Dynamic Control of Runway Configurations and Arrival and Departure Service Rates at JFK Airport under Stochastic Queue Conditions

Alexandre Jacquillat Amedeo R. Odoni
Mort D. Webster
Massachusetts Institute of Technology, Cambridge, MA 02139
alexjacq@mit.edu, arodoni@mit.edu, mort@mit.edu

Abstract

The large delays experienced at major airports, which impose substantial costs on airlines, passengers and society, require the implementation of airport congestion mitigation tools to improve the efficiency and reliability of the air transportation system. This paper presents a decision-making tool to optimize the utilization of airport capacity at the tactical level in the face of operational uncertainty. A novel approach is developed to minimize congestion costs by jointly controlling runway configurations and arrival and departure service rates through the course of a day of operations as a function of observed congestion on the ground and in the air and of meteorological and wind conditions. The approach combines stochastic queue dynamics with a decision-making framework based on dynamic programming. An effective formulation of this problem allows the optimal policy to be found within reasonable time frames. An approximate algorithm is also implemented to enable on-line implementation of the model and is shown to yield near-optimal policies. The application of this model to John F. Kennedy International Airport suggests that the implementation of this dynamic control can result in substantial congestion cost savings, of the order of 15% to 20%. The approach also provides a new way of controlling arrival and departure service rates in dynamic models of airport congestion, thus enhancing the usefulness of these methodologies.

Key words Airport, Capacity, Delay, Runway Configuration, Dynamic Programming, Queuing Model

1 Introduction

Airport congestion is an increasingly important and costly phenomenon worldwide. In the United States, air traffic delays reached an all-time peak in 2007, and their nationwide impact was estimated at over $30 billion for that calendar year, including $8.3 billion in costs to airlines and $16.7 billion in costs to passengers [1]. The main cause of these delays is the demand-capacity mismatches due to lower realized capacity at the airports than scheduling levels. Indeed, the Bureau of Transportation Statistics [6] has reported, over recent years, approximately half of the system-wide delays as

National Aviation System delays, due to heavy traffic volume, to capacity shortages due to non-extreme weather conditions, to inefficiencies in airport operations etc. Moreover, over one-fourth of total delays were due to the propagation of schedule disturbances within the network of operations, themselves primarily created by demand-capacity mismatches at upstream airports. In other words, delays *not* created by demand-capacity mismatches and due to inefficiencies in airline operations, to extreme weather and to security procedures merely account for approximately 25% of the nationwide delays.

The reduction of flight delays in the United States can therefore be achieved primarily by reducing the imbalance between demand and capacity at the busiest airports. This can be done by expanding capacity through the construction of new airports or of new runways, as airport capacity is primarily constrained by the runway system [7]. However, such infrastructure projects require extensive investments and many years to plan and implement, and, more importantly, are hardly feasible in the most densely populated areas because of geographic, environmental, economic and political issues. A second way to reduce the demand-capacity mismatch is to limit peak-hour scheduling levels through the imposition of demand management measures, either administrative (*e.g.*, slot control), or economic (*e.g.*, congestion pricing, slot auctions). It has been demonstrated that such mechanisms could lead to significant delay reductions [16] and could provide operational benefits to air carriers and passengers [21, 20]. However, the implementation of these measures has been limited in the United States due to the constraints on flight schedules they would impose and the resulting potential distortions in airline competition.

In this paper, we examine a third way of reducing the demand-capacity mismatches, namely through improvements in *capacity utilization* at the tactical level to improve operational efficiency. For a fixed schedule of flights on a given day, airport congestion can be mitigated through effective use of available capacity to best operate takeoffs and landings. Airport efficiency has been shown to depend primarily on the runway configuration in use, *i.e.*, the set of active runways on which landings and takeoffs are operated, and on the relative proportion of landings and takeoffs at any given time [19]. We develop an airport congestion mitigation tool that minimizes congestion costs at the tactical level by jointly controlling runway configurations and arrival and departure service rates to best operate scheduled landings and takeoffs. This joint control is most critical when different sets of arrival and departure service rates can be achieved under different runway configurations. This may occur in practice if, for instance, a particular configuration would allocate one runway to arrivals and two runways to departures, while an alternative configuration would allocate two runways to arrivals and a single one to departures.

Given the stochastic nature and the variability over time of airport operations, this control of runway configurations and of arrival and departure service rates is exercised dynamically through the course of the day. Indeed, the formation and propagation of airport queues is inherently stochastic, as queues do not depend solely on the schedule of flights and the availability of airport capacity, but are also shaped by the exact sequencing of arrivals and departures, by the mix of aircraft, by human factors in aircraft operations, etc. Past research has demonstrated that dynamic and stochastic models of airport congestion approximate quite accurately the extents of congestion observed in practice at the busiest US airports [17] and provide higher estimates of flight delays than deterministic ones [13]. Moreover, operating conditions, such as weather and winds, are also stochastic and directly impact airport operations: winds constrain the set of runways that can be used at any time, while meteorological conditions influence the efficiency of airport operations. The control is therefore exercised dynamically throughout the day of operations, as a function of observed arrival and departure queue lengths, of arrival and departure schedules and of observed operating conditions.

In practice, the selection of runway configurations is one major task faced by air traffic controllers and is primarily made on the basis of experience. Ramanujam and Balakhrishnan developed a descriptive statistical model that identifies six factors as the key drivers of these decisions at New York's airports: inertia, wind speed and directions, arrival and departure schedules, noise abatement procedures, configuration switch proximity and inter-airport coordination [18]. Most of these factors are taken into account in the decision-making framework developed in this paper.

From a prescriptive standpoint, the optimization of airport capacity utilization at airports has been an important topic in air traffic flow management research. Gilbo first explicitly considered the problem of strategically allocating airport capacity between arrivals and departures using the representation of the non-increasing relationship between arrival capacity and departure capacity by means of a *Capacity Envelope* [9, 10]. This framework was later extended to allocate available capacity at the multi-airport level in the development of the Ground Delay Program [12, 11]. More recent studies addressed the problem of *jointly* selecting runway configurations and arrival and service rates by introducing different Capacity Envelopes for different runway configurations. Bertsimas *et al.* developed a mixed integer program that solves, at the strategic level, the problem of determining the optimal schedule of runway configurations and of arrival and departure service rates at the beginning of a day of operations [5]. Li and Clarke developed an alternative decision-making framework that dynamically controls runway configurations under stochastic wind conditions [15]. In both cases, it was demonstrated that the joint control of runway configurations and of the arrival / departure balance can in fact lead to significant congestion cost savings. Both of these approaches considered deterministic queue dynamics.

The first contribution of this paper is the development of a new tactical approach to solving the problem of selecting runway configurations and of balancing the service rates of arrivals and departures at major airports under stochastic operating conditions. A stochastic and dynamic queuing model is included in the decision-making framework, so that the control is exercised dynamically during the day as a function of the observed extent of congestion on the ground and in the air. For instance, let us consider a given period of the day when more arrivals than departures are scheduled. Traditional approaches to runway configuration selection might yield a solution where a runway configuration that gives priority to arrivals over departures is used during that period. If, however, a large departure queue is observed on the ground at the beginning of the considered period, while no arriving aircraft are queuing in the air, then it might be beneficial to choose an alternative runway configuration in order to enhance the departure throughput and thus alleviate ground congestion. This dynamic approach might then yield operational benefits in the face of queue uncertainty. Moreover, the stochasticity of operating conditions, including weather and wind conditions, and their impacts on airport operations are also embedded within the decision-making framework developed in this paper.

A second contribution of this paper is the integration of the control of arrival and departure service rates as a part of macroscopic models of airport congestion. Most of the models that quantify the relationships between flight schedules, airport capacity and flight delays, typically exercise no control on the arrival and departure service rates and assume exogenous values of these two quantities [13, 14]. One recent study quantified the impact of runway configuration changes on predicted delays [16], but did so manually and not systematically. The approach developed in this paper, by optimizing the control of arrival and departure service rates at the *tactical* level, allows greater understanding of how these rates are controlled and of how they can be varied during the day of operations as a function of arrival and departure schedules. This understanding can subsequently be used to improve the predictive models of airport congestion at the *strategic* level.

The remainder of this paper is organized as follows. Section 2 formulates the problem as a finite-horizon dynamic programming problem and describes the model inputs, the state variables, the control exercised, the cost function and the system dynamics. Section 3 introduces the dynamic programming algorithm, which can be applied off-line to determine optimal policies. In addition, we present an approximation scheme based on a one-step look-ahead algorithm, which is shown to converge to near-optimal policies. This fast approximate algorithm enables on-line implementation of the model when new information becomes available during the course of the day (*e.g.*, dynamic schedule updates, weather forecasts etc.). Section 4 applies this model to the John F. Kennedy International Airport (JFK), which has experienced the highest average delays nationwide in recent years. We present optimal policies and show that the control scheme can reduce flight delays by an order of 15% to 20%. It is also demonstrated that this approach can be used to improve the selection of arrival and departure service rates in macroscopic models of airport congestion. Finally, Section 5 summarizes the findings of the paper.

2 Model Formulation

We formulate the dynamic control of runway configurations and arrival and departure service rates as a finite-horizon dynamic programming model. A day of operations, between 6 a.m. and 12 a.m., is divided into $T = 72$ periods of length $S = 15$ minutes each. We index these time periods by $t = 1, ..., T$. Observations and decisions are made at the beginning of each period. Note that operations between 12 a.m. and 6 a.m. are not considered in this model since they are typically not capacity-constrained and decisions are based on noise abatement procedures and other environmental concerns.

2.1 Model Inputs

The model takes as inputs the schedule of landings and takeoffs. During each 15-minute period t of the day, we denote by x_t (resp. y_t) the number of aircraft scheduled to land (resp. to take off) at the considered airport.

The model considers the capacity of each of the different runway configurations of the airport. Airport capacity is typically represented by means of a *Capacity Envelope*, which defines the relationship between the *maximal* number of arrivals and the *maximal* number of departures that can be feasibly operated per unit of time [9]. Given the inherent uncertainty in airport operations, this representation has recently been modified to account for throughput stochasticity: the *Operational Throughput Envelope* represents the relationship between the *average* arrival rate and the *average* departure rate in the presence of continuous demand, for a given runway configuration [19]. This representation takes into account the traffic mix observed in practice, such as, for instance, different aircraft types, different sequencing of arrivals and departures, etc. Since airport operations are substantially impacted by weather conditions, we consider two distinct Operational Throughput Envelopes for each runway configuration: the *VMC envelope* (resp. the *IMC envelope*) represents the capacity of the runway configuration in "Visual Meteorological Conditions" (VMC) (resp. in "Instrument Meteorological Conditions" (IMC)).

A schematic representation of an airport's Operational Throughput Envelope for a given runway configuration is provided in Figure 1. Points 1 and 2 represent two sets of average arrival and departure service rates in VMC and Point 3 represents a set of average service rates in IMC. Two immediate observations are noteworthy. First, a higher average throughput can be achieved in VMC than in IMC. Second, the relationship between arrival service rate and departure service rate for a given runway configuration is non-increasing.

2.2 State Variables

Decisions in each period are based on the observed extent of congestion, on the runway configuration previously in use and on observed operating conditions. Operating conditions, including weather and winds, are assumed to be observed at the beginning of each time window and not to change over one 15-minute period. At the beginning of period t, the state is described by the following variables:

- Arrival Queue Length a_{t-1}: Number of arriving aircraft queuing in the air at the end of the previous period

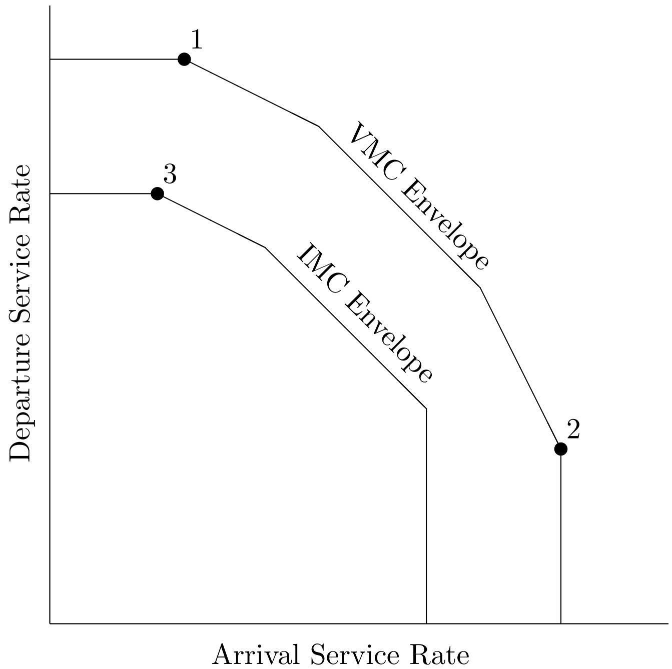


Figure 1: Schematic representation of the VMC and IMC Operational Throughput Envelopes of a runway configuration

- Departure Queue Length d_{t-1}: Number of departing aircraft queuing on the ground at the end of the previous period
- The runway configuration in use at period $t-1$, denoted by RC_{t-1}
- Weather conditions for the next period, denoted by $wc_t \in \{VMC, IMC\}$
- The wind state ws_t, which determines the set of runway configurations that *can* be used at any time. According to FAA-specified safety requirements, a runway can only be used if the tailwind and crosswind do not exceed the thresholds of 5 knots and 20 knots, respectively. A wind state is defined as a set of wind vectors, *i.e.*, a set of wind strengths and wind directions, such that the same set of runways satisfy the threshold requirements. This approach follows the procedure developed by Li and Clarke [15]. Table 1 defines the 13 wind states at JFK by specifying, for each of these states, the runways that meet the safety thresholds and those that do not [1]. For instance, State 1 corresponds to a situation of calm winds, so that flights can be operated on all runways. In contrast, State 9 corresponds to the case of strong winds from the South, in which case runways $4L$, $4R$, $31L$ and $31R$ face above-threshold tailwinds.

[1] These 13 wind states account for observed conditions during 99.7% of the periods between 2007 and 2010. The remaining 0.3% corresponds to rare situations where winds are so strong that no runway orientation satisfies the FAA threshold requirements. Note, finally, that, since the tailwind threshold is lower than the crosswind threshold, the system is never in a state where only the four runways $4L$, $4R$, $22L$ and $22R$ can be used or in a state where only the four runways $13L$, $13R$, $31L$ and $31R$ can be used.

Table 1: Definition of wind states at JFK: set of runways that can be used per wind state

Wind States	1	2	3	4	5	6	7	8	9	10	11	12	13
$4L$, $4R$	✓	✓	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗
$22L$, $22R$	✓	✓	✓	✗	✓	✗	✗	✓	✓	✗	✓	✗	✗
$13L$, $13R$	✓	✓	✗	✓	✓	✓	✗	✓	✗	✗	✗	✓	✗
$31L$, $31R$	✓	✗	✓	✓	✓	✗	✓	✗	✓	✗	✗	✗	✓
Proportion	10.6%	7.6%	16.0%	8.9%	15.6%	2.5%	13.6%	8.1%	8.8%	10.3%	0.8%	0.5%	5.7%

2.3 Decision Variables

The decision exercised at each period has two components:

- The runway configuration for the next period RC_t. This choice is constrained by the wind state: a runway configuration can only be used if each of its runways meets the threshold requirements (Table 1). The set of runway configurations that can be selected when the wind state is ws_t is denoted by $\mathcal{RC}(ws_t)$. For each one of JFK's 8 main runway configurations, Table 2 indicates the set of wind states for which the configuration can be used [2]. When, for example, JFK is in wind state 1, all 8 configurations are eligible for use. The selection of the runway configuration and the observation of the weather state determine, in turn, the Operational Throughput Envelope for the next period.

Table 2: Set of JFK runway configurations that can be selected per wind state

Wind States	1	2	3	4	5	6	7	8	9	10	11	12	13
$13L, 22L\|13R$	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗
$31L, 31R\|31L$	✓	✗	✓	✓	✓	✗	✓	✗	✓	✗	✗	✗	✓
$22L\|22R, 31L$	✓	✗	✓	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗
$4R\|4L, 31L$	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗
$13L\|13R$	✓	✓	✗	✓	✓	✓	✗	✓	✗	✗	✗	✓	✗
$31R\|31L$	✓	✗	✓	✓	✓	✗	✓	✗	✓	✗	✗	✗	✓
$22L\|22R$	✓	✓	✓	✗	✓	✗	✗	✓	✓	✗	✓	✗	✗
$4R\|4L$	✓	✓	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗

- The service rates at which arrivals and departures are served, controlled among the outermost set of achievable service rates for the selected runway configuration and the observed weather conditions. For instance, in the case represented in Figure 1, the decision-maker can decide to operate, in VMC, with the arrival and departure service rates corresponding to point 1 or to point 2 or to any other point on the envelope. The decision-maker in fact selects the arrival service rate for the next period, denoted by $\mu_t^a \in \{0, ..., A_{RC_t,wc_t}\}$. The upper bound of this choice A_{RC_t,wc_t} corresponds to the highest arrival throughput that can be realized in the

[2] A runway configuration is denoted by the set of active runways on which landings are operated, followed by the set of active runways on which takeoffs are operated.

selected runway configuration and observed weather conditions. In turn, the departure service rate, is determined by the Operational Throughput Envelope corresponding to the selected runway configuration and the observed weather conditions. We denote by $\mu_t^d = \Phi_{RC_t, wc_t}(\mu_t^a)$ the departure service rate associated with the arrival service rate μ_t^a when the airport operates in runway configuration RC_t and in weather conditions wc_t.

2.4 System Dynamics

2.4.1 Queue Dynamics

Arrival and departure queues are modeled by means of two dynamic and stochastic $M(t)/E_k(t)/1$ queuing models, which have been shown to describe accurately the queue dynamics observed in practice [17, 14]. The demand processes are modeled as Poisson processes, whose rates are determined by the number of arriving and departing flights scheduled per 15-minute time window. The service processes are modeled as Erlang processes of order k, whose rates are controlled by the decision-maker. A value of $k = 3$ is used in practice [14]. The model is non-stationary: the demand and service rates are time-varying. These rates are modeled as constant over a 15-minute time-window t and are thus denoted by λ_t and μ_t, respectively. Note that the service rates of the arrival and departure processes are not independent from each other since they are jointly determined by the Operational Throughput Envelope. However, the stochastic evolution of the arrival queue is independent from that of the departure queue.

The state-transition diagram of the $M(t)/E_k(t)/1$ queuing system is shown in Figure 2. This relies on the characterization of an Erlang process of order k and rate μ as the succession of k independent and Markovian "stages of work" at rate $k\mu$. The state of the system is therefore defined by the number of remaining of stages of work, denoted by i. The evolution of the queuing system over time period t is described by a system (1) of Chapman-Kolmogorov first-order differential equations. In these equations, we introduce a time index s that varies over the 15-minute time window. In other words, the index s varies, during period t, between $(t-1)S$ and tS, where S denotes the length of a time-window (in this case, 15 minutes). These equations determine the time evolution of the state probabilities $P_i(s)$, which characterize the probability of being in state i at time s. The practical queue capacity is denoted by N. The system is assumed to be empty at the beginning of the day of operations.

$$
\begin{array}{rcll}
\frac{dP_0(s)}{ds} & = & -\lambda_t P_0(s) + k\mu_t P_1(s) & \\
\frac{dP_i(s)}{ds} & = & -(\lambda_t + k\mu_t) P_i(s) + k\mu_t P_{i+1}(s) & \forall i \in \{1, ..., k\} \\
\frac{dP_i(s)}{ds} & = & \lambda_t P_{i-k}(s) - (\lambda_t + k\mu_t) P_i(s) + k\mu_t P_{i+1}(s) & \forall i \in \{k+1, ..., (N-1)k\} \\
\frac{dP_i(s)}{ds} & = & \lambda_t P_{i-k}(s) - k\mu_t P_i(s) + k\mu_t P_{i+1}(s) & \forall i \in \{(N-1)k+1, ..., kN-1\} \\
\frac{dP_{kN}(s)}{ds} & = & \lambda_t P_{k(N-1)}(s) - k\mu_t P_{kN}(s) &
\end{array}
\tag{1}
$$

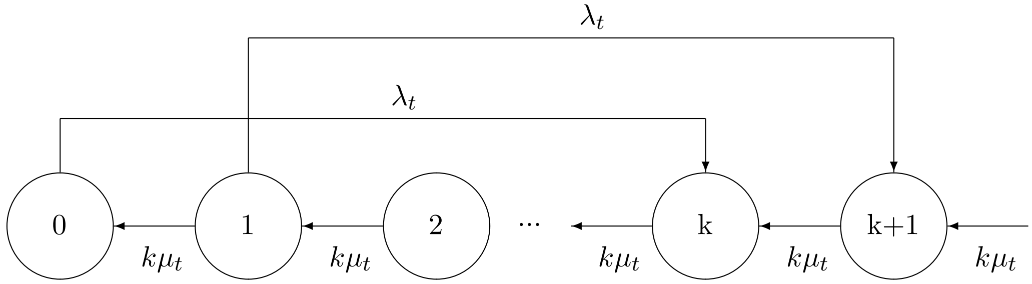


Figure 2: State-transition diagram of the $M(t)/E_k(t)/1$ queuing system

Because of the computational requirements of this system of differential equations, we solve it off-line. We store in a look-up table the transition probabilities over a full time window, *i.e.*, $\mathbb{P}(S_{t+1} = j|S_t, \lambda_t, \mu_t)$, for all j, S_t, λ_t, μ_t, where S_t denotes the state of the system at the end of period t. In this way, we do not have to re-solve this system of equations at each iteration of the dynamic programming algorithm.

Since decision-makers cannot observe the fine-grain state of the system, *i.e.*, the number of remaining "stages of work", but observe the queue length instead, we proceed by aggregation. In other words, we map the state transition probabilities into queue length transition probabilities. The relationship between the states of the system and observed queue lengths is depicted in Figure 3. Since this mapping is not bijective, we make the following assumption: at the beginning of each period, the system is in one of the states $\{mk, 0 \leq m \leq N\}$, *i.e.*, no aircraft is being served at the beginning of a period. The motivation for this choice is that the average service rate of arrivals and departures is typically significantly larger than 1 per period. Under this assumption, the queue length transition probabilities are computed as follows, where q_t (resp. S_t) represents the queue length (resp. the state of the system) at the end of period t, for all λ_t, μ_t:

$$\begin{aligned} \mathbb{P}(q_{t+1} = 0|q_t, \lambda_t, \mu_t) &= \mathbb{P}(S_{t+1} = 0|S_t = q_t k, \lambda_t, \mu_t), \forall q_t, \lambda_t, \mu_t \\ \mathbb{P}(q_{t+1} = j|q_t, \lambda_t, \mu_t) &= \textstyle\sum_{s=1}^{k} \mathbb{P}(S_{t+1} = (j-1)k + s|S_t = q_t k, \lambda_t, \mu_t), \quad \forall j = 1, ..., N, q_t, \lambda_t, \mu_t \end{aligned} \tag{2}$$

2.4.2 Runway Configuration Changes

Switching runway configurations is a challenging operational procedure that requires extensive coordination among several airport stakeholders, including airlines and air traffic controllers. Most importantly, queuing aircraft need to be re-routed, which may lead to operational inefficiencies and to the closure of runway operations for some time. We represent the cost associated with runway configuration changes by a time period of idleness, of length denoted by τ_I, during which arriving and departing aircraft may join the queue at an unchanged rate – determined by scheduling

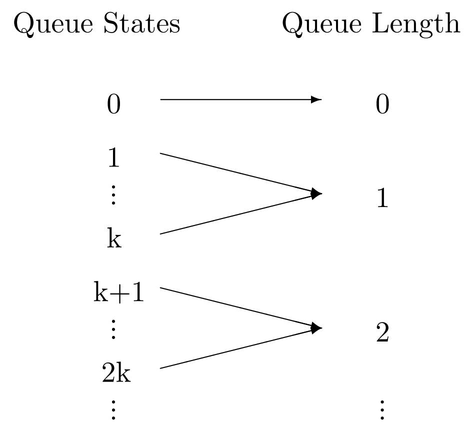


Figure 3: Mapping between queue states and queue length

levels – but no flight is operated. In other words, we assume that, if the decision-maker decides to change runway configurations and chooses arrival and departure service rates $\mu^{(d)}$ and $\mu^{(a)}$ from the Operational Throughput Envelope of the new runway configuration, then the arrival and departure service rates are both equal to 0 for a given time period τ_I, after which the arrival and departure service rates are, respectively, equal to $\mu^{(d)}$ and $\mu^{(a)}$ until the end of the 15-minute time window. This situation is depicted in Figure 4, where a runway configuration change takes place at the beginning of the t^{th} period.

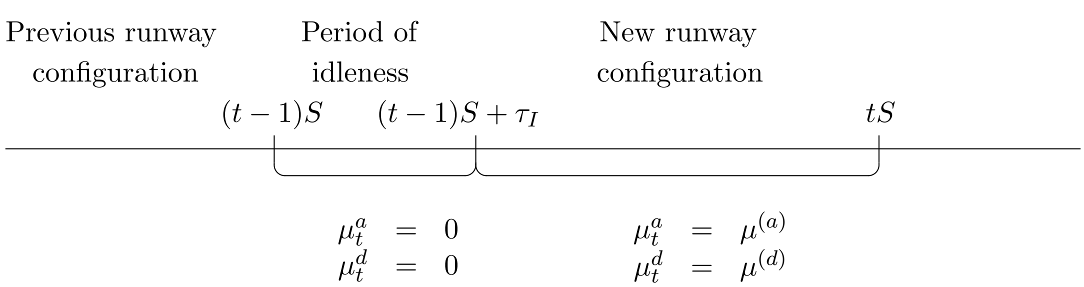


Figure 4: Schematic representation of a runway configuration change at the beginning of period t

Therefore, an inherent trade-off is faced by the decision-maker: if efficiency can be improved by changing runway configurations at a certain time of the day, this comes at the cost of a temporary idleness of the runway system and consequent initial build-up of the arrival and departure queues. The attractiveness of runway configuration changes naturally depends on the length τ_I of the period of idleness.

Note, also, that the duration of the idleness resulting from a runway configuration change may vary as a function of the "proximity" of the two consecutive runway configurations. For instance, simply activating an additional runway for arrivals or departures (such as activating Runway $4R$ when $4L$ is already in use) will typically be less disruptive than a move to a very different configuration that requires a change in the entire flow of arriving and departing aircraft. For this reason, we denote by $\tau_I^{RC_1,RC_2}$ the time of idleness following the change from runway configuration RC_1 to runway configuration RC_2. For all runway configurations RC, one obviously has $\tau_I^{RC,RC} = 0$.

2.4.3 Weather and Wind Dynamics

Weather variability is modeled by means of a two-state Markov chain [14]. We denote by p (resp. q) the transition probability from state VMC to state IMC (resp. from state IMC to state VMC). Its transition diagram is shown in Figure 5 and its transition matrix P is given by:

$$P = \begin{array}{c} \\ VMC \\ IMC \end{array} \!\!\begin{array}{c} \begin{array}{cc} VMC & IMC \end{array} \\ \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array}$$

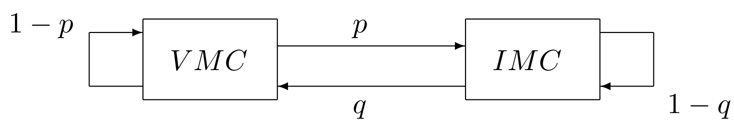


Figure 5: Transition diagram of the weather Markov chain

Note that p and q may vary from one day of operations to another, as a function of perceived meteorological conditions. We consider three types of days:

- "All-VMC days", during which the system stays in VMC throughout the day: $p = 0$
- "All-IMC days', during which the system stays in IMC throughout the day': $q = 0$
- "VMC-IMC days", during which meteorological conditions may vary through the course of the day: p (resp. q) is estimated by its maximum likelihood estimator, *i.e.*, the empirical ratio of the number of transitions from VMC to IMC (resp., from IMC to VMC) over the number of periods in VMC (resp. in IMC).

Similarly, wind dynamics are also modeled as a Markov chain. The transition probability from State i to State j is estimated by its maximum likelihood estimators, *i.e.*, $\frac{n_{ij}}{n_i}$, where n_{ij} (resp. n_i) designates the number of transitions from State i to State j (resp. the number of periods in State i) [15].

2.5 Cost Function

The control strategy aims at minimizing congestion costs, which are typically modeled as a non-decreasing function of the queue length with non-decreasing marginal costs. In this paper, we consider a quadratic cost function of the arrival and departure queue, since the expected total delay scales quadratically with the number of queuing aircraft. Moreover, the costs associated with arrival queues are weighted by a factor α. Arrival delays and departure delays may indeed be expected to have different costs, as arriving aircraft can be more challenging and expensive to hold in queue than departing aircraft. Therefore, a value of α larger than 1 might be used.

The cost function is written as follows:

$$\alpha \sum_{t=1}^{T} a_t^2 + \sum_{t=1}^{T} d_t^2. \tag{3}$$

2.6 Dynamic Programming Formulation

As described in Section 2.3, we denote by $\mathcal{RC}(ws_t)$ the set of runway configurations that can be selected in wind state ws_t and by A_{RC_t,wc_t} the maximal arrival rate that can be handled in runway configuration RC and in weather conditions $wc \in \{VMC, IMC\}$. We denote the cost-to-go function by $J_t(a_{t-1}, d_{t-1}, RC_{t-1}, wc_t, ws_t)$, which represents the expected total cost of being in state $(a_{t-1}, d_{t-1}, RC_{t-1}, wc_t, ws_t)$ at the beginning of period t. The decision-maker minimizes the sum of the expected congestion costs experienced at the end of period t, *i.e.*, $\alpha\mathbb{E}\left(a_t^2\right) + \mathbb{E}\left(d_t^2\right)$, and the future congestion costs from period $t+1$ onward, *i.e.*, $\mathbb{E}\left(J_{t+1}\left(a_t, d_t, RC_t, wc_{t+1}, ws_{t+1}\right)\right)$. The dynamic programming equation is written as follows:

$$\begin{aligned} J_t(a_{t-1}, d_{t-1}, RC_{t-1}, wc_t, ws_t) = & \min_{\substack{RC_t \in \mathcal{RC}(ws_t) \\ \mu_t^a \in \left[0, A_{RC_t,wc_t}\right]}} \Big(\alpha\mathbb{E}\left(a_t^2\right) + \mathbb{E}\left(d_t^2\right) \\ & + \mathbb{E}\left(J_{t+1}\left(a_t, d_t, RC_t, wc_{t+1}, ws_{t+1}\right)\right) \Big), \forall t = 1, ..., T \end{aligned} \tag{4}$$

$$J_{T+1}(a_T, d_T, RC_T, wc_{T+1}, ws_{T+1}) = 0 \tag{5}$$

The arrival queue a_t at the end of period t depends on the number of scheduled arrivals x_t during period t, on the duration of the period of idleness $\tau_I^{RC_{t-1},RC_t}$, if any, on the arrival service rate μ_t^a during period t and on the previous arrival queue a_{t-1}. Similarly, the departure queue d_t depends on the variables y_t, $\tau_I^{RC_{t-1},RC_t}$, μ_t^d and d_{t-1}. A summary of the dependencies described in Section 2.4 is provided below – full lines denote system evolution and dashed lines denote constraints on the decisions.

$$
\begin{array}{rcl}
ws_t & \dashrightarrow & RC_t \\
RC_{t-1}, RC_t & \longrightarrow & \tau_I^{RC_{t-1},RC_t} \\
RC_t, wc_t & \dashrightarrow & \mu_t^a, \mu_t^d = \Phi_{RC_t,wc_t}\left(\mu_t^a\right) \\
x_t, \tau_I^{RC_{t-1},RC_t}, wc_t, \mu_t^a, a_{t-1} & \longrightarrow & a_t \\
y_t, \tau_I^{RC_{t-1},RC_t}, wc_t, \mu_t^d, d_{t-1} & \longrightarrow & d_t
\end{array}
$$

3 Solution Algorithm

3.1 Experimental Setup

We consider the schedule at JFK of Friday, July 13, 2007, one of the busiest days in recent few years. The airport operated in VMC throughout that day. For this reason, the probability p of weather deterioration is set to 0 in this case. The schedule of landings and takeoffs at JFK is obtained from the Aviation Performance Metrics (APM) database [8] and shown in Figure 6.

The Operational Throughput Envelopes of the 8 major runway configurations at JFK in VMC are obtained from the PhD thesis of Simaiakis [19]. We categorize these 8 runway configurations into 4 sets of two runway configurations each:

- Configurations with two arrival runways and one departure runway: $13L, 22L|13R$ and $31L, 31R|31L$. These configurations can achieve the largest arrival rates.

- Configurations with one arrival runway and two departure runways: $22L|22R, 31L$ and $4R|4L, 31L$. These configurations achieve the highest departure rates when few landings are operated.

- Configurations with two independent parallel runways: $13L|13R$ and $31R|31L$. These configurations achieve a higher departure rate than configurations with two arrival runways and one departure runway when few landings are operated, since departures are less constrained by aircraft landing and taxiing in. However, they achieve a lower throughput than configurations with two arrival runways and one departure runway when a large number of landings are operated.

- Configurations with two more closely spaced runways: $22L|22R$ and $4R|4L$. These configurations achieve the lowest service rates.

Figure 7 shows the Operational Throughput Envelope of each of these 8 runway configurations. For each of the 4 sets described above, we plot the envelope of the configuration that achieves the highest service rates with a full line and the envelope of the other configuration with a dashed line. The same color is used for configurations belonging to the same set. In addition, each dot represents the number of scheduled landings and takeoffs per 15-minute period. Note that the scheduling levels exceed airport capacity during a large number of periods. This is likely to lead to significant flight delays. Note, also, that the proportion of arrivals and departures is highly variable over the course of the day: as Figure 6 indicates, significantly more departures than arrivals are scheduled in the morning, while the reverse is true in early afternoon.

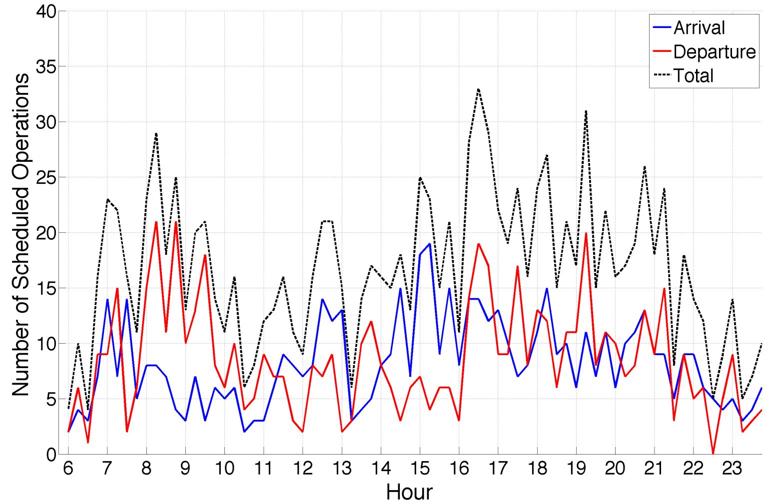


Figure 6: Arrival and departure schedules at JFK on 07/13/2007

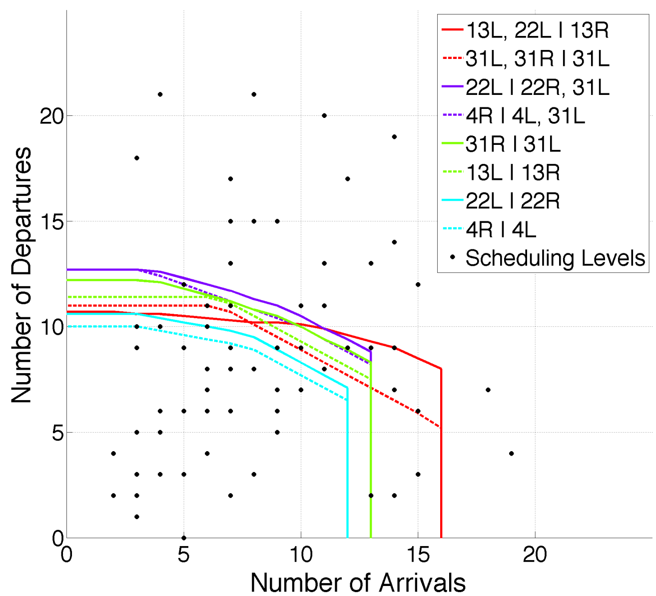


Figure 7: VMC Operational Throughput Envelopes of the main runway configurations at JFK

Unless otherwise specified, we implement this model using the same value of $\tau_I^{RC_1,RC_2}$ for all pairs of runway configurations $RC_1 \neq RC_2$. To simplify notation, we denote this time of idleness by

τ_I in the remainder of the paper. This assumption can be easily modified to include differing values of this parameter for different configuration pairs. We provide an example where this assumption is relaxed in Section 4.1.

The value of the practical queue capacity, denoted by N (Section 2.4.1), is fixed at 30. Note that a small value of N may lead to underestimation of the probabilities of large arrival and departure queues. On the other hand, a large N will increase computational times, as the number of states scales quadratically with N.

3.2 Exact Dynamic Programming Algorithm

First, the exact dynamic programming algorithm is implemented using the solution concept of backward induction [2, 3]. In other words, the optimal policy in the final period, *i.e.*, between 23:45 and 24:00, is computed for all possible states, *i.e.*, for all possible arrival and departure queue lengths that can be observed at 23:45, for all possible runway configurations that can be used in the previous period, *i.e.*, between 23:30 and 23:45, and for all possible weather and wind conditions. This provides optimal costs in the final period as a function of the state of the system at the beginning of the final period. This cost is then used to compute optimal policies in the second-to-last period, as a function of the state of the system at 23:30. This process is then repeated until the optimal policies for all periods have been derived.

The implementation of this algorithm can be performed off-line. The optimal policy can then be applied throughout the day of operations. Nonetheless, the completion time of the exact dynamic programming algorithm is equal to approximately 90 minutes on a laptop computer and thus exceeds the time frame of actual decision-making by air traffic controllers. This might prevent it from being applicable in practice if it needs to be updated on-line when new sources of information become available. For this reason, we have implemented an approximate algorithm that accelerates completion and consequently enables the algorithm to be updated on-line.

3.3 One-Step Look-Ahead Algorithm

The on-line implementation of the model presented in this paper requires very fast execution. Indeed, at the beginning of each period, the selection of the runway configuration and of the arrival and departure service rates must follow very quickly the observation of arrival and departure queue lengths. For this reason, we implement in this section a one-step look-ahead algorithm based on *ex ante* evaluation of the cost-to-go function, obtained from the exact dynamic programming algorithm. The main advantage of this algorithm is that its on-line execution is almost instantaneous and thus well suited to the actual problem faced by the decision-maker.

We consider, for a given day of operations, the off-line solution of the problem, which is computed with the original model inputs, *i.e.*, the original schedule of landings and takeoffs, the original weather transition probabilities, etc. This provides, before the beginning of operations, the optimal policy for each period, denoted by $\left(\widehat{RC}_t, \widehat{\mu_t^a}\right)$, and the expected cost-to-go function, denoted by $\widehat{J}$. If no significant update occurs during the day, then this optimal policy can be applied at the beginning of any time period following the observation of queues and operating conditions. However, this policy might no longer be optimal in the face of dynamic disturbances that may occur during the day of operations. Examples of such disturbances include schedule updates, which might arise from the occurrence of upstream delays in previous flight legs or the initiation of Ground Delay Programs, or changes in weather forecasts that might impact future operations.

In the face of such disturbances, we implement a one-step look-ahead algorithm based on the evaluation of the cost-to-go function $\widehat{J}$ resulting from the original application of the exact dynamic programming algorithm. In other words, we choose, at the beginning of period t, the policy that minimizes expected total costs, assuming that costs from period $t+1$ onward are given by $\widehat{J}$. The policy for the considered period, denoted by $\left(\widetilde{RC}_t, \widetilde{\mu_t^a}\right)$, is therefore determined as follows [4]:

$$\left(\widetilde{RC}_t, \widetilde{\mu_t^a}\right) = \underset{\substack{RC_t \in \mathcal{RC}(ws_t) \\ \mu_t^a \in \left[0, A_{RC_t, wc_t}\right]}}{\arg\min} \left(\alpha \mathbb{E}\left(a_t^2\right) + \mathbb{E}\left(d_t^2\right) + \mathbb{E}\left(\widehat{J}_{t+1}\left(a_t, d_t, RC_t, wc_{t+1}, ws_{t+1}\right)\right)\right) \quad (6)$$

In order to test the performance of this approach, we simulate a schedule disturbance as follows: at each period, we introduce a schedule perturbation randomly sampled from the integers within ε of the original number of scheduled arrivals and departures. For instance, if 10 arrivals and 15 departures are originally scheduled during a given period, then we uniformly sample, for a value of $\varepsilon = 20\%$, the updated number of arrivals (resp. departures) from the five integers between 8 and 12 (resp. from the seven integers between 12 and 18). Note that the expected value of the total number of flights in the updated schedule is identical to that of the original schedule. We then compare the expected cost per stage resulting from the application of (a) the optimal policy with the updated schedule, denoted by $\left(\overline{RC_t}, \overline{\mu_t^a}\right)$, which might be too computationally time-consuming to be determined on-line, (b) the optimal policy with the original schedule, $\left(\widehat{RC}_t, \widehat{\mu_t^a}\right)$, which can be obtained off-line and then applied during the course of the day and (c) the policy produced by the one-step look-ahead algorithm (Equation 6), $\left(\widetilde{RC}_t, \widetilde{\mu_t^a}\right)$, which can be computed almost instantaneously on-line after the schedule update. Results are reported in Table 3, for different values of ε. These results obviously depend on the particular realization of the random schedule perturbation, but nonetheless provide insights into the robustness of the approach.

First, note that the optimal policy computed off-line with the original schedule, $\left(\widehat{RC}_t, \widehat{\mu_t^a}\right)$, performs reasonably well, even after schedule updates. Even for the largest random schedule perturbations, this policy still results in expected congestion costs within 10% of the optimal congestion costs. Moreover, the one-step look-ahead algorithm significantly improves the performance over the original policy: the implementation of policy $\left(\widetilde{RC}_t, \widetilde{\mu_t^a}\right)$ results in expected congestion costs that

Table 3: Comparison of available policies before and after the one-step look-ahead algorithm to optimal policies, for different values of ε

Schedule Considered	Updated Schedule	Original Schedule	Updated Schedule
Algorithm	Exact	Exact	One-Step Look-Ahead
Policy	$\left(\overline{RC_t}, \overline{\mu_t^a}\right)$	$\left(\widehat{RC_t}, \widehat{\mu_t^a}\right)$	$\left(\widetilde{RC}_t, \widetilde{\mu}_t^a\right)$
Available on-line?	No	Yes	Yes
$\varepsilon = 10\%$	Baseline	+3.27%	+0.73%
$\varepsilon = 20\%$	Baseline	+3.39%	+0.77%
$\varepsilon = 30\%$	Baseline	+5.74%	+1.43%
$\varepsilon = 40\%$	Baseline	+6.06%	+1.62%
$\varepsilon = 50\%$	Baseline	+8.43%	+2.81%

exceed minimal costs by an order of only 1% to 2%, for different levels of schedule perturbations.

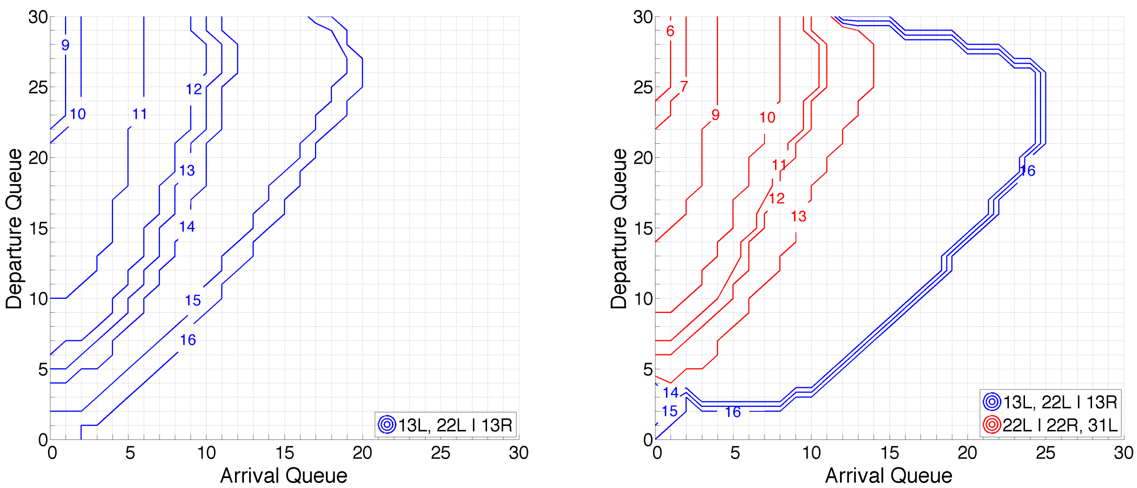
Thus, the one-step look-ahead algorithm offers fast and accurate performance for this problem: the combination of the cost-to-go evaluation, which can be done off-line with the original schedule of flights and, more broadly, with the original model inputs, and a single policy iteration (Equation 6) results in a close-to-optimal policy. This approximation scheme can thus be used as a flexible on-line decision-making tool to help minimize congestion costs by dynamically controlling runway configurations and arrival and departure service rates using the latest information available to decision-makers.

4 Results

We present in this section the results of the dynamic programming algorithm. Section 4.1 characterizes the optimal control of runway configurations and arrival and departure service rates. Section 4.2 shows the result of this control on expected arrival and departure queues and their sensitivity to several model parameters. Then, we quantify the operational benefits resulting from the implementation of the control by comparing, in Section 4.3, the optimal policy to heuristic policies. Last, we present how the control of arrival and departure service rates introduced in this paper can be used at the strategic level in models of airport congestion. To this end, we compare, in Section 4.4, the optimal control to baseline policies, where no control is exercised, and show that the control of arrival and departure service rates can also substantially improve the realism of macroscopic models of airport congestion.

4.1 Optimal Policies and Frequency of Decisions

The optimal policy derived from the dynamic programming algorithm is a function that determines the runway configuration and the arrival and departure service rates at each period of the day and in any state of the system. Figure 8 represents the contours of the optimal arrival rate μ_t^a and the optimal runway configuration RC_t for one specific period of the day t, at 15:30, for a value of $\tau_I = 5$ minutes, if observed weather conditions are VMC and if the airport operates in Wind State 1 (*i.e.*, all runway configurations can be used), as a function of the arrival and departure queue lengths that are observed at the beginning of the time window. Note that the departure rate is not represented here, but it is uniquely determined by the VMC Operational Throughput Envelope of the selected runway configuration. During the period considered, *i.e.*, between 15:30 and 15:45, 9 landings and 6 takeoffs were scheduled, and, more generally, more arrivals than departures were scheduled between 15:00 and 16:00 (Figure 6). Two cases are considered: Figure 8a (resp. Figure 8b) shows the optimal policy at 15:30 when the airport operated in the previous period, *i.e.*, between 15:15 and 15:30, in runway configuration $13L, 22L|13R$ (resp. in runway configuration $22L|22R, 31L$).



(a) $RC_{t-1} = 13L, 22L|13R$

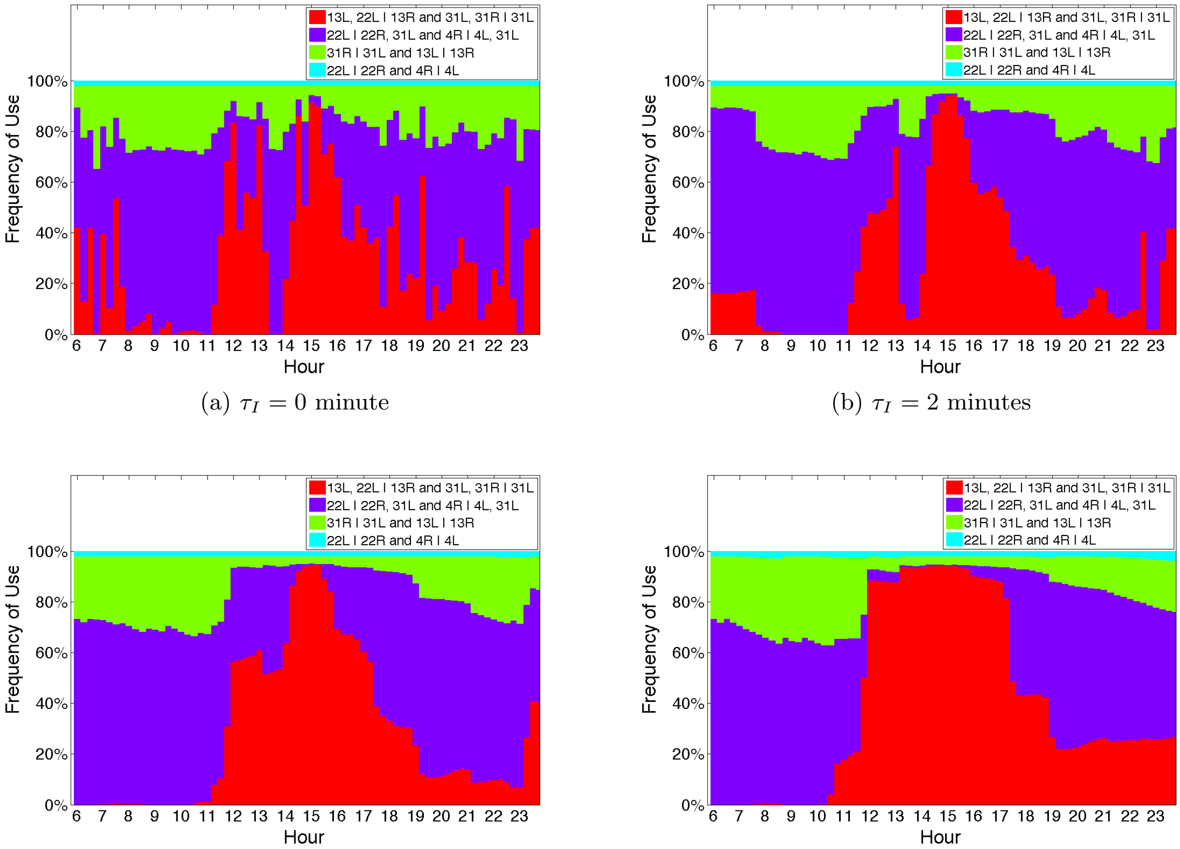
(b) $RC_{t-1} = 22L|22R, 31L$

Figure 8: Optimal runway configuration and arrival rate at 15:30 ($wc_t = VMC$, $ws_t = 1$)

Several observations can be made on the optimal policy. First, the arrival rate is typically non-decreasing as a function of the length of the arrival queue and non-increasing as a function of the length of the departure queue, with the exceptions of some "boundary effects" when queue lengths approach the practical queue capacity N. In other words, the larger the arrival queue, the more available capacity should be allocated to arriving aircraft. Moreover, the optimal policy depends on the runway configuration in use. Indeed, when the airport operates in configuration $13L, 22L|13R$, then the optimal policy is to stay in this configuration, which allocates two runways to arrivals and one runway to departures, in order to best operate the larger number of arriving aircraft during the considered period. In contrast, if the airport operates in configuration $22L|22R, 31L$, then the optimal policy may be either to switch to configuration $13L, 22L|13R$ if the departure queue is small enough, or to stay in configuration $22L|22R, 31L$ otherwise. In addition, whereas the selected arrival rate increases gradually as a function of the arrival queue length in the case where $RC_{t-1} = 13L, 22L|13R$ (Figure 8a), it increases discontinuously from 13 to 16 when $RC_{t-1} = 22L|22R, 31L$ (Figure 8b). Indeed, as the arrival queue becomes large enough, the decision-maker selects the largest arrival rate that can be achieved in the runway configuration in use (13 in this case), and the optimal policy then becomes invariant as the arrival queue length increases. However, when the arrival queue exceeds a certain threshold, then it might become beneficial to switch to another configuration, in this case to configuration $13L, 22L|13R$, if the operational benefits associated with the switch become large enough to offset the costs associated with the time period of idleness following the runway configuration change.

Figure 9 represents the frequency of use of each of the four considered categories of runway configurations defined in Section 2.1 for each of the 72 periods of a day of operations and for different values of τ_I, *i.e.*, different lengths of the time period of idleness following a runway configuration change. As expected, the frequency of use of these different configurations depends on the throughput they achieve (Figure 7). Since the three-runway configurations achieve the highest service rates, they are the most frequently used configurations. In contrast, the two-runway configurations are mostly used in adverse wind conditions when the airport can only operate on a small subset of runways. Nonetheless, configurations with two independent parallel runways are used much more frequently than configurations with two more closely parallel runways, which is due to the significant difference in the capacity of these configurations.

Moreover, the exact timing of use of the different configurations depends on the arrival and departure schedules as well as the evolution of the system through the day of operations. Importantly, note that no runway configuration is used 100% of the time at any period of the day. This indicates that the stochasticity of the evolution of the system has an impact on the optimal control, as suggested by Figure 8. Last, the use of different runway configurations depends substantially on the value of the parameter τ_I. Indeed, in the case where $\tau_I = 0$, runway configuration changes are very frequent to make the best possible use of available capacity. For larger values of τ_I, however, the cost associated with the idleness of the runway system is more likely to exceed the operational benefits associated with switching from one configuration to another and consequently runway configuration changes become less frequent. For instance, if $\tau_I = 0$, the decision-maker should operate, whenever possible, in a configuration with two departure runways between 13:00 and 14:00 to best serve the departure peak at that time (see Figure 6). As τ_I increases, decisions trade off congestion costs with increasing switching costs and therefore depend on the observed number of queuing aircraft on the ground and in the air. As a result, the frequency of a switch between 13:00 and 14:00 becomes smaller. When τ_I is equal to 10 minutes, then it is almost always optimal to stay in a configuration with two arrival runways and one departure runway for the entire period between 11:45 and 17:00.



(a) $\tau_I = 0$ minute

(b) $\tau_I = 2$ minutes

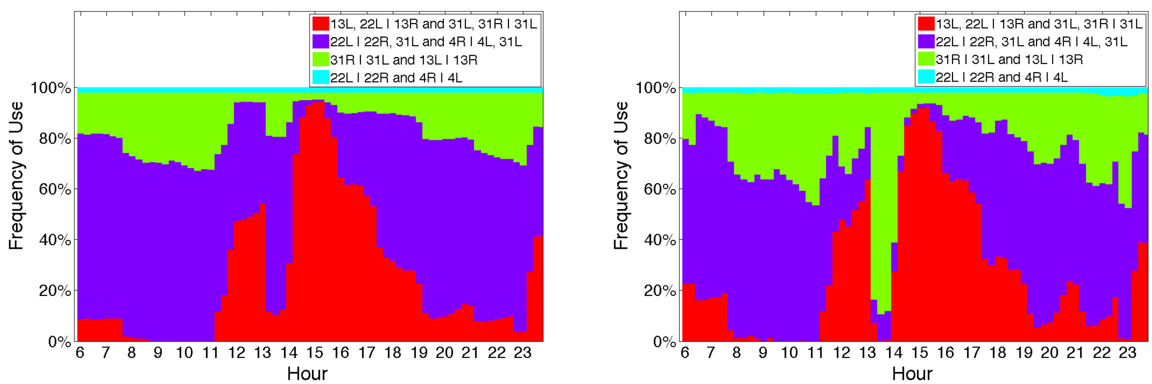
(c) $\tau_I = 4$ minutes

(d) $\tau_I = 10$ minutes

Figure 9: Frequency of use of each runway configuration for $\alpha = 1$ and variable τ_I

As mentioned in Section 2.4.2, the modeling framework developed in this paper enables us to introduce differentiated costs of runway configurations changes as a function of the proximity of configurations. We compare in Figure 10 the use of runway configurations when duration of idleness following a runway configuration change, τ_I, is uniform across runway configuration changes (Figure 10a) to the case where the values of τ_I vary with the runway configuration change considered (Figure 10b). More specifically, we assume in Figure 10b values of τ_I equal to (a) 1 minute if the switch merely disturbs operations by simply adding or removing a third runway (*e.g.*, from configuration $31L, 31R|31L$ to configuration $31R|31L$), (b) 5 minutes if the switch involves a 90-degree reorientation of the flow of aircraft (*e.g.*, from configuration $22L|22R, 31L$ to configuration $31R|31L$) and (c) 10 minutes if the switch involves a 180-degree reorientation of the flow of aircraft (*e.g.*, from configuration $31L, 31R|31L$ to configuration $13L|13R$). The value of τ_I considered in Figure 10a, equal to 3 minutes, is approximately equal to the average value of τ_I in the differentiated case, so the differences between Figures 10a and 10b essentially come from differences in the *distribution* of

τ_I across runway configuration changes. As expected, differentiated costs of runway configuration changes can in fact impact the optimal selection of runway configurations through the course of the day. For instance, it may be optimal to operate between 13:00 and 14:00 in a configuration with two independent parallel runways to best serve the larger number of takeoffs during this period, even though the departure throughput could be increased by using a configuration with two departure runways. This comes from the lower costs of switching from configuration $13L, 22L|13R$ to configuration $13L|13R$ or from configuration $31L, 31R|31L$ to configuration $31R|31L$ than the costs of switching to configuration $22L|22R, 31L$ or to configuration $4R|4L, 31L$. In this case, it may be optimal to increase moderately the departure throughput by using a "closer" configuration than to increase the departure throughput to a greater extent at a larger operational cost. The model implemented in this paper allows flexible calibration by airport operators of these operational costs associated with runway configuration changes.



(a) $\tau_I = 3$ minutes

(b) Differentiated $\tau_I^{RC_1,RC_2}$

Figure 10: Frequency of use of each runway configuration with uniform and differentiated values of τ_I

In addition to the selection of runway configurations, as shown in Figures 9 and 10, the model introduced in this paper controls the arrival and departure service rates at each period of the day. These are represented in Table 4 for six different periods. Note that variations in selected arrival and departure service rates, for a given runway configuration, are solely motivated by differences in prior queue evolution, and depend neither on the runway configuration previously in use nor on weather and wind conditions. For some periods of the day, *e.g.*, at 6:45, 13:30 and 15:15, this decision weakly depends on the state of the system. In these cases, the main control exercised is the selection of the runway configuration, primarily determined by previous runway configurations and wind-related constraints, but the optimal balance of arrivals and departures does not vary substantially from one observation to another. For instance, at 15:15, the decision-maker should most frequently choose the highest arrival rate available, while he should generally select an arrival rate equal to 9 or 10 landings per 15-minute period at 6:45 and an arrival rate equal to 5 or 6 landings per 15-minute period at 13:30. In some other cases, however, the optimal balance of arrivals and departure is highly variable and consequently depends on observed extents of congestion at the time of the decision (*e.g.*, at 17:45, 18:45 and 21:30). In these latter cases, both the optimal runway configuration and the optimal arrival and departure service rates might depend on prior evolution of the system.

Table 4: Policy frequency for six different periods of the day

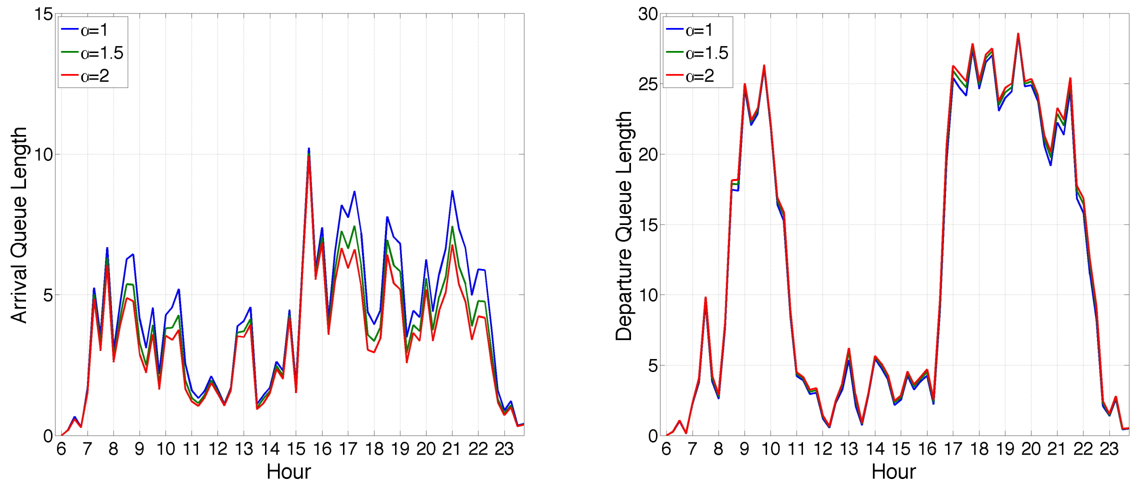
Policy			Time t					
RC_t	μ_t^a	μ_t^d	6:45	13:30	15:15	17:45	18:45	21:30
$13L, 22L\|13R$	16	8.0	–	–	41%	1%	2%	–
	14	9.0	–	–	–	2%	4%	–
	11	9.9	–	–	–	5%	6%	1%
	10	10.1	–	–	–	11%	8%	2%
	9	10.2	–	3%	–	6%	2%	3%
	6	10.4	–	25%	–	–	–	1%
	5	10.5	–	8%	–	–	–	1%
$31L, 31R\|31L$	16	5.2	–	–	46%	3%	4%	–
	15	5.9	–	–	6%	–	–	–
	7	10.7	–	6%	–	2%	1%	1%
	6	11.0	–	42%	–	1%	–	1%
$22L\|22R, 31L$	13	8.8	–	–	–	6%	11%	2%
	12	9.4	–	–	–	3%	4%	2%
	10	10.5	12%	–	–	4%	6%	4%
	9	11.0	38%	1%	–	11%	10%	11%
	7	11.7	–	1%	–	11%	4%	7%
	5	12.3	–	3%	–	–	–	2%
	4	12.6	–	–	–	–	–	7%
$4R\|4L, 31L$	13	8.2	–	–	–	4%	9%	1%
	10	10.0	21%	–	–	4%	6%	5%
	6	11.6	–	–	–	3%	1%	2%
	4	12.4	–	–	–	–	–	6%
	3	12.7	–	–	–	–	–	3%
$31R\|31L$	9	10.5	18%	–	–	1%	1%	3%
	4	12.1	–	–	–	–	–	5%
$13L\|13R$	8	10.5	4%	–	–	–	–	–
	7	11.1	–	–	–	1%	–	3%

These results demonstrate the path-dependency of the control of runway configurations and arrival and departure service rates introduced in this paper. At each period, the optimal policy depends on the state of the system at the time of the decision (Figure 8), which itself depends on previous decisions and on the prior evolution of the system. This includes, first, some deterministic components, such as the runway configuration in use. It also includes some exogenous stochastic components, such as the evolution of weather and wind conditions. For instance, configuration $4L|4L, 31L$ achieves a slightly lower departure throughput than configuration $22L|22R, 31L$ for any value of the arrival rate. Nonetheless, it might be used when the balance of arrivals and departures requires the use of a runway configuration with two arrival runways and one departure runway and when strong winds from the North prevent runways $22L$ and $22R$ from being used. Last, and perhaps most importantly, optimal policies depend on endogenous stochastic components, such as the length of the arrival and departure queues. These observations depend on previous decisions as well as the stochastic evolution of the system. This state stochasticity gives rise to some variability in the optimal control at any period of the day, as indicated in Figures 9 and 10 and in Table 4: both the selection of different runway configurations and the control of arrival and departure service rates depend on prior system evolution. This underscores the importance of considering queue stochasticity within the decision-making framework and its impact on optimal decisions.

4.2 Sensitivity of Arrival and Departure Queues

The control of runway configurations and arrival and departure service rates has a direct effect on arrival and departure queues. Figure 11 shows optimal expected arrival (Figure 11a) and departure (Figure 11b) queue lengths for different values of the weight on the cost of arrival queue α and for a duration of idleness following runway configuration changes equal to $\tau_I = 5$ minutes. In other words, it shows the effect of an increase of the relative cost of arrival queues, compared to departure queues, on expected arrival and departure queue lengths. As expected, larger values of α lead to the selection of policies that increase arrival throughput at the expense of the departure throughput and, therefore to lower expected arrival queue lengths and, reversely, to larger expected departure queue lengths. Note that arrival queues seem more sensitive to variations in α than departure queues: variations in α from 1 to 2 induce changes in peak expected arrival queue length of the order of 2 aircraft in queue, *i.e.*, relative variations of over 20%, and changes in peak expected departure queue length of the order of 1 aircraft in queue, *i.e.*, relative variations of approximately 5%. This difference is likely due to the fact that the slope of the Operational Throughput Envelope at JFK is lower than 1 (Figure 7), so variations in arrival rates induce variations in departure rates of a smaller magnitude.

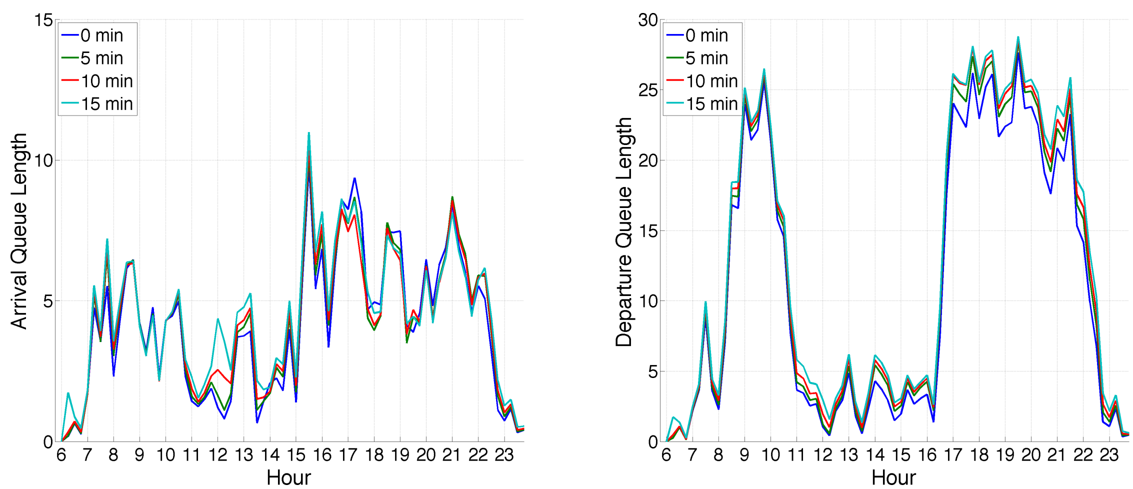
Figure 12 shows the sensitivity of expected arrival (Figure 12a) and departure (Figure 12b) queues to the value of τ_I. Expected arrival and departure queues naturally vary with the duration of idleness following a runway configuration change: as τ_I increases, operations become less efficient and thus congestion costs become larger. At peak hours, the difference might be up to 1 aircraft in the arrival queue and to 2 aircraft in the departure queue. These differences naturally translate into different optimal expected congestion costs: Table 5 indicates that the total expected costs



(a) Arrival queue (b) Departure queue

Figure 11: Sensitivity of expected arrival and departure queues to α ($\tau_I = 5$ minutes)

increase by over 20% when τ_I increases from 0 minute to 15 minutes. Therefore, the efficiency with which airport operators manage to change runway configurations has a significant impact on expected airport congestion costs.



(a) Arrival queue (b) Departure queue

Figure 12: Sensitivity of expected arrival and departure queues to τ_I ($\alpha = 1$)

Table 5: Expected total cost for different values of τ_I

τ_I	Expected total cost
0 minute	Baseline
5 minutes	+9.73%
10 minutes	+15.30%
15 minutes	+21.31%

4.3 Comparison of Optimal Control to Heuristic Policies

We now compare the optimal control of runway configurations and of arrival and departure service rates to two distinct heuristic policies. These comparisons provide estimates of the congestion cost savings that might result from the implementation of the model developed in this paper as a tool for tactical decision-making.

The two heuristics we consider primarily aim at adjusting the arrival rate as a function of arrival demand. The departure rate is, in turn, determined by the Operational Throughput Envelope of the selected runway configuration. In other words, these heuristics tend to privilege arrivals over departures by mostly basing decisions on arrival demand. This is motivated by the fact that, at JFK, the departure rate varies slightly with the arrival rate. Indeed, as previously discussed, the slope of the Operational Throughput Envelope of the main runway configurations at JFK is lower than 1 (Figure 7). Therefore, increasing the arrival rate to match arrival demand induces a relatively small loss in departure throughput. Moreover, this decision might be motivated by practical reasons as well, since it is more challenging operationally to hold arriving aircraft in queue in the air than to hold departing aircraft in queue on the ground.

The first heuristic assumes no cost of changing runway configurations. At the beginning of each period, the decision-maker computes the *effective arrival demand*, denoted by ad_t and defined as the expected number of aircraft that will be ready for landing in the period. It is simply equal to the sum of the arrival queue at the beginning of period t, *i.e.*, a_{t-1}, and the expected number of aircraft that will join the arrival queue in period t, *i.e.*, x_t. He then attempts to match arrival demand by selecting an arrival rate equal to the effective arrival demand, if possible. If the maximal arrival rate of each of the runway configurations that can be used under observed wind conditions is smaller than the effective arrival demand, then he selects the largest arrival rate that can be possibly chosen. He then decides to operate during the following period in the runway configuration that maximizes the expected departure throughput for the selected arrival rate. This heuristic is presented in Algorithm 1.

The second heuristic relies on similar dynamics except that an extensive cost of switching runway configurations is assumed. The decision-maker does not change runway configurations unless he has to do so, because of changing wind conditions. If the current runway configuration can still be used for the next period, then he chooses the arrival rate that matches the effective

Algorithm 1 Heuristic 1: Arrival priority and no cost of runway configuration changes

for $t = 1, ..., T$ **do**
  for a_{t-1} **do**
- Compute effective arrival demand: $ad_t = a_{t-1} + x_t$

    for wc_t, ws_t **do**
      for $rc \in \mathcal{RC}(ws_t)$ **do**
- Define candidate arrival rate with rc: $ar_t(rc) = \min\left(A_{rc,wc_t}, ad_t\right)$

      end for
- Determine the set of candidate runway configurations $\mathcal{RC}_{ar_t}(ws_t) \subset \mathcal{RC}(ws_t)$ that minimize the following quantity: $|ad_t - ar_t(rc)|$ and define new arrival rate μ_t^a as the corresponding value of $ar_t(rc), rc \in \mathcal{RC}_{ar_t}(ws_t)$
- Choose runway configuration $RC_t = \arg\max_{rc \in \mathcal{RC}_{ar_t}(ws_t)} \{\Phi_{rc,wc_t}(\mu_t^a)\}$
- Define new departure rate: $\mu_t^d = \Phi_{RC_t,wc_t}(\mu_t^a)$

    end for
  end for
end for

arrival demand as closely as possible and the departure rate is subsequently determined by the Operational Throughput Envelope of the runway configuration. If, however, the current runway configuration can no longer be used, then he selects the policy determined by the first heuristic. This heuristic is presented in Algorithm 2.

Table 6 reports the relative difference between the optimal expected cost per stage and the expected cost per stage resulting from the application of each of these two heuristics, for distinct values of the duration of idleness τ_I. First, note that the performance of the heuristics depends on the value of the parameter τ_I: Heuristic 2 performs relatively well if $\tau_I = 10$ minutes, while Heuristic 1 performs better if $\tau_I = 0$, which is consistent with the design of these heuristics. Moreover, the optimal control results in substantial cost savings, as expected congestion costs can be reduced by as much as 15% to 20% compared to the best of both heuristics in each case.

Table 6: Relative error of heuristics

Algorithm	$\tau_I = 0$ minute	$\tau_I = 5$ minute	$\tau_I = 10$ minutes
Exact	Baseline	Baseline	Baseline
Heuristic 1	+17.65%	+34.77%	+70.94%
Heuristic 2	+36.54%	+26.08%	+22.19%

The comparison of the optimal control to heuristics suggests that the joint control of runway configurations and arrival and departure service rates can improve substantially the efficiency of airport operations. The model presented in this paper provides a systematic decision-making

Algorithm 2 Heuristic 2: Arrival priority and extensive cost of runway configuration changes

```
for t = 1, ..., T do
    for RC_{t-1} do
  • Determine the set of wind states WS_{RC_{t-1}} in which RC_{t-1} can be feasibly used
        for ws_t ∈ WS_{RC_{t-1}} do
  • Choose same runway configuration: RC_t = RC_{t-1}
            for a_{t-1}, wc_t do
  • Compute effective arrival demand: ad_t = a_{t-1} + x_t
  • Define new arrival rate: μ_t^a = min(A_{RC_t,wc_t}, ad_t)
  • Define new departure rate: μ_t^d = Φ_{RC_t,wc_t}(μ_t^a)
            end for
        end for
        for ws_t ∉ WS_{RC_{t-1}} do
  • Choose policy from Heuristic 1 in all states
        end for
    end for
end for
```

tool that exercises this control dynamically to minimize airport congestion costs under realistic conditions of operational uncertainty. In turn, the implementation of this control can provide significant operational benefits to the system, which we estimate at 15% to 20% of congestion costs, compared to realistic heuristics.

This improvement in airport efficiency represents significant cost savings for airlines, passengers and society. As mentioned in the introduction, the annual delay costs were estimated at over $30 billion for the year 2007 [1]. Moreover, mismatches between demand and capacity are responsible for an estimated 50% to 75% of all flight delays experienced in the United States [6]. Therefore, a reduction in queuing delays of the order of 15% is likely to result in annual savings of the order of $2 billion to $3 billion. Given the disproportionate distribution of delays across airports in the country and the propagation of these delays through the National Aviation System, the implementation of the control developed in this paper at some of the busiest airports in the United States is expected to capture a significant share of these potential delay savings. This would, in turn, save several hundreds of millions of dollars in annual delay costs to airlines and to passengers.

4.4 Use of the Control of Service Rates in Models of Airport Congestion

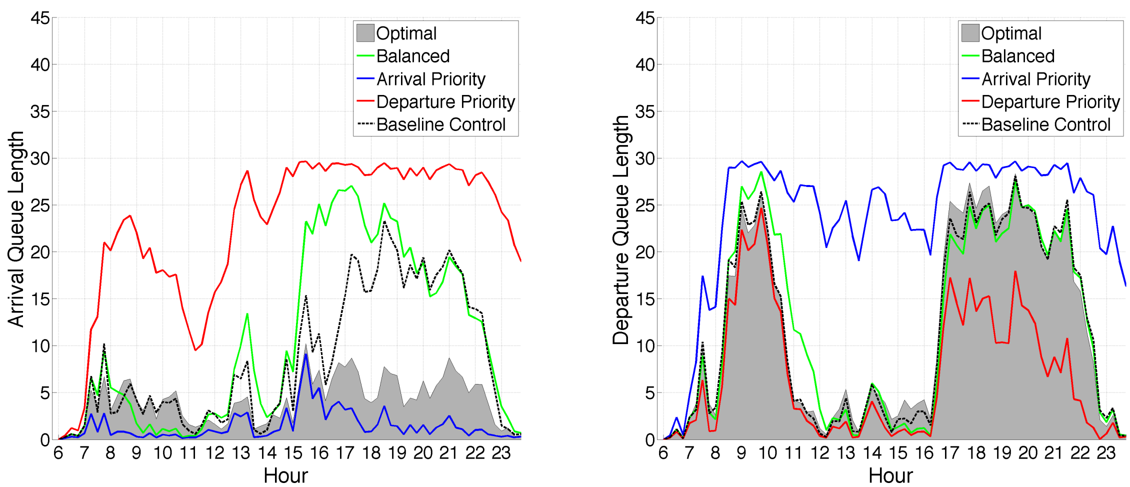
The control of runway configurations and arrival and departure service rates developed and implemented in this paper, in addition to providing a congestion mitigation tool at the operational level, can also be used at the strategic level to improve predictive models of airport congestion by integrating the control of arrival and departure service rates. These models of airport congestion, which quantify the relationships between flight schedules, airport capacity and flight delays, typically consider exogenous values of arrival and departure service rates and thus exercise no control of runway configurations and arrival and departure service rates, or do so manually and not systematically [13, 14, 16]. The limitation of this approach lies in the fact that these service rates are not known in advance but instead dynamically controlled through the course of the day as a function of observed extents of congestion at the airports. The optimization of this control at the operational level, as implemented in this paper, provides a realistic way to select the arrival and departure service rates through the course of the day in models of airport congestion. The combination of this control to the determination of the Operational Throughput Envelope of the different runway configurations at one airport, to the stochasticity of queue dynamics and to the stochasticity of operating conditions therefore provides a new framework to model flight delays that integrates airport operational procedures to airport strategic planning. In turn, the model and the results from this paper can be used to improve the realism of macroscopic models of airport congestion, which can then be used in airline scheduling algorithms, in the design of other congestion mitigation tools and strategies, etc.

In order to illustrate this point, we compare the optimal control of runway configurations and arrival and departure service rates to baseline policies, where no control is exercised and where the arrival and departure service rates are kept constant throughout the day of operations. Three baseline policies are considered: (a) *Balanced Operations*, where the airport serves arrivals and departures at the same rate, (b) *Arrival Priority*, where the highest values of arrival throughput are realized and (c) *Departure Priority*, where the highest values of departure throughput are realized. Corresponding average service rates are provided in Table 7, and are averaged out all runway configurations and traffic mix conditions [19]. We also add a fourth scenario, where the decision-maker minimizes congestion costs (Equation 3) by choosing one of the three baseline policies shown in Table 7 at each period, as a function of the observed arrival and departure queue lengths. This scenario, which we call *Baseline Control*, aims at simulating the most basic form of control that can be exercised. These baseline policies aim at replicating typical choices made in macroscopic models of airport congestion.

Table 7: Definition of baseline policies [19]

Baseline Policy	Arrival Capacity	Departure Capacity	Total Capacity
Balanced Operations	10	10	20
Arrival Priority	16	6	22
Departure Priority	6	12	18

Figure 13 compares the expected arrival (Figure 13a) and departure (Figure 13b) queues resulting from the application of the control presented in this paper, on the one hand, and from these four scenarios, on the other hand. Note, first, that arrival and departure queues are extremely sensitive to arrival and departure service rates: the range between queues resulting from two extreme policies, namely arrival priority and departure priority, is extremely large. Moreover, the optimal control results in significantly smaller expected arrival and departure queues than in the Balanced Operations scenario, *i.e.*, when arrivals and departures are served with identical service rates. Exercising the baseline control, *i.e.*, selecting at each period one of the three baseline policies, results in smaller congestion costs than balancing operations, but still significantly larger ones than when the optimal policy is applied. Even for large values of τ_I, the optimal policy results in congestion costs that are 25% smaller than when the baseline control is applied (Table 8). Last, note that the effects of the control on arrival queues are more important than those on departure queues. This primarily comes from the fact that increasing by one unit the arrival rate results, at JFK, in lowering the departure rate by less than one unit (Figure 7).



(a) Arrival queue

(b) Departure queue

Figure 13: Comparison of optimal arrival and departure queues to baseline policies

Table 8: Comparison of optimal congestion costs to baseline policies

Control	$\tau_I = 0$ minute	$\tau_I = 5$ minute	$\tau_I = 10$ minutes
Optimal Control	Baseline	Baseline	Baseline
Balanced Operations	+80.75%	+64.72%	+56.76%
Baseline Control	+45.17%	+32.29%	+25.90%

These comparisons show that estimated arrival and departure queues are very sensitive to the selection of arrival and departure service rates. In turn, the expected arrival and departure queue lengths obtained from the implementation of the control introduced in this paper differ significantly from queues obtained by exogenously estimating the arrival and departure service rates, as it is typically done in macroscopic models of airport congestion. This suggests that the approach and the results from this paper can also be applied to control arrival and departure service rates in dynamic models of airport congestion at the strategic level, which subsequently improves the predictive power of these modeling methodologies.

5 Conclusion

In this paper, we have presented the first decision-making framework that dynamically controls the selection of runway configurations and of the arrival and departure service rates at a major airport while taking into account the stochasticity of arrival and departures queues and of airport operations. This model solves the problem of allocating airport capacity to arriving and departing aircraft under uncertain operating conditions. We have introduced an efficient dynamic programming formulation that minimizes airport congestion costs and embeds a realistic stochastic model of queue dynamics as well as weather and wind-related uncertainty. The exact dynamic programming algorithm has been implemented and shown to yield optimal policies within reasonable computational times. A one-step look-ahead algorithm that considerably speeds up completion has been proved to produce near-optimal policies. This fast approximate algorithm enables the on-line implementation of the model presented in this paper, which might be critical when new information (*e.g.*, schedule updates) becomes available through the course of the day.

Comparisons of optimal policies to heuristic policies suggest that queuing costs might be significantly reduced by exercising a dynamic control of runway configurations and of arrival and departure service rates at a major airport. Congestion cost savings are estimated at 15% to 20%, which represents annual savings of the order of $1 billion. Moreover, it has been shown that decisions are path-dependent in the sense that they depend on the stochastic evolution of the system, and, in particular, on the stochastic evolution of arrival and departures queues through the course of a day of operations. Results also emphasize the importance of involving all stakeholders to keep transitions between successive runway configurations as smooth as possible: higher efficiency in operating runway configuration switches results in smaller congestion costs.

The model and the algorithms presented in this paper provide an effective decision-making tool to mitigate airport congestion at the tactical level, whose implementation may provide substantial operational benefits to airport stakeholders. Moreover, this approach provides a new framework to control arrival and departure service rates in macroscopic models of airport congestion in the face of operational uncertainty and variability, which subsequently improves the realism of airport congestion modeling methodologies.

Acknowledgments

This research was supported in part by the Federal Aviation Administration as a NEXTOR-2 project. The help of Ioannis Simaiakis in getting the Operational Throughput Envelopes from JFK, is also gratefully acknowledged. The authors are solely responsible for any errors or opinions expressed herein.

References

[1] M. Ball, C. Barnhart, M. Dresner, M. Hansen, K. Neels, A. Odoni, E. Peterson, L. Sherry, A. Trani, and B. Zou. Total Delay Impact Study. Technical report, National Center of Excellence for Aviation Operations Research, October 2010.

[2] Richard Bellman. *Dynamic Programming.* Princeton University Press, Princeton, NJ, 1957.

[3] D. Bertsekas. *Dynamic Programming and Optimal Control*, volume I. Athena Scientific, 3rd edition, 2005.

[4] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, 1996.

[5] D. Bertsimas, M. Frankovich, and A. Odoni. Optimal Selection of Airport Runway Conigurations. *Oper. Res.*, 59(6):1407–1419, December 2011.

[6] Bureau of Transportation Statistics. TranStats. Accessed March 26, 2013. Available at: http://www.transtats.bts.gov/ot_delay/ot_delaycause1.asp, 2013.

[7] R. de Neufville and A. Odoni. *Airport Systems: Planning, Design and Management.* McGraw-Hill, 2003.

[8] Federal Aviation Administration. Aviation Performance Metrics (APM) database: https://aspm.faa.gov/apm, 2013.

[9] E. Gilbo. Airport Capacity: Representation, Estimation, Optimization. *IEEE Trans. Control Systems Technology*, 1(3):144–154, September 1993.

[10] E. Gilbo. Optimizing Airport Capacity Utilization in Air Traffic Flow Management subject to Constraints at Arrival and Departure Fixes. *IEEE Trans. Control Systems Technology*, 5(5):490–503, 1997.

[11] E. Gilbo and K. Howard. Collaborative Optimization of Airport Arrival and Departure Traffic Flow Management Strategies for CDM. In *3rd USA/Europe Air Traffic Management R&D Seminar, Naples, Italy*, 2000.

[12] W. Hall. *Efficient Capacity Allocation in a Collaborative Air Transportation System.* PhD thesis, Massachusetts Institute of Technology, 1999.

[13] M. Hansen, A. Odoni, D. Lovell, T. Nikoleris, and K. Vlachou. Use of queuing models to estimate delays savings from 4D trajectory precision. In *8th USA/Europe Air Traffic Management R&D Seminar, ATM2009*, 2009.

[14] A. Jacquillat. A Queuing Model of Airport Congestion and Policy Implications at JFK and EWR. Master's thesis, Massachusetts Institute of Technology, June 2012.

[15] L. Li and J. P. Clarke. A Stochastic Model of Runway Configuration Planning. In *AIAA Guidance, Navigation and Control Conference*, August 2010.

[16] N. Pyrgiotis and A. Odoni. On the Impact of Scheduling Limits: A Case Study at Newark International Airport. *Transportation Science*, Forthcoming, 2013.

[17] N. Pyrgiotis and I. Simaiakis. An Analytical Queuing Model of Airport Departure Processes for Taxi Out Time Prediction. In *AIAA Aviation Technology, Integration, and Operations*, 2010.

[18] V. Ramanujam and H. Balakrishnan. Estimation of Maximum-Likelihood Discrete-Choice Models of the Runway Configuration Selection Process. In *American Control Conference*, June 2011.

[19] I. Simaiakis. *Analysis, Modeling and Control of the Airport Departure Process.* PhD thesis, Massachusetts Institute of Technology, August 2012.

[20] P. Swaroop, B. Zou, M. Ball, and M. Hansen. Do More US Airports Need Slot Controls? A Welfare Based Approach to Determine Slot Levels. *Transportation Research Part B*, 46(9):1239–1259, 2012.

[21] V. Vaze and C. Barnhart. Modeling Airline Frequency Competition for Airport Congestion Mitigation. *Transportation Science*, 46(4):512–535, 2012.