Stochastic Modeling and Control of Airport Surface Traffic

by

Francis R. Carr

B.S., Engineering
Harvey Mudd College, 1998

Submitted to the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2001

Author ....
Department of Electrical Engineering and Computer Science
September 18, 2000

Certified by ....
Eric Feron
Associate Professor
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Stochastic Modeling and Control of Airport Surface Traffic

by

Francis R. Carr

Submitted to the
Department of Electrical Engineering and Computer Science
on September 18, 2000, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

This thesis develops a stochastic model for airport surface traffic. An analysis of currently-available operations data is presented, and several characteristic behaviors of airport surface traffic are identified. A simple model structure is proposed to capture the observed behaviors, and calibration methods are proposed. These proposals are verified using a Monte Carlo technique. A simple tactical control scheme to control departure congestion is evaluated.

Thesis Supervisor: Eric Feron
Title: Associate Professor

Acknowledgments

Many thanks to Husni Idris, Stephane Bratu, Ioannis Anagnostakis, and Nhut Tan Ho; the best way to critique and create is through dialogue with a savvy colleague. I would like to thank my wife, family and friends for helping keep me mostly sane through my first two years at MIT. I am deeply in debt to Professor Eric Feron, who is willing to take risks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

While air traffic levels in the NAS have risen steadily since deregulation in the 1980's, the development of new airport capacity and the implementation of throughput-efficient procedures for the utilization of existing capacity has lagged behind. As any business traveler can attest, this imbalance has an obvious manifestation in terms of increasing levels of delayed flights, missed connections, cancellations, etc. In addition, the strategic planning and tactical management of air traffic flows is becoming more sensitive and less robust with respect to unforeseen reductions in capacity (typically due to weather). Thus the overall system performance is steadily degrading not only in an average sense, but also in terms of system volatility and controllability.

The development of new airport capacity is typically an enormous undertaking, and hence most research has focused on developing procedures to use existing capacity with greater efficiency. Flow control and strategic management of the en-route airspace is relatively well-developed: ETMS implements airspace monitoring and en-route flow control, and can provide accurate landing times as soon as an aircraft takes off. In the high-density terminal airspace surrounding DFW, CTAS implements efficient trajectory prediction, flow merging, and runway balancing for arrival traffic. The control of aircraft flows in flight is relatively well-established.

In contrast, the control of aircraft flows on the airport surface (between the gates and runways) has not met with such success. This shortcoming is not unreasonable; aircraft on the ground are much safer and more controllable than those in flight, and therefore assume a lower priority. However, the taxi-out and taxi-in processes contribute substantially to the financial, environmental, and uncertainty costs of every aircraft's gate-to-gate travel time. Airport surface congestion is a significant problem at many busy airports.

There is a dearth of models which tractably capture the stochastic nature of airport surface traffic. While there are several well-known models for airport surface traffic (e.g. SIMMOD, TAAM, or the Airport Machine), these models capture the detailed physical layout of the airport and the motion of individual aircraft. The corresponding attention to detail and complexity is useful in many applications, but also makes these models analytically intractable; difficult to calibrate or validate against actual operations data; and unsuited to the development of robust control algorithms.

This thesis presents work related to the development of simple stochastic models and control algorithms for airport surface traffic. An analysis of currently-available data relevant to airport surface operations is presented, summarizing its accuracy, utility, and shortcomings. Based on extensive field observations at BOS, and analyses of recent historical data from BOS, DFW, ATL, IAH, and EWR, several characteristic behaviors of airport surface traffic are identified. A simple model structure is proposed to capture the observed behaviors, and calibration methods are proposed which reflect the limitations of currently-available data. After being calibrated with actual operations data from a given time-period, both the model structure and calibration methods are verified using a Monte Carlo technique: a simulation of the calibrated model structure is driven with actual operations data from a disjoint time-period, a new calibration is derived from the simulation results, and the original and new calibrations are compared. Based on the proposed model structure, a simple tactical control scheme intended to help air traffic controllers mitigate departure congestion is evaluated. To indicate the direction of continuing research, several new characteristic behaviors are discussed which have been identified but not yet successfully modeled. It is important to note that these methods have been developed in an exploratory fashion, and that a robust mathematical development appears premature during this phase of system modeling.

This research is intended to provide a simple and tractable stochastic model for airport surface traffic. The model is acknowledged to be thoroughly incomplete, but it is hoped that such modest beginnings will encourage further development of the model structure and proposed control algorithms. In particular, it is hoped that we have managed to convey the necessity of new sources of detailed historical operations data to fuel this research. Stochastic modeling and control can offer substantial benefits to air traffic controllers, similar to those achieved by the thorough development of queueing theory in communications and manufacturing. In the face of rising traffic demand and volatility, airlines seeking to ensure stable operations and robust risk management may accrue very significant benefits from its successful application.

Chapter 2

Preliminary Data Analysis

2.1 Overview

In practice, the taxi-time of an aircraft operating at a major airport is influenced by a large number of variables, most of which are not directly recorded in the currently-available operations data. Some of these variables include airplane and surface vehicle traffic on the airport surface; the gate and runway assignments of a particular aircraft; the influence of EDCT, DSP, or other slot-based takeoff windows imposed by downstream ATC constraints; and the readiness of the aircraft (approval of flight plan amendments, processing of weight and balance numbers, mechanical problems, etc.). For the purposes of this research, a model is developed which directly captures the effects of airplane surface traffic congestion using operations data from the ASQP database (see section 2.2.1), while attempting to minimize or suitably work around the influence of the other variables.

Gate assignments for each flight are not directly available, but instead are approximated using the airline, since each airline typically controls a physical cluster of gates at a particular airport. The effects of weather conditions and runway configurations are approximated using wind data (to constrain which runways are usable); other weather data (to capture the effects of different operating procedures under inclement-weather conditions); and field observations (to determine ATC preferences among the various possible runway configurations). Currently, for runway assignments, we have no good information sources or approximations. A project is currently underway to obtain accurate gate and runway assignment data for at least one major airport through the assistance of a corporate partner.

At each airport, the tower records and communications logs contain excellent data on downstream restrictions [9], but obtaining this information in statistically significant quantities was judged to be prohibitively time-consuming since the data are not available in an electronic format. Some information on the readiness of each aircraft may be available through CDM, but was not pursued for this research. In addition, operations data for turboprops and smaller regional carriers are noticeably missing from the ASQP records. Collectively, these gaps in the operations data are treated as sources of stochastic noise. The effects of turboprops and regional carriers can be minimized by analyzing airports where they form a relatively small portion of the traffic; note that it would be trivial to include these traffic sources given more complete operations data similar to the ASQP data. A filtering technique has been developed (see Chapter 4) to mitigate the effect of delays due to downstream restrictions and aircraft readiness. Due to these measures, the model is believed to be reasonably robust with respect to these noise sources.

2.2 Data Sources

2.2.1 Flight Operations

This research relies heavily on the Airline Service Quality Performance (ASQP) database. Originally created to help monitor air traffic delays among the major passenger airlines, ASQP provides both scheduled and actual gate-pushback, take-off, landing and gate-arrival times for the jet operations of the 10 largest passenger airlines in the US, including Alaska; American; America West; Continental; Delta; Northwest; Southwest; TWA; United; and US Airways. It is important to note that ASQP does not capture either turboprop operations or the flights of regional airlines, both of which can make up a significant fraction of the passenger traffic at a given air-

Airport	Year	ASQP op's	Ratio of ASQP to CODAS traffic CODAS: jets	CODAS: all
ATL	1998	499,091	74.4%	62.4%
BOS	1997	189,137	73.1%	42.9%
DFW	1997	490,548	89.6%	61.4%
EWR	1998	231,117	70.7%	53.3%
IAH	1998	246,438	76.2%	60.2%

Table 2.1: Fraction of traffic captured by ASQP.

port (see Table 2.1). To some degree, this effect has been accommodated by choosing airports with a relatively small fraction of turboprop or regional carrier traffic.

Scheduled times reported in ASQP are derived from the OAG and CRS databases. Actual times reported in ASQP are derived from two sensors in each equipped aircraft: one sensor detects the release or activation of the landing-gear brakes (corresponding to pushback and gate-arrival events), and another sensor records changes in the weight supported by the landing gear (corresponding to takeoff and landing events). It is worth noting that Southwest Airlines records its ASQP data manually and does not rely on sensors or other automated systems.

All times reported in the ASQP database are rounded to one-minute precision. Although the duration of some operations (notably landings and takeoffs) are on the order of one minute, this resolution limitation can be accounted for in some cases by treating these events as a filtered, time-sampled point process [8]. The accuracy of the ASQP data has been thoroughly tested. Visual observations at BOS have confirmed ASQP pushback times [4]. ASQP takeoff (landing) times have been statistically validated by cross-checking them against high-resolution in-flight radar tracks provided by CTAS at DFW [13]: a threshold location was chosen on the departure path (final approach path) roughly 5 miles from the runway threshold, and the time difference between the ASQP takeoff time (landing time) and the time of threshold crossing (obtained by radar track extrapolation) was computed for all jet aircraft that used that particular runway. The ASQP times closely matched the times estimated from the radar tracks, indicating that ASQP is accurate to within

CODAS weather-classification variables	% of EWR ops VFR	IFR	% of BOS ops VFR	IFR
Meteorological Cond'n	82.2%	17.8%	84.6%	15.4%
Ceiling, Visib., Precip.	81.7%	18.3%	82.1%	17.9%

Table 2.2: Weather statistics at EWR and BOS in 1998.

its one-minute reporting precision.

2.2.2 Airport Operating Conditions

Weather conditions and runway configurations obviously impact the dynamics of airport surface traffic; this conclusion is well-supported by previous research at BOS [20, 16]. In particular, the runway configuration and flight conditions together dictate arrival and departure acceptance rates; surface traffic patterns; and indirectly, the necessity for flow control measures such as ground-delay programs. It is worth mentioning that several of the important results documented in [16] were only discovered and analyzed when runway configurations were properly accounted for.

Fortunately, more than three years of detailed historical weather data are available from CODAS [21]. The CODAS weather data are reported in 15-minute intervals, and include wind speed, direction, and gust; temperature; precipitation and thunderstorm activity; ceiling; visibility; and the reported meteorological flight rules (VFR/IFR). The data set is remarkably complete. For example, at DFW in 1997, only eight 15-minute intervals are missing, and all of the data-fields are complete except for 7% of the temperature data.

To a first approximation, the effects of weather on airport operations can be summarized according to the meteorological flight rules. An analysis of 1998 CODAS weather data for EWR indicates that the meteorological conditions data-field accurately reflects the more detailed data-fields which record ceiling, visibility, precipitation and storm activity (see Table 2.2). For this reason, and to avoid an excess of explanatory factors, this research only considers the IFR/VFR data.

Unfortunately, it is unreasonably difficult to obtain runway configuration data for most U.S. airports[1]. For this reason, a method has been developed which uses the CODAS wind data to *approximate* the runway configuration [1]. While wind conditions alone do not fully determine runway configuration (e.g. at BOS, the potential impact of noise pollution is significant), they may still be used in conjunction with airport layout information to determine which set of runways are operable, and thus constrain which configurations are possibly in use. Based on personal communications with an experienced pilot employed by a major U.S. airline [14], a conservative set of standards has been developed for runway operability under various wind conditions [1]. Under these conservative standards, a runway is considered operable if the crosswind is less than 20 knots and the headwind is positive; otherwise, the runway is considered inoperable. For these calculations, the wind speed is taken to be the maximum of the wind and gust speeds recorded in the CODAS weather data. For variable wind-angles, it is assumed that the wind is perpendicular to each runway, i.e. high-speed variable-direction winds are assumed to shut down all of the runways due to crosswinds[2].

This wind-based method of approximating the runway configuration has not been explicitly validated. However, the results have been implicitly validated at DFW using 5 days of high-resolution in-flight radar tracks provided by CTAS. The radar tracks were analyzed to infer which runways were used for takeoffs and landings, and the radar-inferred runway set was then compared against the weather-inferred runway set. In general, the radar and weather results matched. The general orientation of each runway set was identical, but there were often slight variations between the inferred sets, probably due to the sensitivity of the weather-based analysis to short-term changes in wind speed and direction [1]. Another type of validation was also implemented at BOS. CODAS wind data was used to infer the operable set of runways, and this set of runways was then compared to the (more restricted) set of

[1] BOS is a rare exception to this rule. At BOS, the runway configuration has been monitored and recorded for several years through the PRAS system as part of local efforts to reduce the impact of noise pollution, and these PRAS records were available to support Pujet's research [16].

[2] Closer analysis shows that this assumption has a negligible effect on the final results. For example, variable wind direction only occurs 4% of the year at EWR in 1998, and never during high-wind conditions.

Runway Configuration Departure	Arrival	% of op's in 1997	% of time in error
27-33L	33L-33R	17.4%	2.7%
4R-4L	4R-4L	2.9%	37.4%
22R-22L	22L-22R	3.6%	10.2%
22R-22L	27-22L	32.3%	9.9%
9-4R-4L	4R-4L	26.7%	33.1%
9-4R-4L	4R-15R-4L	4.9%	55.5%
All others		12.2%	N/A

Table 2.3: Validation of weather-segmentation technique at BOS.

runways recorded in the airport records. The CODAS-inferred set was considered to be "in error" when airport records indicated that a particular runway was in operation despite its apparent inoperability due to winds. Table 2.3 records the results. It appears that the weather-segmentation technique is not as accurate for BOS. If a good source of airport runway configuration data does not become available, it may be of future interest to determine more exact wind and weather conditions under which each runway was in operation, in order to refine the weather segmentation methodology.

After the operable runway set was found for each time-interval in the time-period of interest, this information was linked to the IFR/VFR data to classify each time-interval. Except for BOS, all of the airports considered in this thesis have a runway layout that favors two primary runway orientations. As might be expected, for those airports with two primary orientations, the majority of the operations occur under runway sets corresponding to those orientations. Given this result, the possible classifications for those airports are summarized into six groups of characteristic weather conditions, or *segments*: the two primary orientations under VFR conditions, the two primary orientations under IFR conditions, an indeterminate segment where wind information indicated that both primary orientations were operable, and an excluded segment where wind information indicated that neither primary orientation was operable. BOS has a similar classification scheme, although more accurate airport records have been used instead of wind data from CODAS. Section 2.2.3 describes the airport runway layouts for each of these airports, and also describes the groupings of runway operability sets that have been used in lieu of actual runway configuration data.

In summary, our methodology for accommodating the effects of weather and runway configuration on airport surface traffic is to use the CODAS weather data to classify a particular time-period of interest (typically a calendar year) according to weather conditions, and then separately analyze all of the time-intervals corresponding to each particular segment. Explicit runway configuration data are not used, but the effects of each runway configuration are approximated through the association between runway operability and wind conditions.

2.2.3 Airport Runway Geometries and Configurations

We have focused on five major airports: BOS (Boston MA), EWR (Newark NJ), ATL (Atlanta GA), DFW (Dallas/Fort Worth TX), and IAH (Houston TX). BOS is a moderate-sized non-hub airport; EWR is usually thought of as a hub for Continental Airlines, although the Continental flights are not structured into banks [5]; the other three airports are major hubs for Delta, American, and Continental respectively.

As shown in Figure 2-1, the runway and taxiway geometry at BOS is quite complex. Airport records indicate that at least 26 distinct runway configurations were in use at one time or another during 1996 and 1997. However, most of these configurations were not in use for a significant period of time. As shown in Table 2.4, over 85% of the runway operations occurred in only 6 of the possible configurations[3]; none of the remaining 20 configurations captured more than 1.9% of the total operations. For the purposes of this research, only three runway operability sets are considered: a "northeast" set when 4R/L is operable, a "southwest" set when 22R/L is operable, and a "northwest" set when 33R/L is operable. It is worth noting that BOS is somewhat unusual in that a mix of departures and arrivals are often run on the same runway to obtain higher throughputs.

[3]Runway operations data taken from CODAS [21].

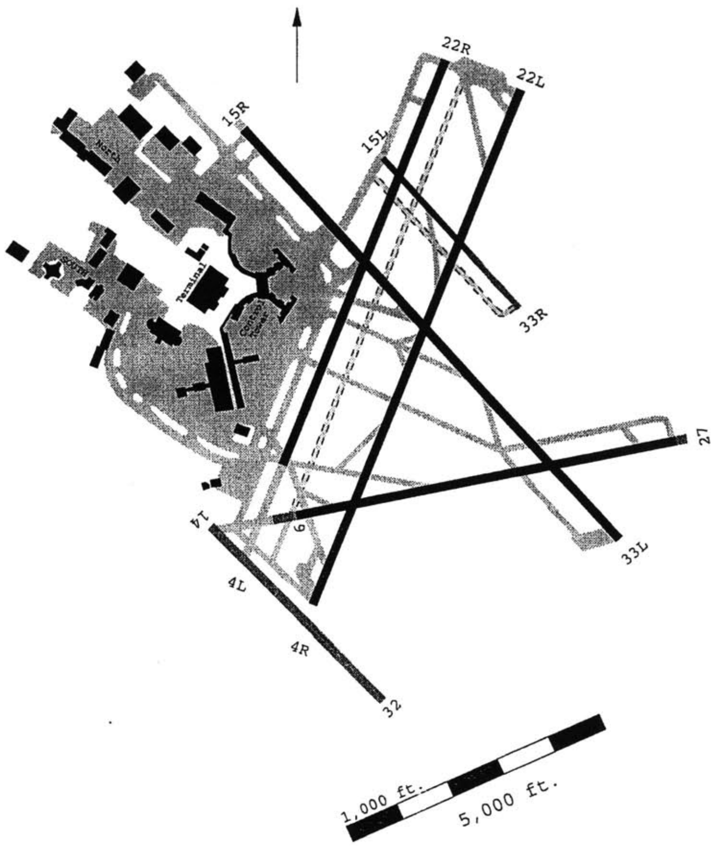

Figure 2-1: Runway layout at BOS.

Runway configuration: Departure	Arrival	% of op's in 1997
27-33L	33L-33R	17.4%
14R-4L	4R-4L	2.9%
22R-22L	22L-22R	3.5%
22R-22L	27-22L	32.3%
9-4R-4L	4R-4L	26.8%
9-4R-4L	4R-15R-4L	4.9%
All others		12.2%

Table 2.4: Heavily-used runway configurations at BOS.

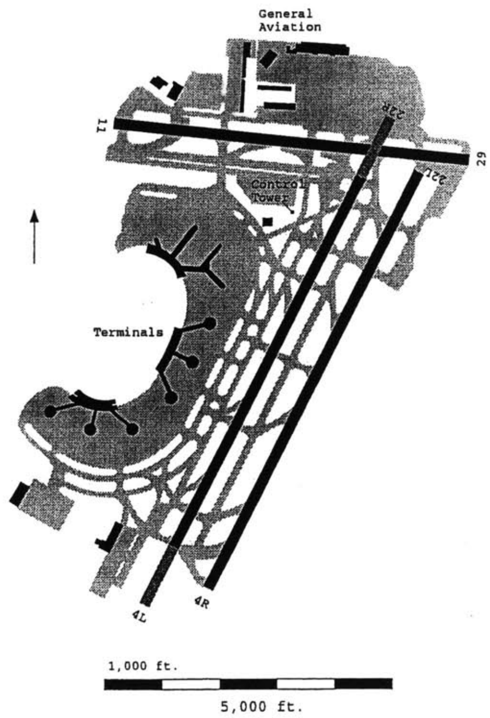

Figure 2-2: Runway layout at EWR.

As shown in Figure 2-2, EWR has a primary pair of runways oriented in a NE/SW configuration (4/22 R/L), and a secondary east/west runway (11/29). According to an interview with the tower supervisor, 11/29 is used less frequently due to multiple runway crossings, e.g. aircraft taking off on 29 have to cross three runways while taxiing and two more while taking off [10]. In addition, 11/29 is used primarily by commuter turboprop traffic operated by Continental Express, and hence these flights are not reflected in the ASQP data. For these reasons, traffic at EWR is approximated as operating on only the primary pair of runways. A "north" set includes the use of 4R/L. A "south" set includes the use of 22R/L. Note that EWR is also somewhat unusual in that arrivals are shunted to the outer runway of the pair, and the inner runway is reserved for departures. This is primarily due to limited taxiway space for departure queues between the runways, and for arrival queues between the inner runway and the terminal buildings [10].

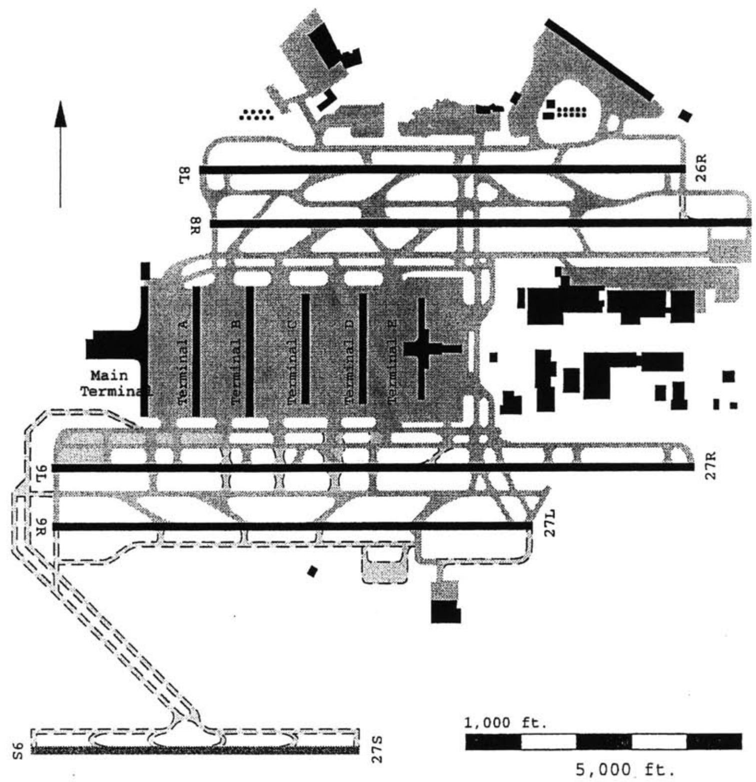

Figure 2-3: Runway layout at ATL.

Surface traffic at ATL was analyzed using data from 1998, at which time the airport had two pairs of runways oriented in a east/west direction as shown in Figure 2-3. The pairs are spaced far enough apart so that simultaneous operations can occur; note that this is not the case for the parallel runway pairs at either BOS or EWR. An "east" set includes the use of any of the runways 8R/L or 9R/L. A "west" set includes the use of any of the runways 26R/L or 27R/L. At any time, several runways are simultaneously available for departure and arrival operations. Typically each runway is assigned to either departures or arrivals, rather than a mix.

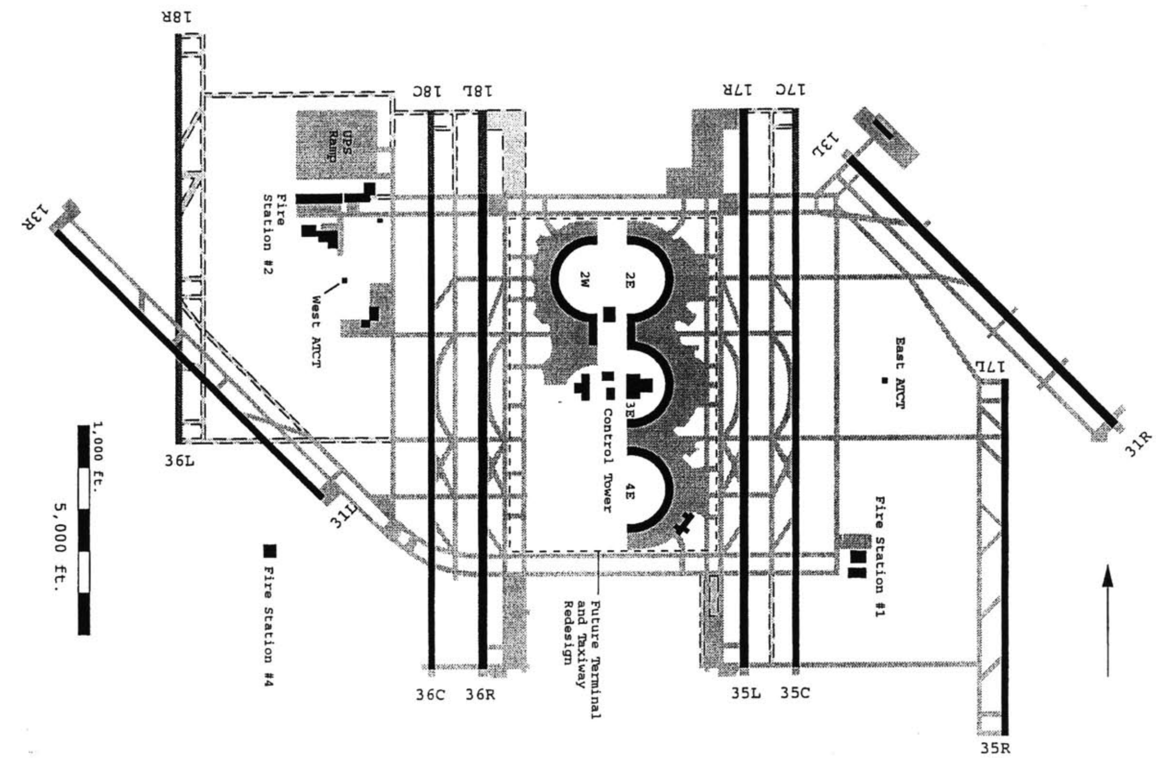

Figure 2-4: Runway layout at DFW.

As shown in Figure 2-4, DFW is oriented in a north/south direction with east and west sides running almost as independent airports [12]. There are two parallel runways and a diagonal runway west of the terminals, and three parallel runways and a diagonal runway east of the terminals. The parallel runways are spaced far enough apart so that simultaneous operations can occur. A "south" set includes the use of any of the runways 18C/L, 13R/L, or 17R/C/L. A "north" set includes the use of any of the runways 31R/L, 36R/L, or 35R/L. At any time, several runways are simultaneously available for departure and arrival operations. Typically each runway is assigned to either departures or arrivals, rather than a mix.

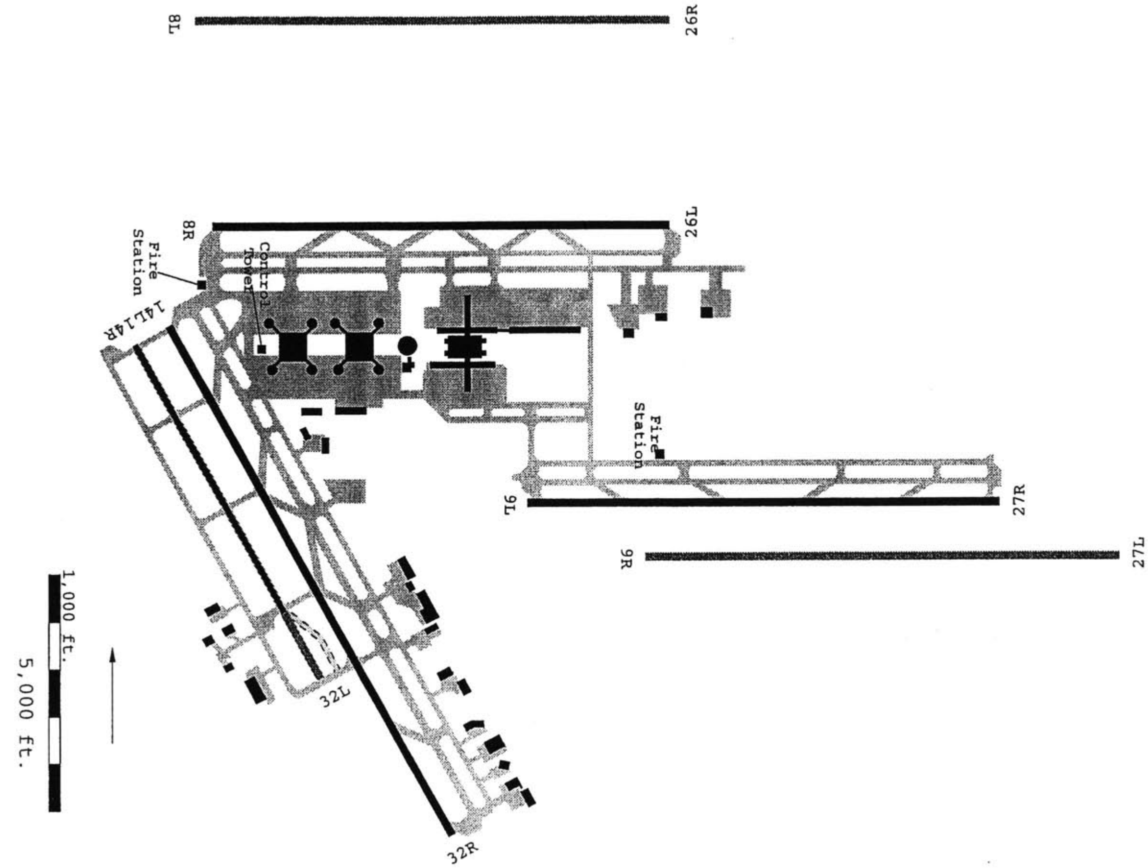

Figure 2-5: Runway layout at IAH.

IAH has six runways, primarily oriented in a east/west configuration as shown in Figure 2-5. The six runways consists of three parallel pairs: two to the north, two to the south, and two diagonal runways to the west. In each pair, the runways are spaced far enough apart so that simultaneous operations can occur. An "east" set includes the use of any of the runways 8R/L, 9R/L or 14R/L runways. A "west" includes the use of any of the runways 26R/L, 27R/L or 32R/L. At any time, several runways are simultaneously available for departure and arrival operations. Typically each runway is assigned to either departures or arrivals, rather than a mix.

Weather segment	Fraction of ASQP operations							
	EWR, 1998		ATL, 1998		DFW, 1997		IAH, 1998	
N/E VFR	31.1%	#1	24.2%	#1	26.6%	#1	31.4%	#2
S/W VFR	37.0%	#2	48.2%	#2	46.4%	#2	16.5%	#1
N/E IFR	9.8%	#3	8.0%	#3	2.8%	#3	5.9%	#4
S/W IFR	5.7%	#4	4.7%	#4	3.7%	#4	2.3%	#3
Indeterminate	11.5%	#5	3.1%	#5	15.0%	#5	41.7%	#5
Excluded	4.9%	#6	11.8%	#6	5.5%	#6	2.2%	#6

Table 2.5: Weather segmentation results (EWR, ATL, DFW, IAH).

Weather Segment	ASQP op's, BOS, 1998	
4R/L VFR	27.5%	#1
4R/L IFR	7.0%	#2
22R/L VFR	32.3%	#3
22R/L IFR	3.6%	#4
33R/L VFR	16.2%	#5
33R/L IFR	1.2%	#6
Other	12.2%	#7

Table 2.6: Weather segmentation results (BOS).

Tables 2.5 and 2.6 show the segmentation results. It is worth noting that these results support anecdotal reports that the south orientation is the primary orientation for DFW. Similarly at ATL, these results are consistent with the account that the west orientation is considered the most efficient [3].

Chapter 3

The Stochastic Model

3.1 Observed Behaviors

Based on extensive field observations at BOS [9], and analyses of historical operations data from BOS, ATL, DFW, IAH, and EWR, several characteristic behaviors of airport surface traffic have been identified. The sources of delay are different for departure and arrival traffic, so they are discussed separately. However, it is important to recognize that the same essential behaviors are present for both traffic flows, and that insights into the modeling and control of one particular sort of traffic are often applicable in a more general setting.

3.1.1 Departure Traffic

The first observation is that the taxi-out time of a given aircraft is strongly linked to the departure surface congestion at the time of pushback. The departure surface congestion (denoted N_D) is defined as the number of departing aircraft that are taxiing out on the airport's surface but have not yet taken off. Figure 3-1 shows how the observed distribution of taxi-out times tends to increase in both mean and variance as N_D increases; similar effects are observed at all of the airports considered. Note the long tail of the observed distributions. Practical modeling experience would suggest that the taxi-out time, arising as a combination of very many independent

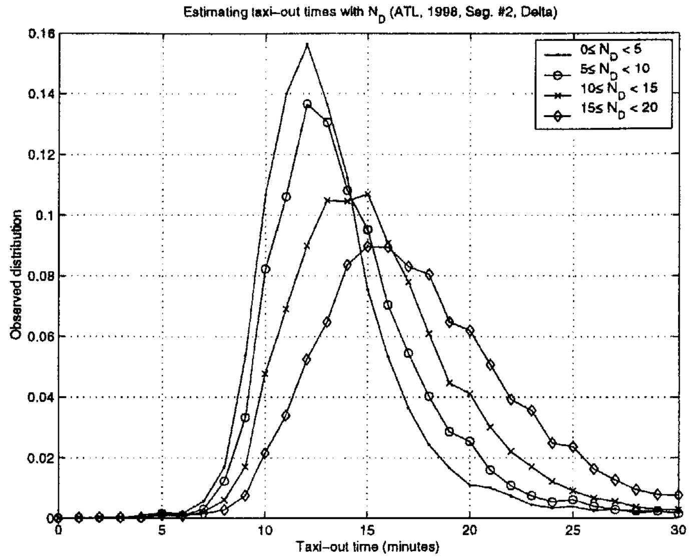

Figure 3-1: Distribution of taxi-out time, estimated using N_D.

sources of delay, should tend to assume a normal distribution; the long tail suggests that additional structure is present.

Once an aircraft reaches the runway, it usually enters a runway queue, and its relative position in the queue becomes fixed. The airport throughput is primarily limited by this bottleneck at the runways [11]. Two types of factors combine to reduce the runway service rate sufficiently for queues to develop [9]. First, there are certain purely physical constraints which must be satisfied, including required wake-vortex separations between the various weight-classes of aircraft, and runway occupancy by landing aircraft. Second, there are the procedural constraints imposed to facilitate downstream flow control. These constraints include takeoff-time windows for flights through congested airspace, and in-trail separations between departing flights headed through the same congested airspace.

The tower controllers are well-aware of these restrictions and are generally able to structure the departure surface traffic to mitigate the necessary queueing delays. In particular, it was shown in [9] that, except for weather at nearby fixes, downstream restrictions did not have a significant effect on throughput. For this reason, the service rate of the runway queue is approximated as constant with respect to the mix of

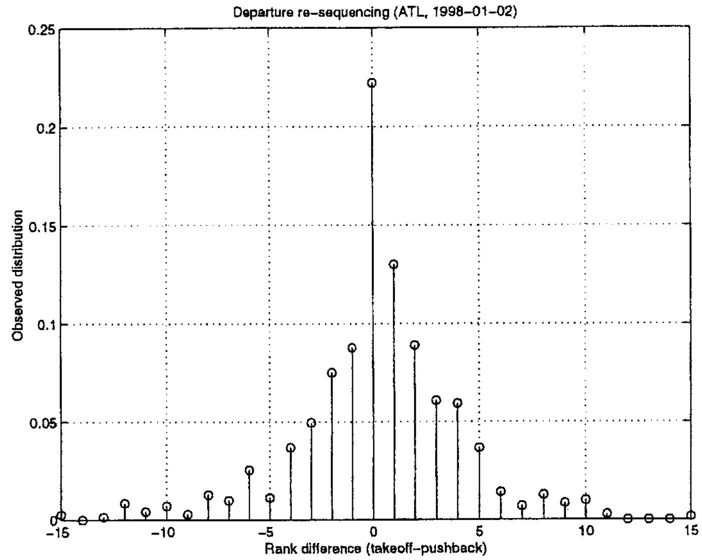

Figure 3-2: Aircraft re-sequencing between pushback and takeoff.

aircraft types and the possible presence or absence of downstream restrictions. Since both causes of service rate reductions are highly sensitive to weather conditions and the runway configuration in use, the service rate of the runway queue is assumed to be dependent on the local weather. Note that weather factors in en-route airspace and at destination airports have not been accounted for, although these are the primary cause of downstream constraints.

Another important observation is that aircraft often take off in a different order from their pushback sequence (see Figure 3-2). This swapping behavior has been thoroughly analyzed in [4]. As a first approximation, most swapping occurs before aircraft reach the runway queues, and hence is solely dependent on the free-flow behavior of the aircraft up to the queues. In the presence of multiple runways and multiple queues in more complex runway configurations, additional swapping can occur after aircraft are physically present in the runway queues; however, this effect is very difficult to capture using currently-available data.

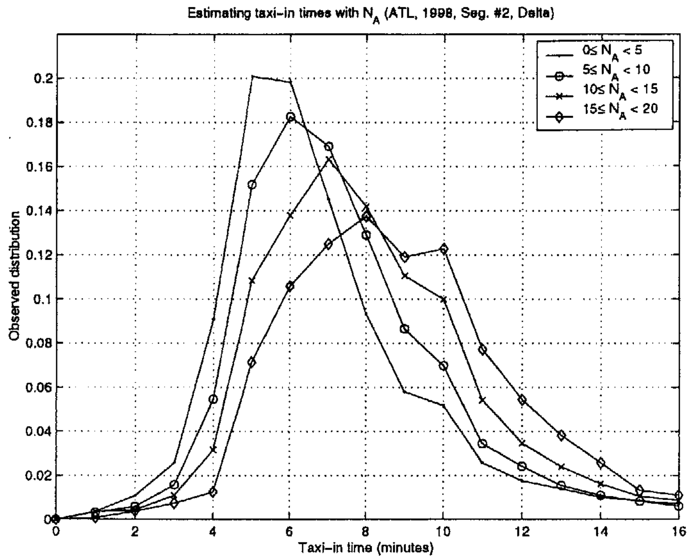

Figure 3-3: Distribution of taxi-in time, estimated using N_A.

3.1.2 Arrival Traffic

Arrival traffic shows many of the same behaviors as departure traffic, although the various observed behaviors have different degrees of influence on delays. Taxi-in time is strongly linked to arrival surface congestion (denoted N_A), although in general, taxi-in times have much lower variance than taxi-out times. Figure 3-3 shows the predictable effect of surface congestion on the distribution of taxi-in times. Some degree of swapping is also apparent (e.g. see Figure 3-4), although typically at a lower level than observed on the departure side. However, in contrast to departure traffic, however, arrival-traffic queueing is purely congestion-induced and can be mitigated by local surface traffic management. Most notably, the absence of downstream constraints accounts for much of the observed difference in volatility between arrival and departure traffic.

3.1.3 Unmodelled Behaviors

There are several other sources of surface-traffic delay which have not been captured by the model presented here. Many airports have regions of limited maneuvering

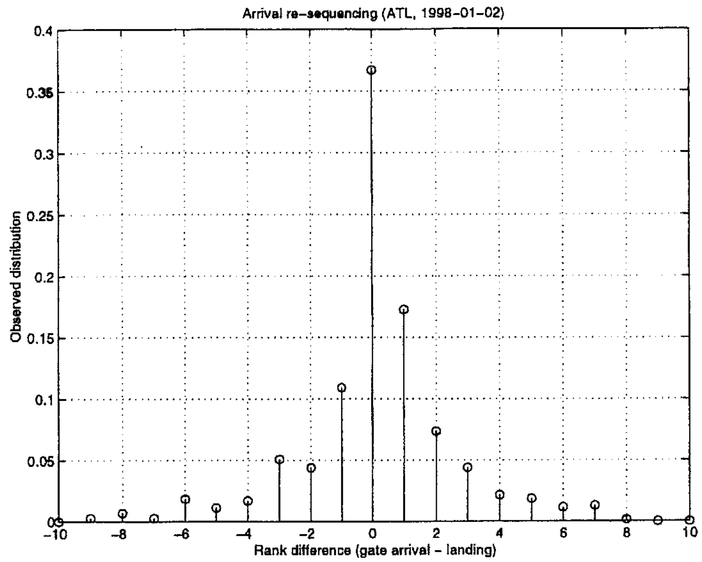

Figure 3-4: Aircraft re-sequencing between landing and gate-arrival.

space, often around the gates and terminal areas. Arrival and departure traffic must be monitored in these areas to avoid deadlocks and substantial holdups, and it is not uncommon for airlines to push back certain flights as soon as possible in order to minimize delays due to pushbacks from competing airlines at nearby gates. Contiguous departures must merge onto the taxiways, and controllers must carefully shuffle taxiing aircraft across active runways. All of these sorts of merge operations are difficult to observe using input-output data, although they are the primary points where controllers can substantially optimize aircraft flows, either by re-sequencing aircraft to minimize later delays due to runway constraints, or by introducing careful gaps in the flows to simplify runway crossings. In addition, these merge operations account for most of the interference between arrivals and departures.

More detailed operations data are required to fully quantify and model these effects. Radar tracks from an airport equipped with an ASDE system would be ideal, but obtaining ASDE data poses several technical and bureaucratic difficulties. Long-term on-site investigations are feasible to implement but highly labor-intensive. One promising possibility is to set up a system of cameras to record airport surface traffic, piping the data to an automated feature-recognition system for analysis and

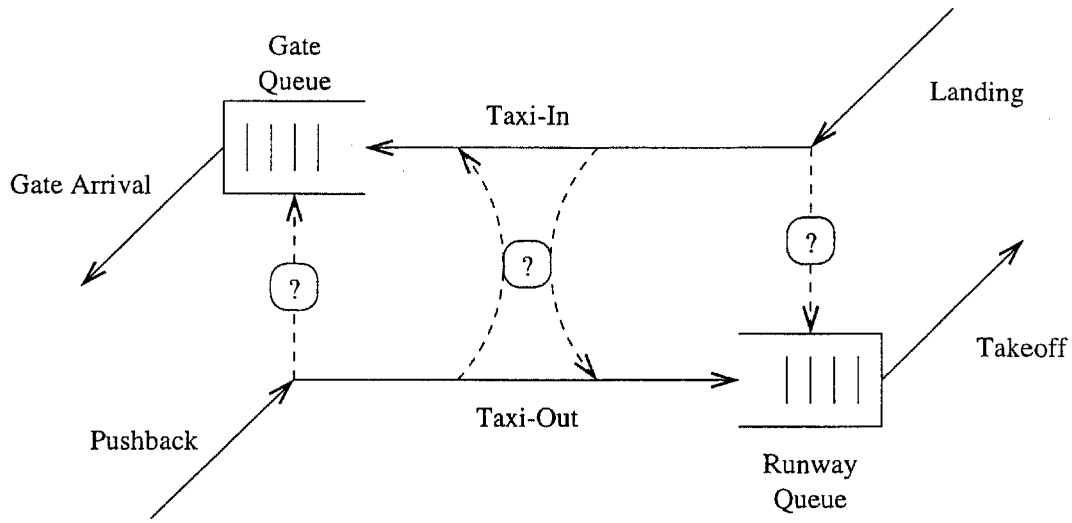

Figure 3-5: Diagram of the proposed model.

archiving. Preliminary investigations of the cost and technical requirements of such a project are underway.

3.2 Proposed Model Structure

A relatively simple model structure is proposed to capture the most significant observed behaviors. Some of the observed behaviors are not currently modelled, primarily due to a lack of detailed operations data with which to perform the necessary model calibration.

When aircraft enter the system at pushback (landing), they are assumed to undergo a period of free-flow travel to reach the runway (gate) queue. The time for each aircraft to reach the runway (gate) queue is termed the nominal or unimpeded taxi-time; these nominal taxi-times are presumed to be independent random variables with distributions that only depend on the airline of the aircraft and the weather-inferred airport operating conditions at the time of pushback (landing). After completing this free-flow period, aircraft are assumed to enter the runway (gate) queue, which behaves in a FCFS manner with some fixed service rate. Figure 3-5 shows the aircraft flows and significant events on the airport surface. The dashed lines and question-marks indicate unmodelled aspects of airport surface congestion which are known to affect taxi-times, including runway occupancy by landing aircraft, gate occupancy by departing aircraft which have not yet pushed back, and taxiway crossings and merges. Currently, we are working to obtain detailed gate and runway assignment data which will enable us to observe and model the gate and runway occupancy effects.

Chapter 4

Calibration (System Identification)

To simplify the exposition, we focus primarily on departure traffic, keeping in mind that the same observed behaviors and model structures also describe arrival traffic. Substantial differences between the two sorts of traffic (where they exist) are duely noted.

4.1 Distribution of Unimpeded Taxi Times

A basic quantity of interest is the distribution of unimpeded (nominal) taxi times, i.e. the distribution of the time it should take an aircraft to traverse the airport surface assuming the absence of delays due to congestion, downstream restrictions, etc. One could argue that a detailed simulation such as SIMMOD could be used to very accurately estimate the average taxi time based on aircraft type, gate location, runway location, and so forth. However, this approach neglects numerous sources of variation in the real-life process, and from a practical standpoint, a more correct approach might be to calibrate SIMMOD against real-life observations.

In previous research [4, 16, 20], it was shown that high departure surface congestion (N_D) at the time of pushback is well-correlated with increased taxi-out times. This result led to a method for estimating the distribution of unimpeded taxi-out times by only considering those flights which push back when N_D is small, and estimating the distribution from this reduced sample. However, this method leads to distributions with very long tails, which must then be artificially truncated. The long tails correspond to flights which push back during periods of very low departure congestion, but then experience non-congestion-induced delays while taxiing out, e.g. mechanical problems or downstream restrictions. It is desirable to eliminate these aircraft from the sample population.

In collaboration with Husni Idris (cf. reference [9]), a method has been developed to better observe the unimpeded distribution of taxi-out times from ASQP input-output data. Each departing flight is assigned an index (denoted N_H) which counts the number of aircraft which take off while the flight of interest is taxiing out on the airport surface. For example, if a flight pushes back at 5:00 and takes off at 5:15, its N_H index is the number of takeoffs which occur during the interval (5:00, 5:15][1].

Intuitively, if a particular flight is delayed on the airport surface due to downstream restrictions, mechanical problems, bureaucratic delays, or other effects unrelated to surface congestion, it will tend to be passed on the taxiway by other aircraft, and its N_H index will be large. If a particular flight pushes back and encounters substantial queueing delays near the runway, then its N_H index will be large due to the large number of other departing aircraft which take off while it waits in the runway queue. Therefore, it appears reasonable to assume that flights with a low N_H index have experienced little delay while taxiing out to the runway, and that the nominal (unimpeded) distribution of taxi-out times can be reliably estimated from their taxi-out times. It is worth noting that the N_H index cannot be calculated when an aircraft pushes back from the gate, and thus its utility for real-time control applications may be limited.

To compare the N_D and N_H indices for a particular departing aircraft, Figure 4-1 outlines the six orderings (A through F) in which two departing aircraft can enter/exit the airport surface. We may immediately discard two of these cases (A and F), since the two aircraft never simultaneously occupy the airport surface. Further, if we are only interested in the queueing and congestion delays imposed on a particular aircraft, case (D) may also be discarded: the aircraft of interest has an earlier position in the

[1]Note that the N_H index is always at least 1, since it includes the takeoff of the flight in question.

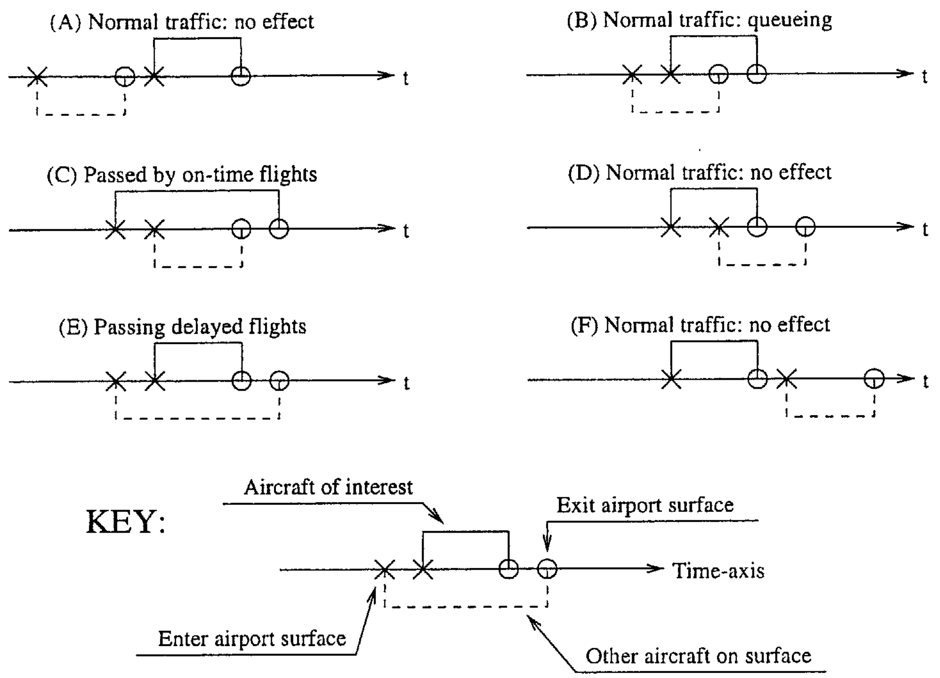

Figure 4-1: The six relative orderings of two departing aircraft.

queue and is unaffected by later departures. From the three remaining cases, we can derive two equations for the N_D and N_H indices:

$$N_D = (B) + (E), \quad N_H = (B) + (C) \tag{4.1}$$

Note that both indices approximate queue-length effects due to case (B). However, the N_D index counts case (E), although the aircraft of interest assumes an earlier position in the queue and thus is not delayed by aircraft which take off at a later time. In contrast, the N_H index counts case (C) and thus better approximates the queue-length delays experienced by the aircraft of interest.

The result of filtering the ASQP records using the N_H index is shown in Figure 4-2. Note that the long tail on the N_D curve has been substantially reduced. In addition, a small number of outliers have been eliminated; approximately 0.1% of the N_D sample had taxi-times in excess of an hour, but all of these flights have been removed from the N_H sample. In Figure 4-3, flights are grouped into separate samples according

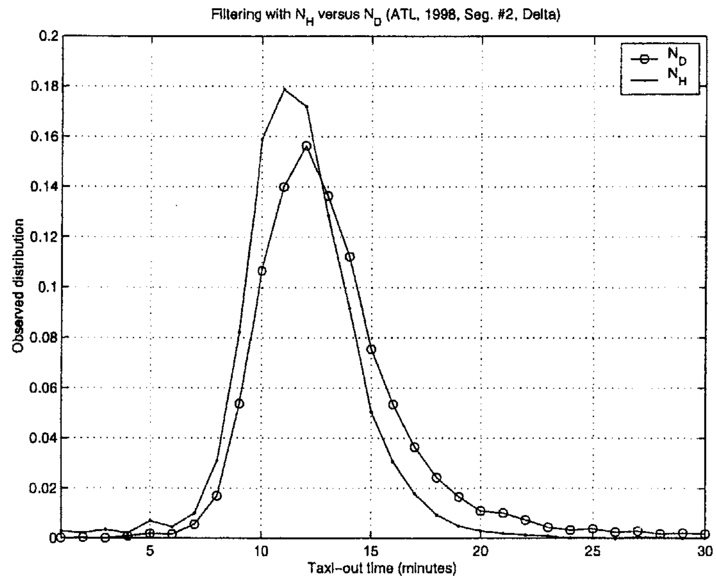

Figure 4-2: Taxi-out times estimated using N_H versus N_D.

to their N_D or N_H indices, the standard deviation of taxi-out times is computed for each sample, and these deviations are plotted versus the indices. Sampling with the N_H index is substantially better at grouping flights with similar taxi-out times.

Representative unimpeded taxi-out time distributions derived using the N_H index are shown in Figure 4-4 for one major airline at three different airports. These curves were derived from operations occurring under IFR conditions. The number of flights used to derive each curve is on the order of 100, since IFR conditions are relatively rare and tend to decrease aircraft throughput rates. For small sample-sizes such as these, log-normal distributions are used to approximate the underlying distribution for simulation purposes (e.g. Figure 4-5). For large sample-sizes where the observed distribution is already sufficiently smooth, the observed distribution is used directly for simulation purposes (e.g. Figure 4-2). From a practical standpoint, experience with the simulations indicates that the model is not particularly sensitive to changes in the distribution.

The N_H index is also well-defined for arrival traffic: for each arriving flight, it is

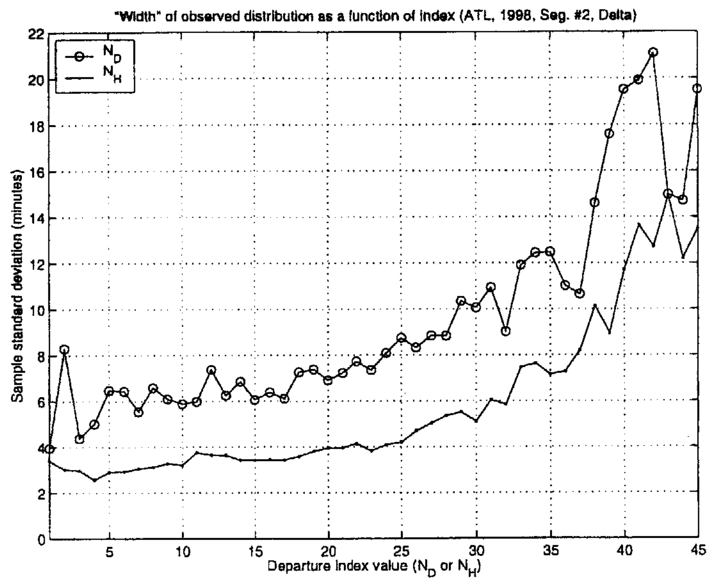

Figure 4-3: Departure sample variance as a function of N_H and N_D.

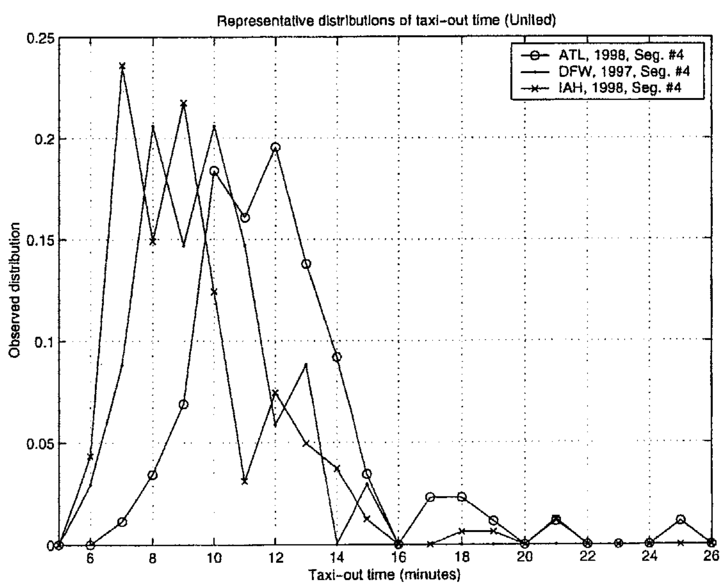

Figure 4-4: Taxi-out times estimated using N_H at three airports.

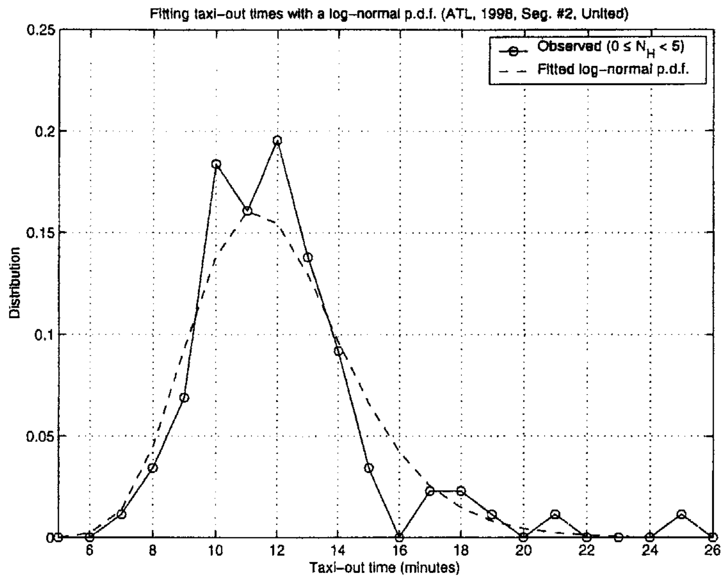

Figure 4-5: Curve-fitting the estimated taxi-out distribution.

defined as the number of gate-arrivals which occur while that flight is taxiing in[2]. In practice, the N_H index is not as useful for filtering out delayed arrival flights as it is for filtering out delayed departure flights. This is to be expected, since downstream restrictions do not impact arrival traffic, and hence one of the major unmodelled sources of taxi-out uncertainty (which can be partially filtered out using the N_H index) is not present. However, it is also important to note that gate occupancy by departing flights is not captured in any of the currently-available operations data, and that this can be a major source of arrival delay, particularly for flights which arrive earlier than planned. Figures 4-6 and 4-7 compare the results of filtering arrivals with the N_H index versus the N_D index. It is interesting to compare these figures with the corresponding figures for departure traffic (Figures 4-2 and 4-3), noting the large differences in mean and variance.

[2]Note that the index again has a minimum value of one, since it is defined to include all gate-arrivals up to and including the gate-arrival of the flight in question.

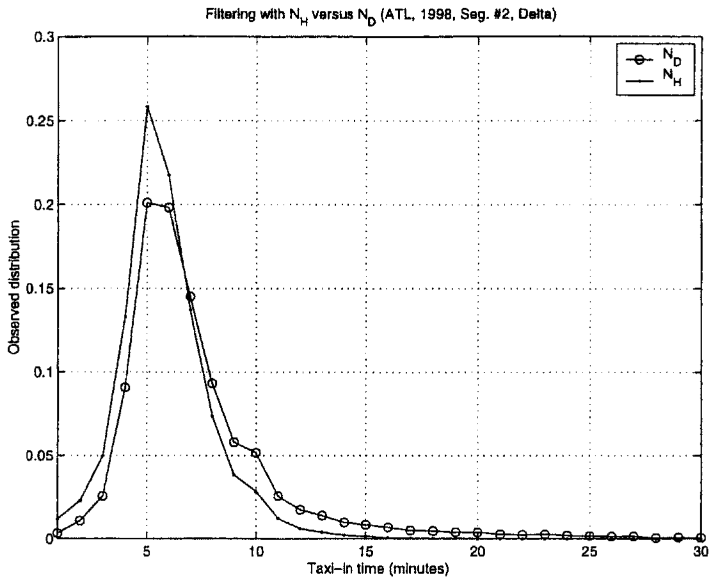

Figure 4-6: Taxi-in times estimated using N_H versus N_D.

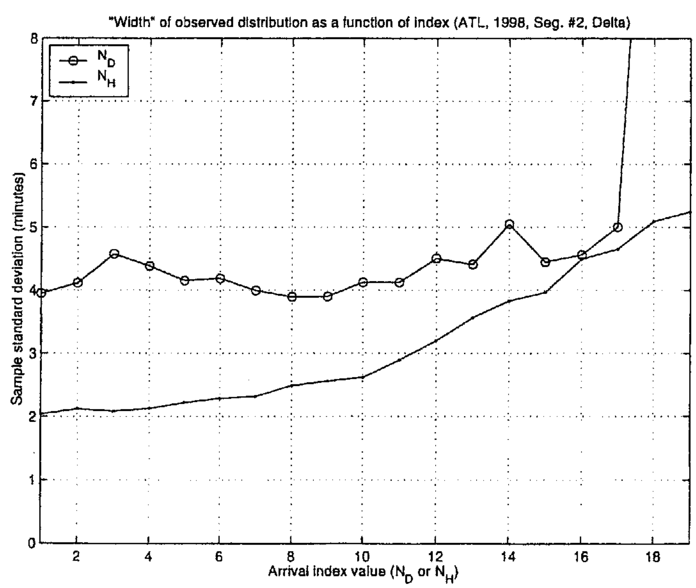

Figure 4-7: Arrival sample variance as a function of N_H and N_D.

4.2 Queueing Service Rate

Queueing is a fundamental feature of airport surface traffic, either of departure traffic near the runways, or of arrival traffic in the physically-constrained areas near the gates. A wide variety of queueing models have been proposed for airport systems, with a corresponding range of complexity in terms of queue dynamics and protocols. Two prominent models for surface queueing are the capacity-blocked networks proposed in [9], and the runway-server absence model proposed and identified in [16]. For the purposes of this research, a relatively simple model akin to the runway-server absence model has been found to satisfactorily describe the available operations data.

The stochastic model for the runway queue behavior is based on the observation that, at a fixed level of departure congestion, the distribution of the number of takeoffs per one-minute interval is well-fitted with a Poisson distribution. Further, as the level of departure congestion increases, the rate of the fitted Poisson distribution increases, until a threshold is reached where further increases in departure congestion levels do not result in increased average takeoff rates[3].

The following explanation is proposed to account for these observations. When a departing aircraft is ready to use the runway, a certain runway service-time is required. In the absence of ATC intervention, the required miles-in-trail and wake vortex separations would seem to indicate that successive runway service-times could not be independent. However, field observations and interviews at BOS indicate that controllers are well-aware of such restrictions and are successful at minimizing the required queueing delays. Therefore it appears reasonable to assume that successive runway service times are approximately independent; this effect should be more pronounced at airports such as ATL and DFW which offer controllers more opportunities to efficiently re-sequence departing aircraft.

[3]Note that at very high levels of departure congestion, the distribution of takeoffs begins to lose its Poisson character; this may be due to a decreasing sample-size (extremely high congestion levels are understandably rare events), or may indicate some as-yet unobserved set of conditions which both induces very high congestion and also influences the runway-takeoff dynamics. It may be of some interest to further analyze the airport conditions, especially local and downstream weather, which occur during periods of extreme congestion.

Further, it appears not unreasonable to assume that the distribution of these independent service-times are exponential. Note that this assumption is difficult to verify directly. Typical runway service times are on the order of a minute, and so the resolution of the ASQP data is not sufficient to observe runway service times directly. In addition, since the departure surface congestion is distributed over the airport surface, it is not possible to directly infer when the queue is heavily loaded. However, an approximate method based on results in [8] has been developed to check the assumption of exponentially distributed runway service times. A particular airport and set of airport operating conditions were chosen (ATL, departures, segment #1). Nominal taxi-out times were calculated using the N_H filtering method described above. To avoid any "edge-effects" due to changes in the weather, a full day of operations under segment #1 was located (October 17, 1998). Assuming that nominal taxi-out times are independent allows us to take the following probabilistic approach to observing the queue:

$$\begin{aligned}
E_n &\equiv \text{event: queue is empty at start of minute } n \\
Pr[E_n] &= Pr[\text{no taxiing aircraft have reached queue}] \\
&= \prod_k Pr[k^{th} \text{ taxiing aircraft has not reached queue}] \\
&= \prod_k 1 - F_k(n - P_k)
\end{aligned}$$

where P_k is the pushback time of the k^{th} aircraft currently in the system, and $F_k(\cdot)$ is the CDF of the nominal taxi-out time of the k^{th} aircraft. Figure 4-8 shows how $Pr[E_n]$ varied on the particular day in question.

During long periods when $Pr[E_n] < 0.02$, the takeoffs in each one-minute interval were assigned to randomly chosen points in that interval; these points were independently and uniformly distributed. This randomization served to smooth out the runway service times in an unbiased manner. From these smoothed takeoff times, runway service times could be estimated. The observed distribution of runway service times (averaged over 100 randomization runs) is shown in Figure 4-9.

The fit to an exponential distribution is quite good, although the probability of

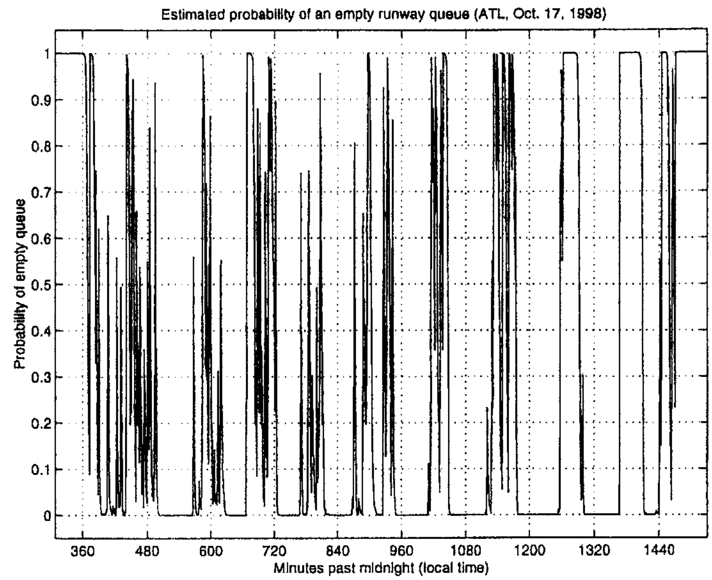

Figure 4-8: Estimated probability of an empty runway queue.

very short or very long runway service times is obviously lower than an exponential distribution would predict. This observation is not surprising, since there is a minimum separation between departing aircraft using the same runway, and the tower controllers attempt to uniformize service times so that excessively large service times are avoided. At this stage of model development, it does not appear worthwhile to replace the analytically useful assumption of exponentially distributed runway service times with a more accurate runway model which incorporates these effects.

Together with the independence assumption, exponential service times directly account for the observation that a Poisson distribution closely fits the observed distribution of takeoffs over one-minute intervals[4]. The last observation (the throughput saturation effect) is explained from the fact that departure surface congestion is spread over the entire airport surface. Considering the sample of one-minute intervals (as derived from ASQP data) which have some fixed level of departure surface congestion, for each interval in the sample there is some probability that departing aircraft were

[4]An interesting and salient theoretical question concerns conditions under which an arbitrary continuous-time point process, when filtered and time-sampled, would also yield a discrete-time Poisson process.

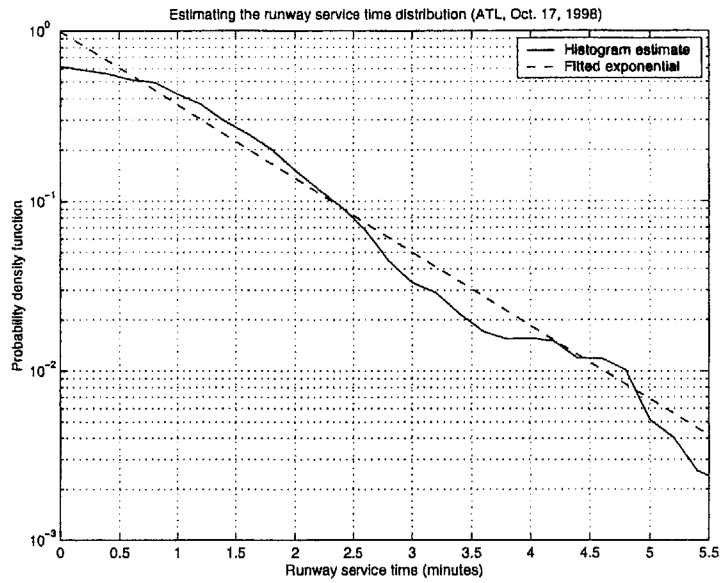

Figure 4-9: Estimating the runway service time distribution.

in fact present at the runway threshold and ready for takeoff. As the level of surface congestion rises, the probability that departure pressure exists on the runway over the entire one-minute interval rises to unity, at which point the maximum achievable runway throughput can be observed.

A type of runway throughput plot has been used to aid in calibrating the queueing component of the model. At each level of departure congestion, a Poisson distribution (with 95% confidence intervals) is fitted to the observed distribution of takeoffs. Then these fitted rates are plotted as a function of the departure congestion level to yield a throughput plot. Additionally, the number of time-intervals at each level of departure congestion is plotted to ensure that sufficient data-points are being used in the fitting process. Several representative plots (Figures 4-10 through 4-17) are included for two of the airports studied in this thesis.

The first pair of plots (Figures 4-10 and 4-11) were made using data from ATL during those intervals in 1998 when the airport was operating with an "east" runway set. Note that the distribution of takeoffs is fitted very well with a Poisson distribution over a wide range of departure congestion levels. It is apparent that the throughput under VFR conditions saturates at a higher level of congestion than the throughput under IFR conditions. Also note that at high levels of congestion, the 95% confidence intervals of the fitted rates do not overlap, reflecting the reduction in airport throughput capacity during periods of inclement weather.

A second pair of plots (Figures 4-12 and 4-13) were made using data from DFW during those intervals in 1997 when the airport was operating with a "south" runway set. We observe effects similar to those seen at ATL. However, note that the throughput at DFW during VFR conditions appears to steadily increase as departure congestion increases; there is no observed saturation effect. In contrast, the throughput during IFR conditions shows a clear saturation effect.

Representative gate throughput curves are also shown. As might be expected, gate throughput appears unaffected by IFR conditions at all of the airports studied (see Figures 4-14 to 4-17). One interesting observation is that the gate throughput can saturate, similar to the saturation effect seen in the departure process.

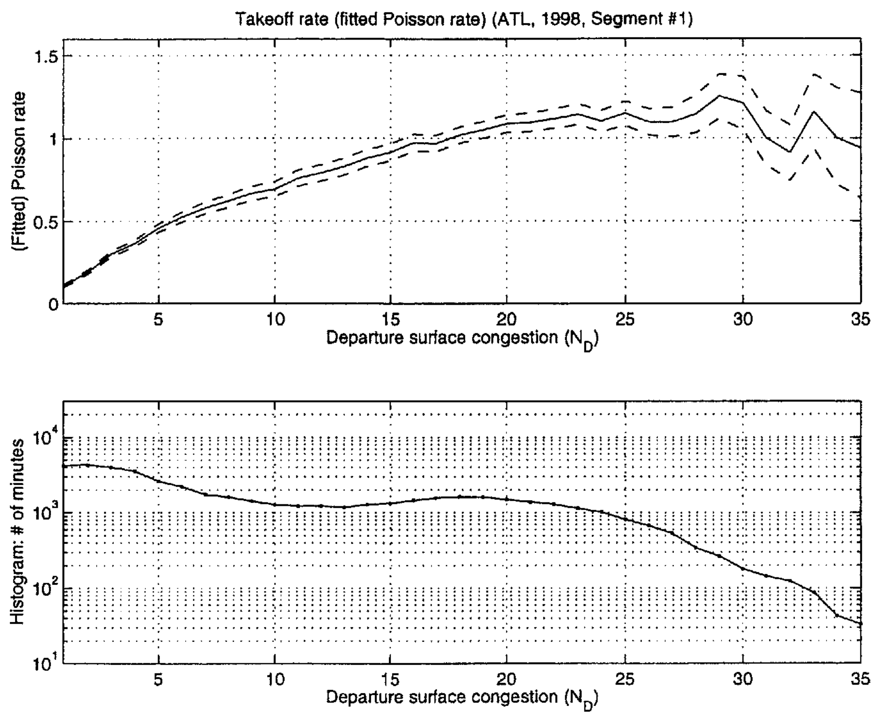

Figure 4-10: Runway throughput at ATL during VFR conditions.

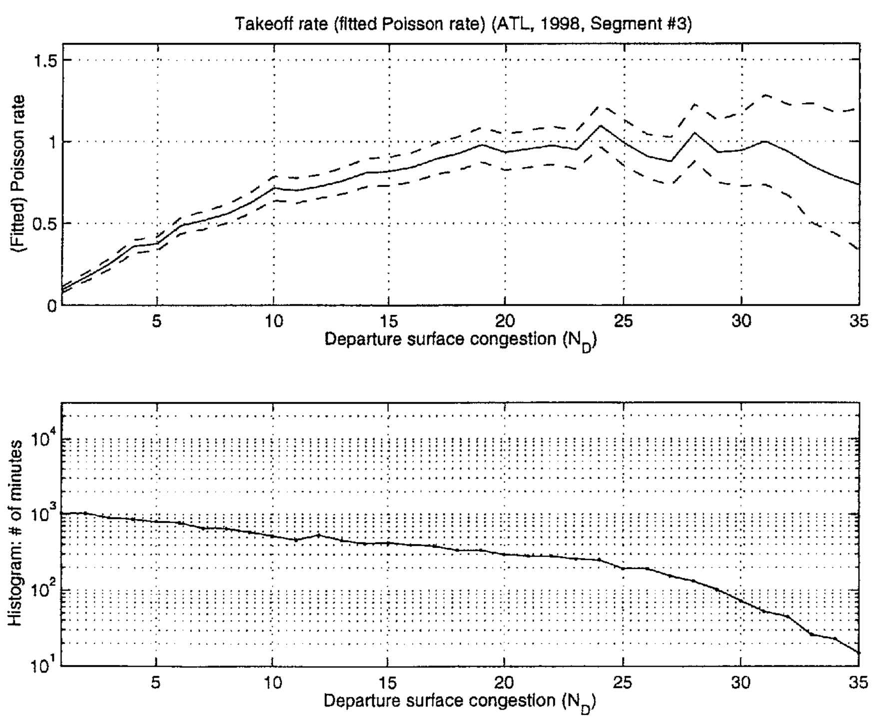

Figure 4-11: Runway throughput at ATL during IFR conditions.

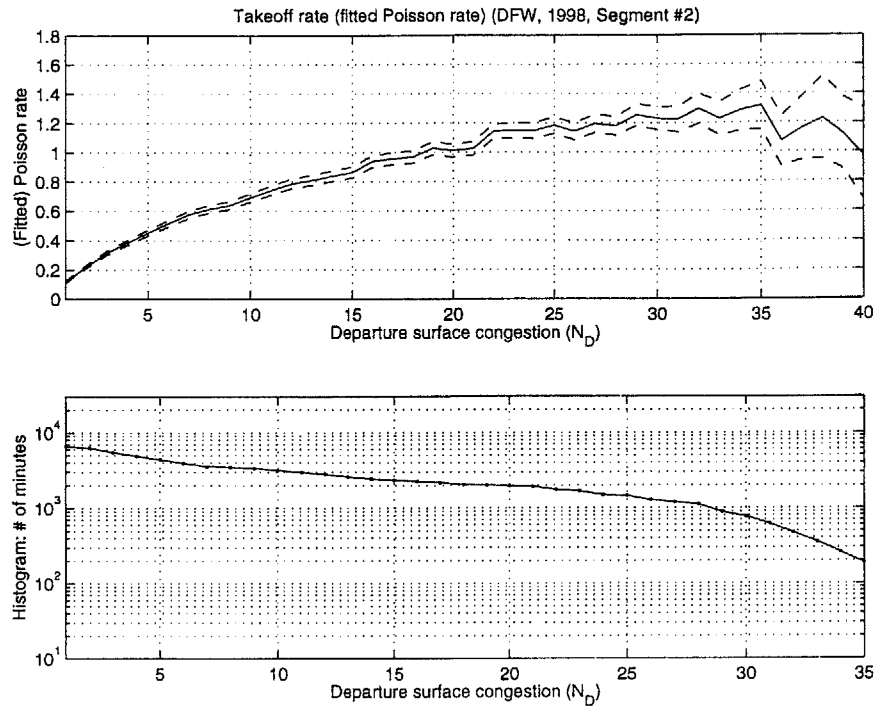

Figure 4-12: Runway throughput at DFW during VFR conditions.

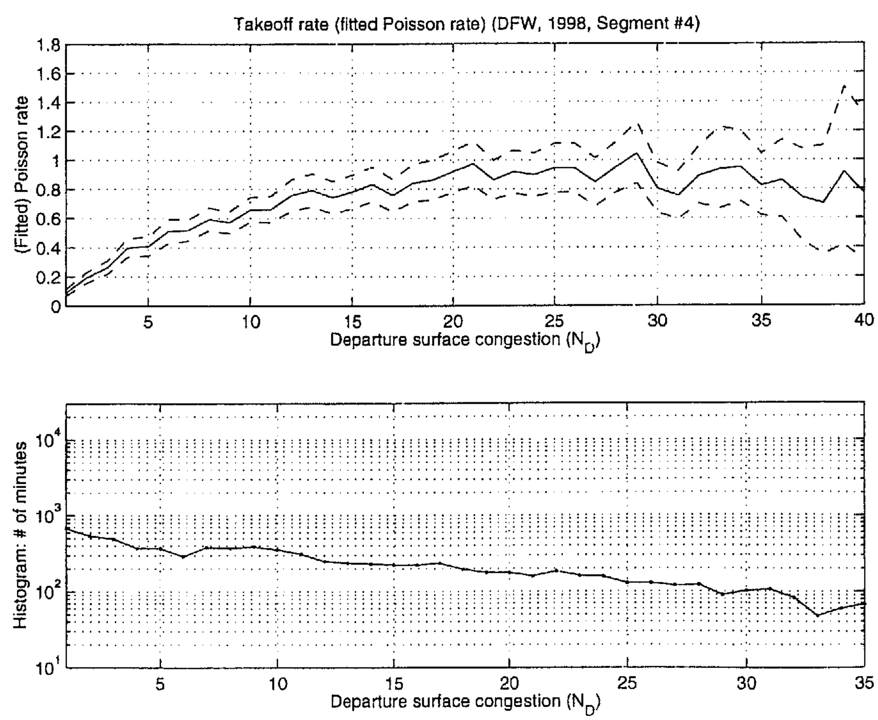

Figure 4-13: Runway throughput at DFW during IFR conditions.

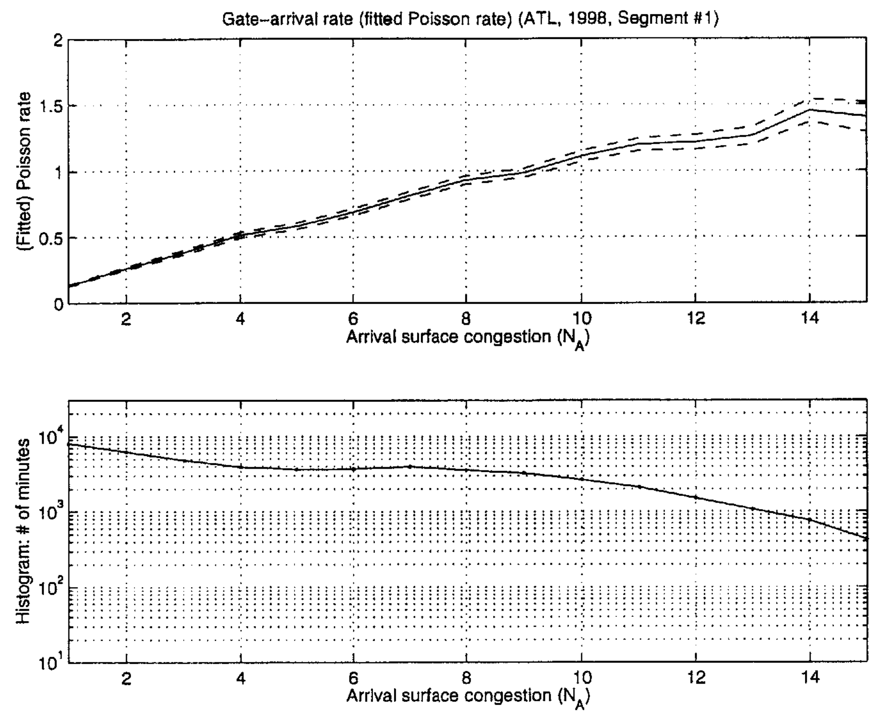

Figure 4-14: Gate throughput at ATL during VFR conditions.

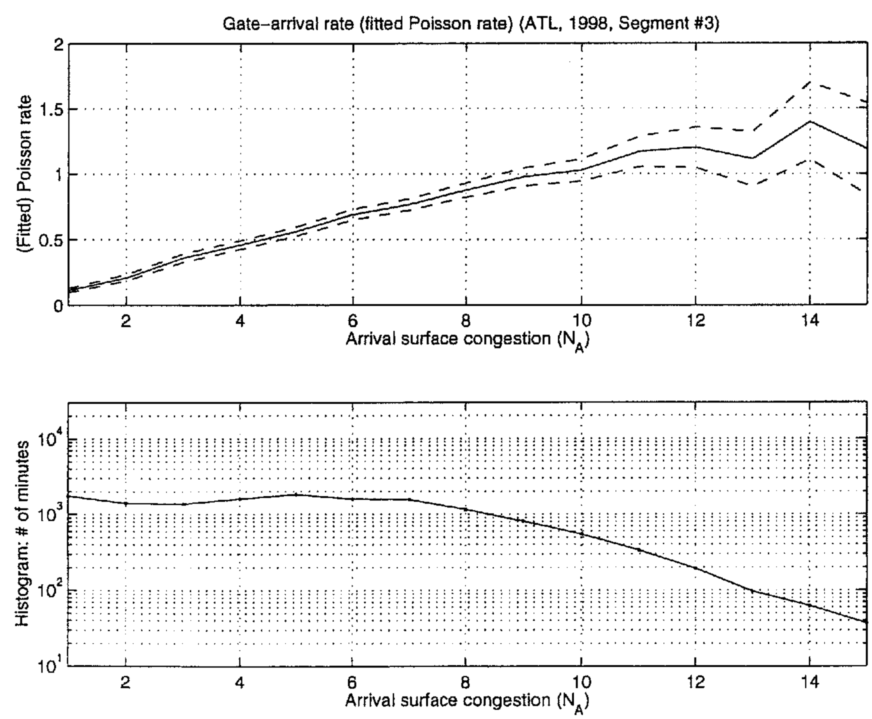

Figure 4-15: Gate throughput at ATL during IFR conditions.

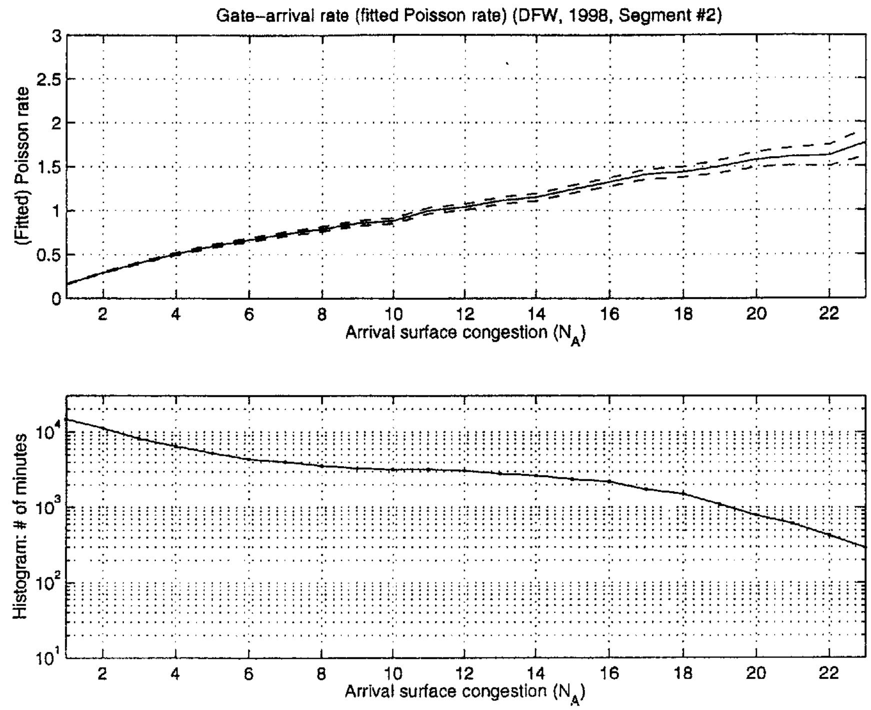

Figure 4-16: Gate throughput at DFW during VFR conditions.

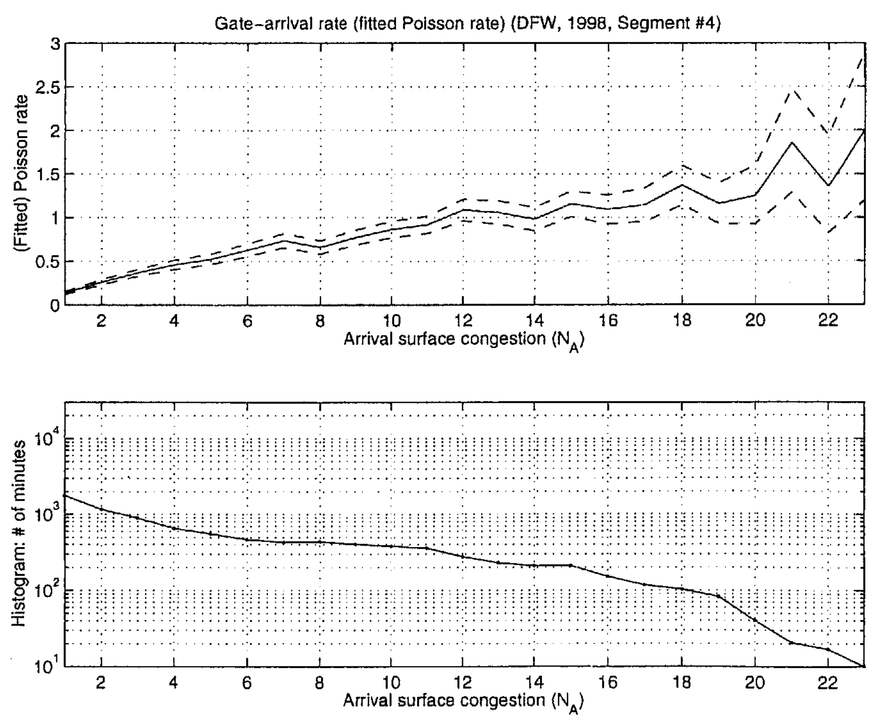

Figure 4-17: Gate throughput at DFW during IFR conditions.

Chapter 5

Validation and Simulation

5.1 Overview

There are two types of validity which must be addressed. First, the proposed model structure itself must be shown to be capable of reproducing, at least in an aggregate or average sense, those behaviors which have been observed in the real-life system. Second, after a particular model has been calibrated using data from a given airport and time-period under specified operating conditions, the accuracy of this calibration must be evaluated.

The basic algorithm chosen to accomplish both of these tasks is as follows:

1. Choose a particular airport operating over a particular time-period (typically a calendar-year) with a particular set of weather-determined operating conditions.

2. Calibrate a model to the airport. Use the techniques described in Chapter 4, and only consider those time-periods *during odd-numbered weeks* when the weather at the airport matched the desired operating conditions.

3. Build a Monte-Carlo simulation of the calibrated model. To accurately capture the effects of schedule-bunching and flight-banks, drive the simulation with the observed system input (either pushbacks or landings). To help avoid statistical artifacts of the data, only consider those time-periods *during even-numbered weeks* when the weather at the airport matched the desired operating conditions.

4. Compare the actual and simulated operations data:
   - the nominal taxi-time distributions
   - the queue throughput curves
   - the error distribution between the actual and simulated time-in-system for each aircraft.
5. Repeat the algorithm for all of the airports and operating conditions of interest.

This algorithm is reasonably robust to statistical artifacts which are not embedded in the calibration methods or the model structure itself. This algorithm does suffer from the deficiency that certain effects are known to be unmodeled, and hence the simulation results will almost certainly fail to match the actual operations data according to standard statistical measures.

5.2 Mathematical Formulation

5.2.1 Continuous-Time

As developed in Chapter 4, the stochastic model exists in continuous-time, and obeys the following formalization:

Nominal Taxi-Time

Define the following variables:

$$
\begin{array}{rl}
A: & \text{number of airlines} \\
\rho_i^j: & i^{th} \text{ plane from the } j^{th} \text{ airline; } i \geq 1, j \in \{1, \ldots, A\} \\
P_i^j: & \text{epoch when } \rho_i^j \text{ enters the airport surface} \\
T_i^j: & \text{taxi-time for } \rho_i^j \text{ to reach the queue}
\end{array}
$$

The T_i^j are i.i.d. random variables for fixed j, while the P_i^j may be either planned or observed times. Note that ρ_i^j enters the queue at the epoch $P_i^j + T_i^j$, allowing for swapping (see Figure 5-1).

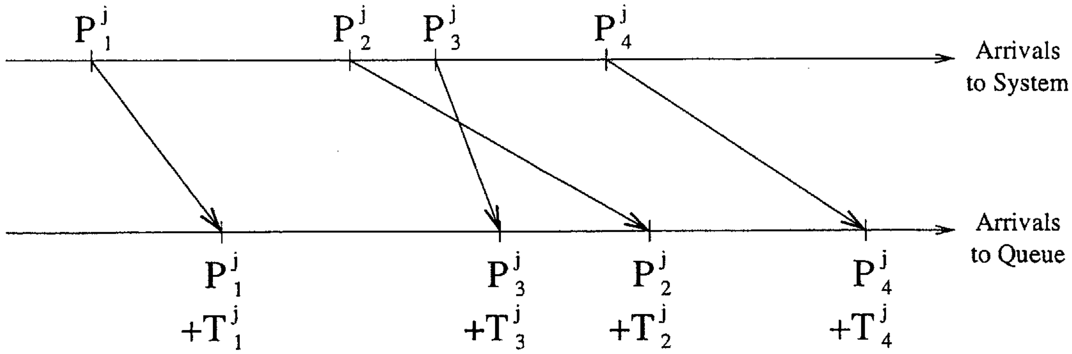

Figure 5-1: Diagram: Aircraft re-sequencing during free-flow travel.

Queueing

Define the following variables:

ϕ_n :	n^{th} plane (from any airline) to enter the queue
S_n :	epoch when ϕ_n enters the queue
R_n :	service-time of ϕ_n in the queue-server
W_n :	waiting time (queueing delay + service-time) of ϕ_n

The R_n are i.i.d. with an exponential distribution. Since the queue is FCFS, the following standard development from queueing theory applies:

$$\begin{aligned} S_n &= n^{th} \text{ element of } \left\{P_i^j + T_i^j\right\}_{i\geq 1}^{j\in\{1,\ldots,A\}} \\ &= \min\left(\left\{P_i^j + T_i^j\right\}_{i\geq 1}^{j\in\{1,\ldots,A\}} - \{S_k\}_{k\in\{1,\ldots,n-1\}}\right) \\ W_n &= R_n + \max\left\{\sum_{k=m}^{n-1}(R_k - \Delta_k)\right\}_{m\in\{1,\ldots,n\}} \qquad \text{where } \Delta_k \equiv S_{k+1} - S_k \end{aligned}$$

Note that $R_k - \Delta_k$ is the *excess* service time of ϕ_k, i.e. that part of the service time which is longer than the interarrival time between ϕ_k and ϕ_{k+1}.

5.2.2 Discrete-Time

It is possible to develop a simulation with very good time-resolution (on the order of seconds), which would accurately capture the above dynamics relative to the shortest time-constants of the system. However, all of the available operations data has a time-resolution of one minute at best. While it is relatively straightforward to adapt our calibration methods to "blur" the operations data using Parzen estimators or similar techniques, the use of an essentially illusory time-resolution does not appear to give extra insight into the behavior of the system, and invites misjudgement about the model's accuracy. Hence, for simulation purposes the following discrete-time approximation is used on a one-minute timescale:

Nominal Taxi-Time

The same variables are defined as before, but P_i^j and T_i^j are assumed to be integer-valued with units of minutes. In addition, the following two variables are defined:

$$
\begin{aligned}
P_n &\equiv \sum_i \sum_j I_{\{n\}}\left(P_i^j\right) \\
&= \text{number of planes entering the system at start of interval } n \\
T_n &\equiv \sum_i \sum_j I_{\{n\}}\left(P_i^j + T_i^j\right) \\
&= \text{number of planes entering the queue at middle of interval } n
\end{aligned}
$$

where $I_A(x)$ is the set-membership indicator for A:

$$
I_A(x) \equiv \begin{cases} 1, & x \in A \\ 0, & \text{otherwise} \end{cases} \tag{5.1}
$$

Queueing

Rather than assign a queue-service time to each aircraft, a certain number of queue-service *opportunities* is generated for each one-minute interval. The number of queue-service opportunities in successive one-minute intervals are assumed to be independent Poisson random variables with rate equal to the average queue-service rate observed under congestion-saturated conditions. This leads to the following dynamics:

$$\begin{aligned}
R_n &\equiv \text{i.i.d. sequence of Poisson random variables} \\
&= \text{number of queue-service opportunities at end of interval } n \\
Q_n &\equiv \max\left(0, Q_{n-1} + T_n - R_n\right) \\
&= \text{number of planes in the queue between intervals } n \text{ and } n+1
\end{aligned}$$

Under high-traffic conditions when the queue is never empty, this formulation is precisely equivalent to assigning each aircraft an exponentially-distributed queue-service time. Further, by Burke's theorem [6], the equivalence also holds under low-traffic conditions if the arrivals to the queue form a steady-state Poisson process at a lower rate than the queue-service rate.

There is a caveat, however. If arrivals to the queue are not Poisson, and the queue becomes empty during a one-minute interval, then the length of time it remains empty does not necessarily have the same distribution as a queue-service time, and the Poisson-opportunities formulation becomes an approximation. Fortunately, this loss of fidelity to the continuous-time model only occurs during less interesting low-traffic periods. For the purposes of this research, the simplicity of the Poisson-opportunities formulation outweighs the loss of fidelity, especially since these models have been developed in an exploratory manner and are certainly not complete stochastic descriptions of the system. It is worth noting that this formulation parallels other discrete-time descriptions of the queue-service process in terms of queue-service opportunities, including the runway-server absence concept used in [16]. For the same reason, these parallel formulations also suffer from the same difficulty during periods when the queue becomes empty.

5.3 Results

The proposed model structure and calibration methods have shown excellent robustness for describing both arrival and departure congestion under a variety of weather conditions at the airports studied in this thesis. Representative validation/simulation results for departure traffic at EWR (a notoriously congested airport) under four types of airport operating conditions are included below to indicate the nature and quality of the validation algorithm results. Table 5.1 records some relevant details describing the simulations. Figures 5-2 through 5-13 illustrate the validation comparisons described in section 5.1.

In general, the model captures the queueing delays in an aggregate sense, as can be seen from the plots of throughput rates and nominal distributions of taxi-out times. Two interesting observations can be made from the comparison of actual to simulated taxi-out times. First, the model is slightly biased towards shorter taxi-out times than those actually observed. This bias is expected, since the model does not capture the effects of downstream restrictions; similar validation results at airports which experience fewer congestion-induced problems show a smaller negative bias. Second, the error between simulated and actual taxi-out times has a large variance. Much of this volatility is due to differences in aircraft re-sequencing between the simulation and the actual operations. Small differences in the actual and simulated taxi-out times which result in different queue-positions are amplified by the queueing effects, especially since the variance of waiting time for the most recent arrival to a FCFS memoryless queue depends linearly on the number of aircraft already in the queue. This amplification is worth noting, since it impacts the controllers' tradeoff between separating flights in the queue to mitigate the effect of in-trail restrictions, and attempting to hit a given takeoff-time window despite the resulting increase in taxi-out time volatility for those aircraft which were re-sequenced further back in the queue.

Data: EWR, 1998, Departures			
Segment index	Queue thr'put (a/c per hour)	# of periods simulated	Average period length (hours)
1	25.2	152	8.5
2	27.0	143	10.8
3	22.5	39	9.8
4	25.8	42	5.0

Table 5.1: Validation/simulation data for 1998 EWR departures.

5.4 Implementation Details

The discrete-time simulation described above was implemented in C using the pseudo-random sequence **ran2** from [15], and standard implementations of random-number generators from [19]. The code was compiled for a Pentium-266MHz processor with 192MB of memory using the Linux-gnu compiler **gcc** (v2.7) with full optimization. A single pass over a full year of operations data took less than a minute to complete.

On the order of ten calls to **ran2** were required to simulate the sample-path of a single aircraft through the system, and simulations were typically run over a year of operations at a single airport; the periodicity of **ran2** ($> 2 \times 10^{18}$) was never stressed. Note that each year was broken into many shorter periods when a particular set of weather conditions prevailed, and that these periods can be assumed to be independent, substantially increasing the effective number of simulated sample-paths.

To avoid possible "edge effects", only time-periods of 2 hours or more were simulated, and only if successive time-periods were separated by gaps of at least 2 hours. In addition, the simulation was started 45 minutes before each time-period of interest (to provide a realistic initial queue population), and the simulation was run at least 45 minutes after the end of each time-period of interest (to allow any and all aircraft which entered the system during the time-period of interest to "flush out" of the system).

Segment #1:

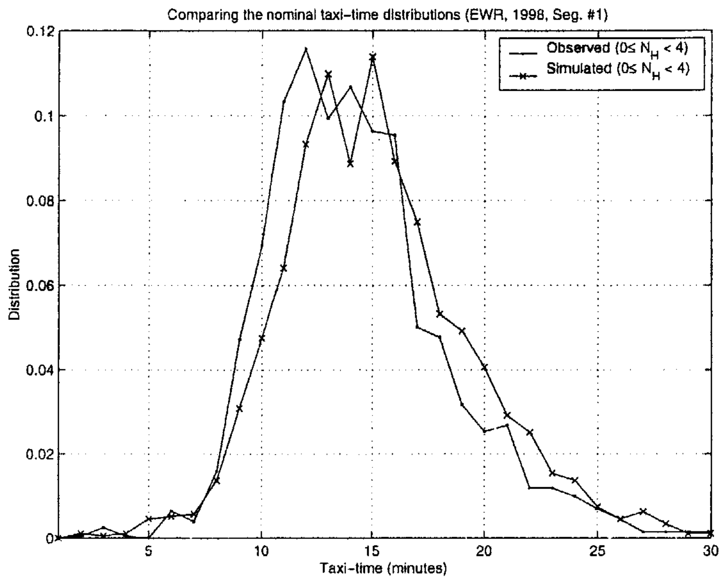

Figure 5-2: Seg. #1: Comparison of taxi-time distributions at low N_H.

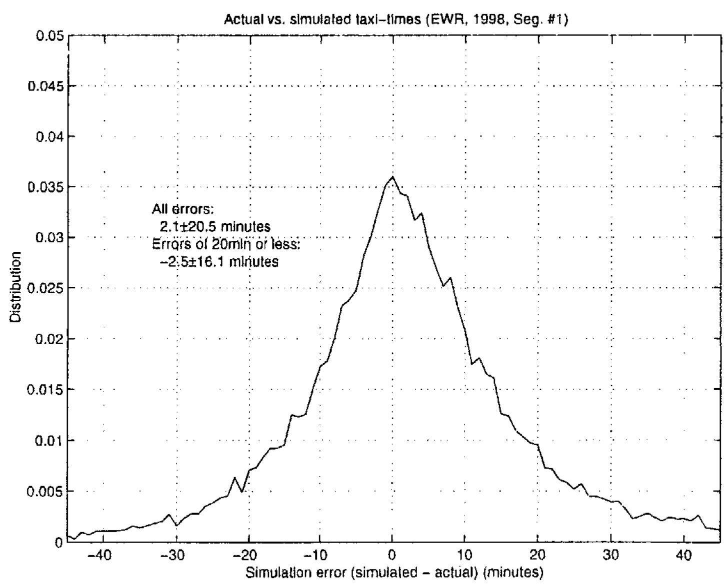

Figure 5-3: Seg. #1: Actual versus simulated taxi-times.

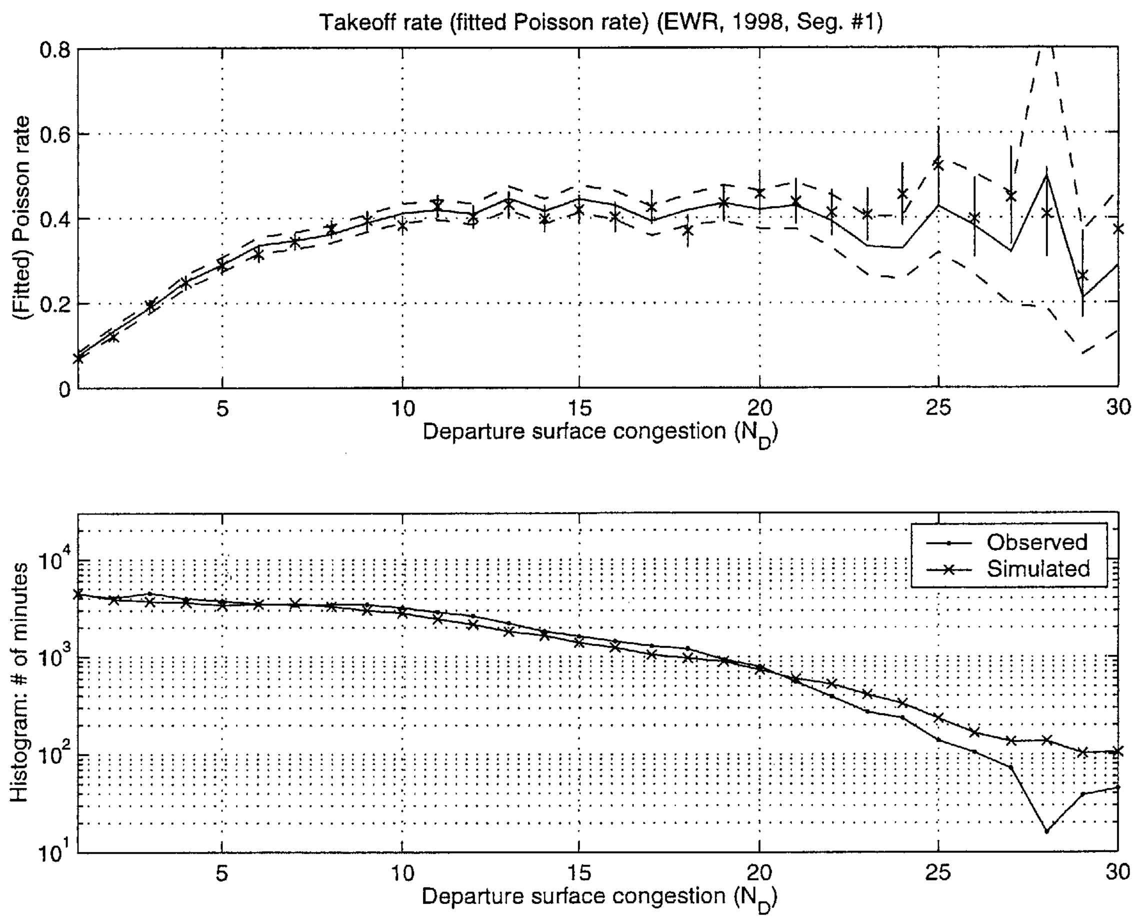

Figure 5-4: Seg. #1: Comparison of throughput curves.

Segment #2:

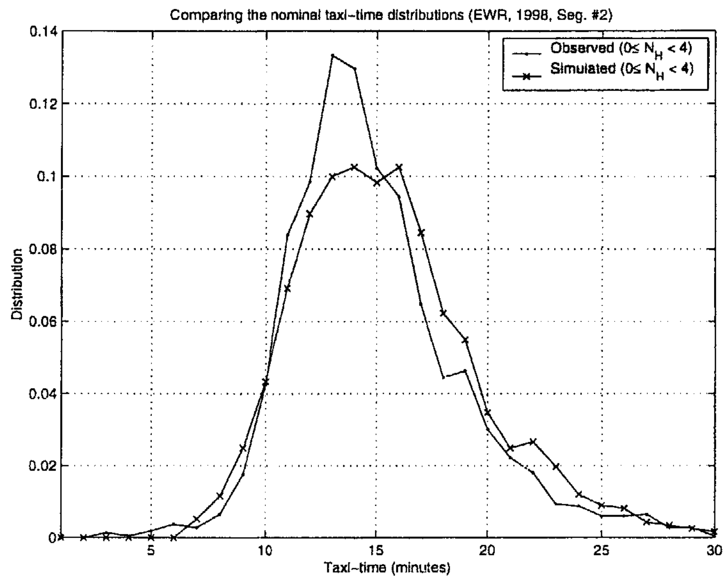

Figure 5-5: Seg. #2: Comparison of taxi-time distributions at low N_H.

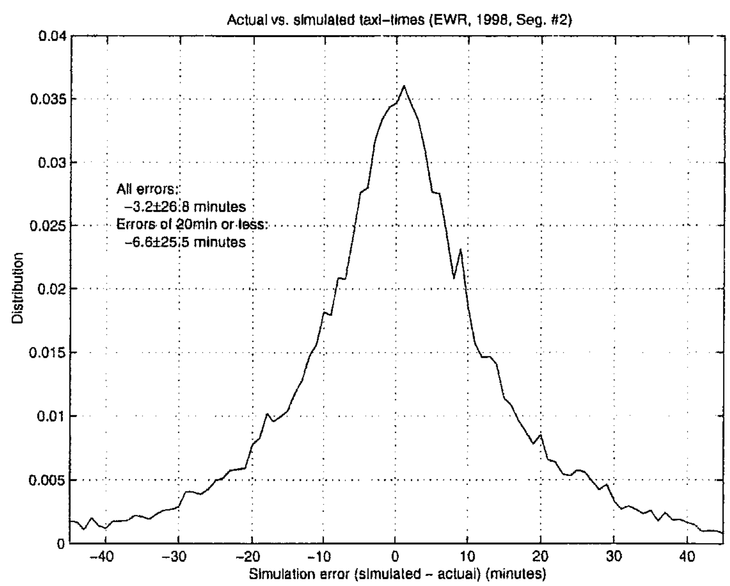

Figure 5-6: Seg. #2: Actual versus simulated taxi-times.

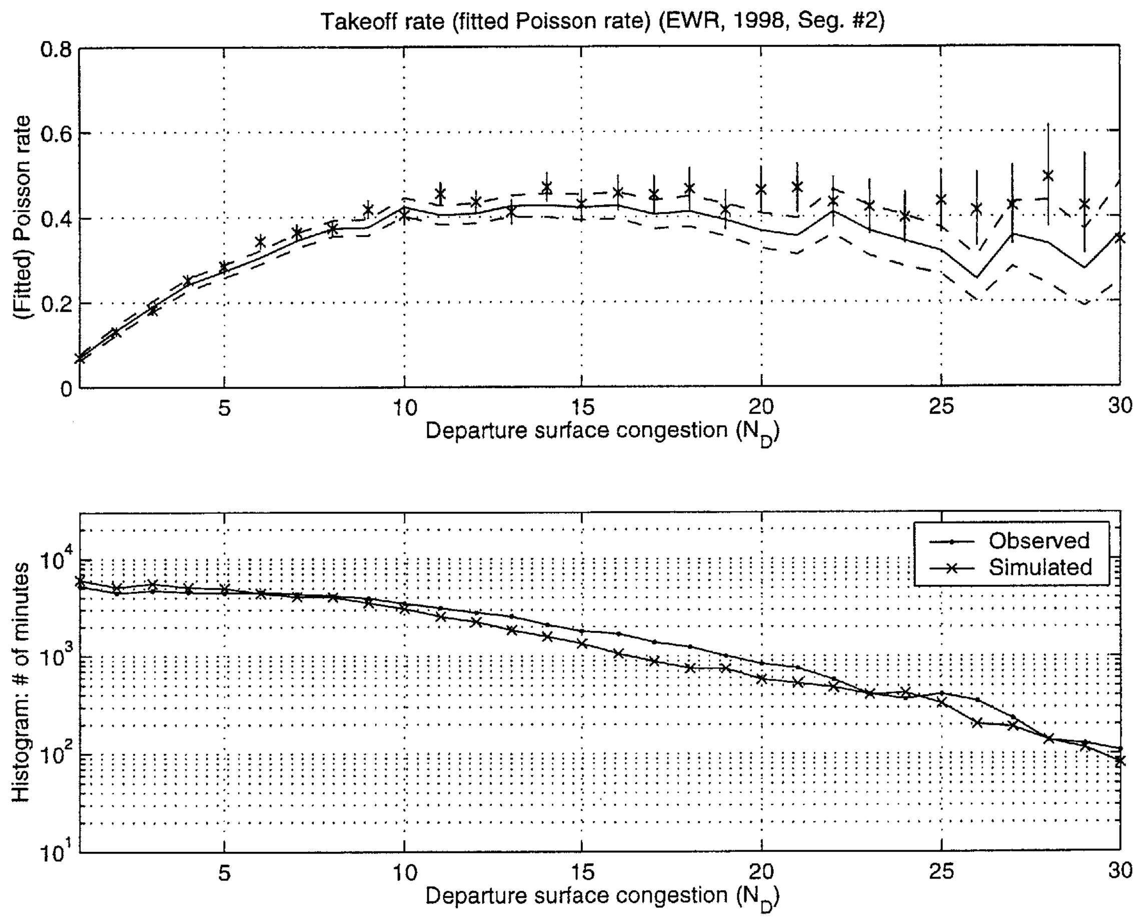

Figure 5-7: Seg. #2: Comparison of throughput curves.

Segment #3:

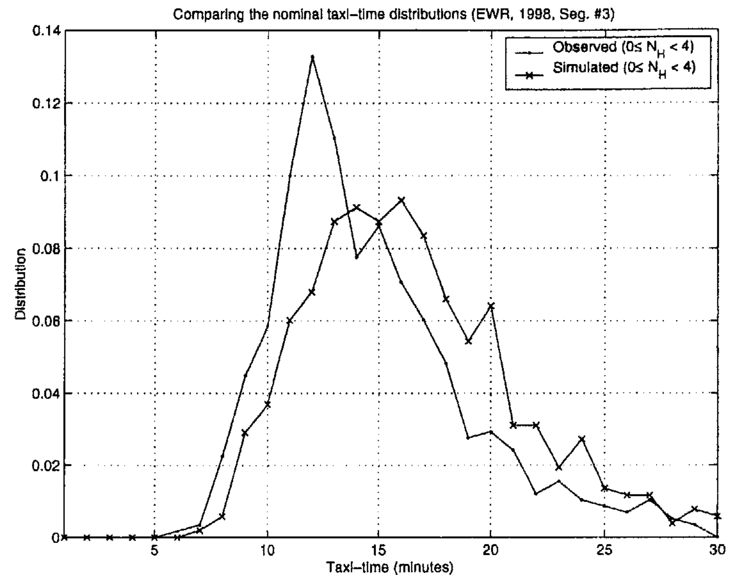

Figure 5-8: Seg. #3: Comparison of taxi-time distributions at low N_H.

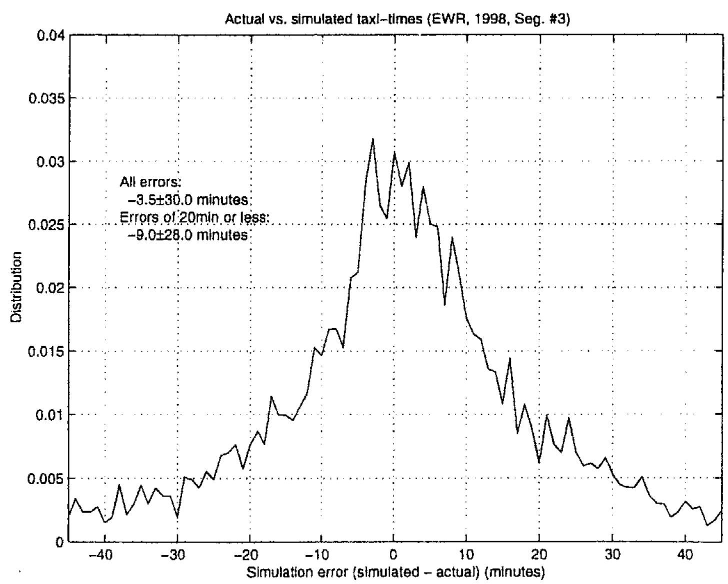

Figure 5-9: Seg. #3: Actual versus simulated taxi-times.

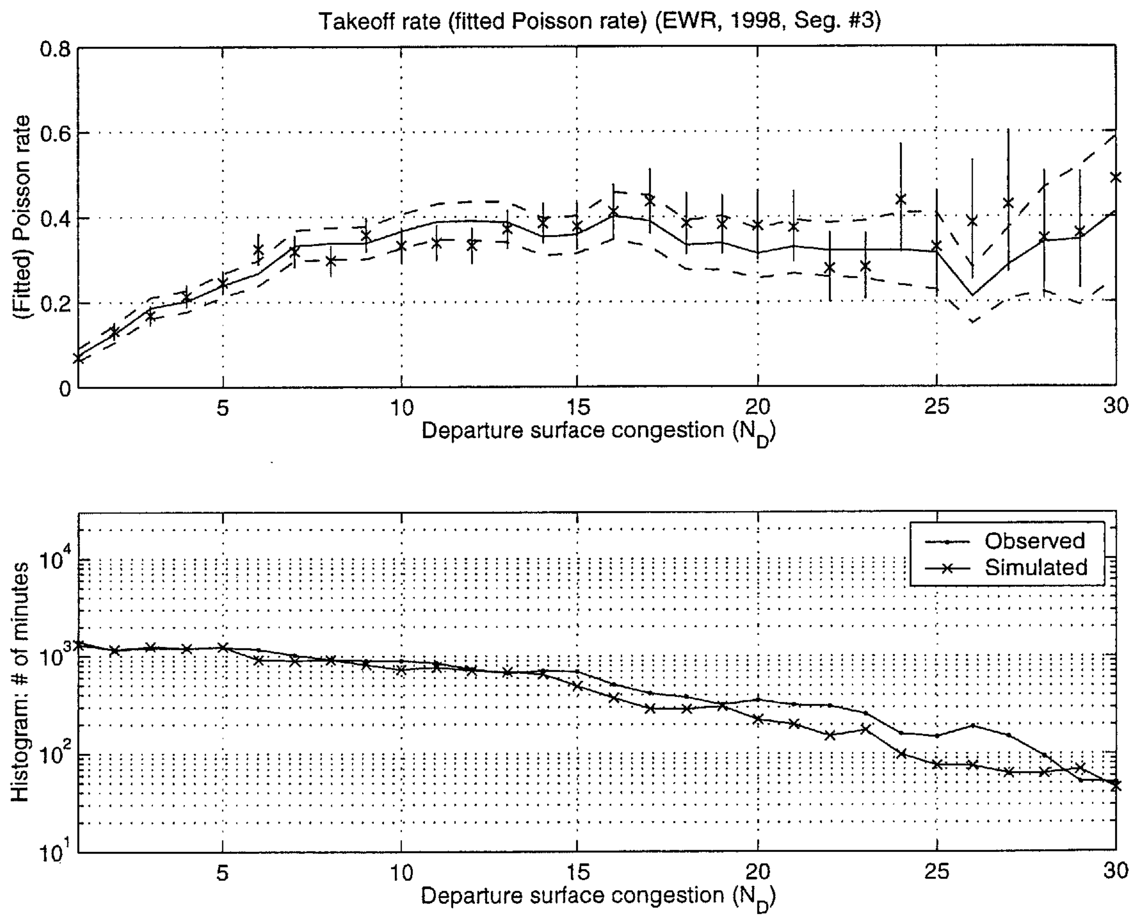

Figure 5-10: Seg. #3: Comparison of throughput curves.

Segment #4:

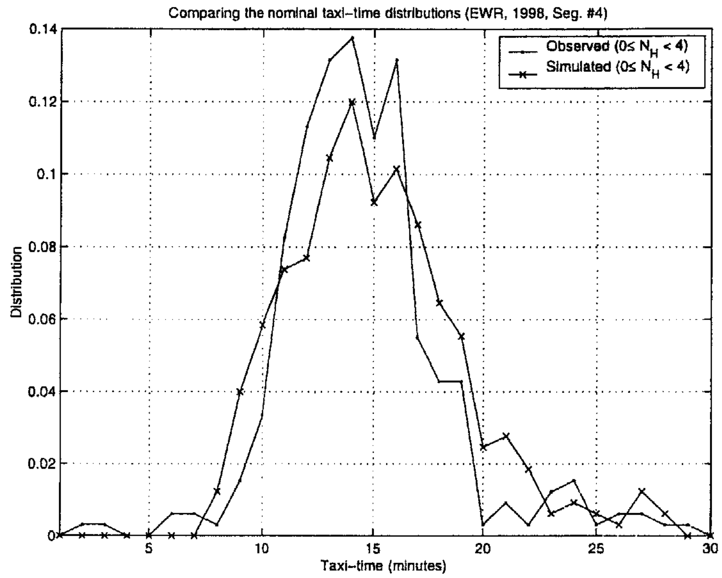

Figure 5-11: Seg. #4: Comparison of taxi-time distributions at low N_H.

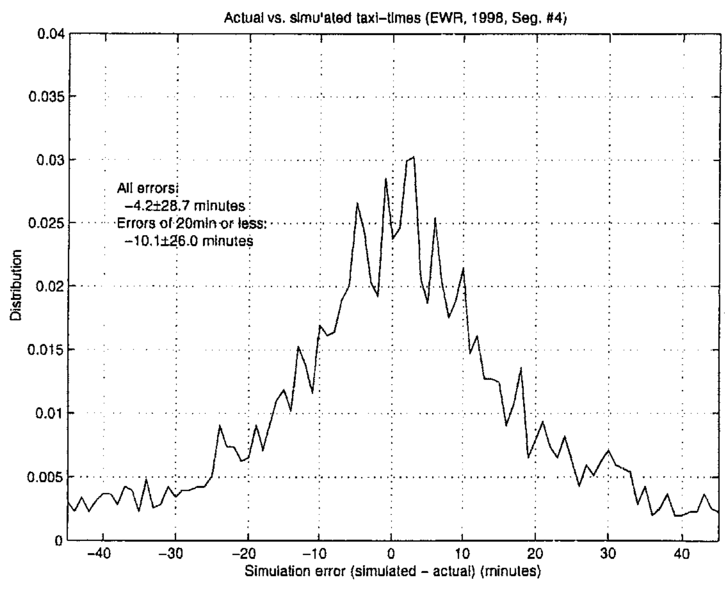

Figure 5-12: Seg. #4: Actual versus simulated taxi-times.

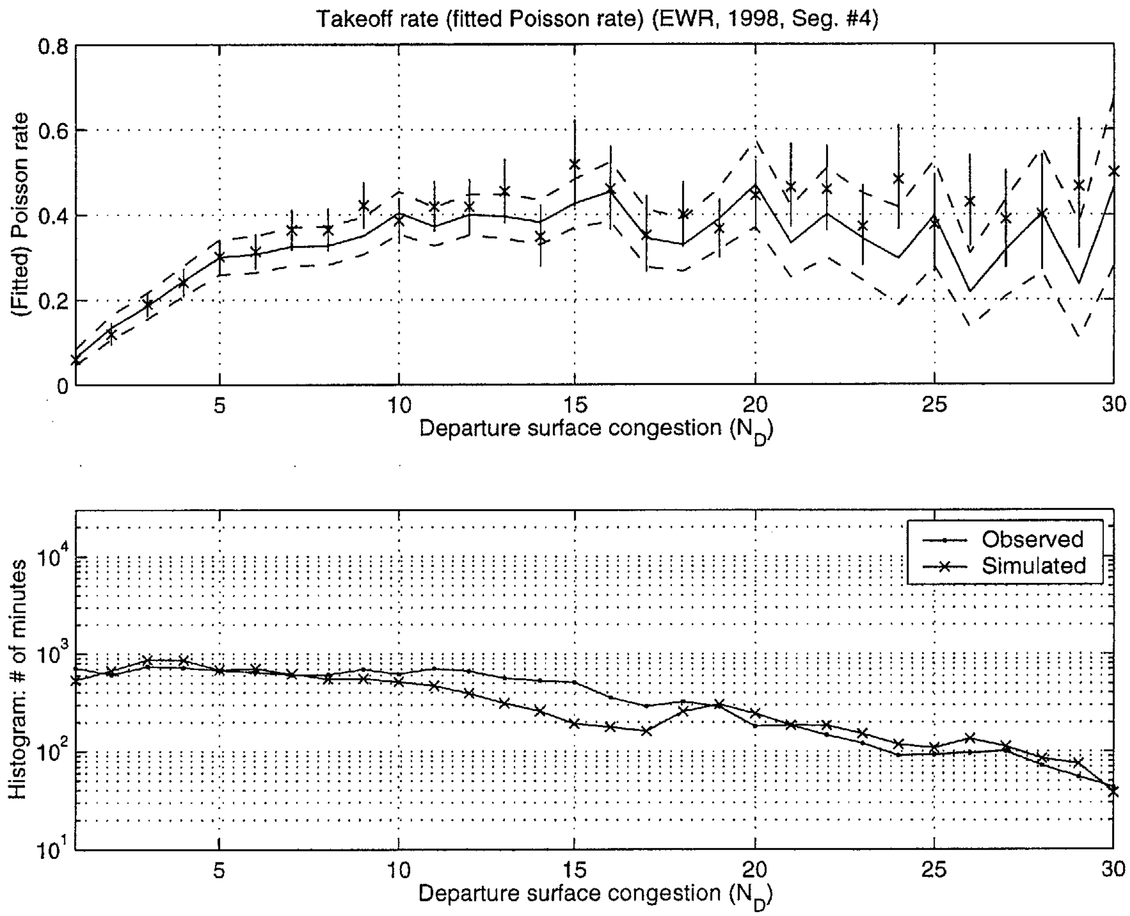

Figure 5-13: Seg. #4: Comparison of throughput curves.

Chapter 6

Runway-Queue Congestion Control

It is apparent from the departure throughput plots that the runway system has a finite capacity. In [16], based on a similar observation at BOS, a control scheme intended to minimize departure congestion and runway queueing was proposed and investigated. It was proposed that departing flights could be held at the gate if N_D exceeded some preset threshold (denoted N_C); these held flights would immediately be given pushback clearance in FCFS order when N_D dropped below N_C. This *N-control* scheme was shown to effectively trade runway queuing delays for gate-hold delays at BOS. Further investigation indicated that even strict adherence to this control scheme would cause only a small increase in the occurrence of gate shortages, and would not substantially increase total delays.

Using the model proposed in this thesis, a variant of the original N-control scheme has been investigated. The original scheme has been modified so that only one airline's flights are held at the gate; all of the other airlines are uncontrolled, althought they still contribute to N_D. This change allows us to test a practical method of implementation which does not require intervention by air traffic controllers or other centralized airport authorities.

Two new airport sites have been evaluated under airport operating conditions associated with above-average congestion and delays. The first site chosen was ATL, a major hub where Delta controlled 84% of the ASQP-recorded traffic in 1998. The second site chosen was EWR, a hub airport where Continental controlled 61% of the

ASQP-recorded traffic in 1998. One significant difference between the two sites is that Continental does not implement a bank structure at EWR, while Delta implements a strong bank structure at ATL. Both airports were evaluated during periods when the airport was running one of its primary runway-sets under IFR conditions.

A departure-process model was calibrated and validated for each airport under the specified airport operating conditions. The model structure was then modified to include a gate-holding queue whose behavior was controlled by the threshold N_C and the departure surface congestion N_D. The behavior of this controlled model was then tested under simulation.

The tradeoffs between runway queuing, gate-hold queuing, and total queuing delays are shown in Figures 6-1 and 6-2. The simulation results suggest that in the case of ATL under the specified conditions, it may be possible to directly reduce runway queuing by 40% without increasing total queuing delay, and further reductions in runway queuing are possible at the expense of increased total delay. At EWR, however, significant reductions in runway queueing come at the expense of larger increases in gate-hold queueing.

Based on the differences between the two airports, this result supports two hypotheses. One possibility is that much of the congestion, delay and inefficiency of airport surface traffic at ATL is due to the flight-bank structure, which pushes the airport runways beyond capacity on a regular basis. In contrast, the demand level at EWR is relatively constant over the course of a day, and the airlines have adapted their schedules to closely match the demand level against the airport capacity. A second possibility is that Continental's 60% traffic-share at EWR is too small to leverage large control benefits, while Delta's 80% traffic-share at ATL is sufficiently large. This is a particularly interesting hypothesis, since it suggests that our simple N-control scheme may be able to take advantage of a threshold effect arising from the nonlinear behavior of queueing systems.

The percentage of flights which experienced some level of gate-holding is indicated in Figures 6-3 and 6-4. Gate shortages impacting arriving flights could be a significant practical obstacle to implementing this control scheme. Unfortunately, gate

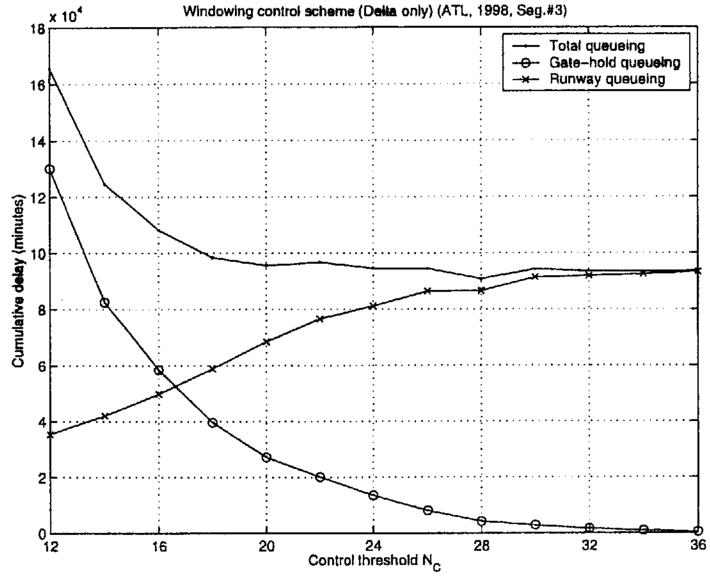

Figure 6-1: Delay tradeoffs at ATL versus the control threshold N_C.

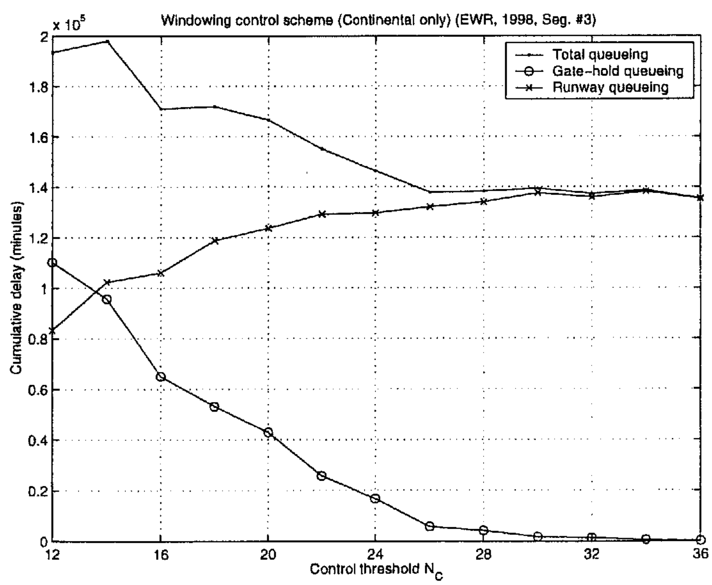

Figure 6-2: Delay tradeoffs at EWR versus the control threshold N_C.

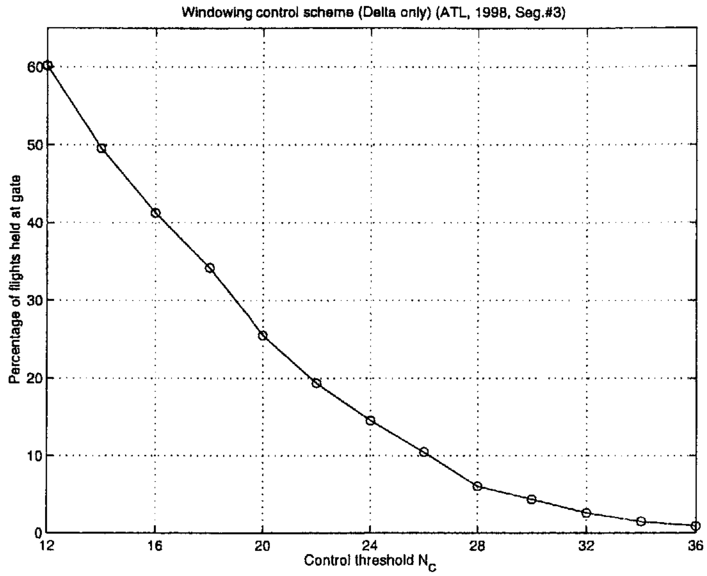

Figure 6-3: Percentage of Delta flights held at the gate vs. N_C (ATL).

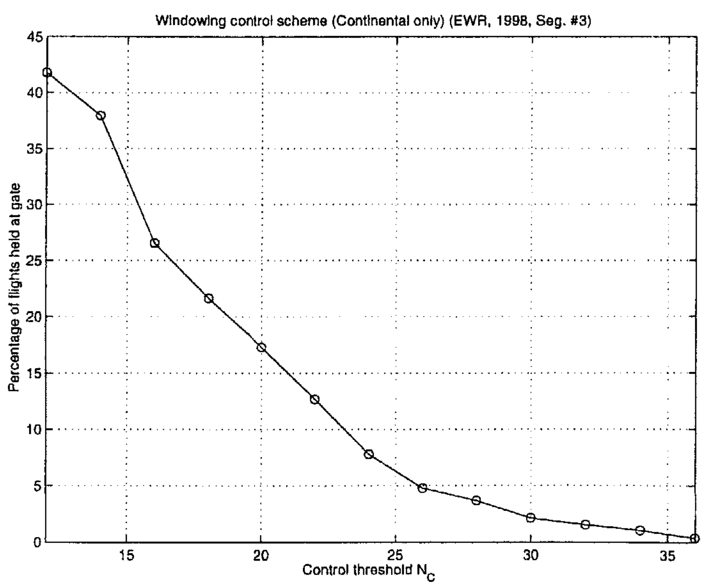

Figure 6-4: Percentage of Continental flights held at the gate vs. N_C (EWR).

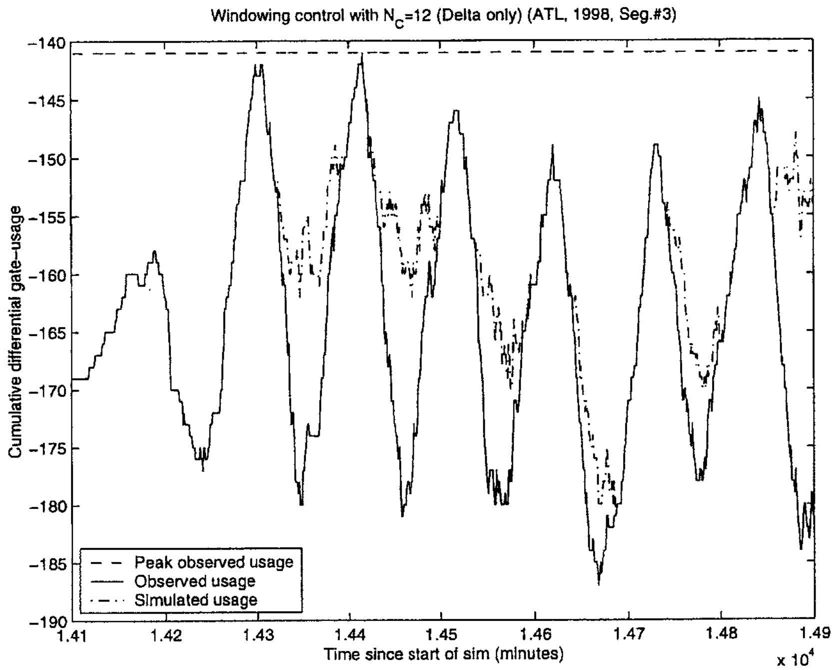

Figure 6-5: Delta gate usage under simulation (N_C=12).

assignment information is not available, and hence gate shortages can only be investigated in the aggregate. To investigate the possible occurrence of gate shortages, the cumulative differential between gate arrivals and pushbacks was calculated for the simulated operations data, and compared with the observed cumulative differential. Surprisingly, at both airports it was found that even with very stiff control thresholds, the simulated gate-usage never rose above the observed daily peak gate-usage (see Figures 6-5 through 6-10). The basic action of the control scheme is to blur out the leading edge of a departure push, spreading it over a longer period of time so that N_D never rises above N_C. At ATL in 1998, the separation between adjacent banks was large enough so that this blurring did not cause overlaps between the banks; at EWR, there is only a single large departure push in the mornings which can be accommodated through the rest of the day.

As may be apparent from the aggregate gate-usage figures, the simulation results indicate that the N-control scheme can cause significant numbers of flights to be held at the gate for periods in excess of 15 minutes. This level of delay is large

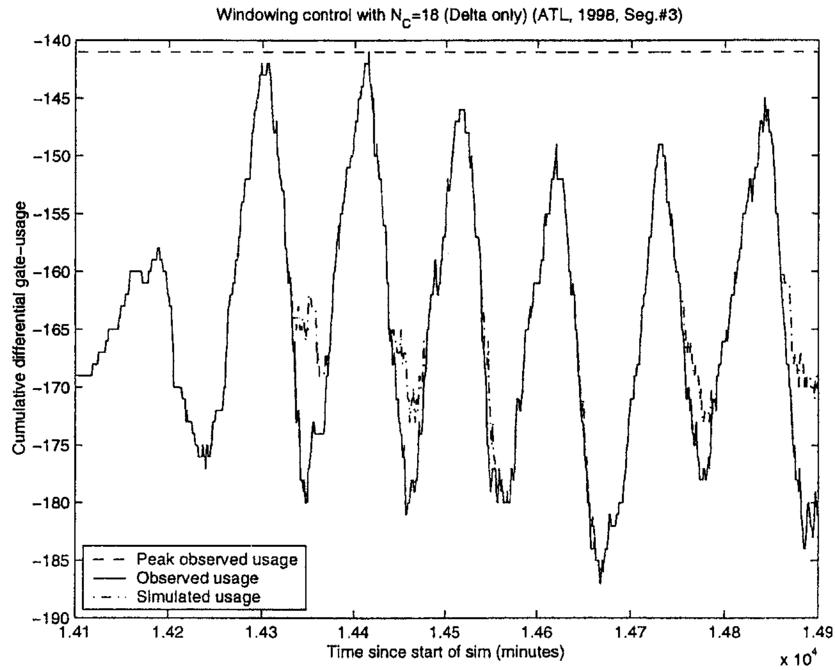

Figure 6-6: Delta gate usage under simulation (N_C=18).

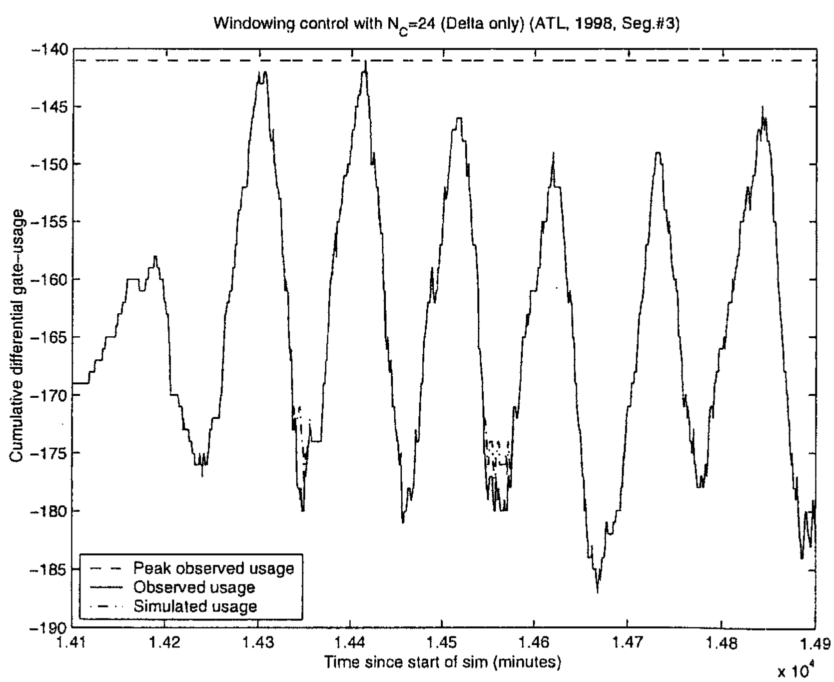

Figure 6-7: Delta gate usage under simulation (N_C=24).

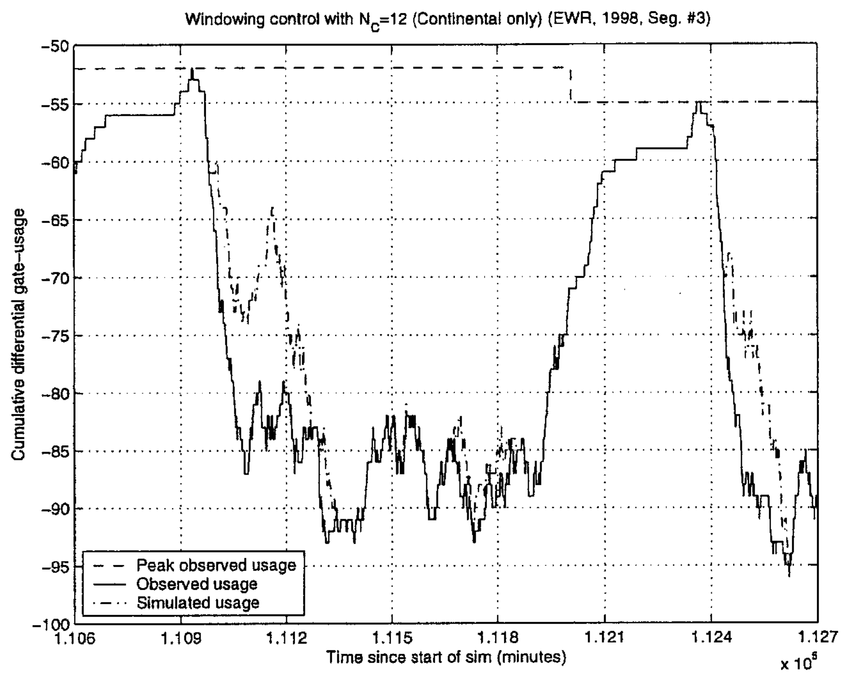

Figure 6-8: Continental gate usage under simulation (N_C=12).

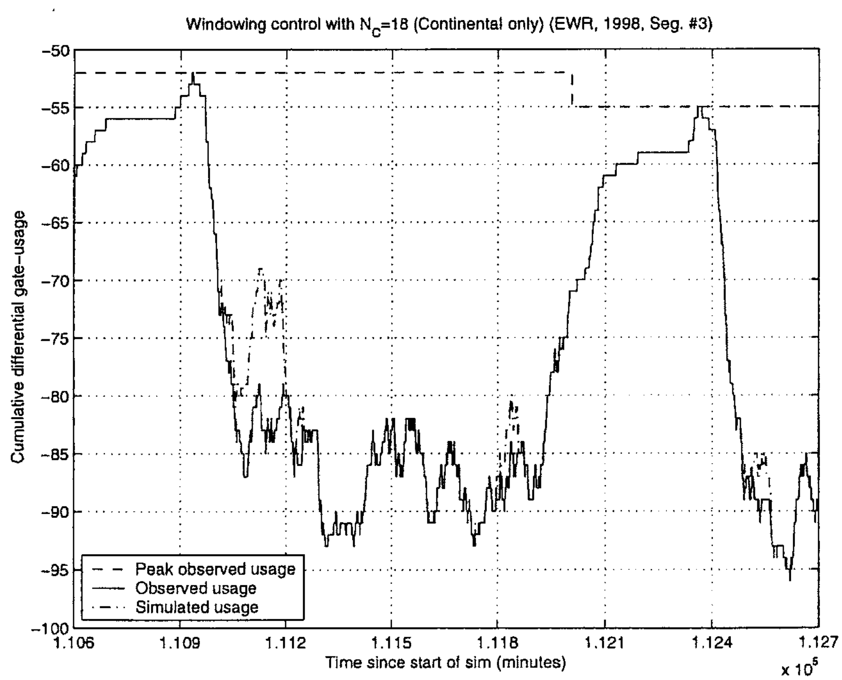

Figure 6-9: Continental gate usage under simulation (N_C=18).

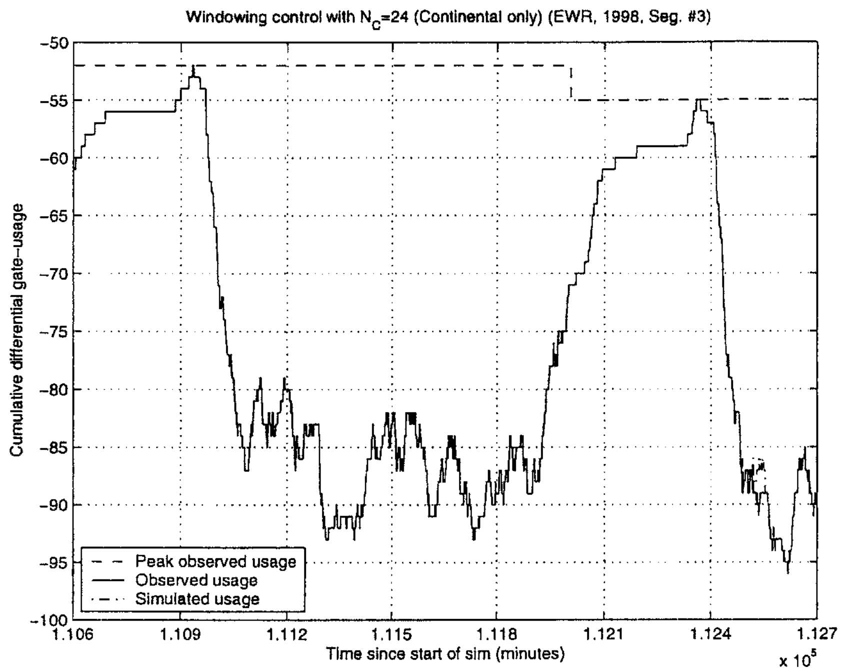

Figure 6-10: Continental gate usage under simulation (N_C=24).

enough to impact published airline on-time statistics. Consumers and the airline industry focus great attention on on-time statistics, and so this situation is untenable. Work is proceeding on refinements of the control law which limit the holding time of any particular aircraft while still obtaining the remarkable reductions in runway queueing delay. In particular, the simple holding queue which currently implements the control law could be replaced by a priority queue, into which each aircraft would be inserted according to its latest reasonable pushback time. However, in order to properly implement and evaluate this or other more sophisticated control schemes, much more thorough operations data is required, including actual gate assignments and a description of reasonable gate-usage "duty cycles".

An interesting (though not unexpected) side-effect of the control law was observed during the simulation. At ATL, the total time-in-system for *all* of the airlines was reduced to some extent by applying congestion-control to the majority-carrier at the airport. This result is very encouraging since it indicates that controlling the queue saturation, even through very rudimentary means, is indeed a valid method for improving the performance of the departure system. In contrast at EWR, no level of control was observed to decrease the total time-in-system by a substantial margin.

Chapter 7

Conclusions

7.1 Contributions of the Thesis

Model structure The basic model structure of a free-flow period followed by a queue has been largely retained from [16]. This research provides substantial improvements to the method of calibrating unimpeded travel-time distributions for the free-flow period, and has both revised and simplified the model structure and method of calibration for the queue.

Applicability It has been demonstrated that a simple stochastic model structure is capable of capturing departure and arrival surface traffic at a variety of airports under a variety of airport operating conditions. In addition, important variations on the control law of [16] have been evaluated at two new airport sites.

Lines of research Throughout the development and validation of the model, several interesting arenas for further investigation have been outlined, including refinement of the model structure; more complex control schemes; and theoretical questions pertaining to the current model. The necessity of gaining new operations data to proceed with several of these lines of research has been argued.

7.2 Work in Progress

7.2.1 Evaluation of N-control

Further investigations of the N-control scheme are underway. The effects of several variables can be tested using the current model and N-control law.

The effects of bank width, duration and spacing can be tested. In particular, the observed departure and arrival demand can be compared with the calibrated runway and gate throughputs to determine how excess demand is related to queueing delays. It may be possible to design reasonable standards for bank structures which mitigate queueing and delays.

One experiment of interest to airlines is to drive the model with the *scheduled* demand (from OAG or CRS), and then compare the predicted and actual taxi-times. It may be feasible to use the stochastic model presented in this thesis as a planning tool to estimate the mean and variance of taxi-times when constructing new airline schedules. Given accurate information about taxi-times, sufficient slack time could be built into a schedule to achieve a desired level of on-time performance; from another viewpoint, the reliability of a proposed schedule could be more accurately evaluated given such information.

The effect of controlling only some fraction of the airport traffic can be evaluated as well. For example, if an airline controls a sufficiently large portion of the traffic at a particular airport, it may be able to accurately control the congestion and resulting queueing delays. However, if the fraction of controlled traffic falls below some threshold, it may not be possible to use that traffic to leverage significant performance benefits from the system.

7.2.2 Improvements to the model structure

Turboprop and regional carrier traffic (which are notably absent from the ASQP database) are currently treated as additional sources of stochastic noise in the system, although in principle the current queuing model can be trivially extended to include both sorts of traffic. Additionally, some aircraft experience significantly longer taxi-out times due to downstream restrictions, and work is underway to accommodate these outliers. It is expected that more complete datasets would allow both of these noise-sources to be properly incorporated into the current model in short order.

It is intuitively obvious that on a very short time-scale there must be some tradeoff between landings and takeoffs on the same runway. This tradeoff is currently treated as an additional source of stochastic noise in the runway behavior, but work is currently in progress to explicitly model this effect in the behavior of the runway queue. In particular, we are investigating several stochastic point-process estimation techniques which have proven useful in the field of neurobiology to study the inhibitory/excitatory coupling between neurons [7, 2]. Runway assignments for arriving and departing flights are required to fully elucidate these effects. Further operations data is also necessary to correctly identify arrival/departure interactions on the taxiways and at the gates.

7.2.3 Theoretical development

Exciting work has begun on theoretical analysis of the model. Again, it is important to note that this research has been conducted in an exploratory manner and should not be construed as mathematically robust. However, the proposed model structure has great promise for describing the stochastic behavior of airport surface traffic, and more substantial theoretical work is warranted.

A question of fundamental interest is how to optimally "feed" a queue through the timed release of traffic. A classic application of this problem, and one of interest to ATC's, is the translation of controlled takeoff times to controlled pushback times, including a robust means of dealing with uncertainties due to local queueing effects. Currently, local ATC's must often implement ground-delay programs which do not necessarily account for taxi-time uncertainty in a robust manner. Another application of interest to airlines is to maximize the time-at-gate of a particular aircraft given fixed landing and takeoff times. Airlines currently prefer to push back flights as soon as possible, both to avoid the imposition of gate-hold delay through ATC programs, and to avoid delays due to a poor rank in the runway queue. However, aircraft at the gate can catch late passengers and missed baggage connections, have lower direct-operating costs, and do not contribute to the ramp and taxiway surface congestion which delays other aircraft.

A successful solution to both these problems could have far-reaching effects, and we are working to formulate both problems as robust optimization problems using our current model structure. We are also working to develop a tractable Markovian representation of the model in order to develop algorithms similar to the Viterbi algorithm or Kalman filtering to dynamically estimate the hidden system-state based on observed input/output behavior. This work leads naturally to substantial theoretical questions of system identification, feedback-based estimation of taxi-out times, and optimal stochastic control of airport surface congestion.

Appendix A

Useful Acronyms

ASDE ATC technology: Airport Surface Detection Equipment (short-range radar for monitoring aircraft on the airport surface during periods of very poor visibility)

ASQP Database: Airline Service Quality Performance (scheduled and actual times of jet operations for 10 major US passenger carriers, covering at least January 1995 to the present)

ATC Air Traffic Control

ATL Airport: Atlanta/Hartsfield, GA

BOS Airport: Boston, MA

CDM Joint ATC/Airline project: Collaborative Decision Making (primarily based on information-sharing between airlines and ATC)

CRS Airline database: Computerized Reservation System (used by ticket agents to make reservations, and by airlines as an accurate tactical schedule over the course of each day)

CTAS ATC technology: Center-TRACON Automatic System

CODAS Database: Consolidated Operations and Delay Analysis System (amalgamation of several FAA-maintained weather and operations databases, covering at least 1994 to the present)

DFW Airport: Dallas/Fort Worth, TX

DSP ATC terminology: Departure Sequencing Program

EDCT ATC terminology: Expect Departure Clearance Time

ETMS ATC technology: Enhanced Traffic Management System

EWR Airport: Newark, NJ

IAH Airport: Houston, TX

IFR ATC terminology: Instrument Flight Rules (when ceilings fall below 1000 feet, visibility falls below 3 miles or inclement weather conditions predominate, only instrument-rated pilots in properly-equipped aircraft can use the airport; certain operating procedures also change)

ICAT Research lab: International Center for Air Transportation, MIT

NAS National Airspace System

OAG Database: Official Airline Guide (published approximately four times yearly; contains the proposed airline schedules for the next several months)

PRAS Database: Preferential Runway Assignment System (BOS runway configurations, covering September 1993 to the present)

SIMMOD Simulation software: Airport and Airspace Simulation Model (stochastic simulation package for modeling network representations of airfields and airspace, developed by the FAA)

TAAM Simulation software: Total Airspace and Airport Modeller (fast-time simulation package for modeling entire air traffic systems, developed by The Preston Group in cooperation with the Australian Civil Aviation Authority)

VFR ATC terminology: Visual Flight Rules (flight rules under good-weather conditions; compare with definition of IFR above)

Bibliography

[1] K. Andersson, F. Carr, E. Feron, and W. Hall. Analysis and modeling of ground operations at hub airports. In *Proceedings of 3rd USA/Europe Air Traffic Management R&D Seminar*, Website: "http://atm-seminar-2000.eurocontrol.fr/", July 2000. EuroControl and the FAA.

[2] David R. Brillinger. Nerve cell spike train data analysis: A progression of technique. *Journal of the American Statistical Association*, 87(418):260–271, June 1992.

[3] J.P. Clarke. Unpublished research work. International Center for Air Transportation, Massachusetts Institute of Technology, September 1999.

[4] Bertrand Delcaire and Eric Feron. Dealing with airport congestion: Development of tactical tools for the departure flows from a large airport. Technical report, Massachusetts Institute of Technology, International Center for Air Transportation, 1998.

[5] Antony Evans. Unpublished research work: "Flattening" of Continental Airlines bank-structure at EWR. International Center for Air Transportation, Massachusetts Institute of Technology, 1999.

[6] Robert G. Gallager. *Discrete Stochastic Processes*, chapter 2, page 202. Kluwer Academic Publishers: Boston, 1998.

[7] Peter Guttorp. *Stochastic Modeling of Scientific Data*, chapter 5. Stochastic Modeling Series. Chapman and Hall, New York, 1995.

[8] Peter Guttorp and Mary Lou Thompson. Nonparametric estimation of intensities for sampled counting processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):157–173, 1990.

[9] Husni Idris. *Observations and Analysis of Departure Operations at Boston Logan International Airport.* PhD thesis, Massachusetts Institute of Technology, 2000.

[10] Husni Idris. Operations at Newark. Personal communications: email, August 2000. Based on ICAT on-site investigations at EWR control tower.

[11] Husni Idris, Bertrand Delcaire, Ioannis Anagnostakis, William Hall, Nicolas Pujet, Eric Feron, R. John Hansman, John Paul Clarke, and Amedeo Odoni. Identification of flow constraints and control points in departure operations at airport systems. *Air Traffic Control Quarterly*, 2000.

[12] SimuFlite Training International. Runway operations at DFW.
Website: "http://www.cefcorp.com/simuflite/"
(Previously: "http://www.simuflite.com/atc/map_frm.html"), 1998.

[13] Katherine Kam. Unpublished research work: Comparison of ASQP times versus CTAS-recorded radar events. International Center for Air Transportation, Massachusetts Institute of Technology, 1999.

[14] Alan Mitkiff. Personal communications: interview, 1999.

[15] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*, chapter 7. Cambridge University Press: Boston, 2nd edition, 1992.

[16] Nicolas Pujet. *Modeling and Control of the Departure Process of Congested Airports.* PhD thesis, Massachusetts Institute of Technology, 1999. Note: these results are also documented (more accessibly) in [17] and [18].

[17] Nicolas Pujet, Bertrand Delcaire, and Eric Feron. Input-output modeling and control of the departure process of busy airports. *Air Traffic Control Quarterly*, 8(1):1–32, 2000.

[18] Nicolas Pujet and Eric Feron. Input-output modeling and control of the departure process of busy airports. In *Conference on Guidance, Navigation and Control*, Portland, OR, August 1999. AIAA.

[19] Sheldon M. Ross. *Introduction to Probability Models*, chapter 11. Academic Press: San Diego, 6th edition, 1997.

[20] R.A. Shumsky. *Dynamic Statistical Models for the Prediction of Aircraft Take-Off Times.* PhD thesis, Massachusetts Institute of Technology, 1995.

[21] Carlton Wine. FAA/Aviation Planning Office: Data Systems. Website: "http://www.apo.data.faa.gov/", 2000.