aviation Investment and Reform Act for the 21st Century, programmed its phase-out, effective in 2007. Due to the unsustainable delays that the resulting unrestricted access led to, the Federal Aviation Administration (FAA) has imposed at New York's John F. Kennedy (JFK), Newark (EWR) and LaGuardia (LGA) airports "flight caps" that limit the number of hourly movements (landings and take-offs) to 81, 81 and 75, respectively.[3] However, these caps are set at higher numbers of movements and are less strictly enforced than at comparable airports outside the United States.

In summary, important differences exist among busy airports around the world in three major respects. First, the number of runways, the main driver of eventual capacity, can range from a single runway to as many as 7 or 8. Second, air traffic control policies and procedures impact greatly the throughput that can be achieved during peak periods by dictating the separations between consecutive movements on the same runway, as well as how simultaneously active runways interact with one another. Third, demand management practices range from largely unconstrained access at most airports in the United States to strict schedule coordination and slot controls at many of the busiest ones elsewhere, with consequent strong impacts on flight scheduling patterns.

1.3. Objectives and outline

This paper provides a comprehensive perspective on airport demand and capacity management. It describes the impact of airport infrastructure, airport operations and flight schedules on airport performance. It identifies briefly the state-of-the-art models supporting airport demand and capacity management decisions, and provides guidelines, based on recent research, concerning methods for enhancing airport performance. For existing airports, it suggests ways to improve current practices, depending on the stage of their lifecycle and their engineering and/or institutional legacy. For greenfield airports under development, it describes a holistic approach to capacity planning, management and utilization.

The structure of the paper follows the three steps of demand and capacity management. Section 2 focuses on the characterization and estimation of airport capacity and identifies its main drivers. Section 3 presents models to quantify on-time performance as a function of demand and capacity, and highlights the strong non-linearities in on-time performance at airports operating close to capacity. Section 4 introduces scheduling models to support, and potentially improve, mechanisms for demand management at airports where demand would otherwise exceed capacity by a large margin during significant periods of time in a day of operations. Results and cited examples offer insights for enhancing the flexibility of the schedule coordination schemes outside the United States, and for introducing limited scheduling adjustments at the busiest US airports. Each of these sections begins by describing relevant airport practices, and then presents the available analytical tools, the main insights derived from recent research, and the implications for policy and practice. Last, Section 5 summarizes the insights from this paper and provides a roadmap toward the management of demand and capacity at busy airports worldwide.

[1] The average airport among the 34 busiest European airports has between two and three runways, while the comparable average US airport has over four runways (Morisset and Odoni, 2011).

[2] For instance, under VFR, simultaneous parallel landings are performed in VMC on pairs of parallel runways whose centerlines may be separated by as little as 800 feet (~244 m). By comparison, under IFR, the separation between parallel runway centerlines must typically be at least 4,300 feet (~1311 m) or, usually, 5,000 feet (~1524 m) to permit simultaneous parallel approaches.

[3] Similar caps are also in effect at Washington's Reagan National Airport (DCA), but currently exceed demand at most times of the day and thus not constraining flight schedules in any significant manner.

2. Airport capacity

Airport operations are enabled by available physical and technological infrastructure. At the same time, infrastructure limitations and safety-related operating procedures constrain the throughput that any airport can achieve. In this section, we define the notion of airport capacity and identify its key drivers. We then present theoretical and computational tools used to estimate airport capacity, and results from capacity comparisons across airports and operating conditions. Last, we discuss the implications of capacity limitations for the planning of airport operations.

2.1. Description

Airport operations involve a range of consecutive processes, from passenger and cargo operations in terminal buildings to aircraft operations at the gates, on the aprons, on the taxiways, on the runways, and in the terminal airspace. Even though throughput restrictions can occur at several stages of these processes, the runway system generally acts as the main bottleneck at the busiest airports. In other words, terminal capacity, gate capacity, taxiway capacity and airspace capacity are generally sufficient or can be expanded (albeit often at high cost) to accommodate the operations that are performed on the runway system. In turn, airport throughput is primarily constrained by the capacity of the runway system.

Airport throughput is often reported in terms of annual traffic statistics, such as the annual number of aircraft movements or of passengers, but these measures do not describe effectively airport capacity because they are mostly demand-driven. For instance, an airport with a capacity of 60 flights per hour and another airport with a capacity of 100 per hour will have the same annual throughput if their daily and seasonal demand profiles are identical and the traffic peaks do not exceed 50 flights per hour. More generally, annual throughput is affected by such factors as demand seasonality, geographical location of the airport (which drives daily peaking patterns) and airport curfews (which reduce the number of hours of airport operations), but all these considerations are not linked to the operating capabilities of the airport.

To address these limitations, it has been necessary to develop more precisely-defined measures that are focused on the amount of traffic the airport can handle during relatively short periods of time when it must operate at full capacity. For runway systems, the most fundamental of these measures is the *maximum throughput capacity*, defined as the "average number of aircraft movements that can be processed per unit of time under continuous demand". Note that this definition concentrates on periods of *continuous demand*, when the airport operates at its full potential to handle a persistent, "non-stop" flow of arrivals and departures. In this way, it isolates the notion of "capacity" from whatever temporal demand patterns prevail at the airport. This also means that maximum throughput capacity can be measured empirically only during peak traffic periods and is therefore typically stated as the *rate* at which traffic movements can be processed over relatively short units of time, e.g., "number of movements per hour" (or per 15 min or per 5 min). Note, as well, that, by defining capacity as an *average* throughput rate, it is implicitly recognized that airport throughput is a random variable whose value may vary depending on a number of factors. Some of the most important among these factors are:

- *number of runways*: All else being equal, the more runways available, the more movements can be operated simultaneously and, in turn, the higher the capacity. As noted in the introduction, airport runway systems range from a single runway up to 6–8 runways.
- *physical layout of the runway system*: The rules and procedures regarding simultaneous operations on multiple runways depend on their layout.[4] Parallel runways with sufficient separations between them allow for independent operations, while, for obvious safety reasons, intersecting runways impose strict constraints on simultaneous operations. In practice, many airports exhibit complex runway system layouts that may include sets of parallel and intersecting runways.
- *runway configuration*: All the runways of an airport are not necessarily active at the same time. Wind conditions and air traffic control capabilities typically enable the simultaneous use of up to 3–5 runways. The *runway configuration*, defined as the set of runways that are active to operate arrivals and the set of runways that are active to operate departures at any particular time, is selected by air traffic controllers, may change several times during each day of operations depending on prevailing weather conditions and other considerations, and impacts the resulting airport throughput.
- *separation requirements*: In order to ensure safe operations, air traffic control procedures impose separation requirements between consecutive movements. For instance, in the United States and under IFR, two wide-body aircraft landing consecutively must be separated by at least 4 nautical miles during final approach. The equivalent required separation between two narrow-body aircraft is 2.5 or 3 nautical miles, depending on the airport. Note that different sets of separation requirements are in use in different countries. For example, many countries use a set of requirements recommended by the International Civil Aviation Organization (ICAO), but many national Civil Aviation Authorities (notably the FAA in the United States) specify their own separation requirements that often differ from those of the ICAO.
- *weather and other operating conditions*: In the United States, the use of VFR, weather permitting, results in shorter, on average, separations between consecutive aircraft than under IFR. As a result, airport capacity can be significantly higher in Visual Meteorological Conditions (VMC) than in Instrument Meteorological Conditions (IMC).
- *mix of arrival and departures*: Separation requirements vary as a function of the type of movement (landing or take-off) and, therefore, the resulting capacity depends on the mix of arrivals and departures at hand. For instance, it may be possible to perform 70 departures and 30 arrivals in an hour with a given runway configuration but not 30 departures and 70 arrivals. The trade-off

[4] Airport operations also depend on the layout of the taxiway system and the location of the terminal buildings.

between the arrival throughput and the departure throughput depends on a runway configuration's layout and associated air traffic control operating procedures.

- *aircraft mix*: Separation requirements vary as a function of the type of aircraft (e.g., wide-body vs. narrow-body)[5] and, thus, the resulting capacity depends on the mix of aircraft types at hand. For instance, all else being equal, more time is required to land a sequence of 10 narrow-body aircraft and 5 wide-body aircraft than one of 15 narrow-body aircraft.

As a result of these dependencies, airport throughput exhibits significant variability. Some of the factors that affect this variability can be controlled (e.g., runway configuration in use), but some others are subject to uncertainty and stochasticity (e.g., weather conditions). Any single-value capacity estimate (e.g., 100 movements per hour) just provides an "expected value" of the capacity and fails to capture this variability. It is therefore advisable to augment, when possible, these estimates with (i) an indication of how they vary under several different operating scenarios, and/or (ii) a characterization of their variability in stochastic terms (e.g., through a probability distribution or, at least, an estimate of variance).

2.2. Analytical tools

Approaches to estimating airport capacity fall into two broad categories: theoretical models and empirical models. Theoretical models of airport capacity were among the first applications of operations research (Blumstein, 1959). They are based on abstract representations of a runway system and associated operating procedures (e.g., aircraft separation requirements). Through a set of mathematical relationships or by simulating a large number of movements, the service times required by each type of movement (e.g., Heavy/Medium/Light aircraft, arrival/departure, etc.) and for every operating scenario (e.g., mix of arrivals and departures, runway configuration in use, etc.) are computed. The capacity of the airport is then obtained from the average service time (i.e., if it takes on average 1 min to operate a movement, then the capacity is 60 movements per hour). Models can range from very simple (e.g., for computing the capacity for departures of a single runway with a single aircraft type) to highly complex (e.g., for computing the capacity of a multi-runway system with complex operating procedures). These approaches enable the approximate estimation of airport capacity under various operating conditions, including hypothetical ones (e.g., a new airport, an additional runway, or even a significant change in separation requirements or traffic control procedures). On the negative side, the abstractions and simplifications of reality that necessarily underlie these mathematical and simulation models cannot fully capture all the operating complexities found in practice.

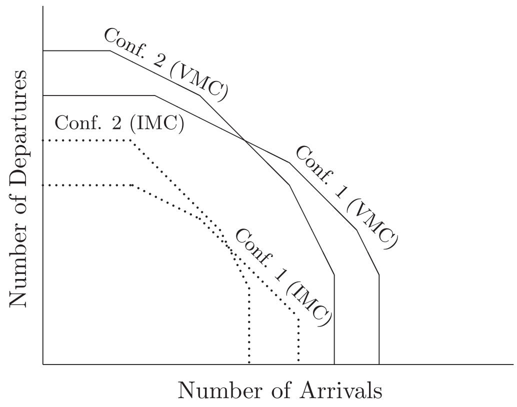
In contrast, empirical models process operations data obtained from historical records (Federal Aviation Administration, 2013) or, in some recent cases, from realistic human-in-the-loop experiments (Barnett et al., 2015). Inputs may include precise take-off and landing times of aircraft at the airport of interest, as well as operating data necessary to identify periods of continuous demand, such as the times when aircraft request to land or arrive at the runway for take-off. Using the scatter plot of the number of departures vs. the number of arrivals per time unit (e.g., per 15-min period), Gilbo (1993) first proposed the characterization of capacity by means of a *Capacity Envelope*, defined as the *maximal* (e.g., 95th percentile) numbers of arrivals/departures that can be feasibly operated per period. This concept was then extended by Simaiakis (2012), who introduced the *Operational Throughput Envelope*, which shows the relationship between the *average* number of departures and arrivals that can be operated per period—more consistent with the definition of capacity as an average throughput. The empirical approach has the great advantage of relying on information that reflects the full complexity of airport operations. On the negative side, it can be applied only where extensive historical databases on airport operations exist.

Fig. 2a shows a typical schematic representation of an airport's Operational Throughput Envelopes that can be computed through either one of the two approaches summarized above. By definition, this representation captures the variations in airport capacity as a function of the mix of arrivals and departures. Several Operational Throughput Envelopes can be computed to reflect the differences in airport throughput in various operating conditions. For instance, Fig. 2a shows envelopes for two different hypothetical runway configurations in VMC and IMC, reflecting the facts that one configuration may achieve a higher arrival throughput but a lower departure throughput than another, and that airport throughput is higher in VMC than in IMC. Fig. 2b shows the estimation of the Operational Throughput Envelope from empirical data from 2007 for New York EWR's Configuration $4R|4L$ in VMC, obtained from the scatter plot of the count of arrivals and departures per 15-min period. Note the significant variability in airport throughput as a function of underlying variations in aircraft mix, operating conditions, human factors, etc.

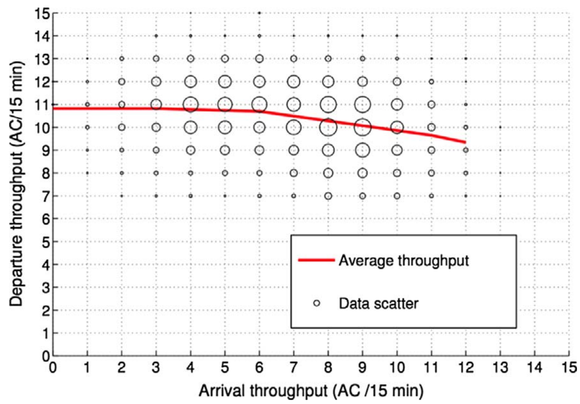
2.3. Insights

The application of theoretical and computational models of airport capacity has demonstrated significant differences across busy airports worldwide. In the United States, a benchmark study of airport capacities by the Federal Aviation Administration (2004) showed that the VFR capacity of the 35 busiest airports ranged from a low of 55–65 movements per hour (e.g., San Diego (SAN)) to a high of over 200 movements per hour (e.g., Dallas/Fort Worth (DFW)). Similarly, IFR capacities range from 50 movements per hour to 150–200 movements per hour. In Europe, several of the top 30 airports declare capacities of the order of 40 movements per hour while Paris Charles de Gaulle (CDG) and Amsterdam Schiphol (AMS) declare capacities of over 100 per hour (Morisset and Odoni,

[5] The actual categories of aircraft classes in use for air traffic control purposes is more refined. For example, ICAO currently classifies aircraft into four classes (Super Heavy, Heavy, Medium and Light) depending on their maximum take-off weight.



(a) Schematic



(b) Application at EWR (VMC, $4R|4L$)

Fig. 2. Representation of capacity with Operational Throughput Envelopes (Simaiakis, 2012).

Table 1
Capacity comparison at airports with similar layouts (Morisset and Odoni, 2011).

Layout	US Airports				European Airports	
	Airport	Weighted	VFR	IFR	Airport	Declared
Single runway	SAN	55	57	49	LGW DUB TXL STR	50 46 48 42
Two closely spaced parallel runways	SEA	76	82	59	DUS MAN NCE	47 59 52
Two pairs of closely spaced, parallel runways	ATL LAX	179 137	184 143	160 121	CDG	112

2011). These major differences are primarily driven by available infrastructure at the airports; not surprisingly, the airports with the largest capacity estimates are those with multiple runways, while airports with capacities in the 35–60 movements per hour range generally have a single runway.

However, the size and physical characteristics of a runway system do not explain all observed differences in airport capacities. To compare the capacities of several airports with similar runway layouts, Table 1 reports the declared capacity of several European airports and the VFR capacity, IFR capacity and "weighted" capacity (i.e., the average value of capacity across all weather conditions) of some US airports (Morisset and Odoni, 2011). Note that the declared capacity of European airports is roughly the same as the IFR capacity of similar US airports.[6] This is consistent with the fact that the vast majority of European airports operate under IFR 100% of the time and can therefore only match the IFR capacity of similar US airports. For schedule coordination purposes, European airports thus base their declared capacity on IFR operating capabilities.[7] As shown in Table 1, US airports achieve significantly higher capacities under VFR than under IFR, because of lower separations between consecutive aircraft and more efficient use of multiple runways. Overall, the use of VFR at US airports, weather permitting (around 80% of the time, on average, at the 35 busiest airports), results in average capacity gains of 26%, as compared to their European counterparts with similar layouts.

These comparisons underscore the impact of the runway system infrastructure, separation requirements and weather conditions on airport capacity. Data on throughput variations *at any single airport* can also be used to identify the impact of runway configuration, the mix of arrivals and departures, and the mix of aircraft, as was done by Simaiakis (2012) at Boston (BOS) and New York's JFK, EWR and LGA airports. First, the representation of airport capacity by means of an Operational Throughput Envelope (see Fig. 2) has demonstrated that any given airport's capacity can vary by as much as 20% depending on the mix of arrivals and departures. For instance, JFK's capacity was estimated at 20 movements per 15-min period under balanced operations (10 arrivals, 10 departures), 22 movements per period when arrivals are given priority (16 arrivals, 6 departures), and 18 movements per period when departures are given priority (6 arrivals, 12 departures). Second, the runway configuration can also induce ~20% variations in arrival and departure

[6] ATL, whose IFR capacity is 40% higher than that of CDG and LAX is an exception, partly reflecting a different mix of aircraft types.

[7] Over the past decade, several European airports have increasingly relied on some use of VFR in good weather conditions and thus have gained capacity. At busy airports such as London Heathrow (LHR), London Gatwick (LGW), Frankfurt (FRA), and Munich (MUC), high demand volumes have motivated the declaration of capacities that are a little higher than the computed IFR capacities.

capacities. For instance, the average departure (resp. arrival) throughput at JFK is estimated at 11 take-offs (resp. 16 landings) per 15-min period in a configuration with one departure runway and two arrival runways, and to 13 take-offs (resp. 13 landings) in a configuration with two departure runways and one arrival runway. Third, airport throughput varies by 1 to 4 movements per 15-min period (depending on the runway configuration in use) as a function of the aircraft mix (e.g., the relative number of wide-body and narrow-body aircraft using the runway system).

2.4. Implications

Characterizing the constraints imposed by airport capacity limitations on air traffic operations is critical to the design of adequate infrastructure plans, operating procedures and demand management schemes at busy airports. It is important to recall that single-value estimates of airport capacity only measure *average* throughput, i.e., the average number of movements per unit of time that can be sustained over long periods of time across a variety of operating conditions. Instead of relying solely on such single-value estimates, airport capacity should be described at a higher level of detail to account for its dependence on several different operating factors through, for instance, the Operational Throughput Envelopes shown in Fig. 2a. Fine-grain estimates should be developed to quantify, at the very least, the variability of capacity with respect to (i) weather conditions (e.g., VMC vs. IMC), (ii) the mix of arrivals and departures, (iii) the runway configuration in use, and (iv) the mix of aircraft types. Theoretical and empirical models support the underlying estimation procedures.

The characterization of capacity also provides guidance for supply-side interventions aimed at enhancing operating capabilities. First, the impact of weather and separation requirements on capacity motivate the systematic investigation of potential improvements in operating practices to increase throughput. This can take place through air traffic control procedures aimed at optimizing the sequencing of arrivals and departures and of different types of aircraft to minimize the average separation between consecutive movements (Balakrishnan and Chandran, 2010; Solveling et al., 2011). Critical, in this respect, are developments in air traffic management infrastructure that enable adherence to the minimum permissible separation requirements between aircraft in the terminal airspace (SESAR, 2012; Federal Aviation Administration, 2014). Similarly, the management of multi-runway systems can result in significant capacity gains and performance improvements (Bertsimas et al., 2011a; Jacquillat et al., 2017). In the longer-term, the design of air traffic control procedures and of improved separation requirements is an important lever for increasing airport capacity while ensuring system safety.[8]

Second, at airports where operating enhancements are not sufficient to scale up capacity to meet demand, capacity increases may be achieved through the expansion of physical infrastructure. Typical interventions include adding a runway, increasing spacing between runways to allow independent operations, and re-designing taxiways. Note that long-term operating capabilities are strongly determined by early-stage airport development. For instance, all current options for expanding significantly the capacity of the London Airport System require enormous investments in infrastructure, have complex and uncertain impacts on local communities and the environment, and will take fifteen or more years to implement (Airports Commission, 2015). For airports under development, this underscores the need for careful long-term traffic forecasts, thorough investigation of the costs and feasibility of different infrastructure options, and flexible infrastructure development plans to avoid the all-too-common risks of “under-building” or “over-building” (Neufville et al., 2013).

3. Airport operations

Given the state of airport infrastructure, the capacity limitations discussed in Section 2 impose constraints on airport operations and, if coupled with strong demand, may result in air traffic queues and flight delays. In this section, we describe the dynamics of airport congestion and review briefly analytical models that estimate delays as a function of flight schedules and airport capacity. We then identify the main drivers of airport on-time performance, and discuss the implications for the planning, management and operations of airport systems.

3.1. Description

Demand for airport access consists of scheduled flights by the airlines and unscheduled ones by general aviation and military aircraft. Airport queues, and resulting flight delays, occur when demand exceeds runway capacity. Departing aircraft will queue primarily on taxiways next to runways, and arriving aircraft primarily in “holding patterns” in the terminal airspace.[9] As mentioned in the introduction, the cost of flight delays in the United States alone was estimated at $32.9 billion in 2007 (Ball et al., 2010). Of these, 25% ($8.3 billion) are direct costs to the airlines, consisting mostly of increased crew, fuel, and maintenance expenses (Zou and Hansen, 2012). The passenger component amounts to roughly 50% of the total ($16.7 billion), and consists of increased travel times, flight cancellations, and missed connections (Barnhart et al., 2014). The remaining 25% are other societal losses, such as lost demand for air travel and broader impacts on the nation’s GDP.

Most of these delays are created by local imbalances between demand and capacity, resulting from excessive scheduling levels

[8] An international effort is currently under way to refine existing separation requirements and adopt them on a global scale at the busiest airports.

[9] In order to reduce fuel burn and other costs of delays, recent procedures aim to absorb departure delays at gates (Simaiakis et al., 2014b,a), and arrival delays in en-route airspace (e.g., through speed control or re-routing) or at the origin airport when a Ground Delay program is implemented (Bertsimas et al., 2011b).

over long periods of time and/or capacity shortages at the airport due to unfavorable (but non-extreme) weather conditions. A second category of delays results from the propagation of operating disturbances in a network of airports. This occurs through two mechanisms. First, once an aircraft suffers a serious delay at an airport, it may take several subsequent flights through a number of airports before the aircraft can "recover" that initial delay and operate back on schedule. Second, flight delays to any number of aircraft of an airline may also lead to changes in schedules of other flights of the same airline, and thus in changes in the dynamics of formation and propagation of delays at airports over the day of operations. A third category of delays is caused by unforeseen disruptions in airline operations (e.g., late passenger boarding, aircraft mechanical problems, etc.).[10] We focus here on the modeling and estimation of the *queuing delays*, i.e., those caused by imbalances between demand and capacity at the airports, as these are also the delays that can be addressed through demand and capacity management interventions. Queuing delays, together with the propagated delays generated by the queuing of aircraft at upstream airports, account for 50–75% of all flight delays in the United States (Bureau of Transportation Statistics, 2013).

3.2. Analytical tools

Models of airport congestion fall into three categories: microscopic, mesoscopic and macroscopic. Microscopic models (almost always simulations) consider each aircraft individually and recreate precisely the physical layout of the airport of interest and of surrounding airspace (Bilimoria et al., 2000; Sood and Wieland, 2003; George et al., 2011) to track the movement of each arriving and departing aircraft, record any delays suffered and inform air traffic control interventions at the tactical level (e.g., aircraft sequencing and spacing). Mesoscopic models compute runway delays, taxi-in and taxi-out times, and other related statistics by using less detailed operational data, such as the runway configuration in use, arrival schedules, and pushback schedules (Shumsky, 1995; Pujet et al., 1999; Simaiakis and Balakrishnan, 2016; Khadilkar and Balakrishnan, 2014). These models typically support the design of air traffic management interventions to optimize flows of aircraft at an airport or in a network of airports (e.g., Ground Delay Programs, speed control, aircraft routing and re-routing, departure metering). Finally, macroscopic models use more aggregate representations of airport operations to generate computationally efficient estimates of flight delays as a function of flight schedules and airport capacity. These models are the most relevant for assessing such strategic interventions as capacity expansion or demand management initiatives.

Macroscopic models of airport congestion require capacity estimates and demand data. Capacity estimates are obtained from the models and data analyses outlined in Section 2. Demand estimates at busy commercial airports are obtained from airline flight schedules. However, published schedules of flights (which correspond to the times when aircraft are expected to depart from or arrive at the gate) do not fully coincide with demand for runway use (which corresponds to the times when aircraft are first available to take off or land). The characterization of runway demand therefore requires additional information such as unimpeded taxi times (i.e., the taxi-out and taxi-in times in the absence of congestion). In addition, macroscopic models of congestion must be able to account for the variability of operating conditions at an airport because of changing weather or winds. In the context of strategic modeling in support of demand and capacity management, this is done through the use of relatively simple models, based on historical records of operations that capture approximately the changes in operating conditions that take place over time at the airport (e.g., the frequency and duration of transitions between VMC and IMC weather).

Most macroscopic models of airport congestion are based on *queuing theory*, and view the airport as a queuing system, as shown in Fig. 3. Service is provided by the runway system, access to which is demanded by arriving and departing aircraft. Any imbalances between demand and service capacity result in the queuing of aircraft. Demand and service can be represented as deterministic processes (Hansen, 2002; Hansen and Hsiao, 2005; Nikoleris and Hansen, 2012) or as stochastic (often, Poisson or Erlang) processes (Kivestu, 1974; Gupta, 2010; Jacquillat and Odoni, 2015b). Stochastic models of demand and service aim to capture the uncertainty and variability associated with all aspects of airport operations. The set of required input parameters depends on the modeling choices, but includes, at the very least, the demand rate (i.e., the average number of aircraft demanding the usage of the runway system per time unit, denoted by λ) and the service rate (i.e., the average number of movements that can be processed per time unit, denoted by μ). The delay models then return various queue-related statistics which, depending on the specific model, can range from simple (e.g., average waiting time per aircraft during a day) to highly detailed (e.g., the probability distribution of the number of arriving and departing aircraft queuing in each period of time throughout the day of operations).

Most recent research has focused on the development of efficient computational methods to solve dynamic and stochastic queuing models of congestion. Most results in queuing theory regard operations in "steady state" i.e., operations after the system has run for a sufficiently long time to reach equilibrium conditions (Larson and Odoni, 1981). Moreover, most of these steady-state results assume that the demand and service rates are constant over time. But airport operations are highly variable over time, as flight scheduling exhibits peaks and valleys and airport capacity may also vary greatly with weather and other operating conditions. The exact solution of queuing models under such time-dependent conditions is very difficult and, even numerical solutions are very computationally intensive. For these reasons, numerical methods have been developed to estimate approximately airport departure and/or arrival queuing statistics using time-dependent analyses (Kivestu, 1974; Odoni and Roth, 1983; Gupta, 2010; Simaiakis and Balakrishnan, 2016; Jacquillat and Odoni, 2015b). The resulting models provide various delay estimates for an entire day of operations in times ranging from a few seconds to a few minutes. More recently, these models have been extended to describe the propagation of delays

[10] These three categories account for the large majority of delays. The remaining ones are due to rare events such as extreme weather or safety- or security-related incidents.

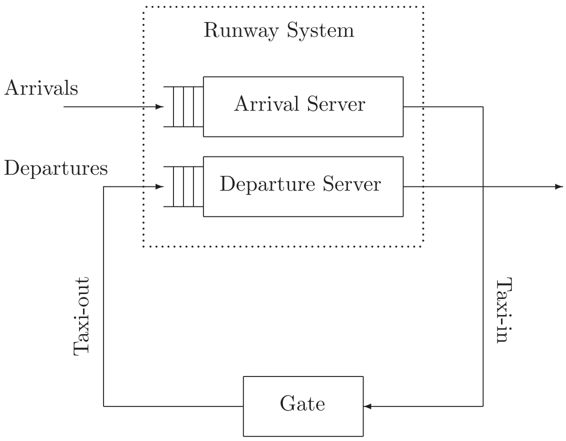


Fig. 3. Representation of the airport as a queuing system.

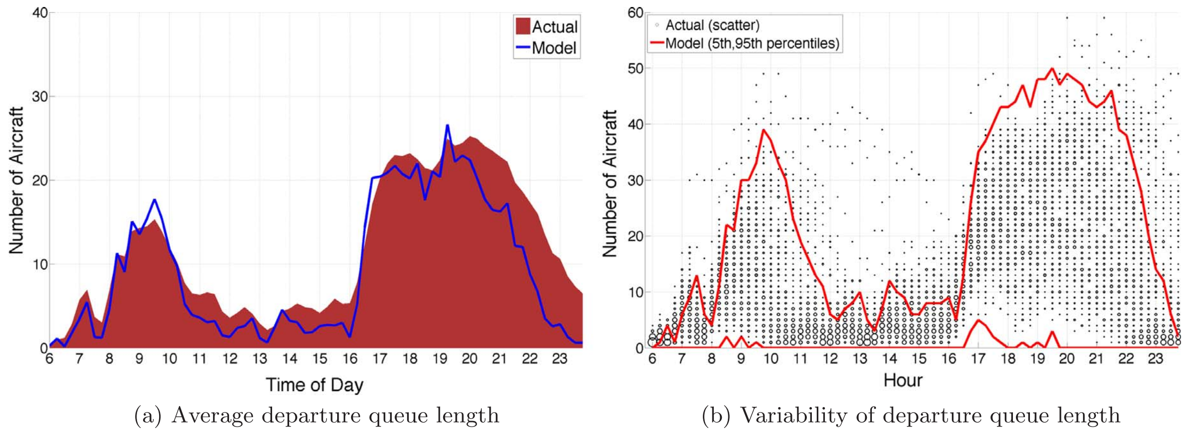


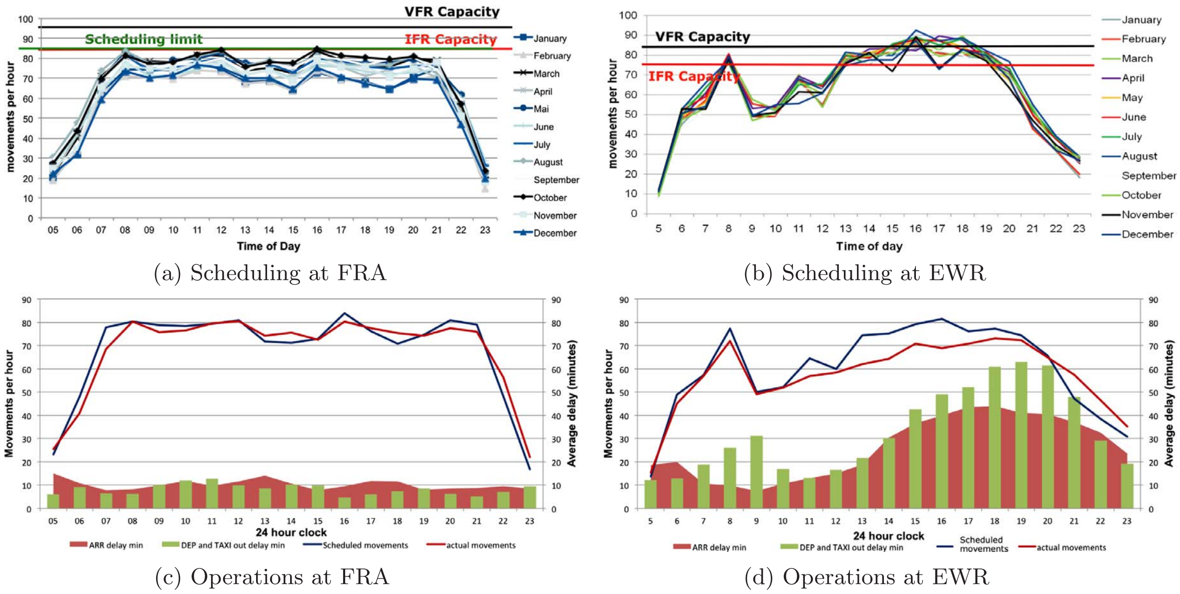
Fig. 4. Average and variability of the departure queue length at JFK in Summer 2007 (Jacquillat and Odoni, 2015b).

in a network of airports (Pyrgiotis et al., 2013). Comparisons with empirical data have shown that the flight delays observed in practice can be estimated with good accuracy by these macroscopic queuing models. Fig. 4 provides an example of model validation with data from New York JFK. It compares the average departure queue length (Fig. 4a) and the range of departure queue lengths (Fig. 4b) observed in practice to those predicted by the model. The good match between observations and model predictions indicates the feasibility of obtaining computationally efficient and reasonably accurate estimates of on-time performance at busy airports (Lovell, 2007; Pyrgiotis and Simaiakis, 2010; Jacquillat and Odoni, 2015b).

3.3. Insights

Queuing models can be applied to the study of the dynamics of formation and propagation of delays over a day at busy airports. Delays obviously occur when the demand rate exceeds the service rate (i.e., when the number of flights scheduled exceeds airport capacity). But delays may also occur when the demand rate is lower than the service rate. These delays are due to stochasticity, i.e., to the random fluctuations in demand and service times, and can be large when demand stays close to capacity over an extended period of time. It is well known from queuing theory that a highly non-linear relationship exists between scheduling levels, airport capacity and on-time performance. In steady-state conditions, the average delay is proportional to $\frac{1}{1-\rho}$, where the *utilization ratio* $\rho = \frac{\lambda}{\mu}$ is defined as the ratio of the demand and service rates. In other words, small changes in flight schedules or in airport throughput (thus, small changes in ρ) can have a disproportionate impact on flight delays when the airport operates close to capacity.

Empirical results have shown that this type of strongly non-linear relationship remains valid when demand and capacity are dynamic, i.e., vary over time, as is the case at airports. To illustrate this point, Fig. 5 shows average monthly schedules of flights (Figs. 5a and b) and average arrival and departure delays (Figs. 5c and d) at Frankfurt (FRA) and Newark (EWR) using data from 2007 when no flight caps were in place at EWR. Note that, as a result of the differences between demand management practices at European and US airports, scheduling levels are much higher relative to capacity and more variable at EWR than FRA. At FRA, the number of flights scheduled per hour remains essentially constant at a level close to the airport's declared capacity, which is roughly equal to its IFR capacity. By contrast, more flights are scheduled at peak hours at EWR than even the optimal (i.e., VFR) capacity of



(a) Scheduling at FRA (b) Scheduling at EWR

(c) Operations at FRA (d) Operations at EWR

Fig. 5. Scheduling and on-time performance at FRA and EWR in 2007 (Odoni et al., 2011).

the airport, while demand falls well below even the IFR capacity at off-peak hours. The result of these practices is that delays are much higher on average at EWR than at FRA. At FRA, average delays remain stable throughout the day at around 8–10 min. In contrast, delays increase rapidly over the day at EWR reaching an average of 40–60 min during the peak afternoon hours, a reflection of the fact that the airport is over-scheduled and cannot keep up with demand. The comparison highlights the point that differences in scheduling patterns (e.g., a 10–20% difference in the number of flights scheduled per hour) can have dramatic effects on on-time performance at an airport.

Similar non-linearities have also been observed at individual airports during time periods when significant changes in flight schedules have taken place. For instance, as part of the phase out of the High Density Rule at New York LGA in 2000, all operations performed by aircraft under 70 seats and operating between LGA and small airports were exempted from slot limits. As a result, demand for airport access increased from 1,050 to 1,350 movements per day in just a few months, creating unprecedented levels of delays and cancellations. To deal with the situation—and as an interim solution—the FAA instituted a limit of 75 on the number of scheduled operations per hour, and carried out a lottery to allocate slots to the airlines. This resulted in a 10% drop in demand (from 1,350 to ~1200 daily movements). This demand reduction, in turn, led to a very large reduction in delays, which fell from an average of 60–80 min during peak afternoon hours to 10–20—an 80% drop (Fan and Odoni, 2002). Significant changes in schedules also took place at LGA, JFK and EWR between 2007 and 2010. The implementation of flight caps in 2008 resulted in the smoothing of demand peaks at JFK and EWR and the economic downturn to a demand reduction of 5–10% at the three airports between 2008 and 2010. Over the same period, average delays declined by an estimated 40–50% at these airports. The application of a queuing model showed that these very significant delay reductions could be largely explained by the comparatively small changes in flight schedules (Jacquillat and Odoni, 2015b).

3.4. Implications

The non-linear relationship between flight schedules, airport capacity and airport on-time performance offers guidelines for congestion mitigation through demand and capacity management. From a supply perspective, airport performance is highly sensitive to even small variations in airport capacity. In the short-term, changes in airport operating conditions can result in significant variations in flight delays. For instance, even brief capacity shortfalls due to temporary weather deterioration can lead to the formation of long queues that will then persist and dissolve very slowly over time—especially when scheduling levels are high such as those depicted in Fig. 5b. In the longer term, significant improvements in airport operating performance can be achieved through capacity increases (e.g., infrastructure expansion) or the enhancement of operating procedures (e.g., optimization of air traffic control and air traffic flow management procedures). Finally, it is important to note that delays also increase with the *variability* of service times (Hansen et al., 2009; Nikoleris and Hansen, 2012). Future air traffic management systems, such as NextGen in the United States and SESAR in Europe, are expected to improve on-time performance at airports not only through increases in average throughput, but also through reductions in service rate variability through more accurate and consistent spacing between consecutive aircraft movements.

From a demand perspective, airport on-time performance is highly sensitive to the volume of flights scheduled in a day, and to the distribution of these flights over the course of the day. First, small increases (resp. decreases) in the number of flights scheduled in a day of operations can lead to disproportionately large increases (resp. decreases) in resulting delays. Second, all else being equal, the more evenly flights are distributed over the day of operations, the lower the resulting delays. From an economic standpoint, this suggests that, at airports operating close to capacity, the marginal cost per extra flight movement at peak hours is very significant and, in fact, much higher than typical landing fees (Carlin and Park, 1970; Hansen, 2002; Fan, 2003). Historically, the combined effect of (i) the sensitivity of flight delays to changes in flight schedules, (ii) growth in air traffic demand, and (iii) the concentration of flights around certain peak periods due to intense airline competition, has led to high levels of congestion during these periods at airports where access is largely unrestricted, such as busy US airports in 2007.

The non-linear relationship between flight schedules and delays provides opportunities for significant improvements in on-time performance through limited scheduling changes (e.g., small reductions in the number of flights scheduled at peak hours, potentially offset by increases in off-peak scheduling levels). In fact, after the 2008-10 decline, demand (as measured by the number of scheduled flights) and, thus, delays have not grown back to the 2007 levels at New York's airports, due to the flight caps in place and to the self-imposed airline scheduling restrictions known as *capacity discipline* (Wittman and Swelbar, 2014). Outside the United States, slot control policies provide an opportunity to capitalize on the non-linear relationship and achieve large delay reductions through comparatively small reductions in demand. At the same time, these schedule-limiting schemes also impose costs by restricting airport access at peak hours. We discuss these questions in more detail in the next section.

4. Flight scheduling and demand management

As noted in Section 3, demand management practices can have significant impacts on airport scheduling and operating performance. In this section, we describe the dynamics of flight scheduling, formalize the trade-offs underlying airport demand management, and present some recent models supporting congestion-mitigating adjustments in airline schedules of flights. The results of these models offer guidelines for enhancing demand management policies at schedule coordinated airports (e.g., at busy airports outside the United States), at airports where access is largely unconstrained (e.g., at the overwhelming majority of US airports), and at airports under development.

4.1. Description

Demand for airport access is primarily determined by airline scheduling of flights, based on managerial objectives (e.g., profit maximization), constraints (e.g., fleet and crew availability, aircraft turnaround times) and social factors (e.g., underlying passenger demand). It involves a number of interdependent airline decisions, ranging from network and route planning through frequency planning and flight timetabling to pricing and revenue management. Over-capacity scheduling at busy airports may occur as a result of growth in air traffic demand and of airline competitive dynamics that create incentives for high schedule frequencies on each origin-destination market, potentially through the use of smaller aircraft (Belobaba et al., 2009; Vaze and Barnhart, 2012a).

To control over-capacity scheduling (at least at busy times of the day), the most common demand management schemes fall into three categories: schedule coordination, congestion pricing and slot auctions (Neufville et al., 2013; Czerny et al., 2008). Schedule coordination (and resultant slot controls) in place at busy airports outside the United States is an administrative, quantity-based scheme: airports declare a quantity of available slots per hour (or other unit of time), and slots are then allocated through an administrative procedure (International Air Transport Association, 2015). In contrast, congestion pricing is an economic, price-based mechanism: congestion tolls are specified for access to the airport, and airlines then schedule their flights based on the resulting cost (Carlin and Park, 1970; Brueckner, 2002). In-between, slot auctions are economic, quantity-based mechanisms: similarly to slot controls, airports "declare a capacity", i.e., specify the number of available slots, which are then allocated through a market-based mechanism (Ball et al., 2006). Alternative demand management schemes have been proposed recently to supplement or replace these three, including (i) hybrid mechanisms, which allocate a fixed number of slots administratively and allocate the remaining slots through an auction or other economic scheme, (ii) secondary trading, which allows buying and selling of slots after an initial allocation has been made (Pellegrini et al., 2012), and (iii) non-monetary targeted scheduling interventions, which adjust flight schedules taking into consideration airline scheduling requests and on-time performance objectives (Jacquillat and Odoni, 2015a, 2017).

Any demand management scheme involves a trade-off between mitigating congestion and maximizing capacity utilization. On one hand, demand management can control peak-hour scheduling levels at busy airports, and thus result in (potentially significant) reductions in flight delays. On the other, any centralized intervention to reduce peak-hour scheduling levels results in some flights being displaced to off-peak hours or not being scheduled at all. Therefore, demand management can provide benefits as reduced congestion costs, but also creates costs for airport stakeholders by constraining airline schedules. Existing demand management schemes can be viewed as varying approaches to resolving this trade-off: slot control policies place a premium on congestion mitigation by setting (generally conservative) limits on flight schedules, while the largely unrestricted access in place at almost all US airports places a premium on capacity utilization by minimizing interference with airline scheduling. More broadly, this trade-off can be resolved by adjusting either the number of slots that are made available or, the congestion tolls imposed for access to the airport.

Note that the demand management interventions may create opportunities for strategic behaviors from the airlines, i.e., potential incentives to provide scheduling inputs that do *not* reflect their true preferences in order to gain a strategic advantage over their competitors (e.g., to reduce the number of their own flights that will be displaced) (Vaze and Barnhart, 2012b; Harder and Vaze, 2017). In the research reported here, these opportunities were not considered explicitly. This is motivated by the fact that, under relatively mild scheduling adjustments (such as the one described in the remainder of this section), there might not be as strong incentives for untruthful behaviors as, for instance, under mechanisms that involve the rejection of some flight scheduling requests.

Nonetheless, the identification and mitigation of such gaming behaviors need to be carefully considered in the design of any demand management mechanism and represent important avenues for future research.

4.2. Analytical tools

The design and implementation of demand management schemes involve scheduling interventions that use, as a starting point, scheduling inputs from the airlines and airport capacity estimates. Scheduling inputs include all the flights scheduled to or from the airport, as well as the preferred departure and arrival times of these flights. They may also include information about connections between flight pairs (e.g., same aircraft performing consecutive flights and/or many passengers typically connecting between two flights). Such connections are central to building and maintaining an airline's network of flights. Capacity estimates can either take the form of "declared capacity" values (i.e., administrative quantities that specify the number of available arrival/departure slots), or, preferably, operating capacity estimates (i.e., estimates of average airport throughput that can be achieved under various operating scenarios). Additional inputs may include stakeholder preferences regarding the trade-off between congestion mitigation and capacity utilization (e.g., flexibility of proposed schedules and/or tolerance of flight delays).

Based on these inputs, scheduling interventions aim to find schedules of flights that satisfy airline scheduling requests as closely as possible, while accounting for demand management rules and procedures, airline scheduling constraints, airport operating capabilities, and desired levels of service. This is typically formalized by minimizing the schedule *displacement*, i.e., the difference between the scheduled times requested by the airlines and the times actually assigned to them, subject to scheduling constraints, network connectivity constraints and demand management constraints:

$$\begin{array}{rl} \text{min} & \text{Schedule displacement} \\ \text{st} & \text{Scheduling constraints} \\ & \text{Network connectivity constraints} \\ & \text{Demand management constraints} \end{array}$$

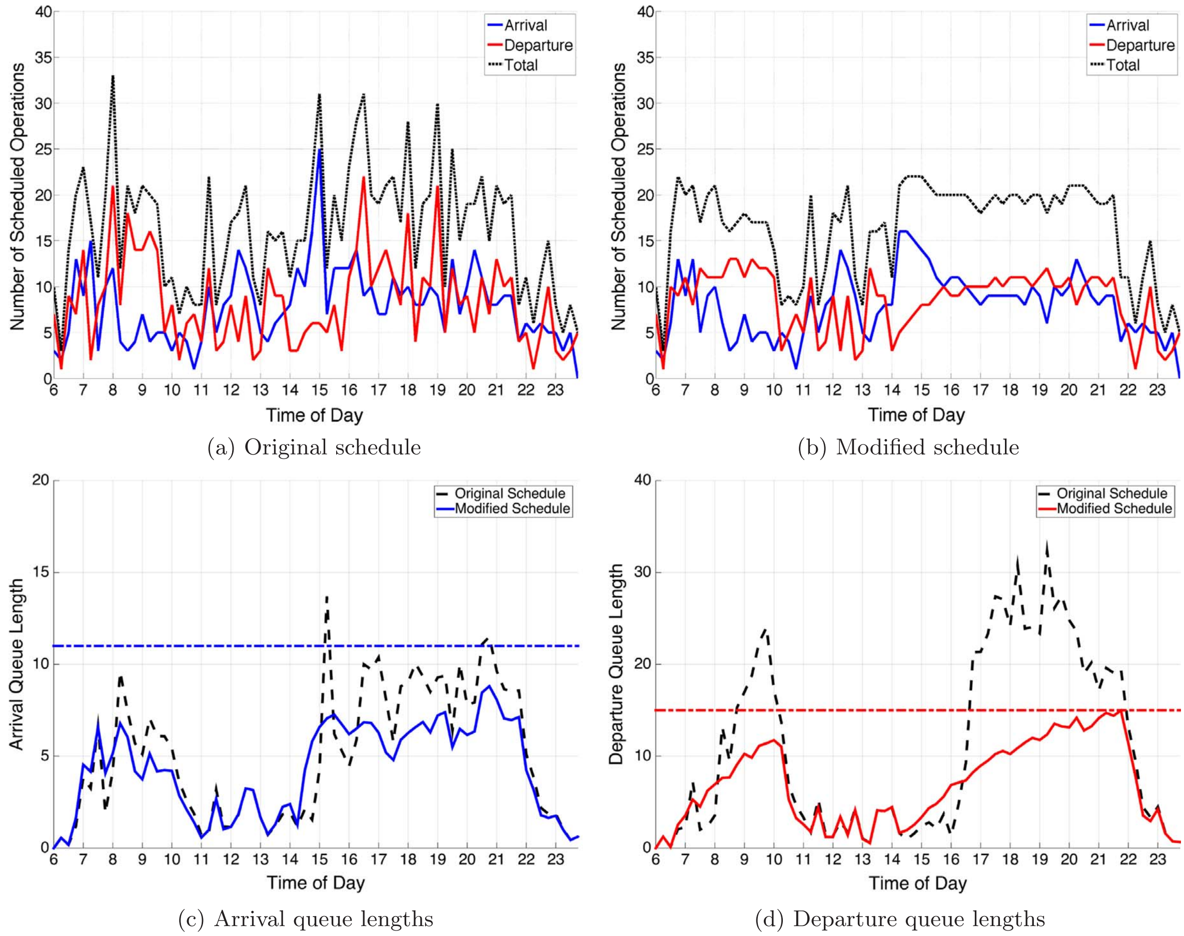
Variants of this general formulation have been applied to the optimization of several different demand management schemes such as slot controls (Zografos et al., 2012), slot auctions (Rassenti et al., 1982) and non-monetary interventions (Jacquillat and Odoni, 2015a). The objective function can be adjusted to capture non-monetary mechanisms (e.g., minimize the total time changes from airline scheduling requests) or market-based mechanisms (e.g., maximize revenues by accepting as many scheduling requests as feasible). Recent research has aimed to include alternative objectives, such as minimizing the number of flights displaced and maximizing inter-airline equity (Ribeiro et al., 2017; Zografos and Jiang, 2016; Jacquillat and Vaze, 2017). The constraints ensure the feasibility of the proposed schedule and its consistency with desired levels of service. The scheduling and network connectivity constraints can ensure that the structure of airline networks and schedules of flights is left unchanged by maintaining all flights scheduled by the airlines or by eliminating selectively some flights, if desired levels of service cannot be satisfied or if some demand management constraints are violated (Swaroop et al., 2012; Pyrgiotis and Odoni, 2016). The scheduling constraints can also capture schedule coordination procedures based on the IATA (or any other) guidelines, e.g., slot bundles and priorities. The demand management constraints can take several forms. The most widely used are slot availability constraints, which impose limits on the number of flights that can be scheduled per unit of time based on the declared capacities of the airports (e.g., "no more than 100 movements scheduled per hour"). These constraints have been recently replaced by on-time performance constraints, which impose level-of-service targets, typically by specifying upper limits on expected arrival and departure queue lengths or, equivalently, on expected arrival and departure delays (e.g., "expected delay should not exceed 20 min at any time of the day") (Jacquillat and Odoni, 2015a). Such on-time performance constraints link explicitly scheduling levels to congestion mitigation objectives, considering the patterns of airport capacity availability. In summary, demand management mechanisms can be supported by advanced models that flexibly optimize the scheduling interventions, based on underlying objectives, rules and procedures.

4.3. Insights

The application of demand management models at busy airports has yielded three major insights. First, optimization models can enhance the efficiency of schedule coordination by accommodating airline scheduling requests better than is typically the case under current practice. Second, at US airports with largely unconstrained access, relatively minor adjustments to airline scheduling preferences can improve significantly on-time performance. Third, the intensity of demand management interventions can be calibrated to select the most desirable trade-off between congestion mitigation and capacity utilization.

At slot-controlled airports, comparisons of optimization model outputs to the actual slot allocation outcomes has demonstrated opportunities for improving existing schedule coordination procedures. Scheduling models can result in significantly better matching of airline scheduling requests than the coordinated schedules observed in practice, while accounting for slot bundles and priorities as specified in the IATA guidelines (Zografos et al., 2012, 2016; Ribeiro et al., 2017). This may suggest that schedule coordination decisions are currently made in practice on an *ad hoc* basis, without accounting for the full set of flight scheduling requests and without performing schedule optimization. Such inefficiencies have motivated some air transportation industry stakeholders to propose alternative strategies to supplement and/or replace existing approaches to schedule coordination (Dot Econ Ltd., 2001; NERA, 2004; Czerny et al., 2008; Madas and Zografos, 2008).

At US airports, research results suggest strongly that over-capacity scheduling at many busy airports can be reduced while still



(a) Original schedule (b) Modified schedule

(c) Arrival queue lengths (d) Departure queue lengths

Fig. 6. Scheduling and on-time performance at JFK on 05/25/2007 under the original and modified schedules (Jacquillat and Odoni, 2015a).

satisfying current levels of demand, and that relatively minor scheduling adjustments can yield significant improvements in on-time performance (Fan and Odoni, 2002; Le et al., 2008; Le, 2006; Swaroop et al., 2012; Vaze and Barnhart, 2012b; Jacquillat and Odoni, 2015b; Pyrgiotis and Odoni, 2016; Jacquillat and Odoni, 2015a). Fig. 6 shows the results of the application of the optimization model of Jacquillat and Odoni (2015a) at JFK, with the flight schedule of May 25, 2007. The top figures show the combined schedule of arrivals and departures submitted by the airlines (Fig. 6a) and a modified schedule of flights obtained with the model (Fig. 6b). The bottom figures compare the resulting expected arrival (Fig. 6c) and departure (Fig. 6d) queue lengths, under the original and modified schedules. Note, first, that the modified schedule is much more evenly distributed than the original one. This is achieved by rescheduling 10%-20% of the flights to or from JFK to later or earlier times by 15 min each (and, in a few cases, by 30 min each), and without eliminating any flights or any connections between these flights. Nonetheless, the resulting schedule is not "flat", but exhibits peaks and valleys consistent with intra-day variations in passenger demand and airline scheduling preferences. This is because the model considers—and satisfies—on-time performance constraints, instead of strict scheduling limits, thus resulting in a schedule of flights that is closer to airline preferences. Second, the expected delays are significantly lower under the modified schedule than under the original schedule. Specifically, peak expected arrival and departure delays are reduced by over 30% and 50%, respectively, and corresponding average arrival and departure delays through the entire day of operations are reduced by 10–20% and 20–40%, respectively. Therefore, in the case where the total daily demand is in line with available capacity at the airport, large reductions in congestion costs can be achieved through limited adjustments that distribute demand more evenly by optimally rescheduling some flights from peak hours to off-peak hours.

Demand management optimization models also provide decision-making support for determining the best trade-off between congestion mitigation and capacity utilization. While Fig. 6 shows the original schedule and only one "modified" schedule, the model can, in fact, generate multiple modified schedules that impose increasingly large schedule displacements and result in increasingly large delay reductions. In a demand management scheme based on slot availability constraints, these different solutions are obtained through the setting of the airport's declared capacity. In a scheme based on congestion pricing, they are obtained through the setting of the airport's congestion tolls. And in a scheme based on scheduling interventions with on-time performance constraints, they are obtained through the setting of the expected queue length targets. Evidence indicates that schedule displacement varies non-linearly with the intensity of the demand management interventions, i.e., a limited increase in declared capacity (or a limited decrease in congestion tolls, or a limited increase in expected queue length targets) can result in disproportionately large reductions in the resulting schedule displacement (Zografos et al., 2012; Jacquillat and Odoni, 2015a). This, combined with the non-linear relationship between scheduling levels and delays, underlines the importance of considering carefully the "intensity" of the demand management measures to be applied at an airport, such as the precise value of the declared capacity, of the congestion tolls or of the targeted maximum expected queue lengths. The scheduling models cited in this section can support these decisions which can be made by relevant national authorities (e.g., the US Federal Aviation Administration, or schedule coordinators outside the United States) or, preferably, through a collaborative process that brings together all the major stakeholders, such as airport operators, the airline community and passenger representatives (Jacquillat and Odoni, 2015a).

4.4. *Implications*

Based on the results outlined in this section, there exist important opportunities to improve the design and implementation of demand management schemes at busy airports worldwide. At schedule-coordinated airports located outside the United States, capacity allocation could be made more efficiently through a more systematic and flexible approach to the setting of slot limits. First, a formal process to declare airport capacities based on demand and capacity patterns would improve and standardize performance across airports. Second, slot limits need not be constant through the whole day of operations but could, instead, permit variations in the number of flights scheduled from one time period to another to match peak-hour demand better and enable airports to recover from potential delays during off-peak hours (as seen in Fig. 6 and discussed, at a more macroscopic level, by Churchill et al. (2012)). Third, relying on hourly limits on the total number of movements may not be sufficient to control effectively potential imbalances between demand and capacity at a more disaggregate level. For example, hourly limits may not prevent the airline practice of scheduling many flight departures "on the hour" and "on the half-hour" for marketing purposes. As practiced at some airports, hourly limits could be supplemented by 5-min, 15-min or 30-min limits (e.g., "no more than 80 movements per hour and no more than 25 movements per 15-min period") and by separate limits on the number of scheduled arrivals and departures (e.g., "no more than 80 total movements, 50 arrivals and 50 departures per hour"). These would make the scheduling limits more consistent with the fine-grain representations of airport capacity shown in Section 2. In the longer-term, the strict scheduling limits currently in place could be replaced by on-time performance constraints that would optimize scheduling adjustments based on congestion mitigation objectives.

In the United States, a more systematic and forward-looking examination of the opportunities associated with scheduling interventions at congested airports is called for. At several among the great majority of airports where no scheduling limits are currently in place, the potential benefits (congestion cost savings) and costs (interference with airline competitive scheduling) of such interventions should be carefully reviewed and discussed by the relevant stakeholders. At the airports where some scheduling limits are already in place, the formalization and optimization of the process of effecting scheduling interventions is recommended, as suggested by the Government Accountability Office (2012) and the US Department of Transportation's Office of Inspector General (2010). This would take place through a careful setting of the scheduling limits based on demand and capacity patterns (as discussed above for the schedule-coordinated airports), an open process through which the airlines can submit their scheduling requests to a central decision-maker (e.g., the FAA, in collaboration with airport operators), and a transparent set of rules and procedures for carrying out any recommended scheduling adjustments. In the longer term, stakeholder alignment regarding on-time performance objectives and the integration of such objectives into the optimization of scheduling interventions could enhance the performance of US airport systems. The scheduling models discussed in this section can provide decision-making support to guide the design of such mechanisms.

For airports under development, these lessons provide guidelines for designing demand management schemes that effectively manage congestion, match airline scheduling preferences, and integrate the objectives, constraints and requirements of various stakeholders into the decision-making process. Essential steps include a careful examination of demand and capacity patterns, the setting of on-time performance objectives with the participation of relevant stakeholders, and the design of scheduling mechanisms supporting demand management measures, if such measures are deemed necessary. Limited political legacy in this context provides important opportunities to design a comprehensive, neutral, and transparent approach to the design of demand management mechanisms that is not necessarily based on the strict schedule coordination approach in place at busy airports outside the United States or on the relatively unconstrained approach in place in the United States. It can be supported by the creation of politically isolated decision-making bodies, and implemented in a collaborative decision-making environment that ultimately brings together the airport, commercial airlines, and other major stakeholders. We discuss this in more detail in the final section.

5. Discussion

This paper has presented a synthesis of the three major operational and managerial drivers of airport systems performance. *Airport capacity* characterizes the average arrival and departure throughput that can be achieved per unit of time under continuous demand. *Airport operations* refer to the air traffic handling procedures for processing arrivals and departures. *Flight scheduling* refers to airport access by the airlines, potentially subject to demand management interventions. In this final section, we summarize the main tools, results and insights presented in this paper, and discuss their implications for airport demand and capacity management, with emphasis on airports under development.

5.1. Summary

Airport demand and capacity management is supported by a suite of models and analytical methods, described in this paper. The application of these powerful tools to cases involving a diverse set of airports offers insights into the strengths and weaknesses of existing practices related to airport capacity, flight scheduling, airport operations and demand management.

First, airport capacity depends on available infrastructure and air traffic operating procedures. It is primarily driven by the size and the physical layout of the runway system. All else being equal, the more extensive the runway system, the greater the capacity. In addition, for any given state of the infrastructure, airport capacity depends on air traffic handling policies (e.g., on separation requirements between consecutive aircraft) and procedures (e.g., those aimed at optimizing aircraft sequencing and spacing). Available infrastructure and existing operating policies and procedures explain most of the significant capacity differences observed between the busiest US airports (with multiple runways and the use of Visual Flight Rules, weather permitting) and most European airports (with generally less extensive runway systems and reliance on Instrument Flight Rules). In addition, the throughput at any individual airport is sensitive to operating factors, such as weather variability and severity, the mix of arrivals and departures, the runway configuration in use, and the mix of aircraft at the airport. Detailed understanding of airport operating capabilities thus requires the estimation of airport capacity *under a broad range of operating scenarios*.

Second, scheduling patterns at airports also exhibit significant differences worldwide, primarily due to differences in demand and capacity management practices. At US airports, the combination of higher capacity levels in VMC (which prevail, on average, for about 80% of the time) and largely unconstrained access results in scheduling levels that are typically in line with (and, sometimes, even higher than) the VFR capacity of the airport. In other words, the *reference* capacity for airline scheduling purposes is the VFR capacity of the airport. In contrast, and because of the use of IFR and strict slot control policies, scheduling at European airports (and most of the other busiest airports outside the United States) is limited by declared capacities that are generally in line with the IFR capacities of the airports. As a result, the utilization of available capacity is usually significantly higher at US airports than at their counterparts in the rest of the world, with significant resulting economic benefits to the airlines, passengers, and society as a whole.

Third, the relationship between flight schedules and airport capacity, on the one hand, and airport delays, on the other, is strongly non-linear. For this reason, air traffic congestion can exhibit large variations as a result of comparatively small changes in airport demand (e.g., in the volume of flights scheduled in a day, or even in the distribution of flights over the course of the day) and capacity (e.g., due to short-term variations in weather conditions or long-term changes in infrastructure). Because scheduling levels per unit of capacity are higher at US airports than at European airports, these non-linearities explain the much larger airport delays—and, thus, much higher congestion costs—observed on average in the US than in Europe. Moreover, at any individual airport, these non-linearities may result in sharp variations in on-time performance over time, in response to evolving changes in demand and/or capacity. For instance, the small reductions in peak-hour scheduling levels at New York's airports due to the imposition of flight caps and the demand reduction in 2008-09 have led to very significant improvements in on-time performance.

Fourth, the design and implementation of any demand management schemes at busy airports involves a trade-off between mitigating airport congestion and maximizing airport capacity utilization. Any demand management mechanism can control the levels of air traffic delays but, at the same time, interferes to some degree with airline competitive scheduling. At schedule-coordinated airports (i.e., the great majority of busy airports outside the United States), the value of the airport's declared capacity is of critical importance in determining the balance between ensuring efficient slot allocation and keeping air traffic delays within reasonable bounds. This has motivated gradual increases in declared capacities at a few major European airports to enable higher scheduling levels. At US airports, substantial gains in airport on-time performance can be achieved through only limited interference with airline competitive scheduling. Overall, recent advances in schedule optimization provide opportunities for flexible approaches to airport demand management that optimize the trade-off between congestion mitigation and maximization of capacity utilization, considering airport operating capabilities and airline scheduling preferences.

5.2. A roadmap for airport demand and capacity management

We conclude by presenting a roadmap of the types of analyses that a major airport under development would expect to undertake in connection with managing its airside demand and capacity at various times during its lifetime from the early planning stage to maturity.

The most fundamental input to the planning process is a forecast of demand over a time-horizon typically spanning a period of thirty to forty years. Estimating airline demand for airport access involves, at its initial stage, anticipating passenger demand for air travel, the airline competitive landscape and airline scheduling decisions. Typically, airport forecasts are expressed in terms of millions of passengers per annum (mppa) over the selected time horizon. For purposes of airside planning, such forecasts must be converted, first, to annual air traffic movements (ATMs) and, subsequently, to peak-hour ATMs. This is because, as noted in Section 2, the true measure of airside capacity is the aircraft throughput (i.e., number of ATMs) that can be performed during periods of continuing demand (i.e., during peak hours). The conversion of mppa to annual ATMs requires forecasts of the types of aircraft with which the airlines will serve the subject airport or, at the very least, estimates of the average number of passengers per ATM in the future. The conversion of annual ATMs to peak-hour ATMs then requires prediction of the (seasonal and daily) peaking patterns of flight schedules at the subject airport over the forecast's horizon. While the topic of airport demand forecasting is beyond the scope of this paper, it is obvious from this discussion that each of the steps in the forecast (generation of annual passenger forecasts, conversion to annual ATMs, conversion to peak-hour ATMs) requires a long series of important assumptions. The forecasts of ATM demand are therefore subject to great uncertainty. The history of airport forecasting is replete with examples in which forecasts have failed by wide margins to anticipate future traffic developments, even over a horizon of ten years or less (Neufville et al., 2013). It is therefore imperative that such demand forecasts be made for a variety of scenarios, reflecting the high variability surrounding the regulatory, competitive and general economic environment of air transport.

Once traffic forecasts have been developed, the following four types of analyses, all drawing on the models and methodologies discussed in this paper, will or may be carried out over time.

1. *Planning for airport infrastructure*: The design of airside facilities aims to develop sufficient operating capabilities to match long-term ATM demand, while controlling the costs of development (including land acquisition costs, design costs, construction costs and other indirect costs such as environmental impact). The ideal outcome is that, at any point in the lifetime of the airport, the runway system and the airfield will have the capability to handle peak-hour ATMs. As discussed in this paper, the runway system is the main driver of airport capacity, but airside infrastructure planning spans a broader range of design decisions, including: the location and general configuration of the terminal buildings; number and size of aircraft gates and of remote stands; number of runways and their layout; and associated system of taxiways. For any planned state of airside infrastructure, the airport's operating capabilities can be characterized through the estimation of airport capacity, using the models and approaches described in Section 2. Because of uncertainties in forecasts, it is advisable that, for airports that may eventually require a multi-runway system, a staged development plan be adopted. Such a plan will provide flexibility, so that airport capacity can grow to match demand over time, while avoiding the risk of over-building at an early stage.
2. *Optimizing air traffic handling policies and procedures*: Once the airside infrastructure is given, the efficiency of airport operations is determined by air traffic management systems, consisting of technological infrastructure (facilities and equipment), trained personnel and a set of air traffic control rules and procedures for operating flights on a day-to-day basis. Objectives here include deploying air traffic management technologies in a cost-effective manner, enforcing adequate separation requirements between aircraft to ensure system safety, adhering to workload constraints and maximizing airport throughput given policy requirements and the capabilities of air traffic management systems. As noted in this paper, the performance and procedures associated with air traffic management systems varies considerably from country to country and it is thus always important to consider local conditions and constraints.
3. *Setting on-time performance objectives*: Demand forecasts and capacity estimates enable the quantification of on-time performance at any subject airport by computing the magnitude and variability of arrival and departure delays over a day of operations, using queuing models of congestion outlined in Section 3. Airports and Civil Aviation Authorities can collaborate with major stakeholders, notably the airlines and possibly passenger representatives, to determine acceptable levels of congestion (e.g., the amount of queuing delay that the airport should not exceed on average). This determines whether the levels of congestion estimated by the queuing models are acceptable or, in contrast, if delay reductions should be targeted.
4. *Establishing demand management rules*: If, at any stage of an airport's development, the actual or forecast levels of delay exceed desired levels of service—and if airport operating capabilities cannot be easily improved—then an additional stage of analysis involves the design of scheduling mechanisms to support demand management interventions. Relevant models and optimization tools were presented in Section 4. The two main decisions involve (i) setting the stringency of demand management (i.e., the intensity of scheduling interventions and of the resulting delay reductions), and (ii) designing and optimizing a scheduling mechanism for allocating capacity to commercial airlines and other airport users based on non-monetary, administrative processes (e.g., schedule coordination, schedule facilitation), or on monetary, market-based approaches (e.g., congestion pricing, slot auctions). In contrast to infrastructure expansion and operating enhancements, which involve upfront investments but then result in gains for all stakeholders, demand management imposes longer-term economic costs on airline scheduling and passenger traveling options, and should thus be carefully reviewed with all relevant stakeholders.

As already noted, each of the above types of analysis may be activated once, more than once or not at all during an airport's lifetime. Continuous monitoring of scheduling and operating performance is required to improve demand and capacity management practices. On the supply side, performance monitoring involves collecting operations data, comparing observed throughput to capacity estimates (under various operating scenarios) and comparing observed delays to on-time performance predictions. On the demand side, it involves observing airline demand and assessing the effects of scheduling interventions, if any, on flight schedules. Any significant deviation from predictions may motivate revisions in operations and scheduling forecasts and, potentially, adjustments in airport policies and practices. For instance, demand increases (resp. improvements in air traffic management systems) may result in significant delay increases (resp. decreases) and may thus motivate stronger (resp. weaker) demand management.

This roadmap provides a holistic approach to the management of airport demand and capacity. At existing airports already operating under physical constraints (e.g., with limited options for infrastructure expansion) or regulatory restrictions (e.g., with constraints on permissible demand management mechanisms), the main challenges involve identifying which options are feasible and most desirable, and adjusting existing policies and practices accordingly. At new airports under development with limited such constraints (e.g., the planned new international airport in Mexico City), the roadmap provides forward-looking, neutral, transparent and flexible guidelines to support the planning of capacity, the optimization of operating procedures and the design of demand management schemes. Through continuous monitoring of scheduling and operating practices, and continuous stakeholder engagement, its application can significantly enhance airport performance over their lifecycles.

References

Airports Commission, 2015. Final Report. Technical Report. Chaired by Sir Howard Davies.

Balakrishnan, H., Chandran, B., 2010. Algorithms for scheduling runway operations under constrained position shifting. Oper. Res. 58 (6), 1650–1665.

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., 2010. Total Delay Impact Study. Technical Report. National Center of Excellence for Aviation Operations Research, College Park, MD.

Ball, M., Donohue, G., Hoffman, K., 2006. Combinatorial Auctions, Chapter Auctions for the Safe, Efficient, and Equitable Allocation of Airspace System Resources. MIT Press pp. 507–538.

Barnett, A., Ball, M., Donohue, G., Hansen, M., Odoni, A., Trani, A., 2015. Collision course? The north airfield safety study at Los Angeles International Airport (LAX). Transp. Res. Part A: Policy Pract. 77, 14–34.

Barnhart, C., Fearing, D., Vaze, V., 2014. Modeling passenger travel and delays in the national air transportation system. Oper. Res. 62 (3), :580–601.

Belobaba, P., Odoni, A., Barnhart, C., 2009. The Global Airline Industry. John Wiley & Sons.

Bertsimas, D., Frankovich, M., Odoni, A., 2011a. Optimal selection of airport runway configurations. Oper. Res. 59 (6), 1407–1419.

Bertsimas, D., Lulli, G., Odoni, A., 2011b. An integer optimization approach to large-scale air traffic flow management. Oper. Res. 59 (1), :211–227.

Bilimoria, K., Sridhar, B., Chatterji, G., Sheth, K., Grabbe, S., 2000. FACET: future ATM concepts evaluation tool. In: 3rd USA/Europe Air Traffic Management R & D Seminar.

Blumstein, A., 1959. The landing capacity of a runway. Oper. Res. 7 (6), 752–763.

Brueckner, J., 2002. Airport congestion when carriers have market power. Am. Econ. Rev. 92 (5), 1357–1375.

Bureau of Transportation Statistics, 2013. TranStats. Accessed March 26, 2013. < http://www.transtats.bts.gov > .

Carlin, A., Park, P., 1970. Marginal cost pricing of airport runway capacity. Am. Econ. Rev. 60 (3), 310–319.

Churchill, A., Lovell, D., Mukherjee, A., Ball, M., 2012. Determining the number of airport arrival slots. Transp. Sci. 47 (4), 526–541.

Czerny, A., Forsyth, P., Gillen, D., Niemeier, H., 2008. Airport Slots. Ashgate Publishing Ltd.

de Neufville, R., Odoni, A., 2013. Airport Systems: Planning, Design and Management, second ed., McGraw-Hill.

Dot Econ Ltd., 2001. Auctioning Airport Slots: A Report for HM Treasury and the Department of the Environment, Transport and the Regions. Technical report.

Fan, T., 2003. Market-based Airport Demand Management – Theory, Model and Applications (PhD thesis). Massachusetts Institute of Technology.

Fan, T., Odoni, A., 2002. Airport demand management: a practical perspective. Air Traffic Control Quart. 10 (3), :285–306.

Federal Aviation Administration, 2004. Airport Capacity Benchmark Report. Technical Report.

Federal Aviation Administration, 2013. Aviation System Performance Metrics (ASPM) Database. Available at: < https://aspm.faa.gov/apm/sys/main.asp > (accessed April 4, 2013).

Federal Aviation Administration, 2014. NextGen Implementation Plan. Technical Report.

George, S., Satapathy, G., Manikonda, V., Wieland, F., Refai, M., Dupee, R., 2011. Build 8 of the airspace concept evaluation system. In: AIAA Modeling and Simulation Technologies Conference.

Gilbo, E., 1993. Airport capacity: representation, estimation, optimization. IEEE Trans. Control Syst. Technol. 1 (3), 144–154.

Gillen, D., Jacquillat, A., Odoni, A., 2016. Airport demand management: the operations research and economics perspectives and potential synergies. Transp. Res. Part A: Policy Pract. 94 (3), 495–513.

Government Accountability Office, 2012. Slot-Controlled Airports: FAA's Rules Could be Improved to Enhance Competition and Use of Available Capacity. Technical Report GAO-12-902.

Gupta, S., 2010. Transient Analysis of D(t)/M(t)/1 Queuing System with Applications to Computing Airport Delays (Master's thesis). Massachusetts Institute of Technology.

Hansen, M., 2002. Micro-level analysis of airport delay externalities using deterministic queuing models: a case study. J. Air Transp. Manage. 8 (2), 73–87.

Hansen, M., Hsiao, C.Y., 2005. Going south? Econometric analysis of US airline flight delays from 2000 to 2004. Transp. Res. Rec.: J. Transp. Res. Board 1915, 85–94.

Hansen, M., Odoni, A., Lovell, D., Nikoleris, T., Vlachou, K., 2009. Use of queuing models to estimate delays savings from 4D trajectory precision. In: 8th USA/Europe Air Traffic Management R & D Seminar.

Harder, R., Vaze, V., 2017. Two-Stage Game Theoretic Modeling of Airline Frequency and Fare Competition (submitted for publication).

International Air Transport Association, 2015. Worldwide Slot Guidelines, seventh ed.

Jacquillat, A., Odoni, A., 2015a. An integrated scheduling and operations approach to airport congestion mitigation. Oper. Res. 63 (6), :1390–1410.

Jacquillat, A., Odoni, A., 2015b. Endogenous control of arrival and departure service rates in dynamic and stochastic queuing models with application at JFK and EWR. Transp. Res. Part E: Logist. Transp. Rev. 73 (1), :133–151.

Jacquillat, A., Odoni, A., 2017. A New Airport Demand Management Approach Based on Targeted Scheduling Interventions. Working Paper.

Jacquillat, A., Odoni, A., Webster, M., 2017. Dynamic control of runway configurations and of arrival and departure service rates at JFK airport under stochastic queue conditions. Transp. Sci. 51 (1), 155–176.

Jacquillat, A., Vaze, V., 2017. Inter-airline Equity in Airport Scheduling Interventions. Working Paper.

Keeney, R., 1973. A decision analysis with multiple objectives: the Mexico City airport. Bell J. Econ. Manage. Sci. 4 (1), 101–117.

Khadilkar, H., Balakrishnan, H., 2014. Network congestion control of airport surface operations. AIAA J. Guid. Control Dyn. 37 (3), 933–940.

Kivestu, P., 1974. Alternative Methods of Investigating the Time Dependent M/G/k Queue (Master's thesis). Massachusetts Institute of Technology.

Larson, R., Odoni, A., 1981. Urban Operations Research. Prentice-Hall.

Le, L., 2006. Demand Management at Congested Airports: How Far are we from Utopia (PhD thesis). George Mason University.

Le, L., Donohue, G., Hoffman, K., Chen, C., 2008. Optimum airport capacity utilization under congestion management: the case of New York LaGuardia airport. Transp. Plann. Technol. 31 (1), :93–112.

Lovell, D., Churchill, A., Odoni, A., Mukherjee, A., Ball, M., 2007. Calibrating aggregate models of flight delays and cancellation probabilities at individual airports. In: 7th USA/Europe Air Traffic Management R & D Seminar.

Madas, M., Zografos, K., 2008. Airport capacity vs. demand: mismatch or mismanagement? Transp. Res. Part A: Policy Pract. 42, 203–226.

Morisset, T., Odoni, A., 2011. Capacity delays and schedule reliability at major airports in Europe and the United States. Transp. Res. Rec. 2214, 85–93.

NERA, 2004. Study to Assess the Effects of Different Slot Allocation Schemes. Technical Report. National Economic Research Associates.

Nikoleris, T., Hansen, M., 2012. Queueing models for trajectory-based aircraft operations. Transp. Sci. 46 (4), 501–511.

Odoni, A., Morisset, T., Drotleff, W., Zock, A., 2011. Benchmarking airport airside performance: FRA vs. EWR. In: 9th USA/Europe Air Traffic Management R & D Seminar.

Odoni, A., Roth, E., 1983. An empirical investigation of the transient behavior of stationary queuing systems. Oper. Res. 31 (3), :432–455.

Office of Inspector General, 2010. New York Flight Delays Have Three Main Causes, but More Work Is Needed To Understand Their Nationwide Effect. Technical Report AV-2011-007. U.S. Department of Transportation.

Pellegrini, P., Castelli, L., Pesenti, R., 2012. Secondary trading of airport slots as a combinatorial exchange. Transp. Res. Part E: Logist. Transp. Rev. 48 (5), 1009–1022.

Pujet, N., Delcaire, B., Feron, E., 1999. Input-output modeling and control of the departure process of congested airports. In: AIAA Guidance, Navigation, and Control Conference.

Pyrgiotis, N., Malone, K., Odoni, A., 2013. Modelling delay propagation within an airport network. Transp. Res. Part C: Emerg. Technol. 27, 60–75.

Pyrgiotis, N., Odoni, A., 2016. On the impact of scheduling limits: a case study at Newark International Airport. Transp. Sci. 50 (1), :150–165.

Pyrgiotis, N., Simaiakis, I., 2010. An analytical queuing model of airport departure processes for taxi out time prediction. In: AIAA Aviation Technology, Integration, and Operations.

Rassenti, S., Smith, V., Bulfin, R., 1982. A combinatorial auction mechanism for airport time slot allocation. Bell J. Econ. 13 (2), 402–417.

Ribeiro, N., Antunes, A., Jacquillat, A., Odoni, A., 2017. An optimization approach for airport slot allocation under IATA guidelines. In: Air Transport Research Society

World Conference.
SESAR, 2012. European ATM Master Plan: The Roadmap for Sustainable Air Traffic Management. Technical Report.
Shumsky, R., 1995. Dynamic Statistical Models for the Prediction of Aircraft Take-off Times (PhD thesis). Massachusetts Institute of Technology.
Simaiakis, I., 2012. Analysis, Modeling and Control of the Airport Departure Process (PhD thesis). Massachusetts Institute of Technology.
Simaiakis, I., Balakrishnan, H., 2016. A queuing model of the airport departure process. Transp. Sci. 50 (1), :94–109.
Simaiakis, I., Sandberg, M., Balakrishnan, H., 2014a. Dynamic control of airport departures: algorithm development and field evaluation. IEEE Trans. Intell. Transp. Syst. 15 (1), 285–295.
Simaiakis, I., Sandberg, M., Balakrishnan, H., Reynolds, T., Hansman, J., 2014b. Demonstration of reduced airport congestion through pushback rate control. Transp. Res. Part A: Policy Pract. 66, 251–267.
Solveling, G., Solak, S., Clarke, J.P., Johnson, E., 2011. Runway operations optimization in the presence of uncertainties. AIAA J. Guid. Control Dyn. 34 (5), 1373–1382.
Sood, N., Wieland, F., 2003. Total airport and airspace model (TAAM) parallelization combining sequential and parallel algorithms for performance enhancement. In: 2003 Winter Simulation Conference.
Swaroop, P., Zou, B., Ball, M., Hansen, M., 2012. Do more US airports need slot controls? A welfare based approach to determine slot levels. Transp. Res. Part B: Methodol. 46 (9), :1239–1259.
Vaze, V., Barnhart, C., 2012a. An assessment of the impact of demand management strategies for efficient allocation of airport capacity. Int. J. Revenue Manage. 6 (1/2), :5–27.
Vaze, V., Barnhart, C., 2012b. Modeling airline frequency competition for airport congestion mitigation. Transp. Sci. 46 (4), :512–535.
Wittman, M., Swelbar, W., 2014. Capacity discipline and the consolidation of airport connectivity in the United States. Transp. Res. Rec.: J. Transp. Res. Board 2449, 72–78.
Zhang, A., Czerny, A., 2012. Airports and airlines economics and policy: an interpretive review of recent research. Econ. Transp. 1, 15–34.
Zografos, K., Andreatta, G., Odoni, A., 2013. Modelling and Managing Airport Performance. Wiley.
Zografos, K., Jiang, Y., 2016. Modelling and solving the airport slot scheduling problem with efficiency, fairness, and accessibility considerations. In: TRISTAN Symposium.
Zografos, K., Madas, M., Androutsopoulos, K., 2016. Increasing airport capacity utilisation through optimum slot scheduling: review of current developments and identification of future needs. J. Scheduling 1–22.
Zografos, K., Salouras, Y., Madas, M., 2012. Dealing with the efficient allocation of scarce resources at congested airports. Transp. Res. Part C: Emerg. Technol. 21, 244–256.
Zou, Hansen, M., 2012. Impact of operational performance on air carrier cost structure: evidence from US airlines. Transp. Res. Part E: Logist. Transp. Rev. 48 (5), 1032–1048.