Robust Decision-Support Tools for Airport Surface Traffic

by

Francis R. Carr

B.S., Harvey Mudd College (1998)
M.S., Massachusetts Institute of Technology (2001)

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2004

Author .....
Department of Electrical Engineering and Computer Science
October 31, 2003

Certified by .....
Eric Feron
Associate Professor of Aeronautics and Astronautics

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

Robust Decision-Support Tools for Airport Surface Traffic

by

Francis R. Carr

Submitted to the Department of Electrical Engineering and Computer Science
on October 31, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Forecasts of departure demand are one of the driving inputs to tactical decision-support tools (DSTs) for airport surface traffic. While there are well-known results on average- or worst-case forecast uncertainty, it is the forecast errors which occur under *best-case minimum-uncertainty* conditions which constrain robust DST design and the achievable traffic benefits. These best-case errors have never previously been characterized.

Several quantitative models and techniques for computing pushback forecasts are developed. These are tested against a dataset of 17,344 real-world airline ground operations covering 3 months of Lufthansa flights transiting Frankfurt International Airport. The Lufthansa dataset includes detailed timing information on *all* of the turn processes, including deboarding, catering, cleaning, fueling and boarding. The dataset is carefully filtered to obtain a sample of 3820 minimum-uncertainty ground events. The forecast models and techniques are tested against this sample, and it is observed that current pushback forecast errors (on the order of $\pm$15min) cannot be reduced by a factor of more than 2 or 3. Furthermore, for each ground event, only 3 observations are necessary to achieve this best-case performance: the available ground-time between actual onblock and scheduled offblock; the time until deboarding begins; and the time until boarding ends.

Any DST used in real-world operations must be robust to this "noise floor". To support the development of robust DSTs, a unified framework called *ceno-scale modelling* is developed. This class of models encodes a wide range of observed delay mechanisms using *multi-resource synchronization* (MRS) feedback networks. A ceno-scale model instance is created for Newark International Airport, and the parameter sensitivity and model fidelity are tested against a detailed real-world dataset. Based on the validated model framework, several robust *dual* control strategies are proposed for airport surface traffic.

Thesis Supervisor: Eric Feron
Title: Associate Professor of Aeronautics and Astronautics

Acknowledgments

I would like to thank my thesis supervisor Professor Eric Feron for his enthusiasm and support throughput this project. Professor John-Paul Clarke has also been an excellent co-conspirator, editor, and general motivating force. Professors John Tsitsiklis and Patrick Henry Winston have been invaluable in their role as "outsiders looking in". Because they have not been immersed in the problem from the very beginning, they have been able to provide a fresh and objective view of the significance, value and presentation of this work.

Several co-workers have contributed significantly to this work, including Ioannis Anagnostakis, Husni Idris, Antony Evans, Georg Theis, Laurence Vigeant-Langlois and Nhut Tan Ho. In a wider scope, MIT can be a lonely place, or at least it was for me until I settled down in ICAT and LIDS and found the students there very welcoming.

The ground event timing data of Chapter 2 is remarkable, and enables an exceptional opportunity for systematic analyses of the turn process using accurate observations collected over a long period of time. I am sincerely indebted to Deutsche Lufthansa AG for this opportunity, and to Georg Theis (previously Senior Manager ALLEGRO, currently a PhD candidate at MIT) for his inestimable role in obtaining and analyzing these data. I would like to thank Arno Thon (Senior Manager ALLEGRO) and Manfred Rosenthal (Manager ALLEGRO) for their hospitality at Frankfurt International Airport, and for their professional and able assistance in obtaining and interpreting the ALLEGRO data.

This work has received financial support from 4 semesters of MIT graduate teaching assistantships; the NASA/Ames Cooperative Agreement #NCC 2-1147 (Departure Planner Project); and the FAA Grant #FAA 95-G-017 (Joint University Program).

Finally, my most humble thanks go to Jen, my long-suffering spouse, who has sacrificed the company of her boon companion over so terribly many lonely evenings and weekends. I look forward to being married again.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since deregulation in 1978, steadily growing demand for air transportation has exposed bottlenecks in the National Airspace System where traffic can easily outstrip capacity. The resulting delays and system instability have driven the development of procedures and automation to accommodate the increased demand. From the viewpoint of Air Traffic Control (ATC), this accommodation has focused on handling an increasing number of flights, from pushback at the gates to landing at a destination airport. From the viewpoint of the airlines, the increasing traffic volume has led to larger and more complex problems of resource scheduling and synchronization in their ground operations. On a tactical timescale of minutes to hours, the principal concern of air transportation is to maintain an efficient cycling of aircraft, crews, passengers and cargo between these two domains of flights and ground operations. This thesis investigates several longstanding open questions regarding the transition between ground operations and flights.

The transition from flights to ground operations has been improving for several years. It is more risky, difficult and expensive to hasten/delay airborne aircraft [DFB02] than departures from the runway [BS00b], and hence the management of airborne traffic has been given a higher research priority. Once flights are airborne, they are readily observable via the Enhanced Traffic Management System, which uses radar data to monitor and predict the progress of each flight through the National Airspace System. The Enhanced Traffic Management System has excellent predictive capabilities up until flights enter the holding stacks near airports [Gor93, Chi93]. Several decision-support tools (DSTs) are beginning to fill the observability and predictability gap from the holding stacks down to the ground, with concomitant benefits for airline ground operations. As part of the NASA Surface Movement Advisor, airlines at Atlanta Hartsfield International Airport (ATL) were given an FAA radar feed covering the ATL Terminal Radar Approach Control airspace. This radar feed allowed the airlines to observe arriving flights leaving the holding stacks roughly 15min before landing. Anecdotal evidence indicates that Delta Airlines (the dominant carrier at ATL) was able

to save $10–30 million yearly by using the improved arrival-time estimates. A similar radar feed has since been installed for airlines at Philadelphia International Airport, Detroit Metropolitan Wayne County Airport, Chicago O'Hare International Airport, Dallas-Fort Worth International Airport, Newark International Airport (EWR) and Teterboro Airport. In more concrete terms, the benefits of improved landing-time estimates have been estimated in [AHAF02]. Forecasts of landing-times at Dallas-Fort Worth International Airport from the Collaborative Arrival Planner, a software tool in the NASA Center-TRACON[1] Automation System, have an empirical standard deviation of ±3min at a 30min forecast horizon. Before the Collaborative Arrival Planner was implemented, similar airline predictions had a standard deviation of ±5min. It was then demonstrated in [AHAF02] that airline ground operations plans which were generated from the reduced-noise forecasts could be implemented with less disruption against actual operations, saving an estimated 150–600min of passenger delay per hour. These savings were *not* due to improvements in the planning process itself, but only in reducing the mismatch between planning and implementation.

One goal of the FAA's Free Flight Phase II program is the development of decision-support tools for airport surface traffic [RTC00]. The role of these DSTs is to automate some of the monitoring, prediction, control and management tasks currently performed by human air traffic controllers responsible for airport surface traffic. The proposed benefits include increased airport throughput, higher efficiency of taxi operations, and improved economic performance for air carriers. These benefits must be achieved without increasing controller workload or sacrificing system safety. At the present time several prototype DSTs which address this goal are deployed and/or undergoing active research and development.

In contrast to the flight-ground transition, the transition from ground operations to flights has had a lower research priority, and mechanisms for automated monitoring of airport surface traffic are still in development [Atk02]. There have been fewer notable successes in developing DSTs to support the transition from airline ground operations into the National Airspace System. NASA Ames Research Center in cooperation with the FAA has developed the Surface Movement Advisor (SMA) system currently in use at ATL and partially deployed at several other major airports [GG97]. The Surface Management System (SMS) is a newer Ames/FAA cooperative project which is presently being developed and field-tested at Memphis International Airport [ABW02]. Both SMA and SMS have been implemented for use in ATC towers and airline stations to provide real-time status information and shared awareness on airport surface traffic. In particular both systems help maintain controller situational awareness with respect to expected future departure demand, runway queue lengths, taxi-out delays and airport departure rates for multiple possible tactical scenarios. The Center for Advanced Aviation System Development at MITRE is developing the Departure Enhanced Planning and Runway/Taxiway Assignment System (DEPARTS), an optimization-based tool which

[1]TRACON: "Terminal Radar Approach Control". Please see Appendix B for a definition; this acronym is both unique to ATC and in standard use.

incorporates current airport conditions, departure demand, taxiing aircraft status, downstream traffic flow restrictions and user preferences to optimally assign and sequence traffic to taxi routes, runways and departure fixes [CCF+02]. The Ground-Operation Situation Awareness and Flow Efficiency (GO-SAFE) concept is currently under development at Optimal Synthesis Inc. [Che02]. This tool has a similar technical focus and optimization-based approach as DEPARTS. However rather than being integrated into the ATC towers and airline stations, GO-SAFE is geared more towards use on the flight-deck to provide precision guidance, navigation and clearance delivery during the taxi process. Under the European programme of Advanced Surface Movement Guidance and Control Systems (A-SMGCS) [EUR95], the German Aerospace Research Establishment (DLR) is currently developing the Taxi And Ramp Management And Control (TARMAC) [B$\ddot{9}$4, VB97]. Because European departure traffic is typically assigned a *departure slot* (a fixed time-window for takeoff), TARMAC assists European ATC in routing aircraft across the airport surface in order to correctly meet the assigned slot. The National Aerospace Laboratory (NLR) is developing another project under A-SMGCS called Mantea (MANagement of surface Traffic in European Airports) [HB98]. In contrast to TARMAC, Mantea first sequences takeoffs at the runway and then works backwards to find optimal routings. Additional tools may also be under development; these demonstrate the range of technical approaches and human-factors integration issues for DSTs for airport surface traffic.

Where the published literature on these DSTs considers real-world implementation issues (as opposed to development of algorithms), there are two commonly-stated engineering assumptions related both to feasibility and performance. The first assumption is the availability of surface surveillance data, e.g. from multiple-sensor systems such as Airport Surface Detection Equipment-X band (ASDE-X). For example, selection of Memphis as a development site for SMS was influenced by the presence of existing infrastructure for the FAA's Safe Flight 21 program which duplicates ASDE-X performance [AB02]. Similarly initial modeling and site-adaptation of DEPARTS focused on ATL due to the availability of infrastructure from the SMA program [CCD+01]. The FAA is currently pursuing advanced surface surveillance through the ASDE-X program, and it is not unreasonable to expect that many airports will have the necessary infrastructure within the roll-out timeframe of current DST projects.

The second assumption is the availability of accurate and timely departure demand forecasts, where departure demand is interpreted as air carrier pushbacks at airports where movement on the ramp is under FAA control, or arrival of aircraft to ramp/taxiway transfer points at airports where ramp movement is under airline control. For example, in the SMS proposal one of the primary reasons cited for the selection of Memphis as an initial site was the availability of partial pushback information from the two major carriers (Northwest Airlines and Federal Express) [WB00]. Similarly the selection of ATL for DEPARTS was influenced by the existing SMA infrastructure [CCD+01].

This thesis presents novel analyses of departure demand forecasts. In contrast to many other

studies of airline ground operations, these analyses focus on the *best-case* pushback predictability rather than average- or worst-case conditions. The best-case predictability corresponds to an upper bound on DST performance, because by definition it is not possible to further reduce uncertainty in the system. An important conclusion of this best-case analysis is that pushback forecast error is lower-bounded away from zero at a standard deviation of several minutes. A comparison with the error in current pushback forecasts shows that the current errors cannot be reduced by a factor of more than 2 or 3.

This unavoidable inherent stochasticity imposes a robustness requirement on any DST for airport surface traffic. To further support the development of robust DSTs, a *ceno-scale* modelling framework is then developed. Ceno-scale models extend previous research on robust mesoscopic modelling to encompass a much wider range of the delay mechanisms observed in airport surface traffic; the increase in model complexity is illustrated in Figure 1-1. A specific ceno-scale model instance is calibrated against real-world operations data and used in Monte Carlo simulations to examine issues of parameter sensitivity and model fidelity. In turn, the concepts underlying ceno-scale modelling suggest several abstract control mechanisms which may be suitable for use in airport surface traffic DSTs.

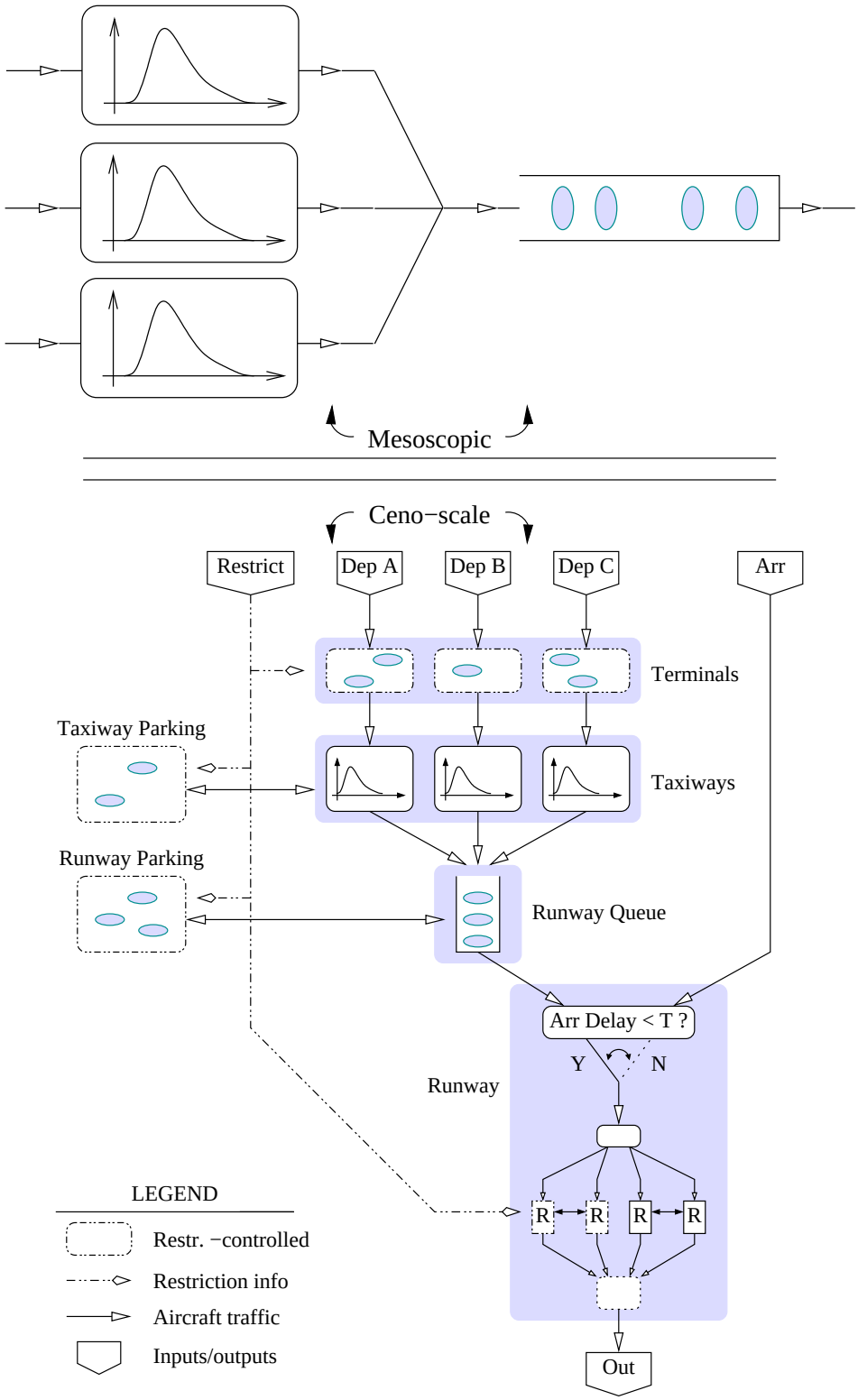

Figure 1-1: Comparison of the previous state-of-the art *mesoscopic* models of [Puj99], and the ceno-scale models developed in this thesis.

Chapter 2

Stochasticity in the Airline Turn Process

2.1 Introduction

Accurate and timely pushback forecasts can potentially improve DST performance, including reductions in both the variability and average duration of taxi-out delays. In contrast to previous work on worst- or average-case characterizations of pushback stochasticity, this chapter presents a novel characterization derived under *best-case minimum-uncertainty* conditions. This minimum stochasticity constrains both the robust design and achievable performance of DSTs. Until recently, such best-case analyses have been possible only in a notional or theoretical sense — no real-world airline operations data were available to support quantitative results. Over the last 6 years, Lufthansa Airlines has been developing their ALLLEGRO project, which provides the most detailed automated monitoring of airline turn operations of any major passenger carrier. Several novel airline turn process models and forecast techniques have been developed and, through collaboration with Lufthansa Airlines' ALLEGRO project, statistically tested against a large and detailed dataset of 17,344 turn operations covering a 3 month period at Frankfurt International Airport.

Uncertainty in pushback forecasts is partly due to the complexity of the associated airline decision processes (which may be amenable to improved models), and partly due to the natural inherent stochasticity of real-world operations (thus placing inherent lower bounds on the quality of any such forecasts). Both probabilistic and structural effects must be considered to derive a lower bound on the minimum uncertainty. Probabilistic effects are mitigated via a two-stage process. First, all analyses are based on a sample of real-world airline turn operations with minimum uncertainty. This sample is derived by filtering a much larger set of operations data to remove any turns which were affected by factors exogenous to the scheduled airline turn processes. Second, a Bayesian *age-*

based forecasting technique is developed which yields optimal time-to-go estimates based on the observed probability distribution of these minimum-uncertainty turns. The structural complexity of the overall turn process is mitigated by focusing on processes which frequently contribute to the critical path. *Status-based* forecast techniques are developed which explicitly take advantage of this critical-path information. Finally, a *combined* forecast is developed which combines both status- and age-based techniques.

In actual practice, the conditions necessary for best-case forecasts are not achieved, and pushback predictions with higher mean-square error over shorter forecast horizons must be used. This imposes a significant design requirement on DSTs: any viable DST must be robust to *at least* the minimum mean-square error of approximately ±5min derived below, and can only rely on pushback forecasts over a *shorter* horizon. It also imposes a limitation on achievable DST performance, which must be considered before investing in DST development. Also, in current operations, ATC has very little real-time information about the airline turn processes [IDA+98]. These analyses clearly show the predictive value of different types of process information. This enables further research on the cost-benefit tradeoffs of investing in the infrastructure needed to make such information available for planning purposes.

2.2 Literature review

2.2.1 Interaction of DSTs with pushback predictions

The sensitivity of DST performance with respect to the uncertainty in pushback time predictions has been analyzed in [CCD+01] and [CCF+02]. The authors of these two studies examined how the performance of the Departure Enhanced Planning And Runway/Taxiway-Assignment System (DEPARTS) varied according to both the *magnitude* of pushback prediction error, and the *horizon* over which predictions are available. An empirical probability distribution was observed for the error between scheduled and actual ready-for-pushback. This empirical distribution was then scaled and shifted to obtain hypothetical distributions (which represented hypothetical improvements in pushback time predictions) with similar *shape* but different *mean absolute deviation*. These hypothetical distributions were then used to simulate pushback time errors and generate randomized inputs for the DEPARTS scheduling algorithm. This enabled a measurement of the performance of the DEPARTS scheduler as a function of hypothesized improvements in pushback predictions. The results showed that perfect prediction of ready-for-pushback at a 10min horizon could reduce the average taxi-out time of each flight by approximately 20sec. However, roughly half of this benefit occurred only when errors were reduced to zero — when errors were simulated at 20% of the empirically-observed level, the average benefit was only 10sec. In an additional set of experiments, the DEPARTS scheduler was given perfect predictions of ready-for-pushback, and the lead-time at

which predictions were available was varied. In these experiments, the decrease in average taxi-out time varied linearly with the lead-time. These results show how the benefits of a DST can be quickly eroded when pushback predictions are noisy and/or not available in a timely fashion. To my best knowledge there is no complementary research which characterizes best-case pushback predictions and the corresponding upper bound on DST performance.

2.2.2 Pushback forecasts

There have been several studies on the stochasticity of pushback times. Statistical analyses have shown that the published departure time is a poor predictor of the actual pushback time. Both structural complexity and inherent natural randomness contribute to this lack of predictability [FHO$^+$97]. Several processes and resources must be synchronized to turn an inbound flight into an outbound flight, and the overall turn process is affected by numerous random factors which are difficult to observe in real-world operations. The general conclusion in the literature is that pushback demand is one of the major sources of uncertainty in airport traffic.

The simplest ("raw") forecast is a flight's scheduled departure time. For large commercial carriers the scheduled pushback is available up to several weeks in advance, while even small commuter and charter aircraft must file a flight plan (including a scheduled departure time) at least 30min prior to pushback. Note that some aircraft do not have a scheduled pushback time, e.g. general aviation aircraft operating in good weather are not required to file a flight plan and may contact ATC for pushback clearance at any time. At most large airports, however, only a small fraction of the total traffic does not have a scheduled pushback time available in advance. Currently air traffic controllers use mental models of upcoming departure demand which are learned through day-to-day experience; these are essentially aggregated raw forecasts. In addition, the US Collaborative Decision Making (CDM) program computes Ground Delay Programs up to 6hrs in advance which are based on raw forecasts, albeit modified by up-to-date airline cancellation plans.

More advanced *model-based* forecasts attempt to reduce structural stochasticity by mimicking airline decisions made during the turn process. Airline decision processes on a tactical timescale have been extensively analyzed. The complexity of a single turn was modeled as a Petri net in [IDA$^+$98], where it was noted that most of the relevant state information is not currently observable outside the local airline station. Recovery models for airline operations on a tactical timescale are currently an important research area, and the interested reader is referred to [YAS$^+$03, RSG$^+$02] for recent literature reviews. However relatively little of this work is devoted to predicting pushback times; the few notable exceptions are given below.

As shown in Table 2.1, the current model-based forecasts have not significantly improved on raw forecasts. This is partly due to the complexity of the airline decision processes, and partly due to a lack of observability into those processes.

Author [citation]	Standard dev.	Mean absolute dev.
Idris [Idr00, p. 144–7]	17.4min	14min
DEPARTS [CCD$^+$01, CCF$^+$02]	12.6min	8.5min
Vanderson [Van00] (raw)	n/a	20–90min
Vanderson [Van00] (model-based)	n/a	$>$ 15min
Andersson [And00] (model-based)	9–19min	n/a

Table 2.1: Summary of previous literature results on pushback predictability.

Raw forecasts

Idris obtained 5hrs of flight-strip data for departure operations at Boston Logan International Airport (BOS) and analyzed how well ATC could predict pushback requests [Idr00, p. 144–7]. He defined pushback time error as the difference between each aircraft's scheduled pushback time and the time ATC was contacted to obtain clearance to leave the gate. Note that this definition is unusual; in most sources of operations data, the request for clearance is not recorded and the actual time of pushback must be used instead. By focusing on the clearance requests, Idris was thus able to reduce stochasticity contributions due to ATC holding aircraft at the gate. In addition, his dataset was unusual because there are flight-strips for *all* aircraft and thus no operations were missing. The observed errors in his dataset had a standard deviation of 17.4min and a mean absolute deviation of 14min.

Two studies on the sensitivity of DST benefits to the magnitude of pushback prediction error were based on raw forecasts for ATL departures [CCD$^+$01, CCF$^+$02]. A unique function of the NASA Surface Movement Advisor system at ATL is automatic recording of flight-plan and flight-strip data. This automated collection enabled the use of a very large dataset which covered August 2000 and (after being filtered to remove pushback forecast errors over one hour) included 18,586 pushback events. The observed errors had mean 6.8min, standard deviation 12.6min, and mean absolute deviation (after removing the mean error) of 8.5min. It was noted that pushback times at ATL have relatively low uncertainty compared to many other US hub airports.

Vanderson analyzed pushback forecasts at horizons of zero to six hours from the US CDM program [Van00]. These forecasts frequently showed errors with mean absolute deviation in excess of 20min even under the best weather and traffic conditions; the mean absolute deviation could exceed 1.5hrs under poor conditions.

Model-based forecasts

Based on an heuristic model of the internal airline decision-making processes of cancellations, swaps, intentional pushback delays and hastening, Vanderson was able to reduce the US CDM prediction error by up to 30% in some cases. In related work, the Aircraft Sequencing Model (ASM) constructed by Andersson was an optimization-based model which minimized passenger delay by modifying

aircraft pushback times [And00]. The modified pushback times were constrained by the arrival sequence and timing, the departure schedule, and gate and crew resources. Andersson selected a hub airport and used actual operations data over the 16:00–19:15 time-period on twelve successive days to show that the ASM yielded pushback predictions with standard deviation of 9 to 19min. An important caveat is that the ASM assumed deterministic inputs, including a perfect forecast of landing times over each 3.25hr period; landing times often show variability on the order of $\pm$10min as enroute aircraft approach the terminal area and encounter congestion and/or holding stacks [GG97].

Previous research has shown that the airline decision processes involved in maintaining a schedule are substantially more complex than can be observed externally [PF00]. There have been several successful attempts to build models which accurately capture the *complexity* of these decision processes, for example in an airline operations center [PF00] or at the station level [And00, Van00]. However, characterizing the contribution of *random* factors to pushback stochasticity has remained until recently an intractable problem. Statistically significant quantities of accurate operations data have simply not been available.

2.3 The value of a well-maintained schedule

Various stakeholders (airlines, the traveling public, ATC, airports, etc.) incur different costs when a pushback prediction is incorrect. It is often difficult to quantify these costs and track the cumulative effect on different stakeholders. In this section, the consistent use of a *square-law* penalty for pushback prediction errors is argued. This functional form is approximately correct for most stakeholders over a tactical timescale. Furthermore, the optimal forecast corresponding to this functional form has the excellent mathematical property of invariance with respect to the particular cost coefficients of each stakeholder. The results in this thesis are derived under this square-law assumption; to support future extensions and generalizations, a framework is also laid out for forecasting based on alternative cost functions.

From the viewpoint of ATC, pushback predictions are simply an input into the planning process. The standard metrics for forecast uncertainty are the standard deviation and mean absolute deviation of the forecast error, where the error is defined as the predicted time of occurrence minus the actual time of occurrence. When only a single forecast is generated, the error is a constant (scalar) over the entire life of the forecast, and such metrics (resulting from square-law or linear penalties for delay) are standard for scalar random variables. It is then reasonable to assume a square-law penalty from the viewpoint of ATC. Note that when forecasts can be updated over time, the error becomes a stochastic process and the penalty for a fixed magnitude of error may change over time as the forecast impacts the planning process at different stages. For the work in this thesis, it is mathematically convenient to assume that the penalty function does not change over time, though the necessary

extensions are worked out in Appendix A.

From the viewpoint of airlines, similar arguments indicate that a square-law penalty is reasonable. However, there is also additional direct evidence for this conclusion. Airlines incur a financial cost due to delayed flights, including both wasted Direct Operating Costs and lost revenue. The losses due to Direct Operating Costs are easy to quantify, and increase proportionally with delay due to additional crew wages and (depending on the situation) excess fuel-burn. There is interesting recent work on characterizing the lost revenue. In high-demand markets where increased delays on one itinerary may drive passengers to substitute alternate itineraries or other forms of transportation, Januszewski [Jan03] derives a marginal price-change of approximately $1USD per fare per minute of additional expected delay. Note that the marginal costs for both Direct Operating Costs and lost revenue are constant for a given flight.

However, delayed flights also induce cascading network disruptions as later flights which share resources (crew, airframe, passengers, etc.) with the first disrupted flight-leg are affected. This propagation effect can be modeled as a *delay multiplier* by which the initial flight delay is scaled to account for the total delay incurred in the entire cascade [BHBR98]. While the exact values of the delay multipliers found in that study cannot be directly translated into other situations (e.g. different schedules, operating conditions, disruption recovery procedures, etc.), two important effects were observed: the additional delay due to the cascade effect (corresponding to the portion of the delay multiplier in excess of unity) tends to increase linearly with both the initial delay and the time-to-go until the end of the operational day. The dependence on time-to-go suggests that delay costs should be linearly discounted as the operational day progresses. More importantly, the dependence on the initial delay implies that even if delays have a constant marginal cost for a single flight, the total cost to the airline increases no slower than the *square* of the initial delay.

The costs due to actual delay have a strong connection to the penalty associated with incorrect forecasts. When delays are not correctly anticipated, there is no chance to mitigate them and the full cost must be borne. In this case a forecast error directly translates into a square-law financial cost. The situation is more complex when delays are forecast but do not materialize. For example, suppose a flight is predicted (and therefore planned) to depart 15min late but is actually ready to depart at the scheduled time. There is an opportunity cost if the flight is then delayed to match the new plan, and there may be a disruption cost (e.g. the push-truck may not be available) if the flight attempts to depart according to the old schedule. In the case when a flight is predicted to depart on-time but in fact is ready early, it becomes even more difficult to evaluate the opportunity/disruption tradeoff. In this thesis the working approximation is that the obvious square-law penalty for unforeseen delays can be reasonably extended to a square-law penalty for forecasting delays which do not materialize.

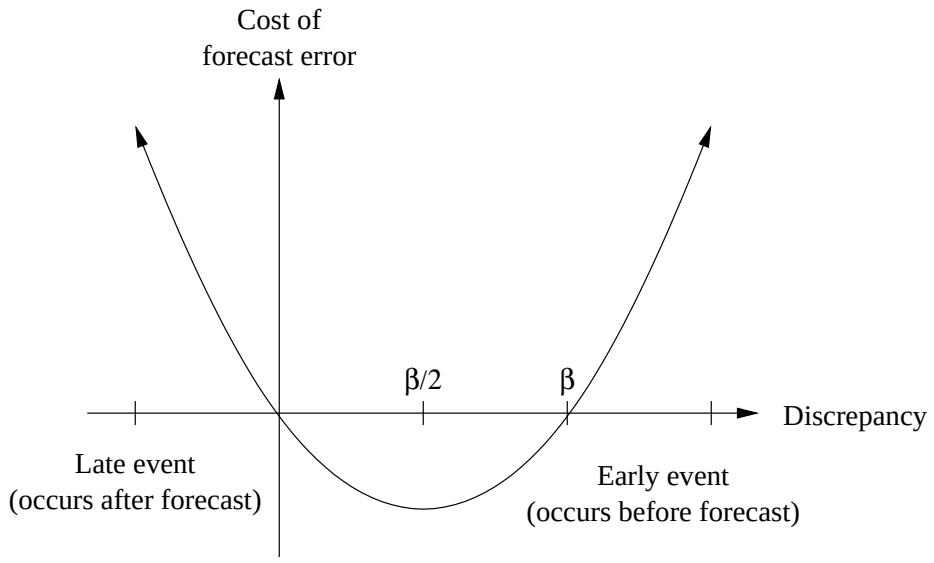

Figure 2-1: Illustration of the cost-function of forecast errors. A square-law penalty has good mathematical properties and reasonable empirical support.

Formally speaking, this thesis assumes a forecast error cost of the form

$$\text{cost} = \alpha \cdot T_{\text{minus}} \cdot \text{deviation} \cdot (\text{deviation} - \beta). \tag{2.1}$$

where α and β are nonnegative constants and deviation is defined as the scheduled minus the actual time of an event. As shown in Figure 2-1, for negative deviations or positive deviations larger than β, the cost is positive as expected. However for small positive deviations, when the event occurs before scheduled but the deviation is smaller than β, the cost is negative indicating a desirable outcome.

This cost structure can be used to produce "optimal" forecasts of pushback times, by minimizing the expected cost of deviations between the actual and forecast time-to-go over the duration of the forecast. This topic is treated in greater detail in section 2.5.2. In addition there is an important conceptual message: cost does not scale in proportion to average delay or haste, and cost increases as the *dispersion* of deviations increases. In fact this notion can be made precise. Suppose two processes with respective deviations $\mathbf{T}_1$ and $\mathbf{T}_2$ have the same average deviation $\mathrm{E}\big[\mathbf{T}_1\big] = \mathrm{E}\big[\mathbf{T}_2\big]$, but $T_1 \leq_{\text{disp}} T_2$ using the definition of stochastic dispersion order $\leq_{\text{disp}}$ from [MS02, p. 40]. Given this order on the deviations, it immediately follows from [MS02, Corr. 1.5.4(a), Thm. 1.7.6(b)] and the convexity of the cost structure that the mean, mean square, mean absolute deviation and variance of the costs must be similarly ordered. In concrete terms, twice the delay can cost four times as much, and a flight which is delayed by 5 ± 1min will always be more expensive than a flight which is delayed by 5min. It is natural for air carriers, ATC and passengers to include buffers and deliberate slack-time in their schedules in order to stay on-time and robust in the face of uncertain

operating condtions. However, there are significant reasons to keep the length of these buffers and all controllable sources of uncertainty under careful scrutiny to avoid rapidly ballooning costs.

2.4 Observing the turn process

While detailed real-time and historical data are available for the different stages of flight, until recently very few data have been automatically collected during ground operations. This lack of observability has blunted efforts to understand and improve these processes. However, as part of the Operational Excellence on-time initiative by Lufthansa Airlines, the ALLEGRO project was initiated: a unique real-time control and analysis system for airline ground operations [The02]. Special focus was laid on the correct definition of target times for each process of the turn. During ground operations on a single aircraft, ALLEGRO's monitoring function collects timestamps for up to 80 distinct types of events and compares them against the respective target times. The use of ALLEGRO increases the transparency of ground events, supports the hub control center in coordinating ground events, simplifies analyses of weaknesses in each process, enables continuous validation of the baseline schedule for ground events, and provides a foundation for internal and external performance agreements. This wealth of data stands in sharp contrast to the minimalist observations of landing, gate-arrival, gate-departure and takeoff times which are currently available for most passenger airline operations.

The improvement in closed-loop control of airline ground operations is particularly important for these analyses. The feedback loop between the processes being monitored and the resulting decision-making and planning has been automated. This improved feedback reduces the stochasticity of managing a large and complex system, resulting in sharper measurements of the minimum inherent uncertainty in the airline turn process.

These analyses are based on a dataset of 17,344 turnaround operations obtained from the Lufthansa OBELISK data-warehouse, of which ALLEGRO is a subsystem. These turns were continental operations on Lufthansa which transited Frankfurt International Airport between 1 Feb 2003 and 30 Apr 2003. For each transit through Frankfurt, several observations were captured in OBELISK:

- Inbound and outbound flight parameters.
  These included aircraft type; flight numbers; connecting airports; scheduled, estimated and actual times for gate-arrival and pushback; scheduled and actual en-route and block times; and standardized delay codes.

- Ground operations data (from the ALLEGRO system).
  Both scheduled and actual ground time intervals were specified (see Figure 2-3, and further

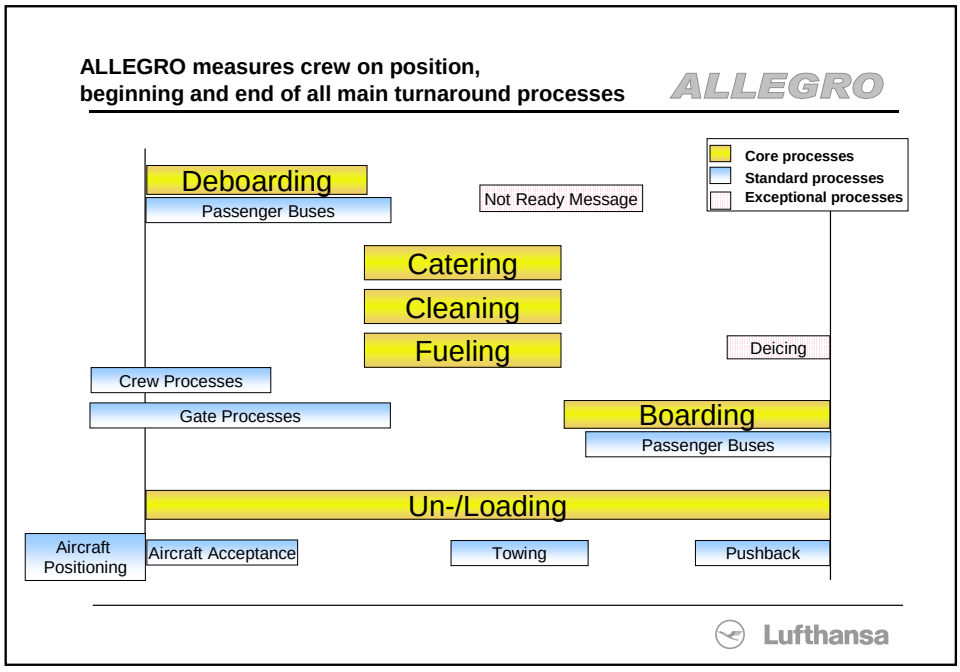

Figure 2-2: Gantt chart of the airline turn process [The02] showing the sub-processes which are automatically monitored by the ALLEGRO system.

subdivided into scheduled and actual start/end epochs for deplaning, cleaning, catering, fueling, and boarding. The type of passenger-loading equipment (either bus or jetway) was also specified.

The timing data for inbound and outbound flights, and scheduled intervals for ground events, were reported with one-minute precision. The timing data for actual ground events were reported with one-minute or one-second precision depending on equippage. The reported numerical precision of these data has been validated [The02]. The delay codes were standard codes from the IATA Airport Handling Manual [Int]. For reference, the general categories of delay codes are reproduced in Table 2.2.

The ALLEGRO system enables novel analyses of how complexity and randomness contribute to pushback stochasticity. Airline decision processes related to cancellations, swaps of crews or airframces, and intentional delay or hastening can be accounted for by examining those turns which were actually operated and the corresponding ALLEGRO target-times. A wide variety of exogenous sources of uncertainty can be filtered out using the delay codes attached to each turn. Thus it is possible to focus attention on just the inherent uncertainty of the airline turn processes.

This minimum uncertainty is an important limiting factor because it implies a corresponding upper bound on the achievable performance of any forecast. To derive this upper bound, the dataset was filtered for *simple turns*, so-called because they occurred under the following set of conditions:

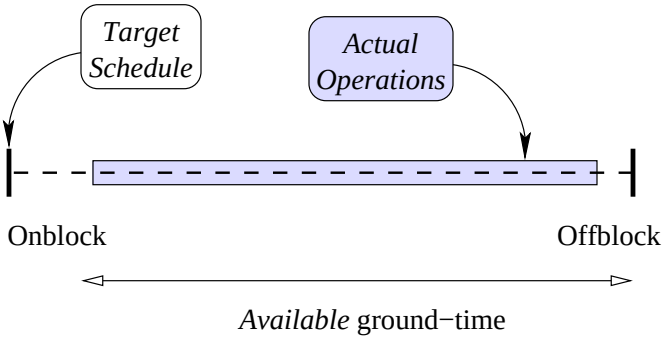

Figure 2-3: Basic terminology for discussing the airline turn process, including scheduled, actual and *available* ground-time.

Classification	ID#'s	Additional description
Airline Internal Codes, Others	1–9	Scheduling, internal records
Passenger and Baggage	11–19	Check-in, boarding, baggage handling
Cargo and Mail, Mail Only	21–29	
Aircraft and Ramp Handling	31–39	Loading, cater/clean/fuel
Technical and Aircraft Equipment	41–49	Unusual maintenance
Damage to Aircraft, Automated Equipment Failure	51–59	Causes of non-scheduled maintenance
Flight Operation and Crewing	61–69	
Weather	71–79	Location affected (origin, destination, etc.)
ATC Flow Management Restrictions, Airport and Government Authorities	81–89	
Reactionary	91–96	Change in rotation(s) to mitigate delays
Miscellaneous	97–99	Industrial actions, provisional codes

Table 2.2: Standard IATA delay codes.

	Board: Bus	Board: Jetway
Deboard: Bus	3820	842
Deboard: Jetway	157	6324

Table 2.3: Sample sizes for predictability analysis.

- *Similar target-times for the turn processes.* No towing occurred; the aircraft departed from the gate where it arrived. Operations occurred on similar aircraft types which require the same scheduled time for each process in the turn. The same type of passenger-loading equipment was used. Changes in target-times due to different available ground-time (scheduled offblock minus actual onblock) were accounted for.

- *Only relevant delay-codes.* Turns with delay codes 40–99 were excluded, *with the exceptions* of code #87 (airport facilities — surface congestion) and code #93 (reactionary — late inbound airframe). Thus turns with delays due to late inbound crews or loads; local or downstream weather; missed takeoff slots; and control imposed by ATC or the airport authority were excluded. Also delays due to abnormal aircraft maintenance requirements were excluded, because those delays did not directly impact the normal turn processes and hence were not directly observable in the ALLEGRO data. The remaining delay codes are specific to aircraft servicing processes.

During actual operations, some causes of delay cannot be predicted, e.g. during the turn, a routine mechanical inspection may discover an unusual problem which requires unscheduled maintenance. In contrast, when working with historical data, these unpredictable sources of uncertainty can be filtered out by censoring any affected turns from the dataset. Hence these analyses yield a *conservative* estimate of the minimum uncertainty; real-world performance is guaranteed to be noisier.

Filtering the original dataset against these conditions yielded a sample of 11,143 turn operations (64.2% of the total sample). Table 2.3 shows how these turns were further subdivided by the type of passenger-loading equipment used on arrival and departure. Note that it is sometimes necessary due to customs or security measures to use different types of passenger-loading equipment on the arrival and departure of the same aircraft; this does not indicate the aircraft was towed or otherwise changed gates. For convenience, the simple turns which deboarded with buses and then boarded with buses will be referred to just as "bus-bus" turns, and likewise for the "jetway-jetway" turns.

A caveat should be noted for the jetway-jetway turns. The ALLEGRO system measures the end of boarding for jetway-jetway turns as the moment when the cabin doors are closed, an event which may be related only indirectly to the actual end of boarding. For example, a flight delayed by ATC might need to have the crew swapped; or if a maintenance delay occured, the captain and maintenance personnel might re-open the cabin doors to perform inspections. At the time these data were collected, the measurement of the end of boarding for jetway turns was being redesigned

in ALLEGRO to ensure a more direct measurement of the actual process. However, the bus-bus turns do not suffer from this potential shortcoming. Therefore, the analyses presented below focus on the 3820 bus-bus simple turns, which yield the cleanest results. The results for the other simple turns have been computed and appear similar, but are not reported here in order to present only the strongest possible case.

2.5 Forecasting

Given the same set of simple turns, a variety of statistical techniques can be used to forecast pushback times. The simplest forecasts are updated only once: once the available ground-time is known, the turn duration is forecast using descriptive statistics such as the mean, median or some percentile of observed turn durations. Any system with monitoring capabilities equivalent to ALLEGRO enables more sophisticated techniques. *Age-based* forecasts use the elapsed duration of a turn to compute a Bayesian estimate of the remaining time-to-go. A conceptually orthogonal *status-based* approach depends on the updated status (not-started/in-progress/completed) of the different processes comprising a turn. The age-based and status-based forecasts rely on different types of process observations and hence are initially considered separately. Finally, a *combined* forecast relying on both age and status information can be derived and calibrated using the available operations data. In the remainder of this section, these different forecasts are formalized and their actual performance against the ALLEGRO dataset is compared.

2.5.1 Forecasts using simple descriptive statistics

Predicting the actual ground-time based on the scheduled ground-time has several drawbacks. In Figure 2-4 the bus-bus simple turns have been binned according to scheduled ground-time. The sample of turns in each bin is then described by a boxplot of the corresponding sample of actual ground-times[1]. From the plot, the minimum ground-time[2] is reasonably apparent. The median of actual ground-time tends to track the scheduled ground-time, indicating that onblock and pushback typically occur on schedule. However the variability of actual ground-time as a function of scheduled ground-time is very high. This variability leads to poor predictions of actual ground-time given scheduled ground-time.

Much of this variability can be compensated for by replacing scheduled ground-time with available ground-time. In particular the effect of early or late inbounds can be accounted for. The resulting

[1]Boxplots are a standard statistical method for robust visualization of scalar data. The "box" in a boxplot covers the interquartile range, with a line through the middle of the box to denote the median. This gives a robust estimate of the central tendency and dispersion of the data. The box has "whiskers" extending to the farthest datapoints within 1.5 times the interquartile range of the median. Observations outside the whiskers are marked individually; these are typically treated as possible outliers.

[2]When the data were measured, the planned minimum ground-time for Lufthansa operations at Frankfurt was 50min.

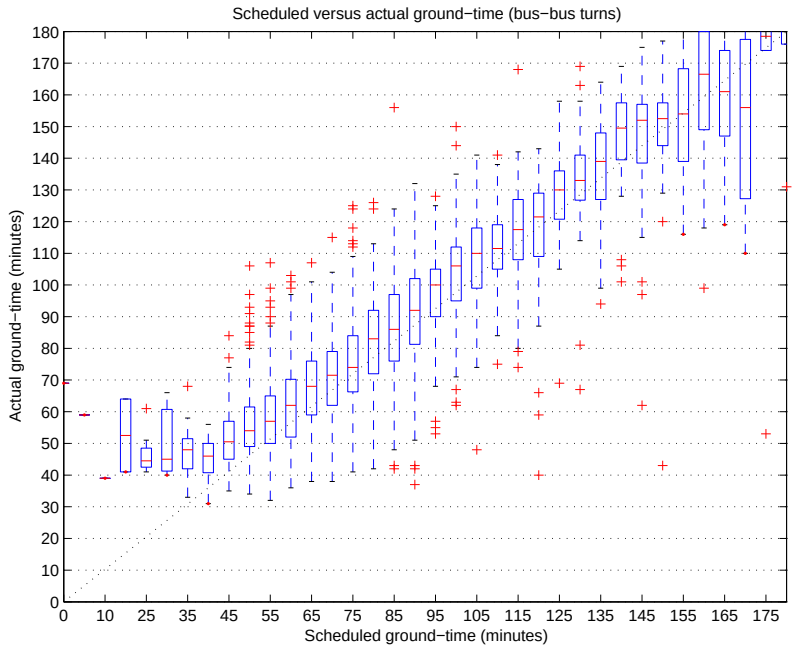

Figure 2-4: Predicting actual ground-time as a function of scheduled ground-time. While the mean error is negligible, the variability is excessive.

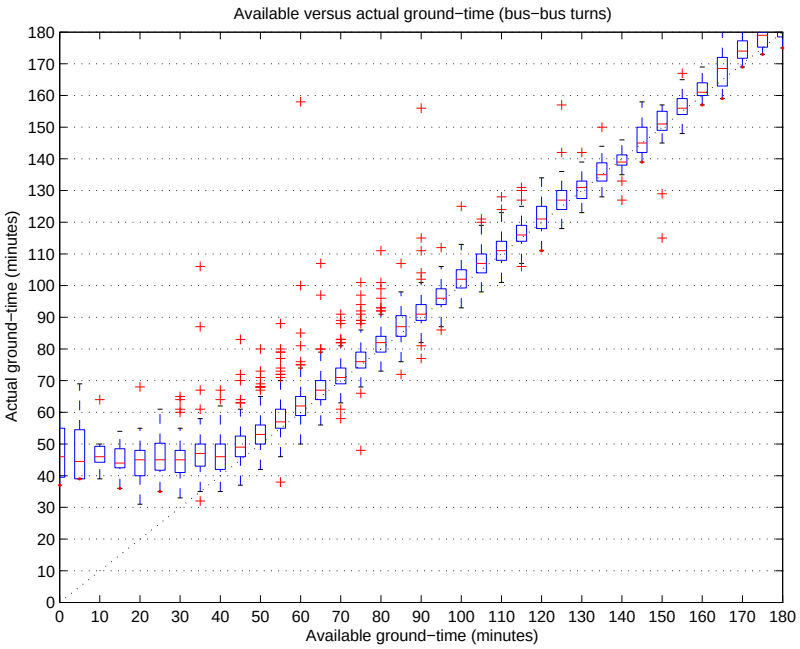

Figure 2-5: Predicting actual ground-time as a function of available ground-time. The available ground-time is a much more accurate predictor.

substantial decrease in variability is shown in Figure 2-5. It is possible to predict actual ground-time with much reduced uncertainty given the available ground-time; even this simple information may thus be very useful for developing DSTs. However it is also important to note that, due to the possible need for cancellations or swaps, the available ground-time cannot always be determined based on the arrival time of an inbound flight and the *scheduled* assignment of this inbound aircraft to an outbound flight. Because the dataset only describes turns which actually occurred, this problem does not affect these analyses. It is an interesting open question to determine how soon an airline station could accurately determine the available ground-time, i.e. how soon an airline station can accurately claim that a particular outbound flight will not be cancelled or swapped.

2.5.2 Bayesian forecasts using elapsed ground-time

The previous section demonstrated that, given the available ground-time, it is reasonable to forecast pushback using the average actual ground-time. This initial forecast can be improved by using the elapsed ground-time (from actual gate-arrival to the current time) to compute updates, especially for turns which are running unusually late and have exceeded the average actual ground-time.

For a turn with some given available ground-time, let the random variable $\mathbf{X}$ denote the actual ground-time. Characterize $\mathbf{X}$ by its complementary density function $G(t) \doteq \Pr(\mathbf{X} > t)$; for convenience the parametrization by the available ground-time is elided. At elapsed time t since onblock, the "perfect" forecast of time-to-go is simply $\mathbf{X} - t$. To approximate this perfect forecast, consider deterministic functions $\tilde{f}(t)$. The instantaneous forecast error $e(t)$ is then defined as the predicted time-to-go minus the actual time-to-go $\tilde{f}(t) - (\mathbf{X} - t)$ as illustrated in Figure 2-6.

The optimal $f(t)$ can then be constructed so that the expected total cost of these errors according to the approximate quadratic cost-function of Equation (2.1) is minimized:

$$f \doteq \arg\min_{\{\tilde{f}\}} \mathrm{E}_{\mathbf{X}} \left[\int_0^{\mathbf{X}} \alpha(\tilde{f}(t) - (\mathbf{X} - t)) \cdot (\tilde{f}(t) - (\mathbf{X} - t) - \beta)\, dt \right]$$

In Appendix A this is solved to obtain

$$f(t) = \mathrm{E}\big[\mathbf{X} - t \mid \mathbf{X} > t\big] + \frac{\beta}{2} \tag{2.2}$$

with instantaneous standard deviation of the forecast error

$$\mathrm{E}\big[\mathrm{cost}(e^2(t)) | \mathbf{X} > t\big] = \mathrm{Var}\big(\mathbf{X} - t \mid X > t\big) - \beta^2/4 \tag{2.3}$$

Note that while f is defined by an integral over a random-length interval, the final result depends on a pointwise-optimal Bayesian estimate of the *remaining life* $\mathbf{L}_t \doteq [\mathbf{X} - t \,|\, \mathbf{X} > t]$. This form is doubly

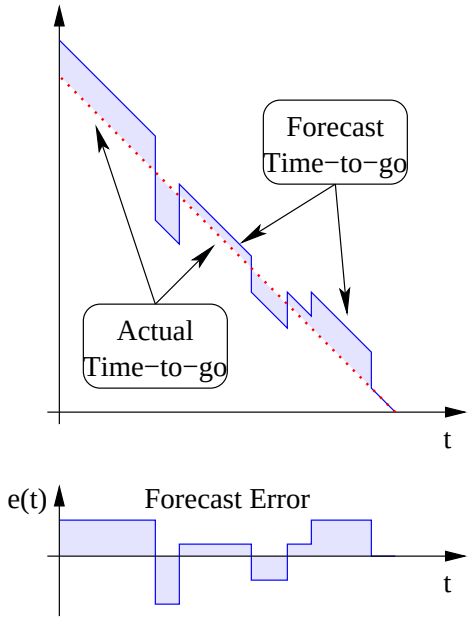

Figure 2-6: Illustrated definition of the instantaneous forecast error $e(t)$.

appealing because it both naturally incorporates observations of the elapsed turn-time $\{\mathbf{X} > t\}$ and minimizes a reasonable cost-structure on the forecast error.

The following theorem is useful for characterizing the remaining life:

Theorem 1 (Remaining Life) *The moments of* $\mathbf{L}_t$ *are given by*

$$\mathrm{E}\left[\mathbf{L}_t^n\right] = \int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1}\, dv. \tag{2.4}$$

A derivation of this result is given in Appendix A. To calculate the optimal forecast (2.2 and its error variance (2.3), it then remains to approximate G given a set of samples of $\mathbf{X}$. Two approximation techniques (nonparametric and parametric) are described below.

Nonparametric estimation

One approach is to estimate G nonparametrically. Given N iid samples $(x_1, \ldots, x_N)$, the standard histogram estimate of G is $\hat{G}(t) \doteq \#\{x_i > t\}/N$. This estimate can be substituted for G to approximate the moments of $\mathbf{L}_t$. In particular the first two moments can be used to approximate

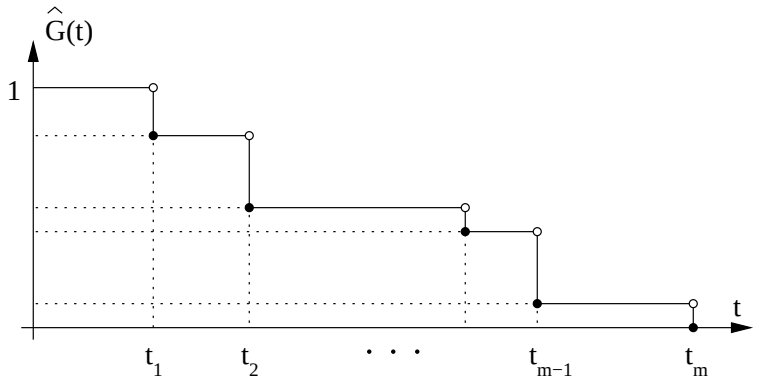

Figure 2-7: Illustration of the nonparametric estimate $\hat{G}(t)$ of the complementary probability distribution.

the Bayesian estimate $\mathrm{E}\big[\mathbf{L}_t\big]$ and its associated mean-square error $\mathrm{Var}\big(\mathbf{L}_t\big)$. From Equation (2.4),

$$\begin{aligned}
\mathrm{E}\big[\mathbf{L}_t\big] &= \int_t^\infty \frac{G(v)}{G(t)}\,dv \\
&\approx \frac{1}{\#\{x_i > t\}} \int_t^\infty \#\{x_i > v\}\,dv \\
\mathrm{E}\big[\mathbf{L}_t^2\big] &= \int_t^\infty \frac{G(v)}{G(t)} 2(v-t)\,dv \\
&\approx \frac{2}{\#\{x_i > t\}} \int_t^\infty \#\{x_i > v\} v\,dv - \frac{2t}{\#\{x_i > t\}} \int_t^\infty \#\{x_i > v\}\,dv
\end{aligned}$$

The integrals can be further simplified because $\hat{G}$ has zero derivative except at a finite sequence of times $\min\{x_i\} = t_1 < \cdots < t_m = \max\{x_i\}$ as illustrated in Figure 2-7. Integrating over the intervals $[t_k, t_{k+1}]$ yields the identity

$$\int_t^\infty \#\{x_i > v\} v^n\,dv = \sum_{k=1}^{m} \#\{x_i \geq t_k\} \max\left\{0, \frac{t_k^{n+1} - \max\{t_{k-1}, t\}^{n+1}}{n+1}\right\} \tag{2.5}$$

where $t_0 \doteq 0$. Note that while the right-hand side of Equation (2.5) appears more complicated than the original left-hand side, it is easier to evaluate from the raw data.

Parametric estimation

An alternative approach is to fit some known parametric distribution to the data and directly compute the integrals in Theorem 1. This approach may be preferred when additional information about the turn-times is available, e.g. when the pushback forecasts are being used as part of a larger queueing model and Erlang distributions are mathematically convenient.

For parametric distributions with smooth complementary distribution functions, the *hazard rate* $r(t) \doteq -\frac{d}{dt} \ln G(t)$ is well-behaved. It is then possible to derive a simple closed-form recursion on

the *dynamics* of $\mathrm{E}\big[\mathbf{L}_t^n\big]$ as a function of t:

Theorem 2 (Hazard-rate Remaining Life Recursion)

For a nonnegative random variable $\mathbf{X}$ *with bounded hazard rate* $r(t)$*, the moments of the remaining life* $\mathbf{L}_t \doteq [\mathbf{X} - t \,|\, \mathbf{X} > t]$ *follow the recursion*

$$\frac{\partial}{\partial t}\,\mathrm{E}\big[\mathbf{L}_t^n(t)\big] = -n\,\mathrm{E}\big[\mathbf{L}_t^{n-1}\big] + r(t)\,\mathrm{E}\big[\mathbf{L}_t^n\big]. \tag{2.6}$$

A derivation is given in Appendix A. The dynamics of the mean $\mu_L \doteq \mathrm{E}\big[\mathbf{L}_t\big]$ and variance $\lambda_L \doteq \mathrm{Var}\big(\mathbf{L}_t\big) = \mathrm{E}\big[\mathbf{L}_t^2\big] - \mathrm{E}\big[\mathbf{L}_t\big]^2$ are of particular interest:

$$\begin{aligned}
\frac{d}{dt}\mu_L(t) &= -1 + r(t)\,\mathrm{E}\big[\mathbf{L}_t\big] = -1 + r(t)\mu_L(t) \\
\frac{d}{dt}\lambda_L(t) &= \big\{-2\,\mathrm{E}\big[\mathbf{L}_t\big] + r(t)\,\mathrm{E}\big[\mathbf{L}_t^2\big]\big\} - 2\,\mathrm{E}\big[\mathbf{L}_t\big]\,\big\{-1 + r(t)\,\mathrm{E}\big[\mathbf{L}_t\big]\big\} \\
&= \big\{-2\mu_L(t) + r(t)[\lambda_L(t) + \mu_L^2(t)]\big\} - 2\mu_L(t)\,\{-1 + r(t)\mu_L(t)\} \\
&= r(t)[\lambda_L(t) - \mu_L^2(t)]
\end{aligned}$$

This yields an ordinary differential equation for computing the mean and variance:

$$\begin{aligned}
y(t) \doteq \begin{pmatrix} \mu_L(t) \\ \lambda_L(t) \end{pmatrix}, \quad y(0) &= \begin{pmatrix} \mathrm{E}[\mathbf{X}] \\ \mathrm{Var}(\mathbf{X}) \end{pmatrix} \\
\dot{y} = F(t, y) &= \begin{pmatrix} -1 + r(t)\mu_L \\ r(t)(\lambda_L - \mu_L^2) \end{pmatrix} \\
\frac{\partial F}{\partial y} &= \begin{pmatrix} r(t) & 0 \\ -2\mu_L r(t) & r(t) \end{pmatrix}
\end{aligned}$$

This ODE can be easily solved numerically.

Performance on the simple turns

Both the nonparametric and parametric approaches were applied against the sample of simple bus-bus turns with 55min of available ground-time. These turns were selected because their available ground-time is close to the minimum planned ground-time: they are neither guaranteed to be late nor guaranteed to have an excess of slack-time. These turns also show minimum uncertainty in the actual ground-time.

The parametric (Gaussian) and nonparametric cumulative distribution functions are shown in Figure 2-8. These cumulative distributions are then used to derive a pair of forecasts for $\beta = 0$ as shown in Figure 2-9. For turn durations where a large number of datapoints are available there is little difference between the two forecast functions. However, for exceedingly long turns which are

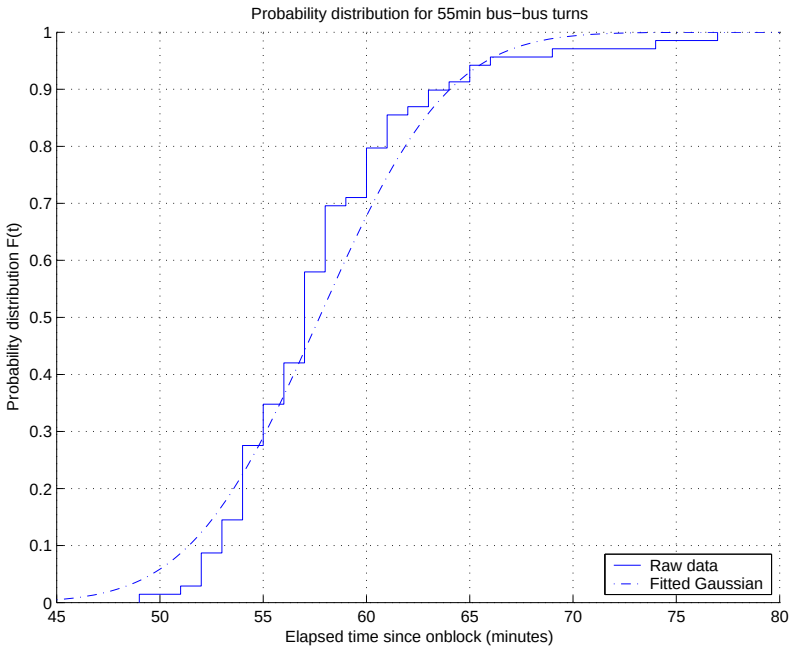

Figure 2-8: Probability distribution of 55min bus-bus turns.

correspondingly rare, the nonparametric forecast must depend on only a handful of datapoints and some deviations are apparent; this is another situation in which a parametric forecast may be useful.

For a specific sample of N turns, it is informative to consider the average instantaneous accuracy $\sigma(t)$ as a function of elapsed time:

$$\bar{e}(t) \doteq \frac{1}{\#\{x_i \geq t\}} \sum_{\{i|x_i \geq t\}} e_i(t)$$

$$\sigma(t) \doteq \sqrt{\frac{1}{\#\{x_i \geq t\} - 1} \sum_{\{i|x_i \geq t\}} \left(e_i(t) - \bar{e}(t)\right)^2}$$

The corresponding forecast accuracies are shown in Figure 2-10. Note that the apparent accuracy of the parametric forecast may be somewhat misleading, becaused it is derived under the assumption that the underlying data is in fact drawn from a Gaussian distribution.

Because these forecasts were derived from a Bayesian minimum-error formulation, and were applied to turns with minimum uncertainty, any remaining uncertainty must either be attributed to inherent randomness in the turn process or lack of observability. The observability issue is treated in the next two sections by considering the other type of observations which are available, namely the *status* of various processes in the turn.

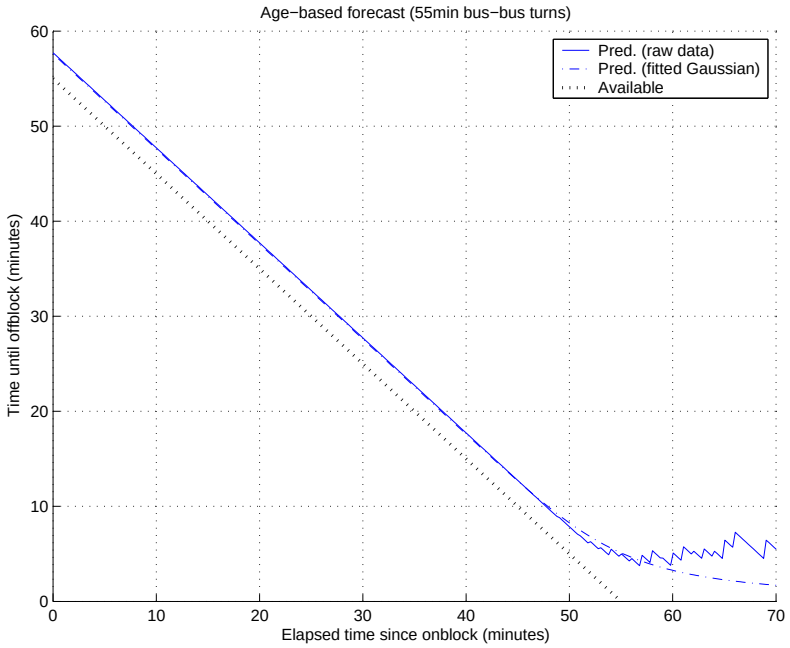

Figure 2-9: Age-based predictor for 55min bus-bus turns.

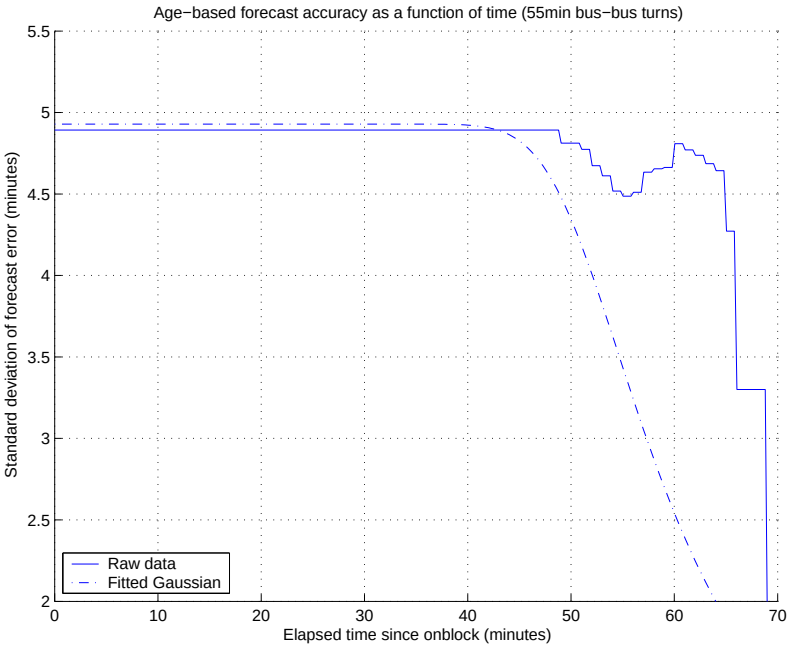

Figure 2-10: Age-based predictor: uncertainty as a function of time. A standard deviation of 5min is approximately 1/3 the magnitude of errors in current pushback forecasts.

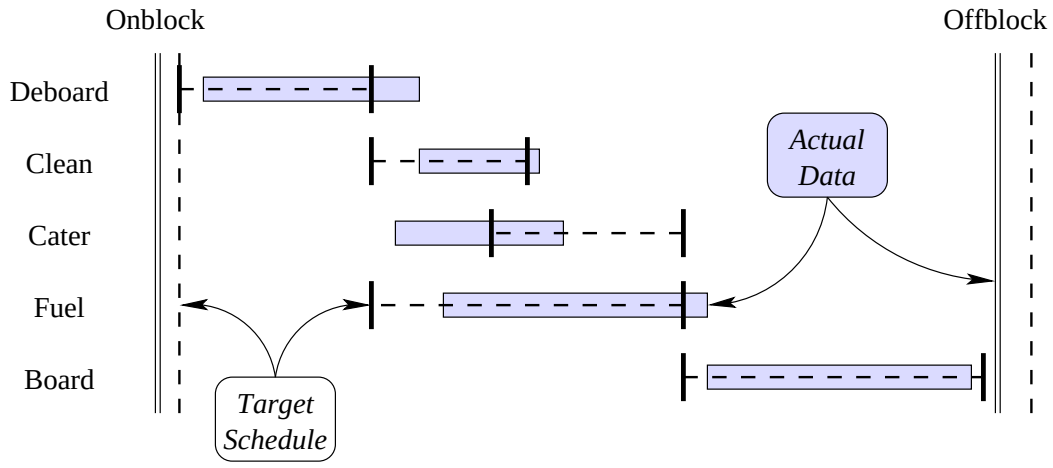

Figure 2-11: Model of the critical-path events in the turn process.

2.5.3 Forecasts using coupled updates of process status

Derivation of the model

Another reasonable approach for forecasting pushback time is to track the status (not-started/in-progress/completed) of the different processes composing a turn. Each process in a turn (e.g. catering) has a characteristic start-time and duration based on the planned ground-time, and typically must occur in sequence with some predecessors (e.g. deplaning) and successors (e.g. boarding). If a process was running unusually late, one would expect this lateness to be transmitted to the successors and thus pushback to be correspondingly delayed. This assumption, that the processes can be divided into a sequence of phases, can be encoded into statistical models of varying complexity. Airline and FAA controllers often use mental models of this form where the cause of an unusually late pushback is ascribed to a particular phase running late [MHV+99, IDA+00, IDA+98, AIC+00].

Based on interviews with Lufthansa personnel, a turn can be approximately divided into three phases: deplaning followed by "servicing" (catering, cleaning and fueling) followed by boarding. This critical-path model is illustrated in Figure 2-11. There are other turn processes including baggage handling, loading of freight, flight crew preparation, etc. but it was indicated that these activities only rarely contribute to the critical path. According to standard procedures, the deplaning process is to be completed before catering and cleaning personnel enter the cabin, and before fuel trucks activate their pumps. Similarly it is standard procedure to begin boarding only after catering and cleaning personnel have left the cabin and the fuel truck has disconnected its hoses. Hence the three phases occur in sequence as illustrated in Figure 2-11.

Because late turns are of much greater importance operationally (early turns are often easy to delay), the servicing phase is designed to model the critical path by focusing on whichever servicing process is causing delay. Therefore, the start of the servicing phase is defined as the *actual* start-time

of the servicing process with the *latest scheduled* end-time; this process has the least slack in the schedule. Similarly the end of the servicing phase is defined as the time when all three servicing processes have completed.

The three phases are assumed to have stochastically independent durations. This assumption is made partly for mathematical tractability, and partly because there is little evidence of real-world mechanisms which would link the durations. In normal operations (i.e. barring medical emergencies or passengers with limited mobility), airline personnel do not enter the cabin to hasten or delay deplaning. The catering, cleaning and fueling processes are performed by subcontractors to the airline, and subcontractor personnel are not under the immediate management of the airline station. Finally, while airline personnel can delay boarding to bring an early flight back onto schedule, it is much more difficult to coerce passengers into hastening the boarding process. Given that the three phases occur in sequence and have independent durations, the expected time-to-go until pushback is solely dependent on the available ground-time, the most recent status update, and the time elapsed since that update.

Note that this model is much simpler than the Petri net representation derived in [IDA+98]. In particular, by focusing on the sequential critical-path processes, straightforward forecast techniques can be developed which account for the available observations. The critical-path model also significantly extends the PERT model of the airline turn process given in [FHO+97, Section 4.2]. The older PERT model was derived before quantitative airline turn operations data were available, and contains 16 tasks connected by 18 links. Through a careful analysis of the currently available quantitative data, it has been possible to set aside several of these tasks, including baggage/freight handling and crew readiness, which typically do not delay the observed turn process. The simplified quantitative PERT diagram[3] is shown in Figure 2-12.

Analysis of average process behavior

To calibrate the status-based forecast, it is necessary to measure the average time-to-go when a process changes status. The raw data is shown in Figure 2-13 for the simple bus-bus turns. It is obvious, especially for turns with lots of available ground time, that some smoothing should be applied to obtain cleaner estimates of the mean time-to-go statistics.

It is reasonable to assume that a turn with more available ground-time would not have a smaller mean time-to-go for any process. The data shown in Figure 2-13 support this assumption. It is also reasonable to assume that the *variability* of the time-to-go would be constant with respect to the available ground-time. This assumption is supported by the conditions under which the data were

[3]The numerical values given in the figure were derived from the sample of bus-bus turns with 55min of available ground-time. Note that the boarding process is defined from the perspective of the gate, i.e. passengers leaving the gate and boarding a bus which takes them to the aircraft. The planned time for all passengers to have left the gate is (offblock-12min). Hence the 14.8min given for link m is *not* unusually large, but in fact agrees with the planned operations.

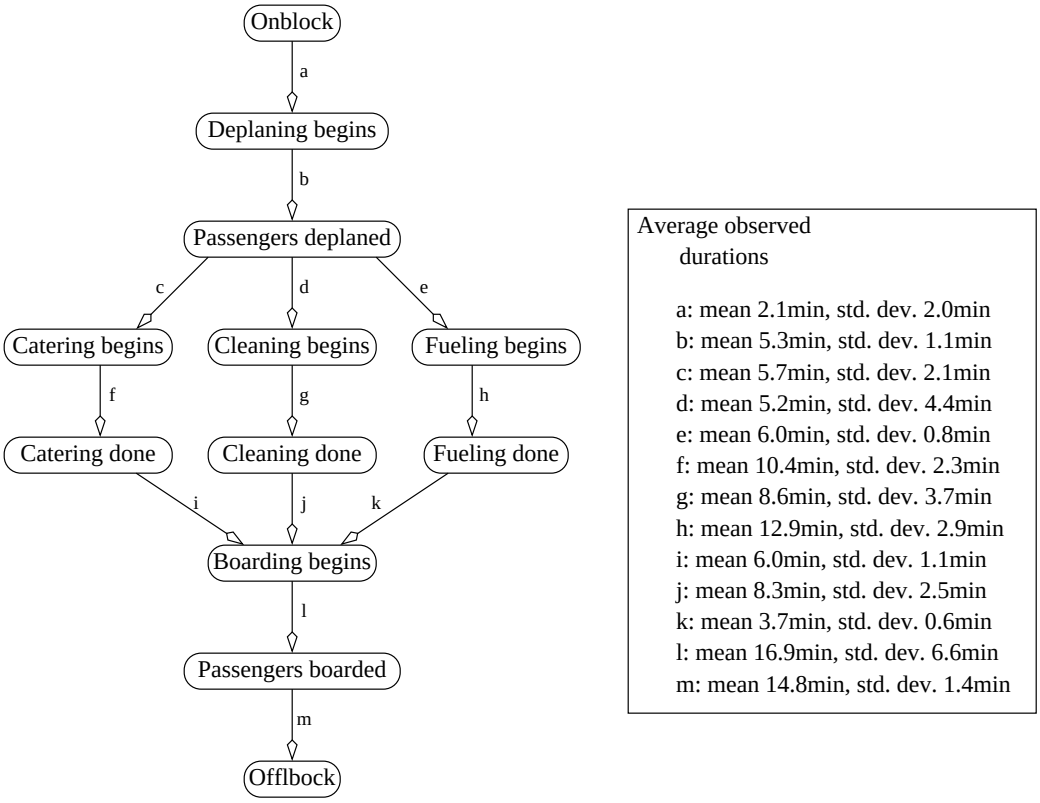

Figure 2-12: PERT representation of the critical-path model (55min bus-bus turns).

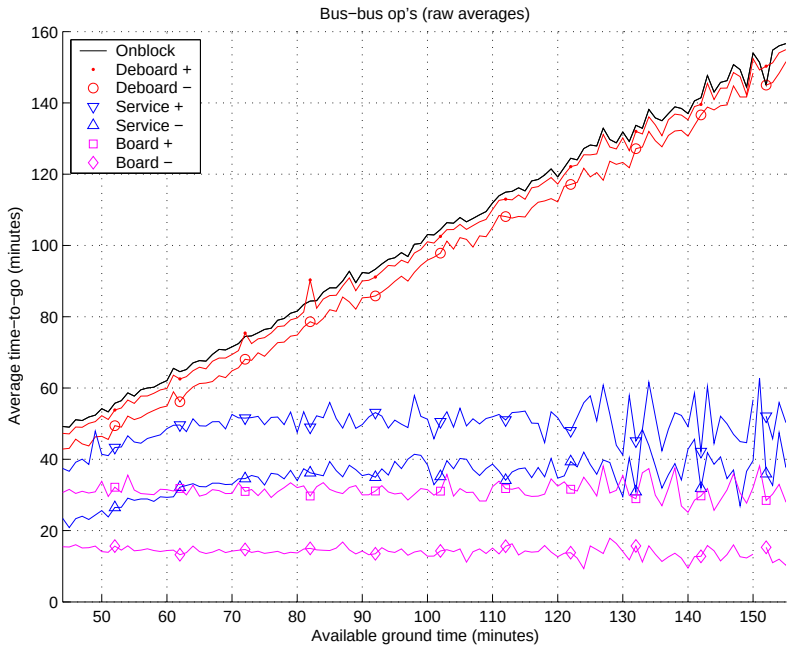

Figure 2-13: Average time-to-go as a function of process status.

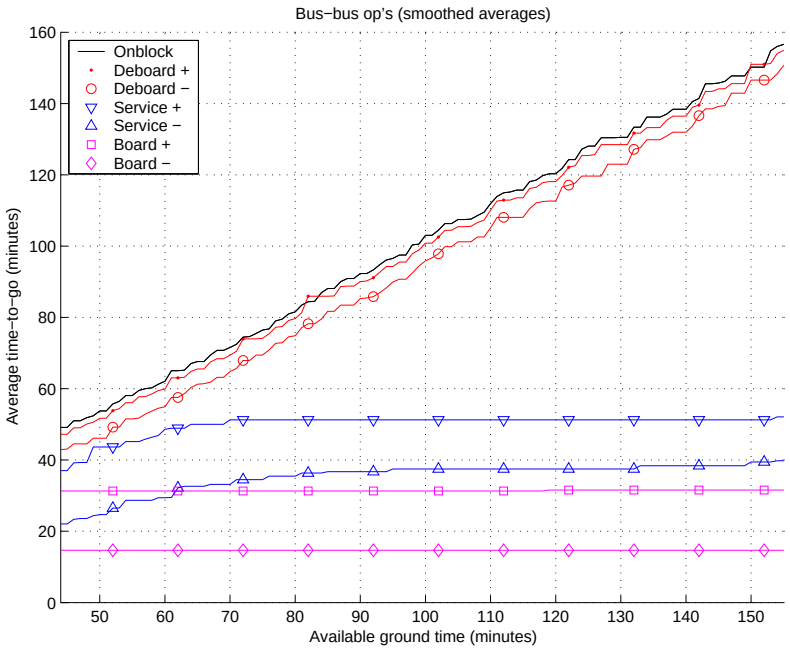

Figure 2-14: Average time-to-go, smoothed via PAVA algorithm.

measured: because the deviation in each process was being monitored and controlled in real-time to adhere to the ALLEGRO target-times, for each process the amount of variation in time-to-go should represent the minimum level of operationally acceptable (achievable) deviation around those target-times. Finally, it is mathematically convenient to assume that the data for a particular process and available ground-time are normally distributed. This assumption was tested by binning the data and applying the Lilliefors test. At 95% significance, approximately 11% of the bins were rejected as deviating from a Gaussian distribution. This result is acceptable given the expected 5% false-rejection rate.

According to these relatively weak conditions, there is a correspondingly robust statistical model which states that, for a turn with available ground-time x, each process k has a duration which is normally distributed with mean $\mu_k(x)$ and variance $\lambda_k(x)$ where $\mu_k(x)$ is non-decreasing and $\lambda_k(x)$ is constant. Under this model, the optimal data-based regression of $\mu(x)$ against x is then given by the PAVA algorithm [RWD98] which essentially smoothes the usual estimated averages $\mu_{\text{est}}(x)$ to enforce the non-decreasing constraint. Note that using a standard regression would implicitly enforce the contrary assumption that turns with different available ground-time have no relationship at all. The smoothing effect can be seen by comparing Figures 2-13 and 2-14. The monotonic regression is able to pool information among turns of similar available ground-time, resulting in a much larger effective sample size and reduced noise.

A bootstrap procedure was used to validate the monotonic regression. The data were binned

according to available time-to-go. From each bin, half the elements (rounded up) were selected. The PAVA algorithm was applied to this new sample of data. This process was repeated 40 times to obtain a set of monotonic regressions, each of which were compared to the original monotonic regression over the full dataset. The discrepancy was surprisingly low, a fact which may be partially due to the uniform sampling across the bins and partially due to the pooling of information between neighboring bins which tends to smooth out large variations.

An intuitive validation is also apparent in Figure 2-14. For short turns, the service process often overlaps with boarding; as expected, airline personnel are attempting to finish everything simultaneously. For turns with available ground-time in excess of the minimum, servicing is usually finished a few minutes before boarding begins. For these longer turns, the critical-path service process is almost always found to be fueling. While the caterers and cleaners can complete their tasks several hours in advance, it is necessary to have up-to-date information on the block fuel before fueling begins. Because the captain decides on the necessary block fuel only 30–40min before scheduled departure, fueling must begin relatively late in the turn to avoid to not have the fuel trucks remain idle for a long time. The recovery of these expected observations from the noisy data indicates that the regression algorithm is indeed behaving correctly.

Forecast algorithm

The status-based forecast is constructed from the monotonic regression as follows. When a turn with available ground-time x arrives onblock, the initial forecast of time-to-go until offblock is just the corresponding average duration. As the turn progresses, the forecast counts downwards at a constant rate of -1sec/sec. When a phase of the turn changes status (starts or stops), the forecast is updated to the average time-to-go for that particular status change, and again the forecast counts downwards at a constant rate of -1sec/sec. This straightfoward update method is sufficient for the majority of turns in the dataset.

There are three conditions in which the straightforward update method is modified:

1. If the predicted time-to-go until offblock is less than the actual time-to-go until the next change in process status, the forecast is not allowed to become negative. Instead the forecast is held at zero.

2. In the real-world operations data, there are some cases in which servicing starts before deplaning ends, or boarding starts before servicing ends. Because the phase which *ended* out of sequence is obviously not blocking the turn, its status-update can be ignored[4].

[4]More generally, to avoid dealing with all the special cases which can crop up due to real-world data deficiencies, for each flight a *maximum-length low-rank increasing subsequence* of the status changes was used for prediction. To compute this subsequence, status changes which were expected to occur earlier were assigned a lower rank. The status changes were sorted according to their time of occurrence, and the longest subsequence of monotone increasing rank was selected. In case of ties, the subsequence with smallest initial rank was selected.

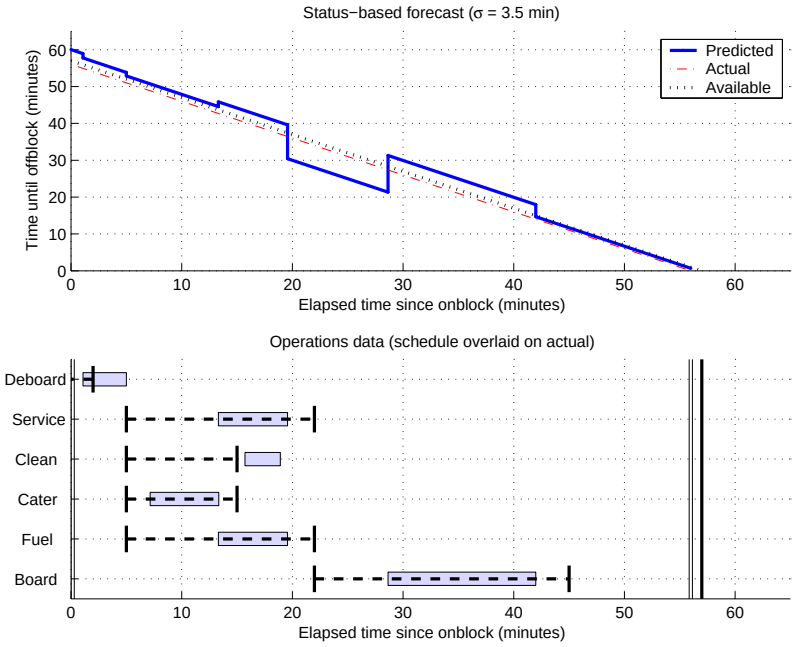

Figure 2-15: Status-based predictor: on-time flight, example #1.

3. If some process starts or finishes unusually early, the straightforward update method may produce a forecast which indicates the turn will take less than the minimum[5] ground-time. When such a case arises, the forecast is instead updated to the minimum *feasible* remaining ground-time (i.e. the minimum ground-time minus the elapsed time). This condition is rarely observed because most turns have more than the minimum ground-time available.

Illustrative examples

Four examples are shown in Figures 2-15 through 2-18. To help characterize the quality of each forecast, an average forecast accuracy σ is computed for each turn:

$$\sigma \doteq \sqrt{\frac{1}{\mathbf{X}} \int_0^{z\mathbf{X}} e^2(t)\, dt}$$

Forecasts for bus-bus flights which had punctual departures are shown in the first two figures; the following two figures correspond to delayed flights. For comparison each plot also shows the forecast based on available ground-time; the "perfect" forecast based on the actual duration; and the scheduled and actual ALLEGRO data.

The average instantaneous forecast accuracy for a subset of the simple bus-bus turns is shown in

[5]Originally this condition was not included, but upon review by experienced Lufthansa personnel, it was noted that the issue of changing the "minimum" ground-time has a large political component which this technical research does not address.

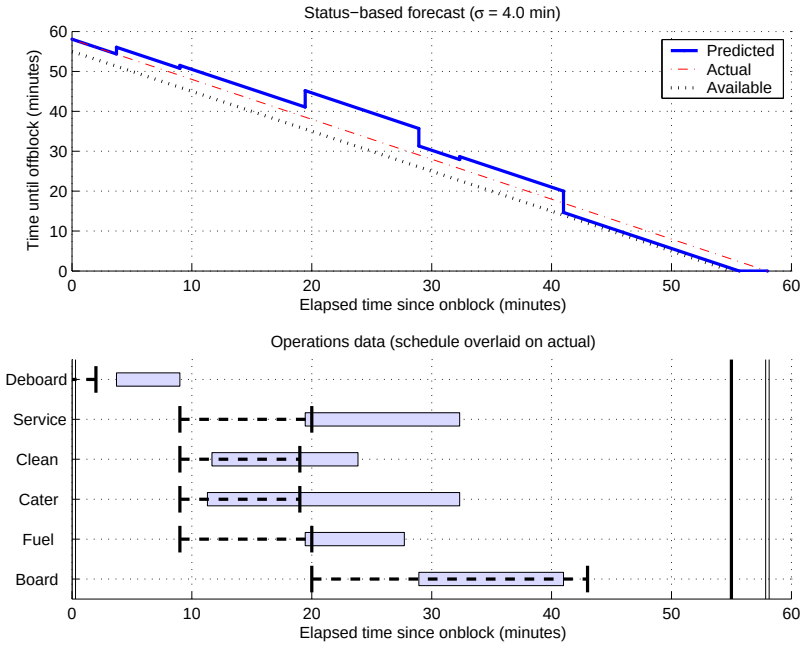

Figure 2-16: Status-based predictor: on-time flight, example #2.

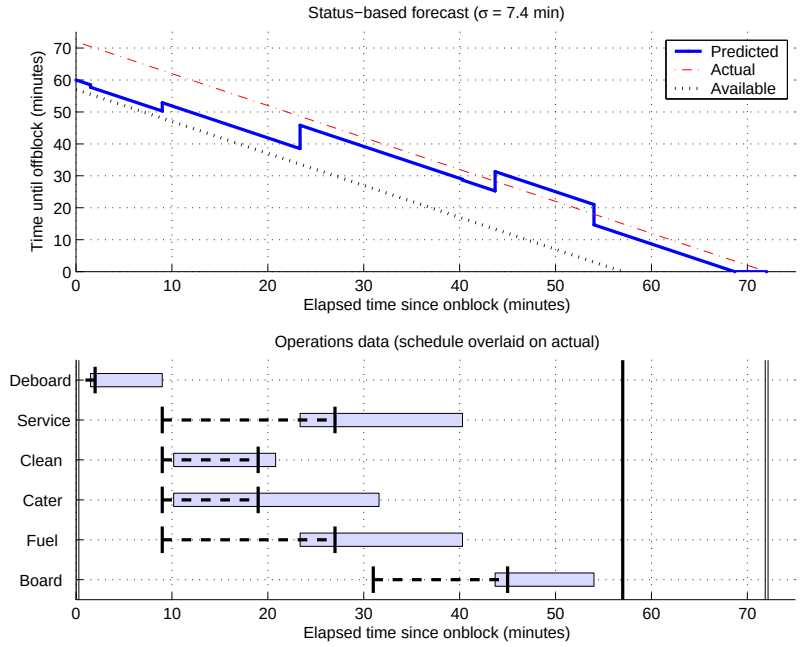

Figure 2-17: Status-based predictor: late flight, example #1.

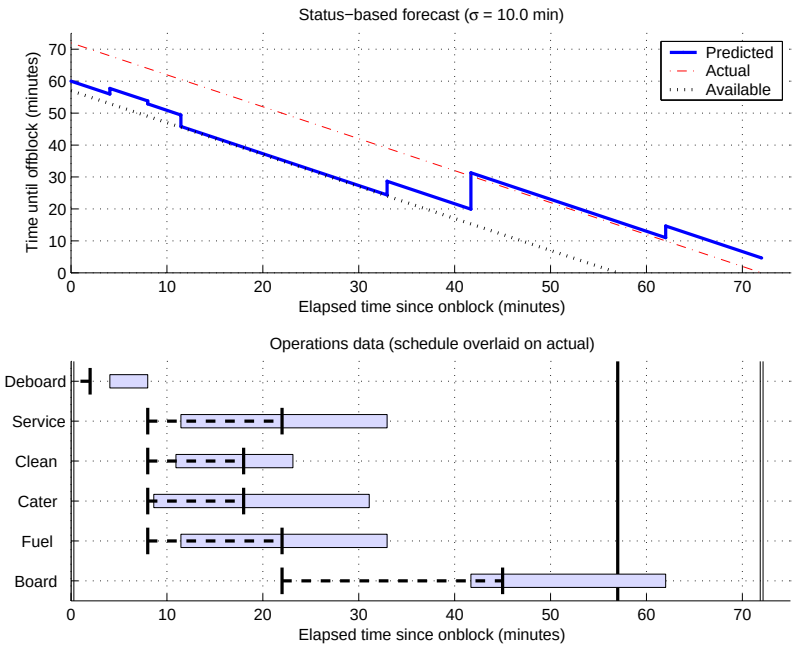

Figure 2-18: Status-based predictor: late flight, example #2.

Figure 2-19. Again a lower bound on the forecast accuracy throughout the turn is plainly apparent. In Figure 2-14 the average time-to-go clearly decreases as successive phases start and finish. However the standard deviation of the forecast accuracy in Figure 2-19 never drops below ±5min.

A third metric for the forecast accuracy is the average instantaneous accuracy as a function of time-to-go (rather than elapsed time). It is natural to expect that the forecast accuracy should steadily improve for forecasts closer to the actual pushback. This expected behavior is observed in practice as seen from Figure 2-20. This result confirms that the status-based predictor is indeed behaving correctly; the lower bound observed from Figure 2-19 is due to the fact that the actual time-to-go is relatively uncertain and of course cannot be observed directly.

Refinement to the best status-based predictor

One very important feature of Figure 2-19 is that the forecast uncertainty *increases* over certain portions of the turn! At first glance this behavior is entirely unintuitive; including additional status observations later in the turn should only decrease the uncertainty. Further analysis provides a simple answer: several of the turn processes have some slack time in the schedule. This is particularly true for longer turns, e.g. the caterers and cleaners may actually finish their jobs several *hours* ahead of schedule given the opportunity. If a process with short duration S can occur anywhere in an interval of length $L > S$, then the actual start and completion times for that process may be uniformly distributed over intervals of length $L-S$, contributing an extra $(L-S)^2/12$ to the observed variance.

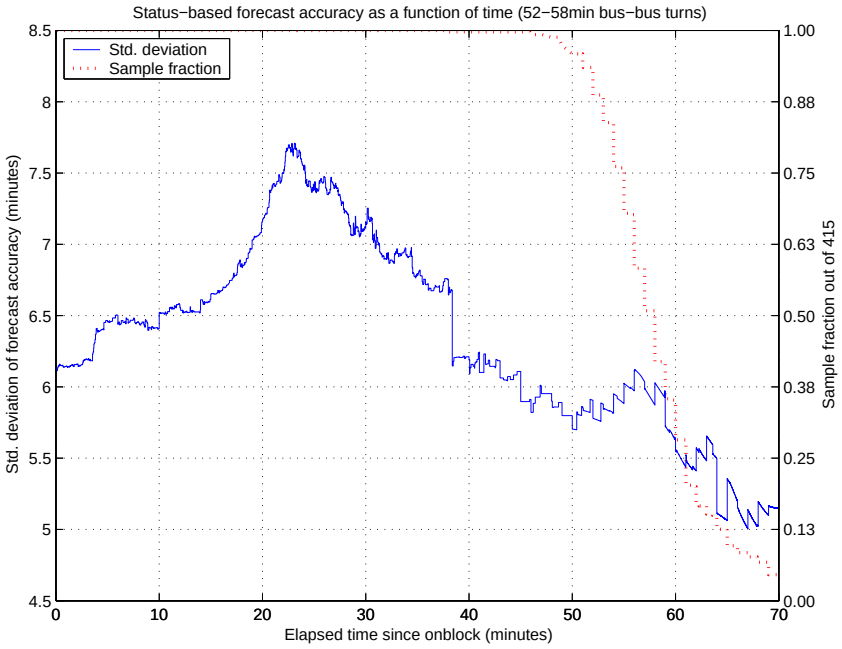

Figure 2-19: Status-based predictor: uncertainty as a function of elapsed time.

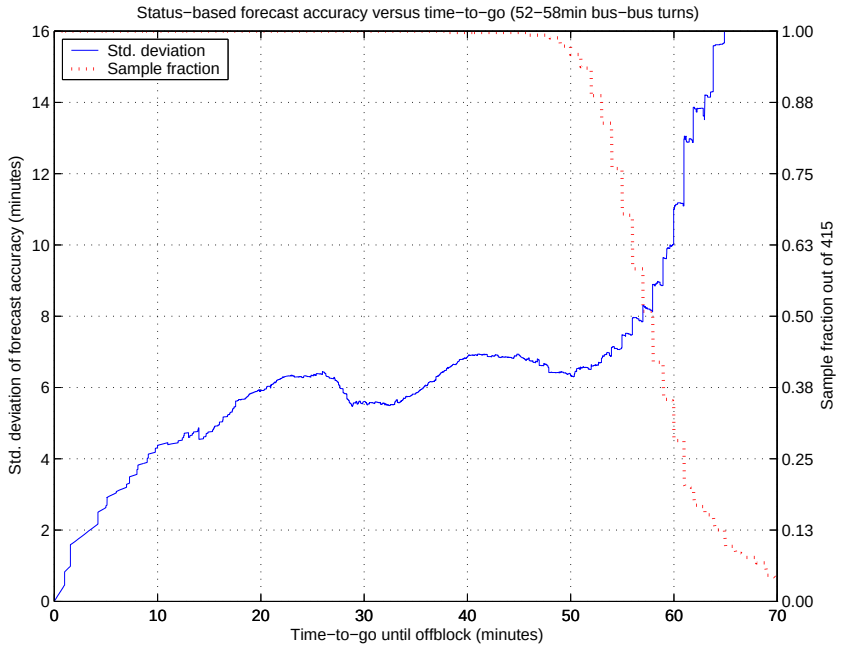

Figure 2-20: Status-based predictor: uncertainty as a function of time-to-go.

To remedy this problem, a statistical search was conducted for the most informative status changes in the turn. It was assumed that the initial onblock event was informative. This left six other processes (deplane, service, cater, clean, fuel, board) with two status changes each. For each of the $2^{12} = 4096$ subsets of these status-changes, a corresponding status-based predictor was tested against the simple bus-bus turns with 52–58min of available ground-time. The forecast uncertainty as a function of elapsed time and time-to-go was computed, and in both cases 4 performance metrics of the forecast uncertainty over 0–65min were recorded, including the minimum, the maximum, the average and the median. Overall this search yielded a sample of 4096 length-8 performance vectors. An optimal performance vector was computed by taking the minimum componentwise of all the performance vectors. The performance vector which best approximated this optimum (according to a scaled minimum-distance approximation) was then determined.

This search yielded a surprising result: the best status-based predictor uses only the initial onblock event, the start of deboarding and the end of boarding. The illustrative examples from the previous section have been re-computed using this status-based predictor and are shown in Figures 2-21 through 2-26. The fact that the servicing processes are *worse* than uninformative in predicting pushback times indicates that either these processes are not sufficient contributors to the critical path to be informative, or that a more complex status-based predictor algorithm is required.

Recall that in the status-based model derived above, the servicing phase is designed to focus on servicing processes which are most likely to lie on the critical path due to tight scheduling requirements. However, it is apparent from Figures 2-21 through 2-24 that the scheduled times for the servicing processes may not be adapting properly to the actual lengths of the processes. In some cases there is significant slack-time in the schedule, at other times the processes may take up to twice as long as expected. From interviews with Lufthansa personnel, the schedule is designed to adapt dynamically based on the available ground-time and the planned servicing requirements. Currently the adaptation rules are defined heuristically by station analysts, and executed automatically within the ALLEGRO system. It is an interesting open question to determine these schedule-times according to robust optimality criteria.

One possibility for a more complex status-based predictor algorithm is to threshold the process times, so that the servicing processes only affect the prediction if they are delayed past some critical point in time. Because the current schedule does not appear to yield informative thresholds, these critical points must be determined empirically. It is another interesting open problem to find efficient methods and/or "optimal" (according to reasonable criteria) solutions for this thresholding problem. Note that finding these thresholds is dual to the problem of improved schedule design. Although the practical impact on improving DSTs for airport surface traffic is not readily apparent, certainly the airlines may be interested in further work along these lines.

Note that this best-case predictor is evaluated using a *symmetric* error cost function — when

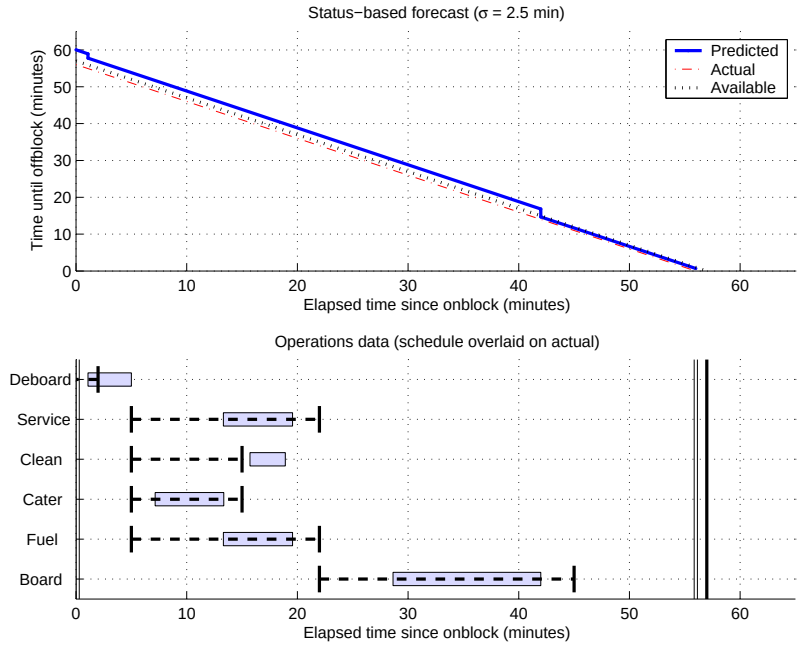

Figure 2-21: Best status-based predictor: on-time flight, example #1.

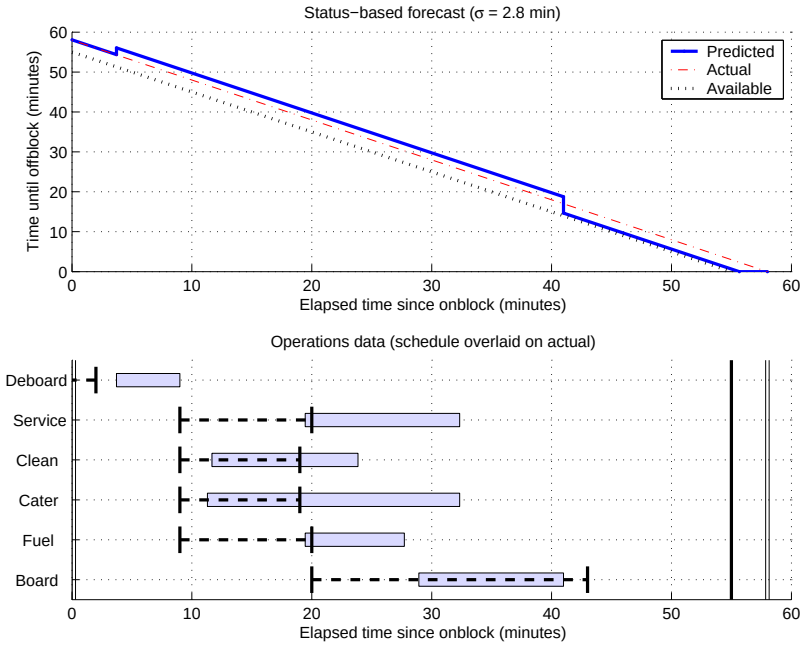

Figure 2-22: Best status-based predictor: on-time flight, example #2.

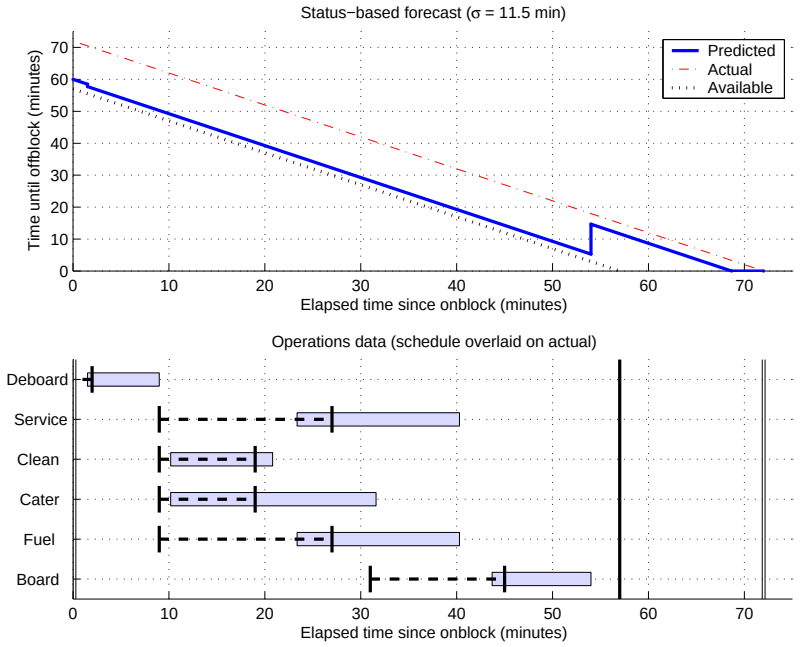

Figure 2-23: Best status-based predictor: late flight, example #1.

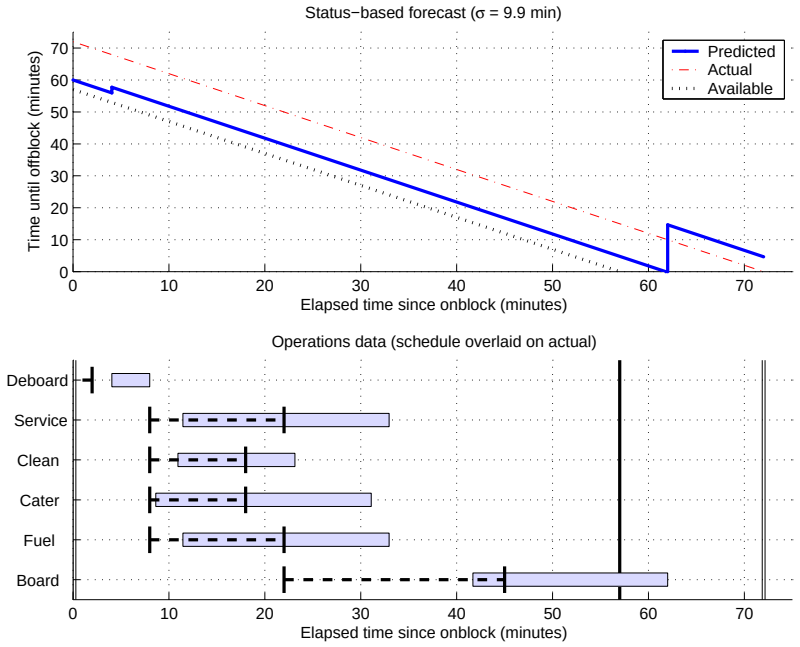

Figure 2-24: Best status-based predictor: late flight, example #2.

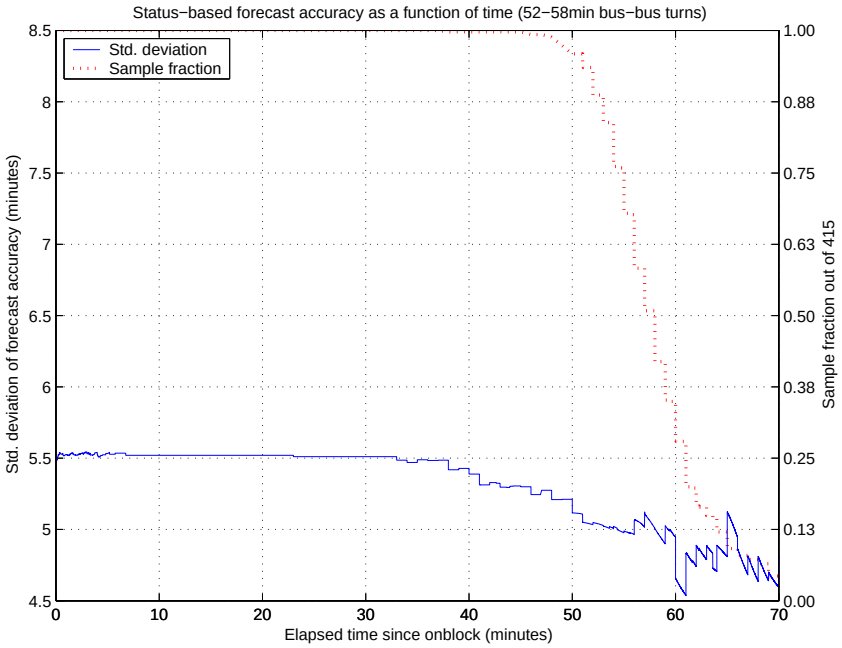

Figure 2-25: Best status-based predictor: uncertainty as a function of elapsed time.

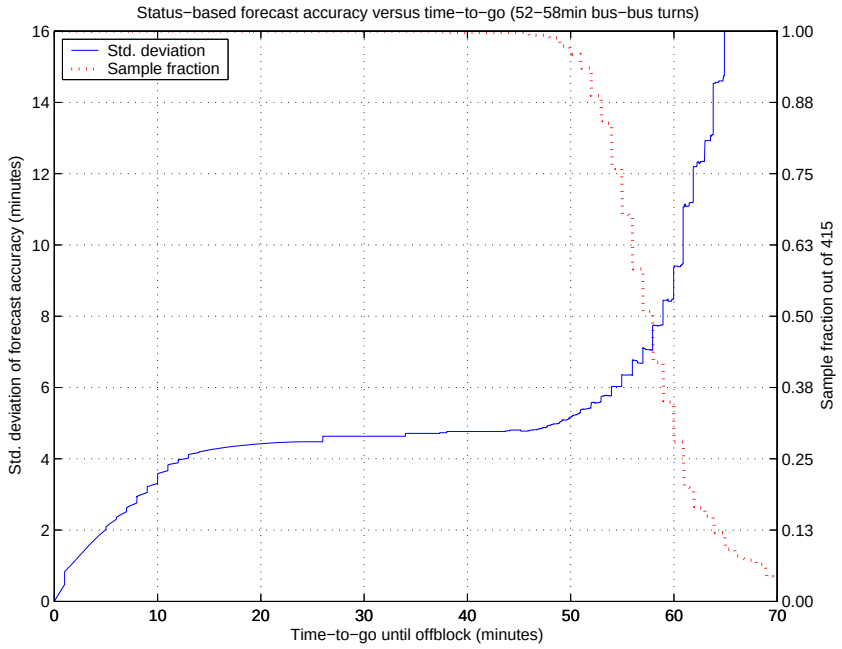

Figure 2-26: Best status-based predictor: uncertainty as a function of time-to-go.

early and late pushbacks have the same absolute deviation from the prediction, the predictor is equally penalized. It may be useful to also have a predictor which focuses soley on pushbacks which occur *after* the predicted time of pushback. Such an asymmetric predictor tends to not surprise the user with bad news. To examine this idea, a similar search was conducted, where all 4096 possible subsets of the status observations were considered, and the resulting predictor performance was evaluated for each possible subset. Instead of the symmetric square-law error cost function, the asymmetric error cost function

$$cost(e(t)) \doteq \max\{0, -e(t)\}^2$$

was used to evaluate the predictor performance[6]. The search revealed an identical best predictor! This effect may occur because there is an upper bound on the haste of a turn, while the delay of a turn is (theoretically) unbounded. Hence the symmetric error cost function would strongly penalize predictors which were unable to forecast large delays, and therefore would favor the same predictors as the asymmetric error cost function. An ongoing research project is to optimize and evaluate status-based predictors using only the sample of turns which were delayed, in order to avoid dilution of the rare delayed datapoints among the numerous on-time datapoints.

2.5.4 Combined pushback forecasts using status and age

Combining status-based and age-based forecasts should result in higher model fidelity and improvements in forecast accuracy. The obvious approach is to interpolate between status updates using the age-based equations. This approach has been successfully applied using the age-based equations and the three status changes (onblock, begin deboarding, and end boarding) from the best status-based predictor. One useful side-effect is that the three special cases for the status-based predictors can be ignored: the age-based equations correctly account for unusually short predictions and/or long turns, and there is no chance that the three status changes which are used will occur out of sequence!

A comparison of the three predictors is shown in Figure 2-27. Again the data are binned according to the available ground-time. For every aircraft in a bin, a combined prediction was computed for each flight, and the mean-square error was computed over the life of that prediction. The plot shows the "typical" (median) mean-square error for each bin. As expected, it is difficult to predict pushback times for turns with less than the best-case minimum ground-time of about 40min.

For turns with one to two hours of available ground-time, the performance does not change substantially, indicating that haste and delay for these flights is not strongly affected by steadily increasing slack in the schedule. For these turns, there is a long period of time between the first two status updates (arrival and the start of deplaning) and the third update (end of boarding). Over this long period of time, there is little risk of the turn ending unusually early, and the age-

[6]Recall that $e(t)$ is the difference between predicted and actual time-to-go, hence this asymmetric cost function only penalizes the predictor when a pushback occurs after the predicted time of pushback.

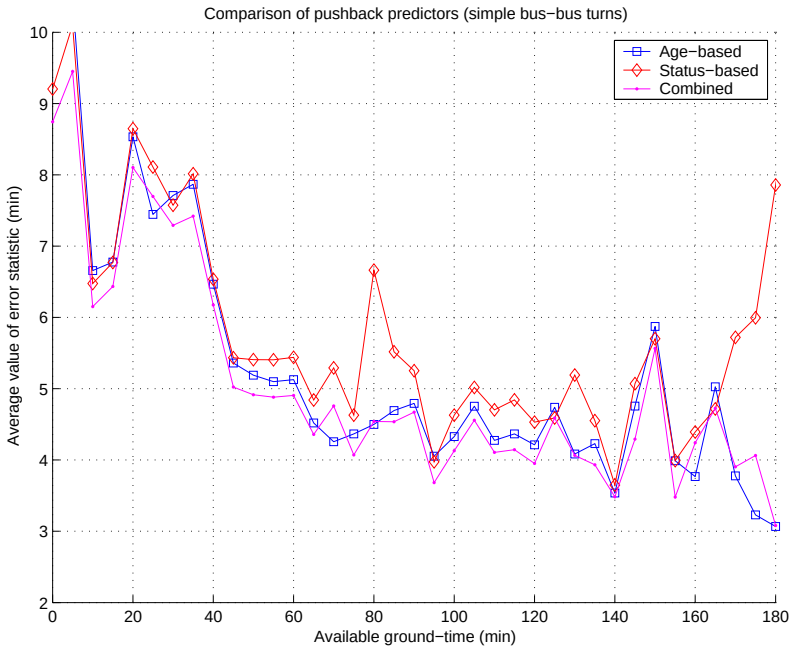

Figure 2-27: Comparison of age-based, best status-based and combined forecasts.

based equations simply decrease the predicted time-to-go at a steady slope of -1. Together these observations indicate that the long prediction horizons observed in Figures 2-10 and 2-25 are also occurring for turns with much longer available ground-times.

2.6 Conclusions

Uncertainty in the airline turn process imposes limitations on pushback predictability, which in turn limits DST performance. The forecast techniques developed in this chapter are the first analyses which quantify the *minimum* of this uncertainty, when operations occur under the best possible conditions. The best-case standard deviation of forecast error for all of the forecast techniques was observed to be lower-bounded at roughly half (±5min) of the average-case standard deviation derived in previous studies. Furthermore, this forecast error did not decrease until only a few minutes prior to pushback. These results have been carefully tested against a very large and detailed set of real-world operations data, which offers more observability into the airline turn processes than has ever been available previously. The obvious conclusion is that robustness issues in DST development will *not* be resolved by future improvements in observability of the pushback process. DSTs must be designed to incorporate robust mechanisms for coping with at least the level of pushback demand stochasticity characterized above.

In general, the stochasticity in airline turn operations appears to be concentrated near the end of the turn process, so that ATC has a forecast of accuracy on the order of ±5min over a long

horizon, with higher accuracies only coming near the end of the turn process. DSTs operating on a long time-horizon, well beyond the 15–30min lookahead typically practiced by ATC, can thus create plans and provide information based essentially on a static forecast of airline operations. These plans must still be robust to changing weather conditions, downstream traffic restrictions, and *error* in predicted airline departure demand, but do not need to cope with changes in airline demand which are not related to weather.

In constrast, DSTs operating on the same time-scale as the human air traffic controllers must adapt to the "sliding delays" observed in Figure 2-9. As a turn's expected pushback time approaches, the absence of expected status changes (i.e. the end, or even the start, of the boarding process) indicates that the turn is likely to be late, but the exact degree of lateness becomes very difficult to predict. In part this is a data deficiency: relatively few turns are delayed in such a manner, and hence empirical prediction becomes difficult. There is also a difficulty in the modeling process itself: while the typical random variables used to model event durations such as lognormal, Weibull or inverse Gaussian all have long monotone decreasing tails, the order of decay for these tails can be quite different. This yields parametric predictions which may be highly sensitive not only to the chosen *parameters* of the model but also to the *selection* of which distribution to use. Due to these difficulties with delayed turns over a short horizon, it is necessary to build DSTs which *do not rely on predicted pushback times* for such turns. Instead, plans must be derived which are invariant with respect to both actual pushback times and time-varying pushback demand for delayed turns. The development of such robust DSTs is a topic which is taken up in the following chapters.

Chapter 3

Ceno-Scale Modelling of Airport Surface Traffic

3.1 Introduction

With remarkable day-to-day repeatability, human controllers in charge of airport surface traffic solve dynamic sequencing and scheduling problems characterized by uncertain inputs, partially observed states, multiple objectives, and complex constraints. This general problem is referred to herein as Surface Operations Management (SOM) following Anagnostakis' proposed classification scheme [Ana03]. Sequencing and scheduling problems have a long history in operations research and it is well-known that even simple instances can be NP-hard to solve analytically; SOM is not a simple instance.

The research literature on SOM has tended to focus on either characterizing SOM as a whole, or developing solutions for challenging sub-problems. The motivation in either case is to further the development of useful DSTs by identifying sources of inefficiency, investigating algorithms and software to assist controllers with relevant tasks, and evaluating the potential benefits. In this vein, several building blocks for *ceno-scale*[1] models are developed in this chapter. These encode several of the most important mechanisms by which delay and inefficiencies accrue in airport surface traffic.

Ceno-scale modelling has been derived with several goals in mind. It is designed to be usable within the limitations of currently-available operations data. Therefore it follows the *mesoscopic* modelling tradeoff as outlined in [Puj99]. This tradeoff balances the use of deterministic model

[1] If the reader will permit some tongue-in-cheek humor... Pujet used the term *mesoscopic* to distinguish his work from previous *macroscopic* models (e.g. queueing models which do not explicitly encode **any** airport-specific delay mechanisms) and *microscopic* models (e.g. large-scale simulations which encoded every meter of pavement) [Puj99]. It did not seem correct to simply co-opt his terminology, but in this research area the macro/meso/micro linguistic axis is filled to capacity. In a bold conceptual leap, we have drawn inspiration from the Protero/Paleo/Meso/Cenozoic eras of paleontology.

structures to encode effects which are observable through currently-available operations data, and the use of stochastic distributions to represent both inherent random effects and unobserved system complexity. It is also designed to complement the mental models used by human air traffic controllers. The relevant timescale for ceno-scale modelling is a few minutes to several hours. This complements the tactical time-scale of human air traffic controllers, who often operate within a 15min lookahead window. Ceno-scale models have a modular structure, where each component represents one of the major mechanisms of delay and inefficiency observed in airport surface traffic. This structure aids in comprehension and possible future acceptance by human air traffic controllers who use mental models with similar structure; in adaptability to different airports with differing operational character; and in the ability to separately calibrate the various components for use in Monte Carlo simulations. The utility of ceno-scale modelling is that a very wide range of the delay mechanisms observed in airport surface traffic can be described using only a simple low-complexity extension of existing mesoscopic [Puj99] models. The design rationale is to *approximately* capture the traffic dynamics, incorporating those dynamics which are known to be significant such as queueing congestion and takeoff sequencing with respect to downstream restrictions, while judiciously choosing to model the real-world system at the same level of detail as represented by available operations data.

3.2 Background

It is important to understand SOM from a broad viewpoint before tackling specific research problems. There are numerous examples in the literature where a subproblem of SOM has been isolated, abstracted out, and solved — all to yield a solution with no practical impact on real-world operations. For example, deterministic sequencing of takeoffs at the runway resembles a standard operations research problem. A steady stream of "new" takeoff sequencing algorithms have been generated starting in the late 1970's [Dea76, Psa79] and continuing to the present [BKSA00, BKSA03]. Essentially none of these deterministic formulations have been employed in real-world control of departure traffic. In the larger context of SOM, takeoff sequencing is a fundamentally stochastic problem which requires robust solutions involving a partnership between DST computation and human controller flexibility, for example [Ana03, AC02, AC03, ACBV01, AB02, AW02]. With the intent of developing useful and practical results, this section reviews some of the existing work on characterizing SOM, including a meta-survey of existing models for airport surface traffic; a generalized conceptual model of delay and control mechanisms; and mesoscopic queueing models which are historical precursors of ceno-scale modelling.

3.2.1 Review of existing models

There already exists a wide variety of models for airport surface traffic. Several authors have written literature reviews for this area. Three such surveys which cover the same corpus from three contrasting viewpoints are summarized here. The principal capabilities and shortcomings of a variety of software tools are surveyed in [OBD+98]. Although that survey is several years old, the tools it describes are still in practical use for day-to-day traffic control on strategic timescales of 1–6 hours; longer-term infrastructure planning on timescales of months to years; and research analyses. In Anagnostakis' PhD thesis [Ana03], a pair of comprehensive literature reviews are presented, one devoted to SOM in general and one to the subproblem of Runway Operations Planning. Note that Anagnostakis' work is recent and covers the state of the art. Finally, Pujet opens his PhD [Puj99] with a review of stochastic mathematical models for airport surface traffic. These models are typically derived from observations of real-world operations and then instantiated as Monte Carlo simulations, which in turn drive many of the software tools described in [OBD+98].

Starting from these surveys, an in-depth literature search has been conducted. From this in-depth search, it is apparent that pre-existing models do not simultaneously fufill the ceno-scale modelling requirements of robustness to the limitations of currently-observable data; operation on a tactical time-scale; and accurate representation of the identified delay mechanisms for airport surface traffic. Ceno-scale modelling fills this niche.

3.2.2 Conceptual model for airport surface operations

There are several mechanisms by which uncertainty and delay accrue in airport surface traffic. Through an intensive investigation at BOS based on a variety of data-sources, Idris was able to identify and provide anecdotal evidence of the full range of such mechanisms [Idr00]. These mechanisms can be partitioned according to the location of the aircraft when the delay is accrued: during the taxi processes, at the runway queues, or while waiting for an available gate. In addition, because ATC control processes must be synchronized with the physical motion of aircraft, it is possible for aircraft to experience delays when ATC becomes overworked. One of the design goals of ceno-scale modelling is to encode as many of these identified mechanisms as can be observed using currently available operations data.

If the airport surface was completely empty, an aircraft could proceed unimpeded between the gates and runways. Even in this hypothetical situation, several stochastic factors could still perturb the taxi-time. Delays can be caused by mechanical problems discovered during pre-flight checks[2]; late delivery of weight-and-balance numbers from the airline operations center; differing aircraft and pilot performance; etc. Similar causes of delay can afflict arrival traffic. Note that these causes

[2]Several pre-flight checks, e.g. engine-function tests which require the engines to be running, are performed while the aircraft is taxiing to the runway,

of delay are not associated with warning signs which can be observed in advance, and hence are modeled as random effects.

After a departure completes its unimpeded taxi-time from its gate to its assigned runway, it waits in a queue. These queues develop because the runway is the primary location where ATC copes with downstream restrictions, arrival/departure interactions, runway crossings by taxiing aircraft, etc. [Idr00, p. 92–114]. There are other places on the airport surface where departure delays due to ATC may manifest, but the runway queues account for the majority [IDA$^+$98, IDA$^+$00]. In a symmetric fashion, after an arrival completes its unimpeded taxi-time, the primary cause of delay is lack of available gate resources. For both arrival and departure traffic, the ceno-scale model explicitly encodes these queueing processes as contentions for aircraft- and process-specific sets of limited resources. Note that these causes of delay, although confounded with random unimpeded taxi times, are at least conceptually amenable to forecasting. In particular, the resource requirements, resource contentions and resource limitations can be estimated before aircraft physically arrive at the runway queues. This idea is worked out in more detail in the next chapter.

Idris also identified a novel *parallel* queueing mechanism by which delays can accrue. As an aircraft moves physically across the airport surface, control and responsibility for that aircraft must be transferred through the ATC system. Each aircraft has a corresponding *flight strip* in the ATC tower which is literally handed between the various controller positions; possession of the flight strip indicates who has control and responsibility. For an aircraft to move across the airport surface, its flight strip must also move through parallel queues in the control process. Thus aircraft become susceptible to delays caused by excessive controller workload. Aside from Idris' conceptual models, none of the models found in the literature search of section 3.2.1 explicitly capture this effect. One property of the ceno-scale model is that the concept of resource contention used to encode queueing processes at the runways and gates is inherently parallel. Hence Idris' novel parallel queueing mechanism can be easily encoded in a ceno-scale model without introducing additional complexity.

Idris described the overall system of airport surface traffic as a queueing network controlled by blocking (abbreviated herein as "QNCB"), with blocks due to both the finite physical capacity of the system (e.g. of a physical segment of the taxiways) and the intentional blocks applied by air traffic controllers [Idr00]. The network structure of a QNCB is intended to reproduce the physical queues of aircraft which develop on the airport surface, while the queue-length constraints and the ATC clearance request/delivery processes form the blocking part of a QNCB. The reader is referred to [ICBK02, Fig. 2] for an example of the numerous surface queues and blocking points which can form at BOS. There is corresponding research on the design of DSTs based on the QNCB concept [AIC$^+$00].

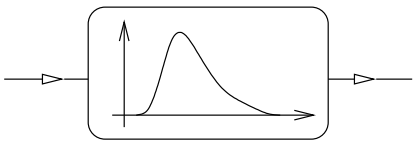

Figure 3-1: Basic component of a mesoscopic model: Unimpeded stochastic taxi-time.

3.2.3 Mesoscopic modelling of airport surface traffic

There are some difficulties with using a conceptual QNCB model for mathematical or simulation-based analyses. To calibrate a QNCB so it represents a particular airport, the capacities of the taxiway network must be estimated from available operations data. Currently the available data do not reflect aircraft positions on the airport surface explicitly (i.e. as observations from surface surveillance radar) but only implicitly (i.e. as verbal clearance requests and deliveries) [Atk02]. Assuming that capacities can be inferred from these verbal messages, it is still necessary to infer the implicit ATC control strategy from the blocking actions distributed across the various physical queues of aircraft. To date only heuristic methods have been used to infer the ATC control strategy. At the present time, lacking automated tools for collection and analysis of the required data, QNCB calibration requires an extraordinary amount of fieldwork. In fact Idris did take this approach in his thesis [Idr00]: he invested several hundred hours in the ATC tower at BOS. On a number of occasions he was able to obtain and analyze several hours of operations data and thus derive anecdotal support for the delay mechanisms he described.

There are alternative modeling approaches which choose a different tradeoff between abstraction of the full system complexity; fidelity between the model equations and intuitively understood behaviors observed in the real-world system; and robustness of the model state-space against real-world complexities and lack of observability. In particular, the difficulties with calibrating and using a QNCB model may be sidestepped, at the price of reduced model fidelity, by adding another layer of abstraction. To describe an airport, a QNCB model replicates many times the same basic network component of a finite-capacity queue with blocking. A large number of physical parameters and a complex distributed control algorithm must be specified. However, under typical operating conditions, most of these network components are not saturated and do not actively constrain the flow of traffic. It is therefore reasonable to lump together the cumulative effects of queueing congestion, physical travel-time of aircraft, minor changes to standardized routings, etc. throughout the network into just a handful of network elements. Similarly, while there are multiple control points distributed across the airport surface, much of the ATC control activity is intended to cope with downstream flow restrictions and arrival/departure interactions. These constraints are primarily met due to ATC control actions applied at the runways.

After simplifying both the traffic network and the control strategies, a reasonable first-order

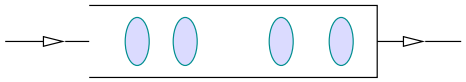

Figure 3-2: Basic component of a mesoscopic model: Simple FIFO queue.

model of airport surface traffic consists of a stochastic distribution for the unimpeded taxi-time (Figure 3-1), followed by first-in first-out (FIFO) queues that represent the performance-dominating physical queues of aircraft which form near the runways (Figure 3-2). This type of *mesoscopic* model has been thoroughly worked out in [ACFH00, Car01, Puj99], including calibration algorithms for working with actual operations data, validation via Monte Carlo simulation, and application to several airports. Mesoscopic models directly encode the most important mechanisms for delay, while factors which are not automatically recorded in the available operations data are lumped together and represented as stochastic effects. The ceno-scale model extends the class of mesoscopic lumped-element models to handle a larger subset of the delay mechanisms identified by Idris et al.

3.3 Abstract ceno-scale model components

In this section, two types of ceno-scale model components are developed. These model components interact with the FIFO aircraft buffers used in mesoscopic modelling, and permit a significant increase in modelling power for a modest increase in model complexity. The discussion in this section is fairly abstract. A variety of concrete examples are given in section 3.4 to demonstrate how these ceno-scale model components can be used to capture a wide range of observed delay mechanisms.

3.3.1 Resource contention

The servicing/blocking behavior of aircraft buffers has been modelled in several ways. The QNCB model focuses on the process of ATC clearance delivery. In this case, the human air traffic controller creates an artifical resource scarcity to modulate traffic flow rates. Queueing models (e.g. mesoscopic models) often model blocking with a single stochastic server at each runway. In this case, while there may be multiple physical resources which are causing a traffic bottleneck, the use of just a single server enforces the first-come first-served dynamics which are typically observed. Models for optimization of departure sequences usually focus on deterministic spacing constraints between pairs of aircraft. A sequencing model may focus on one or several types of procedural limitations on the minimum spacing between each pair of aircraft, e.g. due to wake-vortex restrictions (between adjacent aircraft only) or downstream flow restrictions (between all pairs of aircraft). These models are variations on a common theme, each focusing on a particular limiting resource and the corresponding timing restrictions.

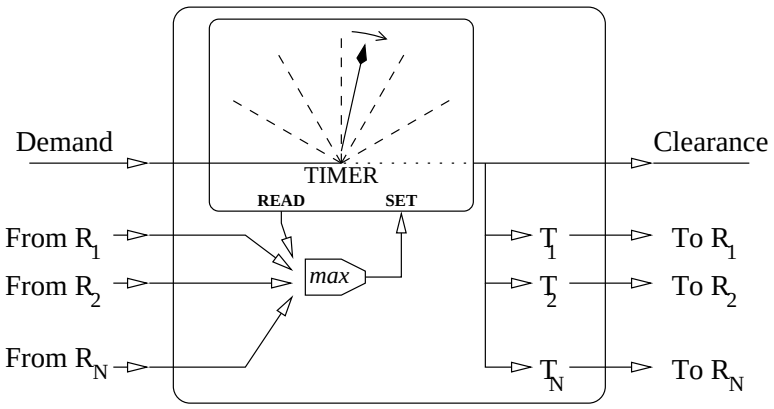

Figure 3-3: A conceptual schematic of the workings of a single MRS resource.

A more general concept of *multi-resource synchronization* (MRS) encompasses these specific models. In a ceno-scale model instance, the output of each FIFO aircraft buffer is regulated by a resource-controlled MRS gate. As will be shown in section 3.4, such MRS gates can encode a wide range of the dynamics observed in airport surface traffic. For example, by considering several resources simultaneously, the runway throughput effects caused by ATC workload limitations, queueing, and deterministic spacing constraints can be merged into a single framework. This modelling power is achieved using only a single prototypical resource model. The generality is derived from connecting together collections of these simple resources into feedback networks.

A notional diagram for a single MRS resource is shown in Figure 3-3. Each resource has two types of inputs, shown on the left-hand side of the diagram, and two corresponding types of outputs, shown on the right-hand side of the diagram. The first input is *demand*, which signals that an aircraft is attempting to use this resource. The remaining inputs are feedback signals, either via a self-loop or from other resources in the feedback network. The internal state of a resource is represented as a count-down timer. The timer indicates the time-to-go until the resource is next available for use. When the timer is at zero and an aircraft demands to use the resource, a *clearance* is immediately delivered on the first output. Simultaneously, the resource sends *consumption* signals on its remaining outputs into the feedback network. These consumption signals go into the feedback inputs of the connected resources. Each resource examines its input consumption signals and the value of its timer, computes the maximum of these values, and resets its timer to this maximum value.

A simple example is shown in Figure 3-4. A single resource is connected to itself in a feedback loop. Whenever an aircraft receives a clearance from this resource, the resource sends itself a consumption signal T, which blocks itself from delivering a clearance for T time-units. Acting in isolation, this resource behaves like a single queueing server of either deterministic character (T is constant) or stochastic character (T is chosen i.i.d. from a given probability distribution).

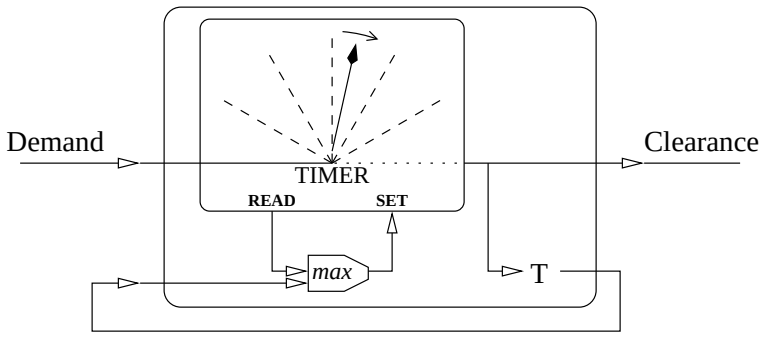

Figure 3-4: Simple MRS example: A single MRS resource with self-feedback which models a queueing server.

Larger feedback networks are used to model more complex systems. The network shown in Figure 3-5 could be used in a departure runway model. In this case, there are three resources which constrain takeoffs. The first two resources HEAVY and LARGE/SMALL correspond to the standard FAA aircraft weight classes. A Heavy aircraft will demand the HEAVY resource. After a Heavy aircraft is cleared for takeoff, the HEAVY resource blocks itself for 90sec, and also blocks the LARGE/SMALL resource for 120sec — these times correspond to the necessary minimum spacing due to wake vortices. Similarly the takeoff of a Large/Small aircraft will demand the LARGE/SMALL resource and block all Large/Small and Heavy takeoffs for 60sec. In addition, there is a resource Fix-EXMPL corresponding to a traffic metering fix. This resource imposes a minimum 300sec delay between successive flights through the EXMPL fix.

To build a MRS gate, input and output synchronization logic are wrapped around a resource feedback network. A notional diagram and an expanded logic-circuit diagram of such a gate are shown in Figure 3-6. The logic-circuit diagram shows the input logic (a demultiplexer), the resource feedback network, and the output logic (an array of identical circuits, one per resource). Each aircraft attempting to pass through the gate has a characteristic set of demanded resources. The input demultiplexer examines the aircraft's set of resources and signals any corresponding resources inside the gate. In order to pass through, the aircraft must obtain clearance from *all* of its resources which are present in the gate. As soon as the last required clearance is delivered, the aircraft is immediately processed through the gate.

It is useful to also give an algebraic formulation. Each resource R is described by a 3-tuple (last-used, next-avail, spacing). Both last-used and next-avail mark event epochs: last-used is the previous time when R delivered a clearance, and next-avail is the next time when R will be able to deliver another clearance. Both of these epochs are initialized to $-\infty$. The consumption outputs from R are described by spacing, a mapping between resources to which R sends consumption signals, and the duration of time for which R consumes those resources. A MRS gate G is simply a set of resources $\{R_{G_1}, R_{G_2}, \ldots, R_{G_n}\}$. An aircraft A is also a set of resources $\{R_{A_1}, R_{A_2}, \ldots, R_{A_m}\}$.

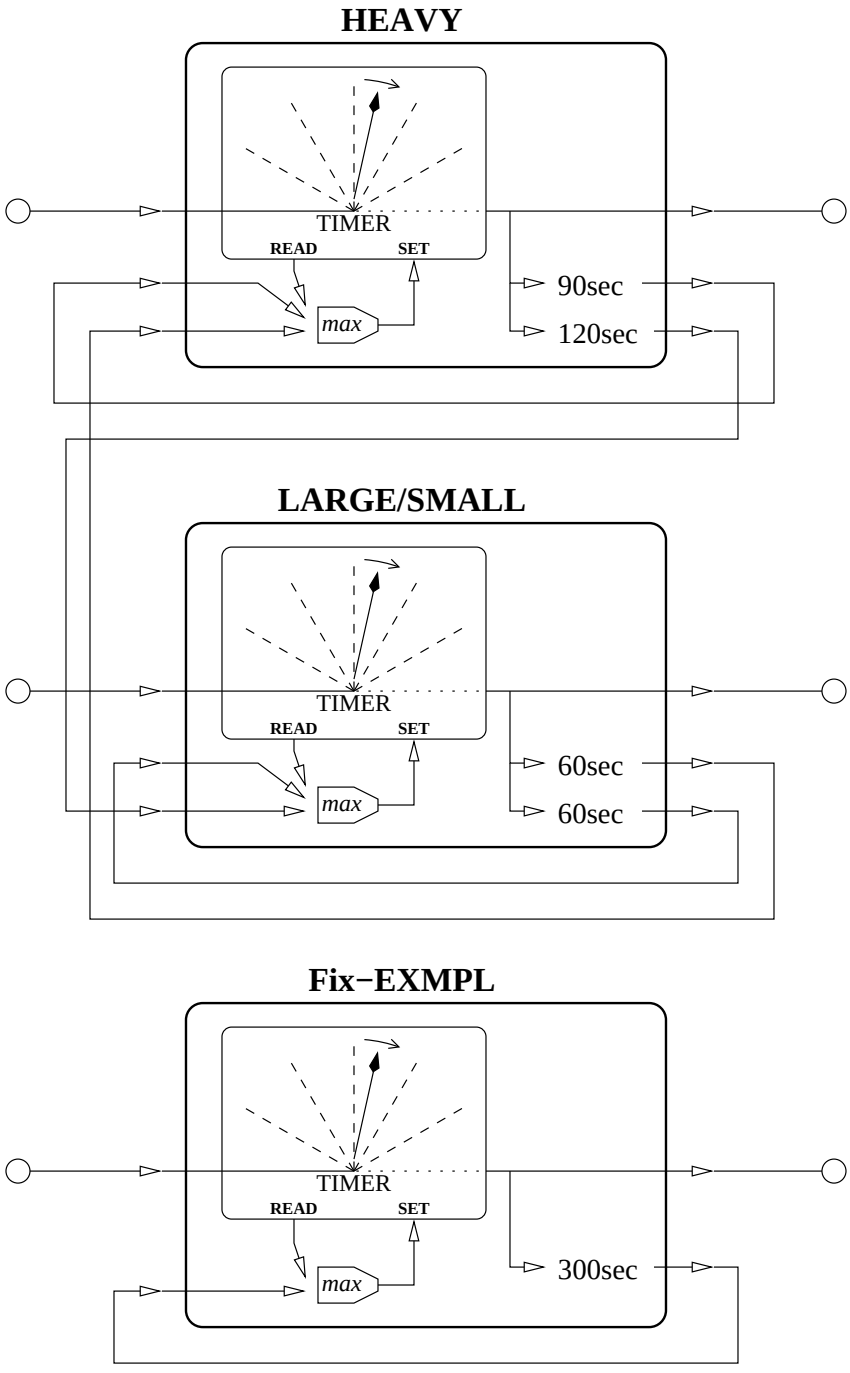

Figure 3-5: Complex MRS example: A feedback network of MRS resources which represent runway restrictions due to both wake vortices and departure fix throughput.

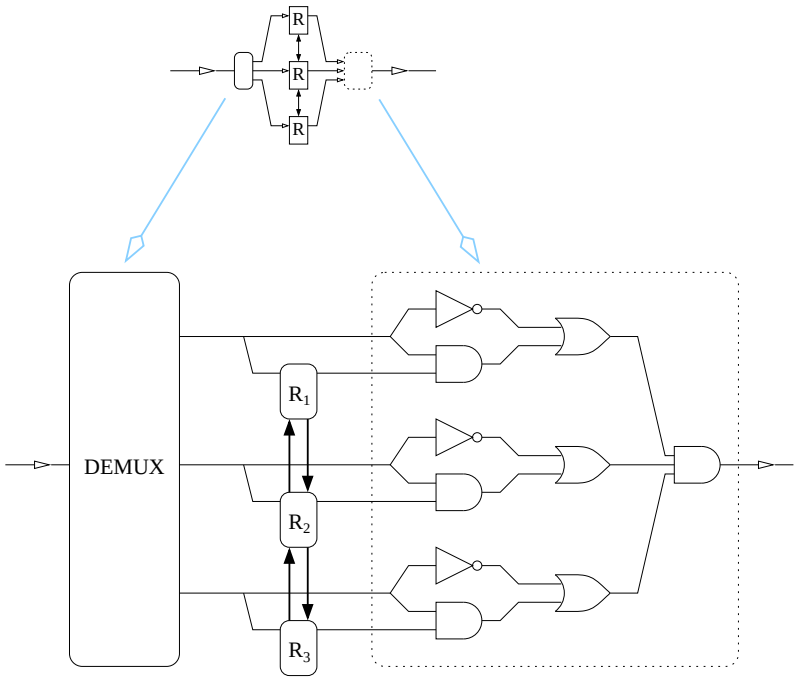

Figure 3-6: Input/output synchronization logic of a resource-controlled gate.

Given this algebraic formulation, Algorithm 1 describes the process of sending aircraft A through MRS gate G.

Algorithm 1 Process aircraft A through MRS gate G.

clearance-delivery $\leftarrow -\infty$
for all $R \in \mathrm{A} \cap \mathrm{G}$ **do**
  clearance-delivery $\leftarrow \max\{\text{clearance-delivery}, R.\text{next-avail}\}$
for all $R \in \mathrm{A} \cap \mathrm{G}$ **do**
  R.last-used $\leftarrow$ clearance-delivery
  for all (resource $\mapsto$ separation) $\in R$.spacing **do**
    resource.next-avail $\leftarrow \max\{\text{resource.next-avail}, \text{clearance-delivery} + \text{separation}\}$

One advantage of the MRS concept is ease of use. As shown in Algorithm 1, it is easy to determine when an aircraft can pass through a gate. Furthermore, when the spacing requirements change for resource R, it is easy to correctly update the times when resources consumed by R will next be available since R.last-used is known. Such an update of the spacing requirements is orthogonal to any update of aircraft states, e.g. due to re-routing when an aircraft is assigned a new set of resources. As shown in section 3.4, a wide variety of different traffic behaviors can be encoded via MRS. Despite this modelling power, when an MRS-based model is instantiated, the MRS dynamics can be compactly described using three matrices of coefficients: resource-aircraft pairings, resource-gate pairings, and resource-resource feedback. Hence the instantiation of MRS-based models with complex behavior does *not* require correspondingly complex simulation software

or optimization packages.

There is a diverse corpus of techniques which can be used to analyze MRS-based models. The specialized topic of timed inhibitory Petri nets may find application. More general results from analyses of idempotent $\{\max, +\}$ algebras, including a full theory of the corresponding linear algebra [BCOQ92], are relevant and have been previously applied to scheduling applications. MRS constraints can certainly be expressed as mixed-integer programming constraints. The ceno-scale model encodes a very wide variety of airport surface traffic dynamics using the MRS concept, and hence enables analyses using all of these previous research results.

3.3.2 Sequencing constraints

MRS gates permit two types of ATC control actions: changing the spacing requirements for a resource, and/or adding extra delay after the gate has cleared an aircraft. Finer-grained control is achieved by re-sequencing aircraft in the FIFO buffer which precedes the gate. In this section, two primary types of re-sequencing constraints are described.

Precedence constraints are standard in scheduling problems. Such constraints are equivalent to a poset or directed acyclic graph $(\mathrm{V}, \preceq_E)$ where V is the set of constrained aircraft and $i \prec_E j$ implies that $i \in \mathrm{V}$ must be processed before $j \in \mathrm{V}$. Note that $\preceq_E$ is transitively closed. For precedence constraints, there are well-known efficient algorithms for computing sequence feasibility, transitive closure and reduction, and minimum start-times. In airport surface traffic, precedence constraints arise when multiple taxiway buffers are merged together at an intersection [BS00c], or when a taxiway buffer is too narrow to permit aircraft to pass each other

The other type of re-sequencing constraint limits the *magnitude* of changes to an aircraft's relative rank in a sequence. This type of constraint has been analyzed with respect to airport arrival traffic by Dear and Sherif, who coined the phrase Constrained Position Shifting (CPS) [DS89, DS91]. Dear and Sherif considered symmetric CPS, where both the maximum and minimum change in sequence have the same upper bound. This is appropriate for airborne arrival traffic because each aircraft can only be delayed a finite amount. Trivizas later analyzed CPS[3] applied simultaneously to both airport arrival and departure traffic [Tri98].

In general, one must also consider asymmetric CPS. For example, departure traffic waiting at a runway to take off can be delayed indefinitely, while arrivals waiting to land on the same runway can only be delayed a finite amount. Even considering departure traffic in isolation, asymmetric CPS which allows very limited advancement in the takeoff queue and much larger delays can effectively capture ATC "fairness" constraints, for example when a small handful of restricted departures are set aside in penalty boxes while the remaining unrestricted departures are allowed to proceed normally.

[3]Trivizas used the phrase Maximum Position Shift. However, Dear's work appears to predate Trivizas', hence the acronym CPS is used herein.

It is important to note that CPS and precedence constraints are *not* equivalent. Consider the permutations of (1, 2, 3, 4) in which no element moves more than 3 backwards nor more than 1 forwards. The permutations which meet this constraint are shown in Table 3.1. Note that for any

(1, 2, 3, 4)	(1, 2, 4, 3)	(1, 3, 2, 4)	(1, 3, 4, 2)
(2, 1, 3, 4)	(2, 1, 4, 3)	(2, 3, 1, 4)	(2, 3, 4, 1)

Table 3.1: Counterexample showing that Constrained Position Shifting and partial-order constraints are *not* equivalent.

pair of elements (x, y), there are permutations in Table 3.1 where y occurs before x. Hence it can be inferred that this CPS constraint is not equivalent to any precedence constraint, except possibly a complete antichain. But a complete antichain on 4 elements would admit *any* permutation as feasible. This counterexample proves that the two types of constraints are distinct.

It is also important to note that scheduling under CPS has exponential computational complexity, just as scheduling under precedence constraints. In a sequence of N aircraft, let $CPS(N, f, b)$ count the number of sequences in which no aircraft has advanced more than f slots, nor been retarded more than b slots. For the highly-constrained case $f = b = 1$, an obvious induction argument shows that

$$CPS(N, 1, 1) = CPS(N - 1, 1, 1) + CPS(N - 2, 1, 1)$$

Hence the number of feasible sequences increases exponentially with N in the highly-constrained case. Relaxing the constraints only yields a larger number of feasible sequences to consider.

As with the MRS concept, by focusing on two simple classes of sequencing constraints, the cenoscale model allows a wide range of theoretical analyses to be applied. Precedence constraints are of course ubiquitous in scheduling problems. CPS has not received as much attention in the literature. As noted above, several authors in ATC research have used CPS to model fairness constraints and practical operational limitations on re-sequencing. The *displacement problem* of optimizing a sequence subject to a constraint on the disruption relative to another fixed sequence (typically the previously planned sequence) has also been considered [BKSA03]. There is a remarkable paucity of general theoretical research on CPS, particularly asymmetric CPS. It appears that even general enumeration algorithms are not known; [Leh70, DGH99] were the only references which could be unearthed[4], and both focus on the symmetric case. The analysis of algorithms for sequencing under both CPS and precedence constraints is an unappreciated open problem, one with direct bearing on ATC research.

[4]Many thanks to Peter Boothe for assisting in the search!

3.4 Natural dynamics of the runway queues

In this section, the ceno-scale building blocks are used to model a wide variety of the behaviors and delay mechanisms observed in airport surface traffic. The issues discussed include:

- multiple objectives which have been identified as important to ATC and the airlines;
- spacing constraints between runway operations imposed by downstream flow restrictions, arrival/departure/runway-crossing interactions, and procedural requirements on physical separation of aircraft;
- effects of ATC workload in the parallel queueing process; and
- re-sequencing constraints in the runway queue(s) and in merging multiple taxiway queues onto the runway.

There are also three effects which are known to be significant but are not explicitly encoded in the ceno-scale model. Two of these — time-windowing constraints and physical queue capacities — are difficult to observe in currently available operations data. Lacking direct observations, either these effects must be treated as additional sources of stochastic noise, and/or it must be verified that the ceno-scale model is applied under conditions where these effects do not significantly impact traffic. The third effect — the strategies used by ATC to obtain high-performance takeoff sequences — is treated in the next chapter by further extension of the ceno-scale model.

3.4.1 Terminology

In general, the formulation of traffic dynamics at a runway follows the lines of a standard resource-constrained project scheduling problem. A set of runway operations must be scheduled to use a runway system. Runway operations can include takeoffs, landings and crossings; for convenience these will be referred to simply as **r-op**'s. Similarly a runway system can include variations ranging from a single independent runway to a complex set of interacting parallel and intersecting runways; the key point is that when a r-op occurs on one runway in the system, this occurrence imposes constraints on the other r-op's in the same runway system.

Once the sequence of r-op's is established, a schedule is specified by giving the start-time of each r-op. This start-time corresponds to the start of runway "occupancy", which can have different physical interpretations for different types of operations. For example, the start-time might mark the actual start of a takeoff roll, the verbal delivery of a runway crossing clearance from ATC, or the point when a landing aircraft on final approach is 2nm from the runway threshold.

3.4.2 Objectives

Air traffic controllers dynamically balance a variety of objectives against each other. It is typically the case that no sequence/schedule is simultaneously optimal for all of the desired performance metrics. For example, ATC may try to ensure that the runway queues are never empty so that a high runway throughput can always be maintained, while environmental costs dictate that runway queueing should be minimized [Puj99]. Encoding all of the objectives into a single objective function for a mathematical program is not realistic. In real-world operations, certain objectives are evaluated dynamically and thus influence ATC decision-making on a tactical time-scale, whereas other objectives have been incorporated as constraints on standard ATC operating procedures. Conceptually such tradeoffs resemble classical mathematical programming formulations where constraints may be encoded as violation-penalties in the objective function, or minimum performance in a given objective function may be encoded as a hard constraint. However, human air traffic controllers often have the flexibility to dynamically alter this tradeoff. As much as possible, ceno-scale modelling is designed to reflect the typical tradeoff(s) as well as this real-time flexibility.

Maximize the runway throughput

Interpretations of throughput vary among airports and between air traffic controllers. Variations include maximum r-op's per hour, minimum latest start-time for departures currently in the queue, maximum runway utilization, etc. In mathematical programming formulations of runway sequencing problems, minimizing the latest start-time is one of the most popular objectives. However there are certainly cases in which throughput is throttled back from its maximum rate: Idris notes that controllers at BOS frequently select departure sequences which do not *maximize* throughput, but only maintain it at a level sufficient to match departure demand [Idr00, p. 107]. The chosen sequences are instead optimized to avoid overloading air traffic controllers in the downstream airspace with excessive workload. In general, ceno-scale modelling is amenable to any one of these objectives.

Balance the throughput

The runway throughput of arrivals, departures and crossings must be matched to the demand for each type of r-op. This is an issue of providing "fair" service. Also it is important to avoid congestion and deadlock due to overloaded buffers, both in airborne holding stacks and in taxiway queues on the ground. For example, when departure queues become sufficiently long at BOS, ATC may switch to a modified runway-usage procedure called "Accelerated Departure Process" in which arrival traffic is delayed for a short period in the air and departure queues are rapidly flushed from the system. In general the balancing of throughput is sufficiently important to be treated as a structural property of airport surface traffic, although at airports such as Atlanta-Hartsfield International Airport, Dallas-Forth Worth International Airport and Memphis International Airport which have

two semi-independent runway systems, DSTs which dynamically balance the split of departure traffic between the runway systems have shown some benefits [ABW02, AW02]. This type of fairness constraint can be encoded in the ceno-scale model as a CPS constraint on both arrival and departure r-op's at the runway, to restrict undue deviations from the "natural" order of r-op occurrence.

Minimize delays

As with "throughput", there are also multiple interpretations of "delay". In addition, the penalty incurred due to delays varies depending on the situation, and there can be sharp thresholds in the penalty function. For example, the environmental and direct operating costs of taking delays at the gate are lower than when delays are taken while taxiing, *unless* an arriving aircraft also needs to occupy that gate. Landing aircraft cannot stop in mid-air and only carry a limited (typically 30min) supply of extra fuel, so anticipated delays above 30min require an expensive diversion to a backup airport. If a departure is assigned a time-window or *slot* for takeoff but is delayed, finding a new slot can require lengthy negotiations involving local ATC, downstream ATC, the pilot and the airline operations center. In the ceno-scale model, delay minimization is treated implicitly by attempting to constrain the throughput balance and thus avoid excessive delays for any given class of r-op's. Time-windowing constraints on the takeoff times for specific flights are some of the most difficult constraints for human air traffic controllers to meet; the ceno-scale model explicitly does *not* attempt to capture the complex strategies which human air traffic controllers may implement.

Maintain fairness

The presence of fairness is typically indicated by a lack of pilot complaints about unfair ATC practices; this is difficult to capture analytically. In addition, some "tit-for-tat" behavior is considered fair, i.e. strict fairness may be maintained on average between different airlines but violated on a per-aircraft basis. Again this type of constraint is expressed in the ceno-scale model as CPS on r-op's in the runway queues.

Enable airline objectives

Airlines have a variety of internal objectives. Currently ATC is responsive to a subset of these, primarily respecting the very high value of international flights, and the need for certain flights to depart in a timely manner to make curfews at downstream airports [EC02]. However, incorporating additional airline scheduling objectives into the control of airport surface traffic is one of the most challenging problems in DST design. Airline objectives can vary among airlines and over time according to corporate policy; a classic example is the shift in Continental Airlines' focus towards on-time departures at the impetus of CEO Gordon Bethune [Bet99]. Like the ATC objectives listed above, airline objectives are not necessarily conflict-free. On a tactical hour-by-hour timescale, airline

stations are primarily concerned with maintaining their schedule [DH02]. When the schedule breaks down due to weather, mechanical problems, etc., the secondary tactical objective is to minimize lost revenue by flying the "most valuable" operations. Measurement of flight "value" involves a number of factors including the number of passenger and crew connections, the required ground-time for those connections, the required ground-time to prep the aircraft for its next flight, and priority rankings among flights based on anticipated revenue and delay propagation through the remainder of the schedule [DH02]. The primary mechanisms by which DSTs could incorporate airline objectives is to selectively weight the different r-op's operated by the same airline when re-sequencing decisions are made. Ceno-scale models capture the constraints on re-sequencing and hence support this type of re-sequencing optimization.

3.4.3 Spacing constraints

Local interactions among arrivals and departures

The simplest type of spacing constraints are procedural restrictions on the time or distance between successive r-op's [FAA02]. The full set of rules is complex, and considers factors such as the physical performance, weight, wake-vortices, planned and escape trajectories, etc. of the two successive r-op's. These are minimum separation criteria; the air traffic controller working the Local position typically adds extra padding as an uncertainty buffer, and in some cases there are established procedures for ATC to reduce spacing below minima for short periods of time or for pilots to voluntarily waive certain minima. However in principle, the minimum fixed spacing (procedural or practical) between two successive r-op's can be found by table-lookup. Bolender and Slater consider single-resource formulations of runway sequencing subject to such pairwise constraints [BS00a, BS00c], together with precedence constraints on merging parallel queues of aircraft. They schedule to minimize the latest takeoff time via heuristic or genetic algorithms, although typical departure queues at major US airports are on the order of 20 to 50 aircraft and could easily be solved using current algorithms for directed traveling-saleman problems (TSP). Pairwise spacing constraints have already been shown to be easily expressed using MRS.

Downstream interactions between departures

Despite its appealing resemblance to a classical problem like TSP, single-resource formulations are inadequate for runway operations in general. In addition to local procedural constraints, ATC at the airport must also cope with downstream spacing restrictions. This is particularly the case for Miles-in-Trail (MIT) and Minutes-In-Trail (MINIT) spacing restrictions imposed by downstream ATC to reduce flow rates, because these spacing constraints are too large to meet solely by controlling airborne traffic. Note that these constraints are not specified at the runway but at a downstream

navigational fix. There is an important consequence of this distinction: because MIT/MINIT restrictions act pairwise between successive aircraft at the downstream fix, they must be imposed at the runway between aircraft traveling through that fix, *which may be separated by multiple intervening r-op's*. For example, the normal MIT spacing is 5nm, but downstream ATC may impose larger spacings of 10, 20 or 40nm. Suppose that a 40MIT restriction is active at the nearby departure fix EXMPL. If a leading aircraft takes off towards EXMPL, then there should be a delay of approximately 15min (equal to 40nm at a speed of approximately 240kt at the nearby departure fix) until the next departure towards EXMPL. This is sufficient time for at least two or three other r-op's[5]. However, the TSP-like formulation only allows constraints between aircraft adjacent in the sequence, in which case the "optimal" solution may only insert a single r-op between successive takeoffs towards EXMPL. The TSP formulation results in either substantial wasted time and/or the need for controllers to manually patch or ignore the generated sequences.

Beasley et al make this distinction between *successive* and *complete* separation constraints in their work [BKSA00, BKSA03]. Note that the use of complete separation constraints breaks the isomorphism to directed TSP. However, a MRS gate correctly represents complete separation. To encode downstream restrictions, as well as different separations for diverging versus non-diverging successive aircraft, each fix and/or group of fixes is treated as a resource in the gate. Distance-based constraints such as MIT restrictions are converted to time-based constraints in this representation. This conversion is standard (human air traffic controllers use several standard rules-of-thumb) and can be computed by considering the speed at which an aircraft is expected to be traveling when it reaches the downstream fix where the restriction is applied.

Simultaneous r-op's

Runway crossings can become a significant source of controller workload and taxi delays [Idr00, p. 103–10]. To some extent runway crossings form a separate traffic stream which can be modeled using pairwise separation minima identical to the local interactions among arrivals and departures. A shortcoming of this approach is that runway crossings can be bunched together to save time. There is an acceleration-delay penalty for the first crossing aircraft as it accelerates from a standstill to its taxi speed. If several crossing aircraft accelerate simultaneously, then the second and later crossings do not incur the same acceleration-delay penalty as the first. This effect is significant — the acceleration-delay penalty observed at BOS is approximately 1.7 times larger than the time required for each additional crossing after the first [Idr00, p. 110] — and controllers explicitly take advantage of this fact. Note that the acceleration-delay penalty cannot be modeled using a single-resource formulation since statically-computed pairwise separation minima on a single runway resource do not reflect the state history. In addition, the same runway may be intersected along its length

[5]These aircraft inserted into the departure sequence are termed "splitters" for obvious reasons by ATC in BOS [Idr00, p. 201–2].

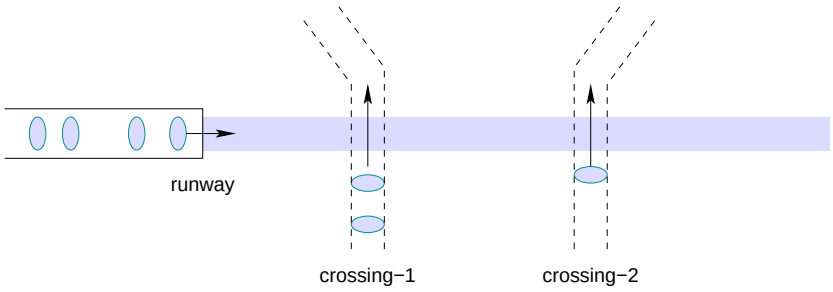

Figure 3-7: If only a single resource is used to represent each runway, it is impossible to model common situations such as a runway which is intersected by several parallel runway crossing queues.

at several points by different taxiways. In this case controllers not only bunch together runway crossings at a single intersection but also use all of the intersections simultaneously.

Encoding runway crossings in a ceno-scale model provides a good example of the modelling power of MRS. Each location where the runway is intersected by runway crossing queues is represented as a resource, e.g. crossing-1 and crossing-2 as shown in Figure 3-7. These resources consume themselves, and this self-feedback is used to enforce separation between successive runway crossings in the same queue. This separation is taken as the *marginal* separation between successive runway crossings, i.e. the acceleration-delay penalty discussed previously is not included. To model the acceleration-delay penalty, the runway resource is modified to consume the crossing resources for a time equal to the acceleration-delay penalty.

This modification causes a minor discrepancy between the model and reality as shown in Figure 3-8. In the model, the first runway crossing is delayed for a period of time equal to the acceleration-delay penalty, while in reality the first runway crossing begins immediately after the preceding landing/takeoff. This discrepancy is acceptable, because it is easy to translate between the two domains, and in both domains the first runway crossing occupies the correct amount of time.

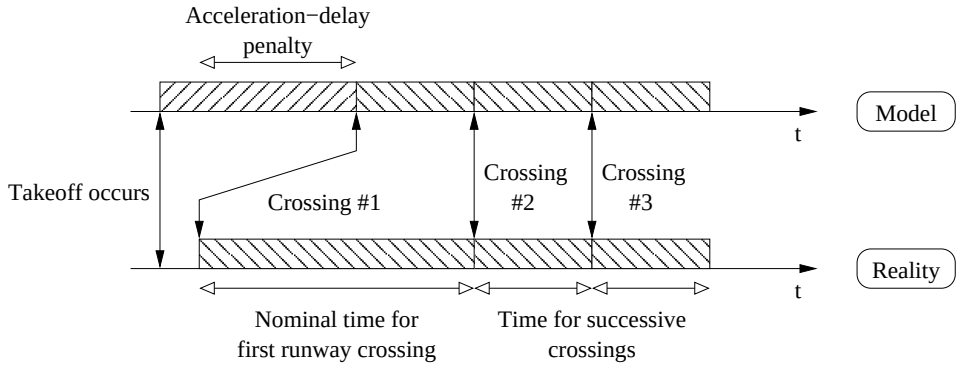

Figure 3-8: The MRS-based model does not *perfectly* represent the acceleration-delay penalty observed with "bunched" runway crossings, but the discrepancy does not affect the traffic dynamics.

Runway crossings in separate queues do not affect each other, but as each runway crossing occurs, the demand for crossing-1 or crossing-2 consumes the runway resource demanded by landings and departures. Thus with three resources, a wide variety of separation constraints between runway crossings, landings and takeoffs have been encoded using exactly the same framework which encodes wake-vortex separations, downstream flow restrictions, arrival-departure interactions, etc.

3.4.4 Causes of sequencing constraints

There are a number of causes for sequencing constraints in airport surface traffic. It may be the case that along a given segment of the taxiways it requires an extraordinary effort to re-sequence aircraft already queued on that segment. Usually this occurs when there is insufficient space for one aircraft to pass another along the given segment and there are no standard alternative routes going forwards toward the runway, so that any re-sequencing involves switching a departing aircraft nearer the head of the queue over to an arrival taxi-path and sending it back towards the gates. Note that such retrograde situations can certainly occur, for example when a mechanical problem is discovered during final preflight checks and the afflicted aircraft cannot take off but must taxi back to the gates for repair. However in normal operations both controllers and pilots avoid re-sequencing along constrained taxiway segments, and this imposes natural precedence constraints among aircraft queued on those segments. As mentioned previously, traffic may also be subject to fairness constraints. A typical fairness rule is that departures must be serviced by ATC in nearly first-come first-served order. Obvious exceptions are allowed, for example when the aircraft at the head of the queue has a long wait due to downstream restrictions and several other aircraft can be allowed to depart first without delaying the delayed aircraft.

3.4.5 Constraints with a partial representation

Time-windowing

In addition to sequencing and spacing constraints, departure traffic in the runway queues may also be subject to time-window scheduling constraints where the takeoff of a constrained flight must occur within a specified window of opportunity. Examples include the Departure Sequencing Program (DSP); Approval Request (APREQ) procedures; and Expected Departure Clearance Times (EDCTs). DSP is currently used to help merge departure traffic from the 4 major metro New York airports into enroute flows; APREQs are an older version of DSP still in use nationwide [DH01]. EDCTs delay the pushback of departures which are held at the gate by Ground Delay Programs. Time-window constraints may be difficult to encode in a ceno-scale model because the corresponding ATC strategies can be difficult to identify in practice. In many cases, however, time-window constraints also have a smaller aggregate effect on airport surface traffic than downstream restrictions.

DSP and APREQ constraints are particularly difficult challenges for ATC. The width of the time-windows (only 3min for DSPs [DH01, p. 18]) are of the same magnitude as the minimum takeoff spacing between r-op's, so that the constrained departure misses its time-window if it is incorrectly sequenced forwards or backwards more than a couple slots. For this reason, a robust ATC strategy is to form a separate FIFO buffer of departure aircraft which are subject to time-window constraints. This additional queue must be physically located very near the runway so that each upcoming time-windowed departure is always immediately available. The need to maintain an extra physical queue can severely limit the use of this strategy at older space-constrained airports. In that case, ATC may need to use much more complex sequencing and routing strategies, and these strategies can be difficult to infer from available operations data.

The severity of time-window constraints is often smaller than the severity of downstream restrictions, where severity is measured in terms of the overall impact on airport throughput and taxi-out delay. In [Idr00, chap. 3] it was observed that flow-restrictions caused by local weather at the airport's departure fixes tended to produce the greatest deviations from average throughput and thus the greatest delay. Interviews with BOS air traffic controllers in the tower and TRACON support this observation:

- Non-local restrictions impact a smaller fraction of traffic than local restrictions. In particular, EDCT restrictions for departures to a given downstream airport may cause flights to be cancelled outright, thus achieving the maximum impact on that particular flight, but relatively few departures are destined for any particular airport.

- DSP and APREQ constraints are often imposed to merge a departure into an available traffic slot either enroute or at a destination airport, and thus also affect a smaller fraction of the total departure traffic.

From a modelling perspective, when ATC can develop a secondary runway queue of time-windowed departures, this strategy is easy to encode in a ceno-scale model instance and the time-window constraints typically have little impact on aggregate throughput and delays. But when this strategy is infeasible, the corresponding strategies are more difficult to observe and encode, and the time-window constraints may have a larger overall impact. Ceno-scale modelling may not be a correct approach in such cases.

Physical buffer capacities

The physical capacities of taxiway segments form another constraint on airport surface traffic. Lack of sufficient buffering space can severely constrain airport surface traffic. For example, during interviews with airline station managers at BOS, it was commonly noted that traffic bottlenecks due to overloaded buffering areas near the gates can cause delays. Capacity constraints can also

become a significant problem for runway crossings, especially when aircraft must cross closely-spaced parallel runways with only enough room for a queue of two to four crossings between the runways. Rather than push crossings into every available gap in the departure/arrival sequences, a robust ATC strategy is to minimize the acceleration-delay penalty by bunching several crossings together and widening the natural gaps in the departure/arrival sequences to make room. Examining the research literature, management of runway crossings is apparently a much less glamorous problem than scheduling takeoffs and landings, although recent work is changing that perception [Ana03, AC02, AC03, ACBV01].

The primary reason capacity constraints are not encoded in the ceno-scale model is that, although in practice some robust ATC strategies have been observed for coping with these constraints, these strategies are highly sensitive to the actual trajectories of individual aircraft across the airport surface. It is reasonable to lump together similar routing decisions and random effects into an unimpeded taxi-time if all aircraft must eventually traverse the same runway queue; a DST design which is based on such an approximation will be robust to minor modeling errors. However, an accurate encoding of strategies with respect to capacity constraints would require accurate aircraft trajectories on the airport surface. Calibration/validation data is not currently available to create models with the necessary fidelity.

As an aside, note that while deterministic departure sequencing is a popular but impractical problem in the literature, there are other unrecognized and practical opportunities for deterministic optimization to improve departure traffic flows. Under conditions of extreme downstream restrictions, ATC and the airlines can convert much of the airfield into parking space for delayed flights. The efficient *packing* of the available capacity, with aircraft parked in parallel sequences that allow rapid deployment after downstream restrictions are lifted, has not been considered in the literature to date.

3.5 Conclusions

The goal of ceno-scale modelling is to aid in the development of robust DSTs for airport surface traffic. Such DSTs require traffic models which accurately represent a wide variety of the delay mechanisms which are observable in available operations data, and have a flexible structure which is well-adapted to human air traffic controller mental models and to variations between different airports. The ceno-scale framework laid out in this chapter embodies several significant advancements over existing mesoscopic models.

Building on previous research in mesoscopic modelling, the addition of MRS gates and two complementary types of sequencing constraints only adds a modest amount of complexity to the basic mesoscopic modelling components. The increase in modelling power fully outweighs the additional

complexity. Note that both of these new model components are designed to reflect actual observable constraints on airport surface traffic, and hence can be calibrated using available operations data and easily adapted to different airports. A wide variety of traffic dynamics can be encoded using these ceno-scale components, including Idris' novel parallel queueing mechanism and several types of fairness constraints. Several models of runway sequencing which have been used in previous research can be unified in a single ceno-scale model which encodes all of the constraints due to arrival/departure interactions, runway crossings at multiple intersections along the runway, pairwise separation minima and downstream flow restrictions.

There are numerous opportunities for further research using ceno-scale models. Theoretical problems include analysis of the behavior of a FIFO buffer controlled by a MRS gate, and the development of algorithms for sequencing through such a gate under both precedence constraints and CPS. Practical extensions of the ceno-scale modelling framework in this chapter include encoding Land And Hold Short operations, interactions in multi-runway systems, and identification and modelling of more complex ATC strategies. It is of particular interest to analyze how ATC deals with more complex problems such as time-window and queue capacity constraints which are not currently encoded in the basic ceno-scale framework. Some of these open problems are treated in the next chapter, which includes a ceno-scale model for EWR which is calibrated and validated against available operations data, and several types of corresponding analyses.

Chapter 4

Robust modelling: EWR, June 29, 2000

A ceno-scale model instance has been developed and experimentally validated in collaboration with a case-study conducted at EWR on June 29, 2000 [CECF02, EC02]. On the day in question, the spacing restrictions imposed at the EWR runways were severe and subject to frequent change, and large departure queues developed. Ceno-scale models are designed to capture the resulting effects on surface traffic, and hence the EWR case-study provides a good test case. In this section, after the model instance is developed and calibrated, it is used in Monte Carlo simulations. These simulations use the actual pushbacks and downstream restrictions observed on June 29 as inputs. The taxi-out times between gates and runways are modeled using stochastic distributions; in addition, the effects of additional stochasticity in pushback times is modeled. From these simulations, the expected runway queues and takeoff times for each departing flight are determined. The model-based results are then compared against the actual operations to test the parameter sensitivity and fidelity of the ceno-scale model.

4.1 Data sources and field observations

As part of the EWR case-study [EC02], an unusually complete description is available of the traffic operations on June 29, 2000 at EWR and in the surrounding airspace. The FAA's CATER system at EWR recorded both the airport operating conditions and timing data for individual departures. The CATER data on airport operating conditions includes local weather and the runway configurations which were in use in response to the weather. For each individual departure, the CATER timing data includes the time when pushback clearance was requested and the time when takeoff occurred. Logbooks from the EWR ATC tower recorded the downstream restrictions which were in effect.

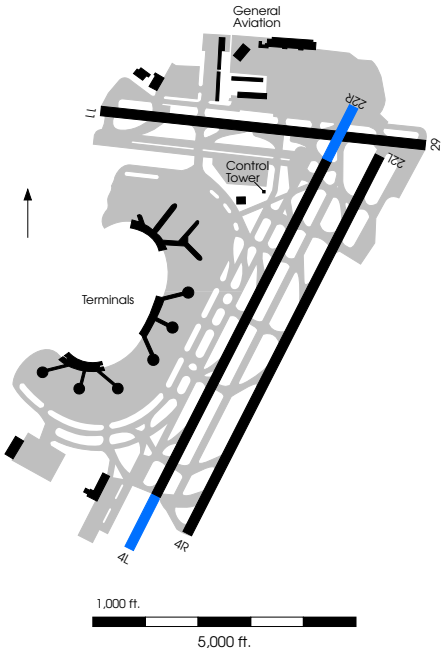

Figure 4-1: EWR has a simple runway layout — typically only a single runway queue develops, and runway crossings are not a significant source of delay [CEFC02].

Flight strips were photocopied by hand and provide a secondary source of information on departure operations. Data from the Airline Service Quality Performance (ASQP) database is also available for June 2000; it provides a validation of the CATER and flight strip data for departures, as well as flight data on arrival traffic. When combined together, the CATER data, logbooks, flight strips and ASQP data yield a snapshot of one day's worth of airport operations with much higher detail than is typically available for research analyses.

As shown in Figure 4-1, EWR has a primary pair of dependent 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, runway 11/29 is used only infrequently [EC02]. Analysis of the CATER data shows that less than 4% of the recorded traffic on June 29 used runway 11/29. Under typical conditions at EWR, arrivals are shunted to the outer runway of the pair, while departures are taken off the inner runway. This traffic pattern is caused by limitations on taxiway space, both for departure queues between the runways, and for arrival queues between the inner runway and the terminal buildings. Due to the simple runway configurations and taxi-paths, only a single runway queue develops in each of the two possible runway configurations. This permits the use of a simple traffic network in the ceno-scale model. In particular, it is not necessary to infer the ATC strategy for routing traffic across the airport surface. Because departures use the inner runway and arrivals only need to cross the far end of that runway, there are also no significant runway crossing dynamics.

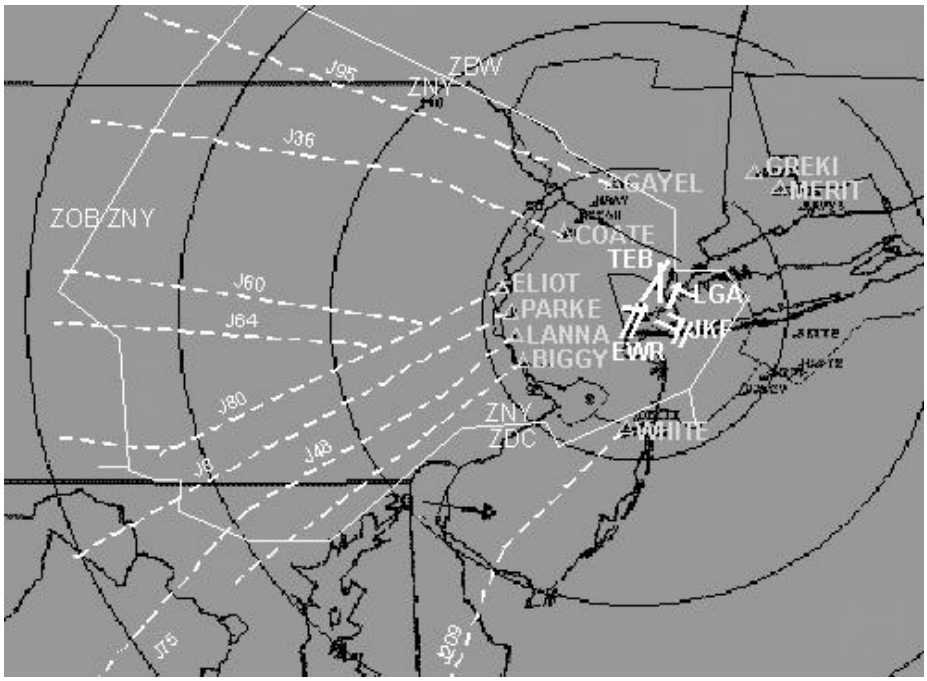

Figure 4-2: Delays at EWR are frequently caused by restrictions in the neighboring airspace [EC02].

After takeoff, U.S. domestic traffic leaving EWR is first sent through one of three groups of navigational fixes — North, South and West — as shown in Figure 4-2. There is also a large volume of international departure traffic sent through the East and South fixes. During the day in question, many of the departure fixes were impacted by downstream restrictions. The impacted North fixes included GAYEL and COATE, which serve destinations in the Northwest U.S. such as Minneapolis-St. Paul International Airport, Chicago O'Hare International Airport, Seattle-Tacoma International Airport and San Francisco International Airport. The impacted West fixes including ELIOT, PARKE, LANNA and BIGGY serve from the Midwest U.S. through the South U.S., ranging from Denver International Airport and Los Angeles International Airport down to Charlotte/Douglas International Airport and Tampa International Airport. The impacted South fix WHITE serves eastern Florida down through Central and South America. Finally the impacted North fixes MERIT and GREKI serve the Northeastern U.S. and Europe. A description of the downstream restrictions relevant to the ceno-scale model is given in Appendix C. Note that the downstream restrictions given in the Appendix are often above 20MIT; subject to frequent changes; and often leave only a few routes open to departing traffic leaving EWR. The ceno-scale model is designed to capture the resulting traffic effects with much higher fidelity than previous models.

Even when a departure fix was subject to the most severe 40MIT downstream restrictions, it was observed that impacted departing flights were still able to leave the airport. Human controllers in the FAA tower and TRACON had sufficient control authority to meet these downstream restrictions

at the runway by re-sequencing of the departure queue, assignment of extra delays to the departure at the head of the queue, and vectoring flights in the departure airspace. However, there were also periods on June 29, 2000, when some departure fixes were completely closed. When a downstream fix was completely closed, the affected departures were subject to a modified set of rules [CECF02]:

1. Affected flights which had not yet received clearance to taxi were held at the gate until the fix closure was lifted.

2. Affected flights which were taxiing out to the runway were pulled into penalty boxes or into temporary parking areas on the airport surface. Affected aircraft could not make progress towards the runway, but also did not lose ground by being sent back to the gates.

3. Affected flights in the runway queue were similarly pulled out of the queue and held in staging areas. However, when the fix closure was lifted, the affected aircraft joined the end of the runway queue; their previous position in the queue was not reserved or saved.

An interesting open question concerns the threshold severity of downstream restrictions which *effectively* closes a departure fix. It has been observed that severe spacing restrictions can combine with high traffic levels, overloading the flexibility to meet such restrictions via re-sequencing before takeoff or vectoring after takeoff [EC02]. There must be some threshold between the minimal 3MIT spacing required of all departure aircraft, and the effectively unbounded spacing of a closed fix. It is possible that very large fix-spacing requirements above the 40MIT sometimes encountered are not observed because they effectively close the fix under normal departure traffic mixes. It may also be the case that such large spacing requirements are workload-intensive to impose at the runway, and hence their absence from the operational record primarily represents a human-factors limitation.

4.2 Model dynamics

Based on the available operations data and the observed traffic dynamics, the ceno-scale model instance illustrated in Figure 4-3 has been developed. The model is intended to accurately represent all of the observed traffic dynamics except for physical capacity limitations.

Departing aircraft enter the system when the pilot requests clearance to taxi. These inputs correspond to the "Dep" blocks at the top of the figure. If the departure fix of a departing flight is not closed, clearance to taxi is assumed to be delivered immediately; immediate clearance delivery has been observed previously in [Car01, Puj99]. After receiving taxi clearance, each aircraft undergoes a stochastically independent unimpeded taxi-out time. As shown in Figure 4-1, there are three main terminals A, B and C at EWR. The departure traffic is segregated by departure terminal so that the probability distribution for unimpeded taxi-out time can vary depending on the travel distance from each terminal to the runway queue. To avoid overfitting the June 29 data, the probability

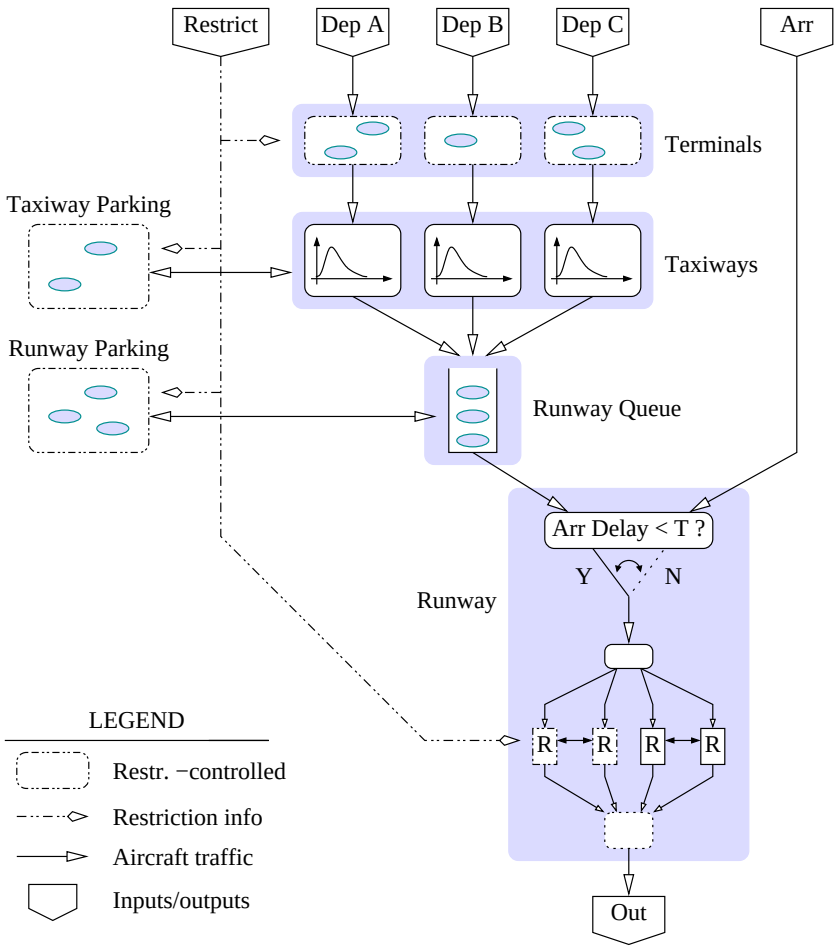

Figure 4-3: Ceno-scale model of EWR arrival/departure traffic.

distributions for unimpeded taxi-out time are derived from May and July ASQP data using the techniques developed in [Car01] for mesoscopic modelling. Factors such as communications delay with the tower, lack of proper weight-and-balance numbers, interactions with other taxiing aircraft, etc. are not observed in the available operations data and therefore are lumped into the stochastic unimpeded taxi-out time. Once a departing flight completes its stochastic unimpeded taxi-out time, it enters a FIFO buffer which represents the departure queue at the runway. After reaching the head of the queue, departures are subject to the separation criteria given in Appendix C. The MRS gate controlling the runway ensures complete separation in the technical sense of [BKSA00, BKSA03]. The start of a departure r-op is modelled as the time when the departure begins its takeoff roll. Departures leave the system after takeoff.

An arriving aircraft enters the system when it is 2nm from the runway threshold. According to [FAA02], after an arrival passes inside the 2nm mark, no departure can start its takeoff roll until the arriving aircraft has landed and sufficient separation between the arrival and the following

departure can be ensured. In contrast, if the departure begins its takeoff roll just before the arrival passes inside the 2nm mark, then the departure can accelerate away and off the runway and thus ensure separation with the following arrival. This asymmetry in operations explains the large asymmetry in Appendix C between successive arrivals and departures depending on their relative ordering. Because all arrivals use the same runway, arrival traffic is not segregated into multiple streams like the departures.

Due to the restriction that arrivals cannot be significantly delayed, the sequencing of r-op events at the runway requires additional consideration in the Monte Carlo simulation. A standard event-based simulation is insufficient. Instead, a "lookahead" step has been implemented to ensure that departures do not delay arrivals more than T seconds. Rather than a single global event-queue, three separate event-queues — arrivals, departures, and changes to the downstream fix restrictions — are maintained. These event queues are iterated through asynchronously. At each step the earliest event from each queue is attempted. If the earliest event is a departure, and it is determined that performing the departure would delay the earliest arrival by T seconds or more, then the arrival is performed instead.

As shown in Figure 4-3, there is a separate set of dynamics for departing aircraft which are blocked by downstream restrictions. If a departure cannot take off from the airport, its progress is temporarily suspended. Departures still at the gate are not allowed to push back while their departure fix is closed. Departures in the unimpeded taxi-out phase are stopped while their departure fix is closed, i.e. the duration of the fix-closure is added to their unimpeded taxi-out time. Departures in the runway queue are removed from the queue when their departure fix closes so that unblocked departures can still make progress; the removed departures are added back to the end of the queue when their departure fix re-opens. Note that a closed departure fix is conceptually similar to an infinite spacing requirement between aircraft which need to use that fix. Therefore it is possible to encode these modified rules using MRS and some extra sequencing logic. However, because this ceno-scale model instance is evaluated using Monte Carlo simulation, it has been decided to simply implement these modified rules directly in the software as a special-case shortcut.

4.3 Analyses

This ceno-scale model instance has been implemented in a Monte Carlo simulation. The actual pushback demand from the flight strip data, and the actual arrival demand from ASQP, are used to drive the simulation as exogenous inputs. Fix assignments for each departing flight are taken from the flight-strip data. The downstream restrictions affecting EWR were obtained from TRACON logs from 11:00 to 19:00 EDT on the day in question, and these simulation results focus on that period. To approximate any residual queues from the lightly-restricted morning traffic, the simulation is

started two hours before the 11:00–19:00 period. The simulation also runs for two hours after the 11:00–19:00 period to remove any additional edge effects. Unless it is indicated otherwise, each plot and statistic reported below is based on 256 independent simulation runs. Convergence appeared reasonably fast and variance-reduction techniques were not employed.

4.3.1 Sensitivity to numerical parameters

The probability distributions of taxi-out time were calibrated from operations data which did not overlap the data from the case-study. It was found that a shifted lognormal distribution was a reasonable fit to the calibration data. The shift was 3.5min, and the unshifted lognormal distribution corresponded to a Gaussian distribution with mean $\mu = 1.7$ and standard deviation $\sigma = 0.40$. To test the robustness of the model with respect to these fitted parameters, a set of simulations were run in which the parameters were varied around the fitted values. All 36 combinations of

$$(\mu, \sigma, \text{shift}) \in \{1.6, 1.7, 1.8\} \times \{0.38, 0.40, 0.42\} \times \{3.0, 3.5, 4.0, 4.5\}$$

were used as test-cases, with each test-case being comprised of 100 independent simulation runs. Figures 4-4 through 4-6 summarize the results. The dark line in each plot is the mean queue-length as a function of time for the nominal parameters. The grey band in each plot shows the mean queue-length plus/minus one standard deviation as obtained from simulating with the perturbed parameters given at the top of each plot. Most of the plots appear similar, indicating that the model is reasonably robust to minor errors in the numerical parameters.

4.3.2 Model fidelity

The data shown in Figure 4-7 are drawn from the simulation test-case with nominal fitted parameters. Every departure is plotted as a single point, where the x-coordinate gives the time that departure was ready to leave the gate, and the y-coordinate gives the difference between actual takeoff time and the mean simulated takeoff time. The vertical axes are bounded at the 2.5% and 97.5% percentiles of this difference. Note that many more departures have a positive difference, indicating that the time at which the aircraft actually took off was later than predicted according to the simulation.

An examination of the input data shows that many of these high-difference departures were headed to airports with Approval Request (APREQ) restrictions. This shows the impact of such time-windowing takeoff constraints on ATC performance. The ceno-scale model, without any sort of sequencing logic beyond that required to minimize arrival delays, performs comparably with the observed operations from 11:00 to approximately 14:00. However, when APREQs begin to impact certain airports, real-world ATC performance declines and the affected flights accrue delays much greater than predicted by downstream spacing constraints.

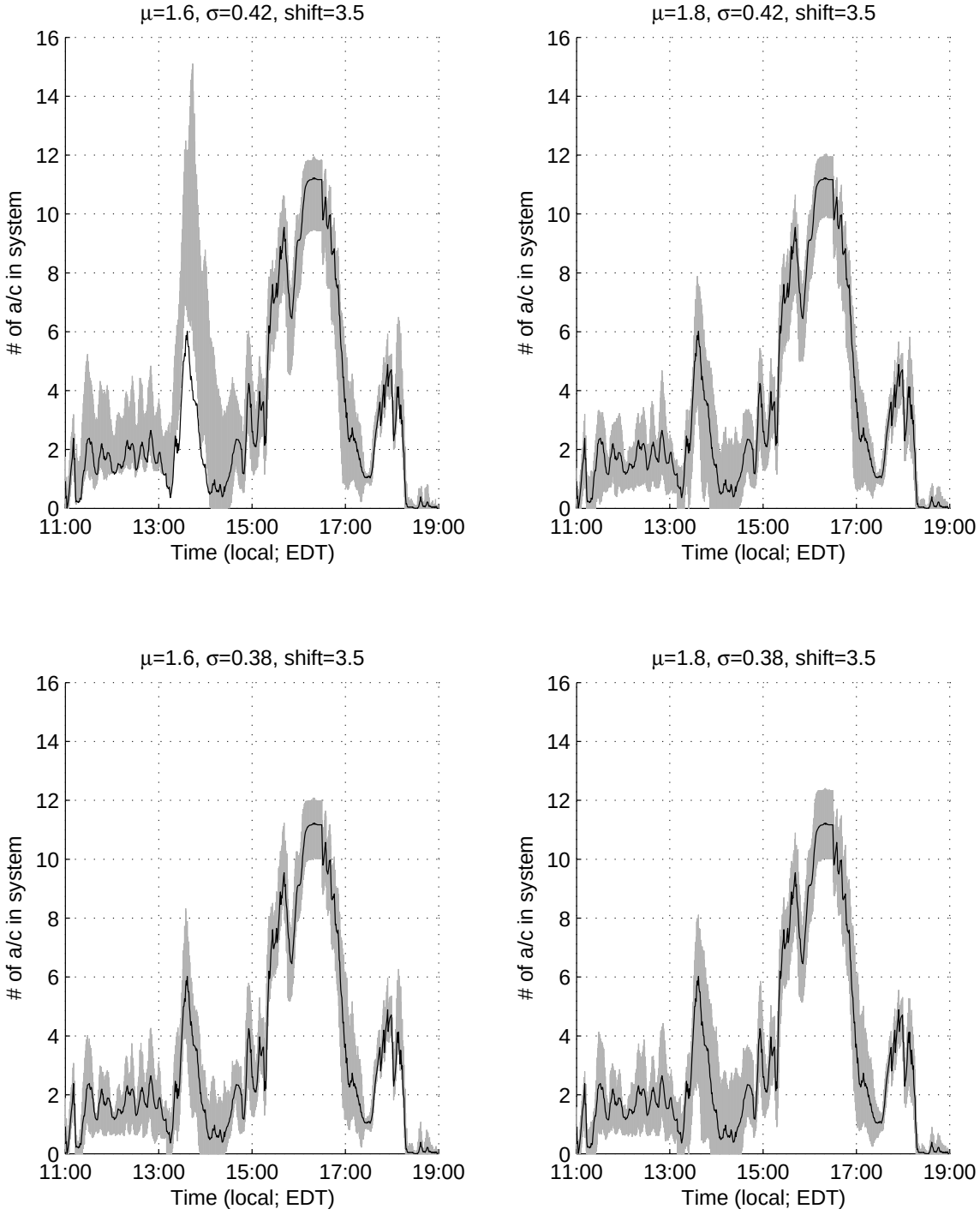

Figure 4-4: Variability due to μ and σ parameters.

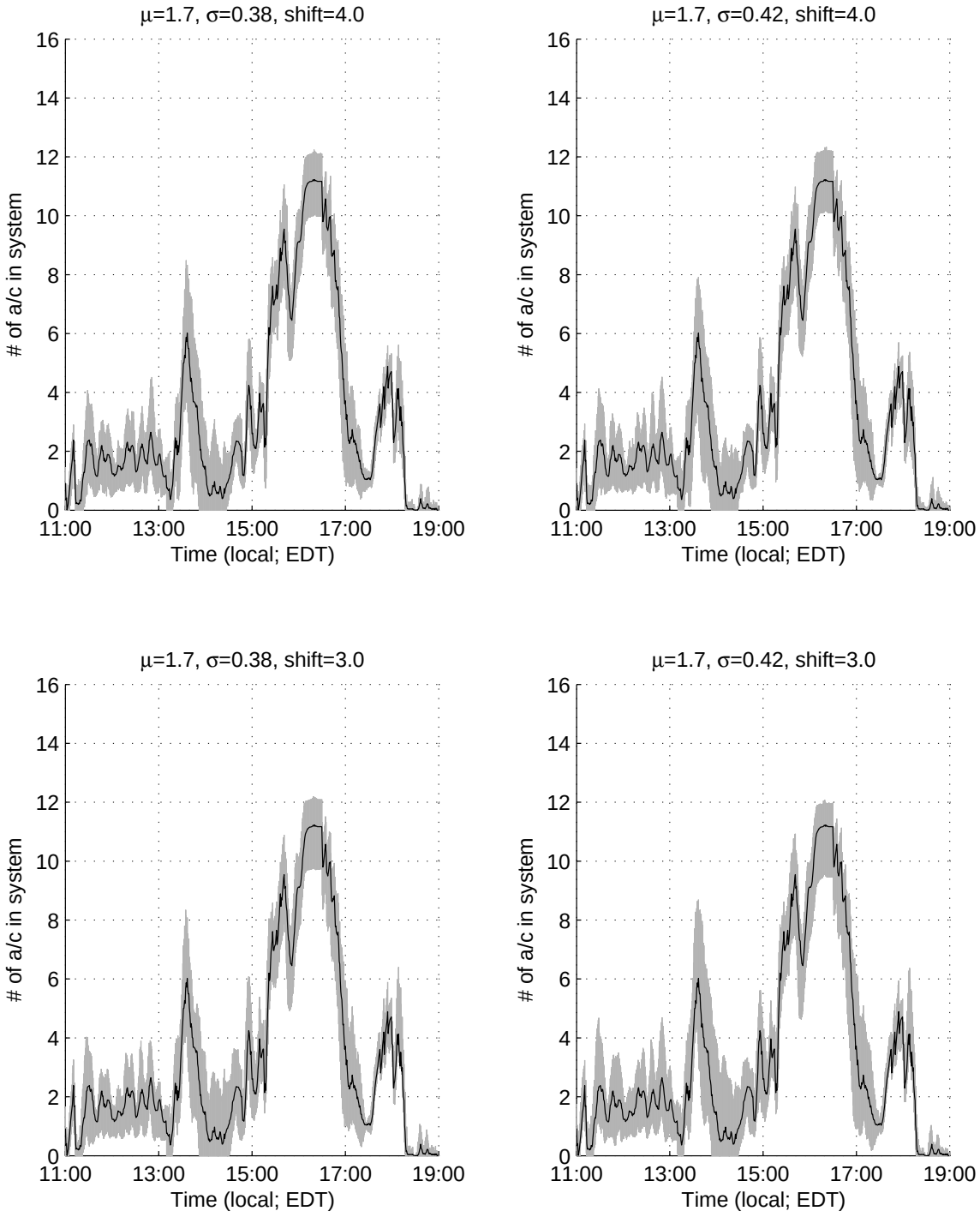

Figure 4-5: Variability due to σ and *shift* parameters.

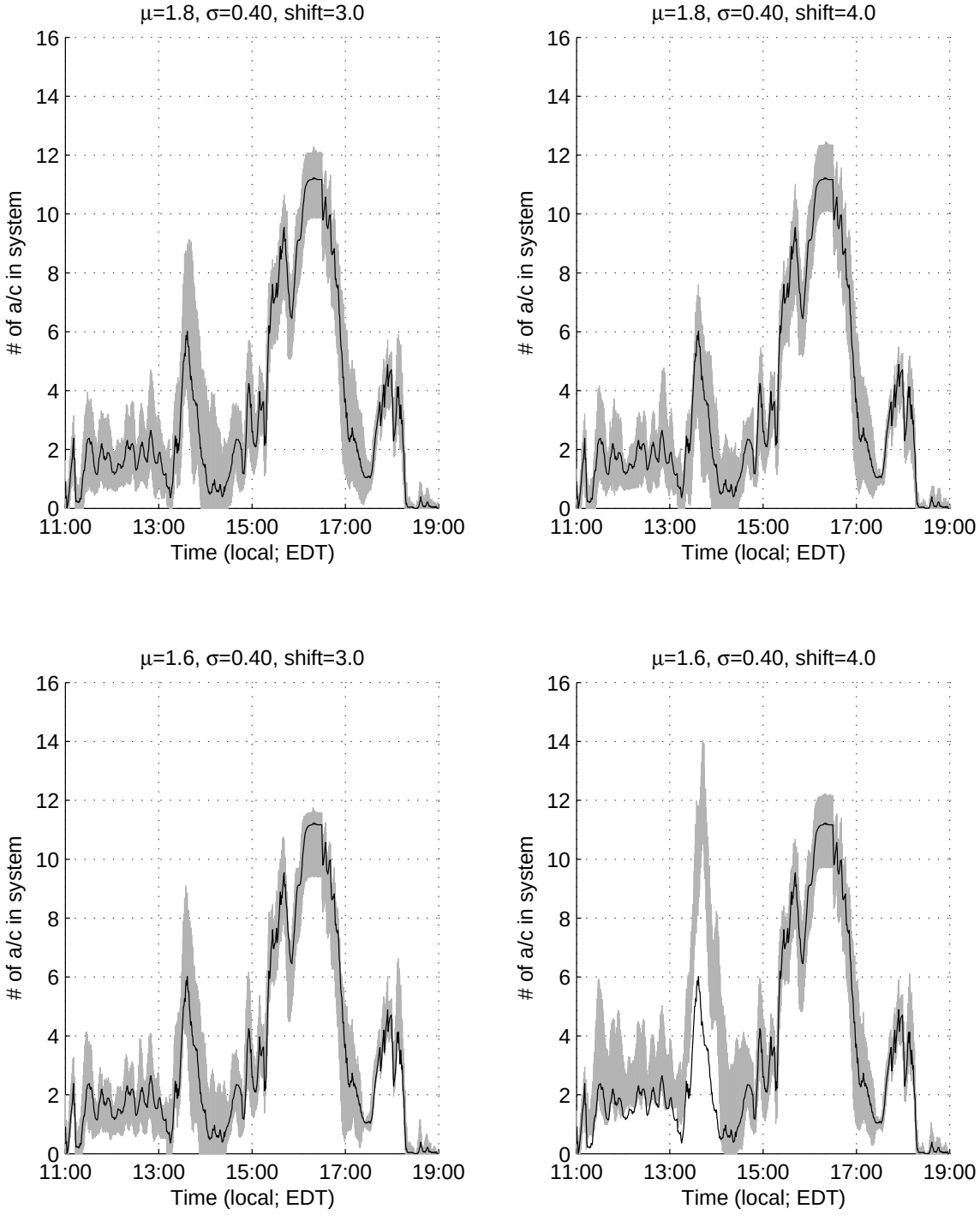

Figure 4-6: Variability due to *shift* and μ.

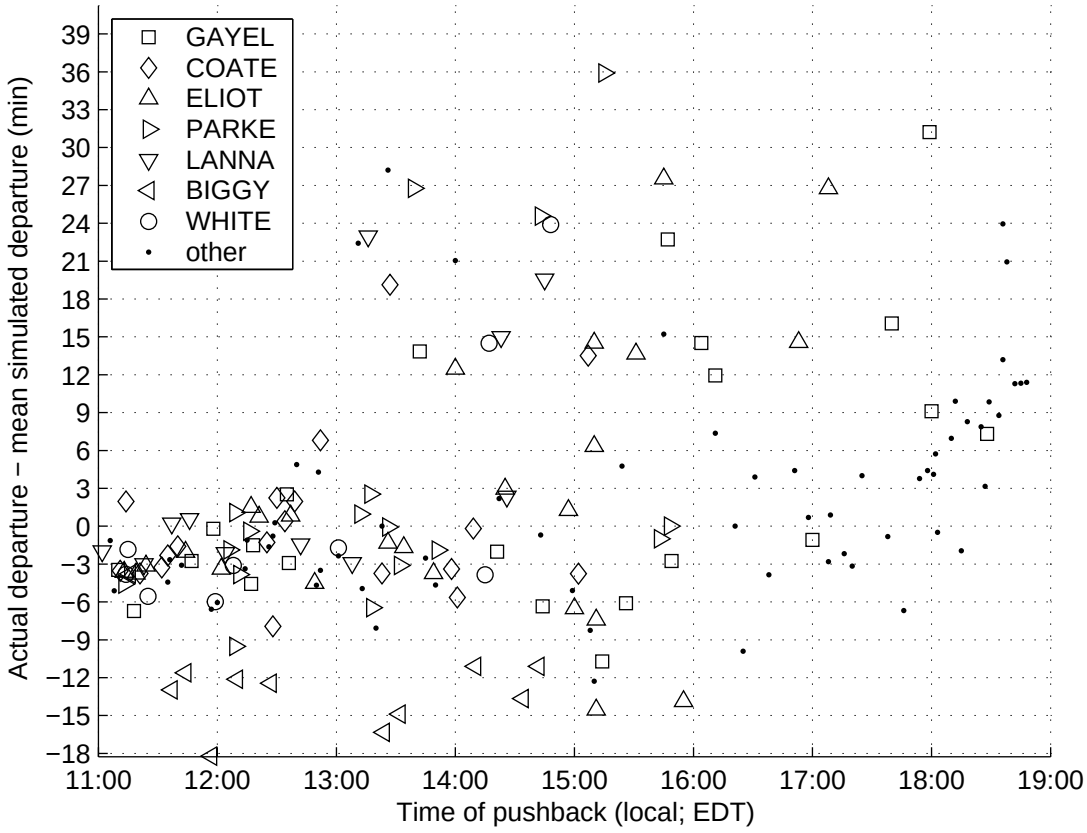

Figure 4-7: Model fidelity as a function of pushback time and departure fix.

There are other possible explanations for the discrepancies. In the simulation, there is no delay between changes in downstream restrictions, and the process of blocking and/or unblocking departure traffic. In reality, ATC may be suffering from a substantial switching penalty as departure fixes are opened and closed. In particular, if aircraft are not properly parked on the airport surface, there may be a significant delay before ATC can access those parked aircraft and re-insert them in the traffic flow. This packing problem deserves further research attention, with the caveat that operations data to support such research is not currently available.

Note that these Monte Carlo simulations test the ceno-scale model in an open-loop environment. Differences between the simulated and actual operations data are not used in a feedback loop to bring the ceno-scale model closer to reality. Figure 4-7 shows a trend of increasingly large discrepancies at later times in the simulated time-period. This may be due to a small bias in the model parameters which slowly accrues over the course of the simulated period of operations.

4.3.3 Effect of pushback prediction uncertainty

It is of interest to evaluate the effect of pushback prediction uncertainty on the model fidelity. This can be done in the Monte Carlo simulation by perturbing the pushback request time for each departure. A set of simulation runs was conducted in which each pushback request time was given an i.i.d. perturbation, modeled as a Gaussian random variable with mean 0.0min and standard deviation 5.0min, to simulate the best-case pushback predictability. The results are shown in Figure 4-8. The result is almost identical to Figure 4-7, indicating that the ceno-scale model is relatively insensitive to best-case perturbations.

A second set of simulation runs was also conducted in which the Gaussian perturbation was changed to have standard deviation 15.0min. This larger standard deviation represents the larger pushback time uncertainty from current departure forecasts. In this case, significant changes are apparent relative to Figure 4-7.

4.4 Robust sequencing

The ceno-scale model is relatively robust with respect to inaccurate parameters and modest pushback prediction uncertainty. This suggests that DSTs based on the ceno-scale model structure should inherit some of this robust character. In this section, two prototype DST algorithms are outlined for future investigation.

The Virtual Queue (VQ) control scheme was first proposed in [FHO$^+$97]. The basic idea is to maintain a notional queue of departing aircraft according to their current status and expected position in the takeoff queue. By conceptually extending the physical queue of departures at the runway, the VQ would act as a single representation of the system state suitable for use as a

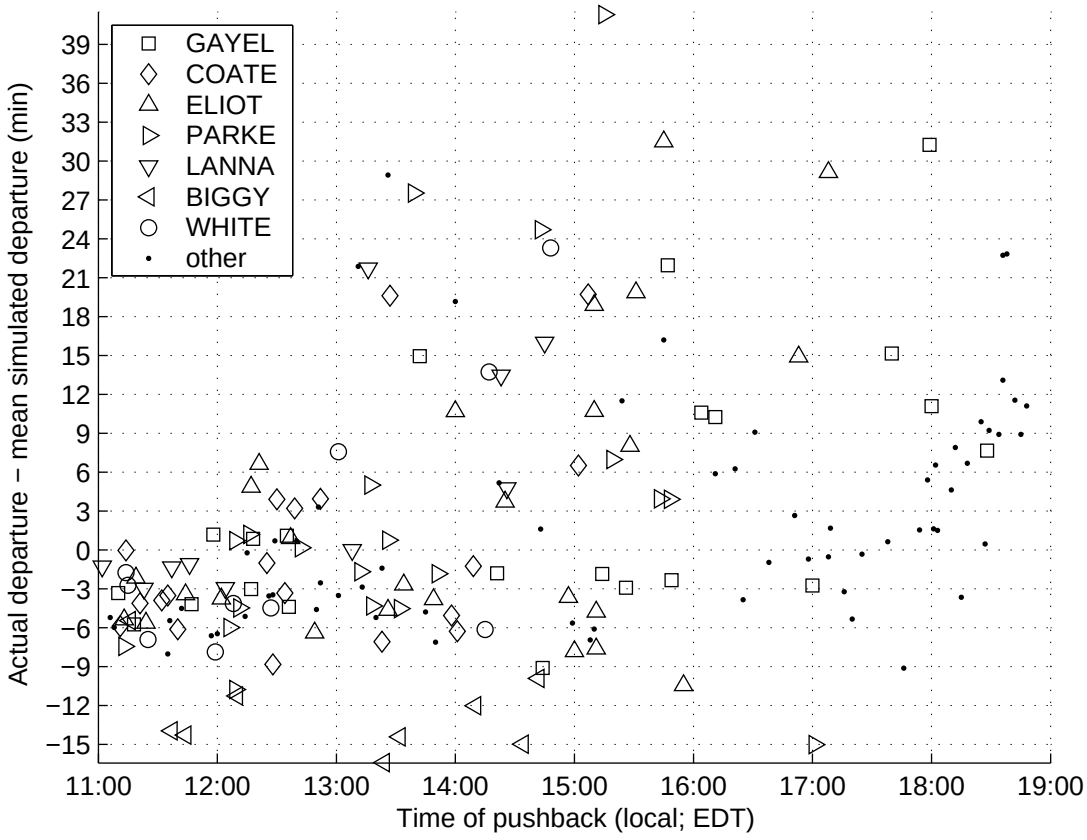

Figure 4-8: Model fidelity under small (±5min) pushback time inaccuracy.

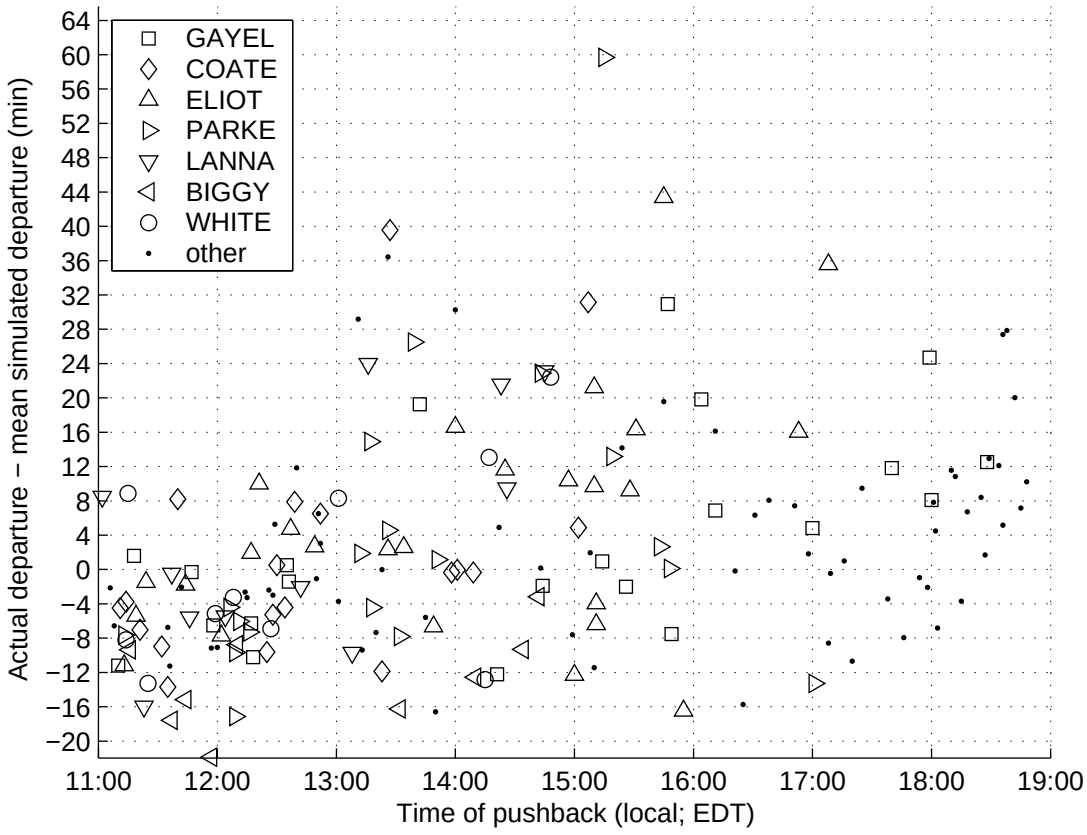

Figure 4-9: Model fidelity under large (±15min) pushback time inaccuracy.

central plan which could synchronize airport operations. The MRS concept suggests a possible implementation of a VQ control system. Instead of maintaining a single VQ for a single runway resource, multiple VQ's should be maintained, one per resource in the MRS gate which regulates runway traffic. A single departing aircraft would hold multiple positions in this multi-VQ system. Each such position would indicate the earliest time when the corresponding resource would be available to deliver a clearance to that aircraft. Especially under severe downstream restrictions, it is possible that an aircraft will not be able to receive a clearance until after it finishes taxiing to the head of the runway queue. In this case, it could be determined in advance that the aircraft should wait at the gate. Gate delays are less expensive in terms of operating and environmental costs [Puj99].

This idea of virtual multi-queues suggests a more fundamental control strategy. In the EWR case-study, it was observed that Continental Airlines responded to severe downstream restrictions by *presequencing* departing aircraft on the runway surface [EC02]. This presequencing strategy was conceptually simple: departing aircraft which needed to use the same downstream navigational fix were grouped together into separate queues. This allowed ATC to easily select aircraft to meet the availability of each fix. In contrast, the other airlines at EWR did not have the necessary infrastructure to presequence or aggregate their departures. In that case, the departure at the head of a FIFO buffer was often delayed waiting for its fix to become available, while several following aircraft in the same buffer which could have departed were unnecesarily delayed. A conceptually similar aggregation-based sequencing algorithm has also been developed in[Ana03, AC02, AC03, ACBV01]. However, in those works, the focus has been on aggregation according to the standard FAA aircraft weight classes.

At a higher level of abstraction, these aggregation-based sequencing strategies attack the sequencing problem by first solving a *dual* problem to determine which resources are limiting traffic flows. Once this dual problem is solved, the primal sequencing problem is greatly simplified: rather than sequencing N distinct departures, only a handful of distinct *classes* of departures must be sequenced, where departures in the same class are effectively identical. Given a MRS gate and an anticipated mixture of aircraft demanding certain types of resources, the dual problem is to determine which resources in the MRS gate are most likely to constrain the runway throughput. A useful first-cut result would be to determine the threshold of downstream restriction severity when aggregation according to weight-class is approximately as correct as aggregation according to departure fix.

Chapter 5

Contributions

The development of decision-support tools (DSTs) for airport surface traffic faces several challenges. The lack of accurate forecasts for airline departure demand is a major and longstanding problem in this area. Through collaboration with one of the best-equipped passenger carriers, Deutsche Lufthansa AG, this thesis has been able to investigate and characterize such forecasts under *best-case minimum-uncertainty* conditions. Even after carefully filtering out a sample of 3820 real-world turn operations expected to exhibit minimal uncertainty, the standard deviation of forecast error for all of the forecast techniques developed herein is lower-bounded at approximately ±5min. It has also been statistically demonstrated that only 3 real-time observations are necessary to achieve this forecast error under best-case conditions: the available ground-time (from actual onblock to scheduled offblock); the time when deboarding begins; and the time when boarding ends. In the ATC research community, it has often been assumed that accurate pushback forecasts will necessitate extensive real-time observations of airline turn processes, and that such forecasts will have essentially zero error. Under quantitative testing, neither assumption appears correct.

Achievable DST performance is fundamentally limited by this inherent stochasticity. More importantly, any viable DST must be ready to cope with *at least* the inherent uncertainty characterized in Chapter 2. To support the development of robust DSTs, a robust modelling framekwork called *ceno-scale* models are developed in the following two chapters. Ceno-scale models fufill several goals related to robust modelling:

- adapted to the limitations of currently-available operations data;
- capable of encoding a much wider range of observed delay mechanisms than existing models of similar complexity;
- built on a low-complexity framework which can be easily translated to various airports and human air traffic controllers.

The ceno-scale modelling framework introduces only two new model structures into pre-existing mesoscopic models: *multi-resource synchronization* (MRS), and two complementary types of sequencing constraints. These extra structures are shown to be capable of encoding a very wide range of observed airport traffic delay mechanisms, and are amenable to several different analytical techniques, yet add only a modest amount of conceptual complexity. There are many interesting theoretical and practical problems which arise from these simple structures; one advantage of the generality of ceno-scale models is that future research will be easy to generalize and apply to many different airports. EWR was chosen as an excellent test-case for ceno-scale modelling based on the availability of a detailed "snapshot" of one day's operations at that airport, allowing tests of the parameter sensitivity and model fidelity to be performed.

5.1 Identified open problems

Given their importance in maintaining an efficient and reliable air transportation system, it is remarkable that ground operations are not more transparent to both air traffic controllers and airline stations. Several decision-support tools for airport surface traffic are now in development, and there is published research linking the potential ATC benefits of these tools to the availability of accurate and timely pushback forecasts. Airlines also stand to benefit financially from these ATC improvements through reductions in either the variability and/or average duration of ground delays. However, airlines with the necessary infrastructure to provide accurate pushback forecasts are the exception rather than the rule. To date only a few carriers have internally justified the business case to support infrastructure investment. It is worth noting that while current DST development and deployment has been heavily dependent on these well-equipped carriers, the extension of these DSTs to sites beyond the initial prototype airports may be significantly handicapped by a lack of high-quality pushback forecasts, an issue which has received little treatment in the literature. In this area, one of the thesis results — that relatively few measurements are necessary to achieve such forecasts — indicates that it may be possible to use the *boarding-pass readers* already installed by many airlines to take a large step towards providing such forecasts. This issue is certainly worthy of further investigation.

There are several other airports which could immediately be encoded as ceno-scale model instances. These include Washington-Dulles International Airport, San Francisco International Airport, Atlanta-Hartsfield International Airport, and Dallas-Fort Worth International Airport. The primary limitation is that airports amenable to ceno-scale models should not be strongly influenced by time-window constraints on departure takeoff times nor by taxiway capacity constraints. At airports such as Atlanta-Hartsfield International and Memphis International where prototype surface surveillance infrastructure is already available, it will be straightforward to develop and calibrate

ceno-scale model instances. The issues of parameter sensitivity and model fidelity must be tested at a wider range of such sites.

Finally, ceno-scale models have pointed the way towards a promising primal/dual formulation of runway operations sequencing. Aside from its practical value and theoretical interest, the idea of using DSTs to solve problems which are dual to the problems being solved by human air traffic controllers may be worthy of investigation from the viewpoint of human-machine interfaces.

Appendix A

Derivations for Bayesian age-based forecasts

A.1 Optimal age-based forecasts

Consider a nonnegative random variable $\mathbf{X}$. When $\mathbf{X}$ is the lifetime of an object or process, $\mathbf{L}_t \doteq [\mathbf{X} - t \,|\, \mathbf{X} > t]$ is the *remaining life* given survival past time t. It is of interest to forecast the remaining life via a deterministic function $f(t)$ using observations of the form $\mathbf{X} > t$. We will assume that forecast errors $f(t) - (\mathbf{X} - t)$ have an instantaneous cost given by $c_f(t, \mathbf{X})$. Our aim is then to minimize the expected integrated cost over the life of the forecast

$$J(f) \doteq \mathrm{E}_{\mathbf{X}} \left[\int_0^{\mathbf{X}} c_f(t, \mathbf{X}) \, dt \right] \tag{A.1}$$

Because the integration interval depends on $\mathbf{X}$, the integral is nonlinear with respect to $\mathbf{X}$ and cannot be directly interchanged with the expectation operator. Let $F(x) \doteq \Pr(\mathbf{X} \leq t)$ be the cumulative distribution of $\mathbf{X}$. Then $J(f)$ can be explicitly written out as

$$J(f) = \int_0^{\infty} \int_0^{x} c_f(t, x) \, dt \, dF(x)$$

Assuming the conditions of Fubini's theorem for improper integrals are satisfied, the integrals may be exchanged to yield

$$\begin{aligned} J(f) &= \int_0^{\infty} \int_t^{\infty} c_f(t, x) \, dF(x) \, dt \\ &= \int_0^{\infty} \mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] \Pr(\mathbf{X} > t) \, dt \end{aligned} \tag{A.2}$$

Thus integration and expectation have been interchanged, albeit with both new limits of integration and new conditioning on $\mathbf{X} > t$. Note that to minimize $J(f)$, it suffices to minimize $\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big]$ pointwise, i.e. by treating $f(t)$ as a free parameter at each instant t.

A.1.1 Quadratic error costs

In section 2.3 it is argued that pushback prediction errors should have a quadratic cost. Thus in the case when forecast errors at any time can disrupt planning activities, it is reasonable to model the instantaneous cost as quadratic:

$$c_f(t, \mathbf{X}) \doteq \alpha(f(t) - (\mathbf{X} - t)) \cdot (f(t) - (\mathbf{X} - t) - \beta) \tag{A.3}$$

where α and β are fixed nonnegative location/scale parameters. Then the expectation may be written out as

$$\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] = \alpha\left(\mathrm{E}\big[(f(t) - (\mathbf{X} - t))^2 \mid \mathbf{X} > t\big] - \beta\,\mathrm{E}\big[f(t) - (\mathbf{X} - t) \mid \mathbf{X} > t\big]\right)$$

Using the shorthand $\mathbf{L}_t = [\mathbf{X} - t \,|\, \mathbf{X} > t]$ and $f(t) = f$,

$$\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] = \alpha\big(f^2 - 2f\,\mathrm{E}\big[\mathbf{L}_t\big] + \mathrm{E}\big[\mathbf{L}_t^2\big] - \beta f + \beta\,\mathrm{E}\big[\mathbf{L}_t\big]\big)$$

Then the optimal (minimizing) value of f may be computed via

$$\begin{aligned}\frac{d}{df}\,\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] &= \alpha\big(2f - 2\,\mathrm{E}\big[\mathbf{L}_t\big] - \beta\big) = 0 \\ \implies f(t) &= \mathrm{E}\big[\mathbf{X} - t \mid \mathbf{X} > t\big] + \frac{\beta}{2}\end{aligned} \tag{A.4}$$

By inspection the optimal expected instantaneous cost of forecast error is

$$\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] = \alpha\,\mathrm{Var}\big(\mathbf{X} - t \mid \mathbf{X} > t\big) - \alpha\beta^2/4 \tag{A.5}$$

Substituting back into Equation (A.2) we obtain

$$\begin{aligned}J(f) &= \int_0^\infty \big(\alpha\,\mathrm{Var}\big(\mathbf{X} - t \mid \mathbf{X} > t\big) - \alpha\beta^2/4\big)\,\Pr(\mathbf{X} > t)\,dt \\ &= \alpha\int_0^\infty \mathrm{Var}\big(\mathbf{X} - t \mid \mathbf{X} > t\big)\,\Pr(\mathbf{X} > t)\,dt - \frac{\alpha\beta^2}{4}\,\mathrm{E}\big[\mathbf{X}\big]\end{aligned} \tag{A.6}$$

A.1.2 Discounted quadratic error costs

Forecast errors which occur long before the actual event itself may contribute relatively little to the total error cost, assuming there will be sufficient time to correct the corresponding planning errors. In this case a reasonable approach is to multiply the presumed instantaneous error cost by a discount factor $e^{-\zeta(\mathbf{X}-t)}$ where $\zeta > 0$ is a fixed parameter:

$$c_f(t, \mathbf{X}) \doteq \alpha(f(t) - (\mathbf{X} - t)) \cdot (f(t) - (\mathbf{X} - t) - \beta) \cdot e^{-\zeta(\mathbf{X}-t)} \tag{A.7}$$

Using the same shorthand as before, we obtain

$$\begin{aligned} \mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] &= \alpha \left(\mathrm{E}\big[(f^2 - 2f\mathbf{L}_t + \mathbf{L}_t^2)e^{-\zeta\mathbf{L}_t}\big] - \beta\, \mathrm{E}\big[(f - \mathbf{L}_t)e^{-\zeta\mathbf{L}_t}\big]\right) \\ \frac{d}{df}\, \mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] &= \alpha \left(2f\, \mathrm{E}\big[e^{-\zeta\mathbf{L}_t}\big] - 2\, \mathrm{E}\big[\mathbf{L}_t e^{-\zeta\mathbf{L}_t}\big] - \beta\, \mathrm{E}\big[e^{-\zeta\mathbf{L}_t}\big]\right) \end{aligned}$$

Setting the derivative to zero yields

$$\begin{aligned} f(t) &= \frac{\mathrm{E}\big[\mathbf{L}_t e^{-\zeta\mathbf{L}_t}\big]}{\mathrm{E}\big[e^{-\zeta\mathbf{L}_t}\big]} + \frac{\beta}{2} \\ &= \frac{\mathrm{E}\big[(\mathbf{X} - t)e^{-\zeta(\mathbf{X}-t)} \mid \mathbf{X} > t\big]}{\mathrm{E}\big[e^{-\zeta(\mathbf{X}-t)} \mid \mathbf{X} > t\big]} + \frac{\beta}{2} \end{aligned} \tag{A.8}$$

After wading through a bit of algebra we also obtain

$$\mathrm{E}\big[c_f(t, \mathbf{X}) \mid \mathbf{X} > t\big] = \alpha \left(\mathrm{E}\big[\mathbf{L}_t^2 e^{-\zeta\mathbf{L}_t}\big] - \frac{\mathrm{E}^2\big[\mathbf{L}_t e^{-\zeta\mathbf{L}_t}\big]}{\mathrm{E}\big[e^{-\zeta\mathbf{L}_t}\big]}\right) - \frac{\alpha\beta^2}{4}\, \mathrm{E}\big[e^{-\zeta\mathbf{L}_t}\big] \tag{A.9}$$

Similar results would hold for a wide class of discount factors. A reasonable qualification for any discount factor is that it should be a nonnegative monotonically decreasing function of $(\mathbf{X} - t)$, equal to 1 when $\mathbf{X} - t = 0$. Then assuming the existence of the simple forecast (A.4) (i.e. assuming that the required expectations converge), any such discount factor could be substituted for $e^{-\zeta(\mathbf{X}-t)}$ in (A.8) and (A.9).

A.2 Remaining Life Theorem

From the results of the previous section, it is of interest to compute the moments of $\mathbf{L}_t$. In this case it is more convenient to characterize $\mathbf{X}$ by its complementary distribution $G(t) \doteq \Pr(\mathbf{X} > t)$. For any nonnegative random variable $\mathbf{Z}$ one has the identity [Gal98, p. 8]

$$\mathrm{E}\big[\mathbf{Z}\big] = \int_0^\infty G_{\mathbf{Z}}(t)\, dt. \tag{A.10}$$

For $n \in \mathbb{N}^+$ and $t, \tau \geq 0$, $\mathbf{L}_t^n$ is nonnegative with complementary distribution function

$$\begin{aligned}
\Pr(\mathbf{L}_t^n > \tau) &= \Pr\big((\mathbf{X} - t)^n > \tau \mid \mathbf{X} > t\big) \\
&= \Pr(\mathbf{X} > \tau^{1/n} + t \mid \mathbf{X} > t) \\
&= \frac{G(\tau^{1/n} + t)}{G(t)}
\end{aligned}$$

Applying (A.10) yields the desired result:

$$\begin{aligned}
\mathrm{E}\big[\mathbf{L}_t^n\big] &= \int_0^\infty \frac{G(\tau^{1/n} + t)}{G(t)}\, d\tau \\
&\cdots \begin{cases} v = \tau^{1/n} + t, & dv = n^{-1}\tau^{(1-n)/n}\, d\tau, \\ \tau = (v-t)^n, & d\tau = n^{+1}\tau^{(n-1)/n}\, dv = n(v-t)^{n-1}\, dv \end{cases} \\
&= \int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1}\, dv
\end{aligned}$$

A.3 Hazard-Rate Remaining Life Recursion

Recall Liebniz' integral rule:

$$\frac{\partial}{\partial z} \int_{a(z)}^{b(z)} f(x,z)\, dx = \int_{a(z)}^{b(z)} \frac{\partial f}{\partial z}\, dx + f(b(z), z)\frac{\partial b}{\partial z} - f(a(z), z)\frac{\partial a}{\partial z}$$

For the case $n = 1$,

$$\begin{aligned}
\frac{\partial}{\partial t}\, \mathrm{E}\big[\mathbf{L}_t\big] &= -1 - \frac{\dot{G}(t)}{G(t)} \int_t^\infty \frac{G(v)}{G(t)}\, dv \\
&= -1 + r(t)\, \mathrm{E}\big[\mathbf{L}(t)\big]
\end{aligned}$$

while for $n > 1$,

$$\begin{aligned}
\frac{\partial}{\partial t}\, \mathrm{E}\big[\mathbf{L}_t^n\big] &= -n \int_t^\infty \frac{G(v)}{G(t)} (n-1)(v-t)^{n-2}\, dv - \frac{\dot{G}(t)}{G(t)} \int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1}\, dv \\
&= -n\, \mathrm{E}\big[\mathbf{L}_t^{n-1}\big] + r(t)\, \mathrm{E}\big[\mathbf{L}_t^n\big]
\end{aligned}$$

It is necessary to make two separate derivations because different terms from the right-hand side of Liebniz' integral rule are contributing in each case. Except in the trivial condition $\mathbf{L}_0 = 0$, however, $\mathrm{E}\big[\mathbf{L}_t^0\big] = \mathrm{E}\big[1\big] = 1$ and a single formula suffices:

$$\frac{\partial}{\partial t}\, \mathrm{E}\big[\mathbf{L}_t^n(t)\big] = -n\, \mathrm{E}\big[\mathbf{L}_t^{n-1}\big] + r(t)\, \mathrm{E}\big[\mathbf{L}_t^n\big].$$

Appendix B

Acronyms and abbreviations

APREQ ATC terminology: Approval Request

APREQs are a type of restriction applied to airport departure traffic. When a departing flight is subject to an APREQ restriction, it is not allowed to take off until ATC at the airport has obtained permission from the ATC in charge of the surrounding airspace. Used to ensure that the departing flight can be merged into airborne traffic flows.

ASQP US Department of Transportation database: Airline Service Quality Performance

Describes the flights operated on jet aircraft for 10 major US passenger carriers, covering January 1995 to the present. Each flight record contains the scheduled and actual times for the gate departure, takeoff, landing, and gate arrival of that flight.

ATC ATC terminology: Air Traffic Control

ATL Airport: Atlanta/Hartsfield International, GA

BOS Airport: Boston Logan International, MA

CDM Management structure in US ATC: Collaborative Decision Making. CDM brings together airline and ATC representatives on a daily basis to negotiate joint plans for coping with airport capacity reductions (usually due to inclement weather).

CPS Thesis shorthand: Constrained Position Shifting

A type of scheduling/sequencing constraint. See section 3.3.2 for a definition and discussion.

DEPARTS Prototype DST for airport surface traffic: Departure Enhanced Planning and Runway/Taxiway Assignment System

The Center for Advanced Aviation System Development at MITRE is developing DEPARTS, an optimization-based tool which incorporates current airport conditions, departure demand,

taxiing aircraft status, downstream traffic flow restrictions and user preferences to optimally assign and sequence traffic to taxi routes, runways and departure fixes [CCF+02].

DST Technical term: Decision-support tool (automation to assist human decision-making)

EWR Airport: Newark International, NJ

FIFO Queueing terminology: First-In First-Out

MIT ATC terminology: Miles In-Trail

Definition: "a specified distance between aircraft, normally in the same stratum associated with the same destination or route of flight" [FAA95]

MINIT ATC terminology: "Minutes In-Trail"

Definition: "a specified interval between aircraft, expressed in time" [FAA95]

MRS Thesis shorthand: Multi-Resource Synchronization

MRS is a modelling concept which is well-adapted to encoding the delay mechanisms observed in airport surface traffic. See section 3.3.1 for a definition and discussion.

QNCB Thesis shorthand: Queueing Network Controlled by Blocking

Originates from the conceptual models developed in [Idr00]. See section 3.2.2 for a definition and discussion.

r-op Thesis shorthand: runway operations

Includes arrivals, departures and runway crossings.

TRACON ATC terminology: Terminal Radar Approach Control. Each major airport has a TRACON which controls airborne traffic immediately around that airport.

Appendix C

Downstream restrictions at EWR on June 29, 2000

These data describe the downstream restrictions as they evolved at EWR on June 29, 2000. The data in Table C.1 are derived from [FAA02, Wei80] and are valid for closely spaced runways operating under Visual Flight Rule (VFR) conditions.

Leading Aircraft:	Following Aircraft:					
	D, H/B757	D, L/S	A, H	A, B757	A, L	A, S
Dep, Heavy or B757	1.5	2.0	0.8	0.9	0.9	1.3
Dep, Large or Small	1.0	1.0	0.8	0.9	0.9	1.3
Arr, Heavy	1.2	1.2	1.2	1.7	1.7	3.0
Arr, B757	1.0	1.0	1.1	1.2	1.2	2.4
Arr, Large	1.0	1.0	1.0	1.0	1.0	1.8
Arr, Small	0.8	0.8	0.8	0.9	0.9	1.3

Table C.1: Separation minima for closely spaced VFR runways (minutes).

- Time: 6:30
  ELIOT: 10MIT
  PARKE: 15MIT
- Time: 7:00
  GAYEL: 10MIT
  COATE: 10MIT
  BIGGY: 10MIT
  WHITE: 20MIT
- Time: 7:20
  ORD: 20MIT
- Time: 8:00
  PARKE: 20MIT
  BIGGY: 20MIT
  ATL: 20MIT
- Time: 9:18
  BIGGY: closed
- Time: 9:30
  GAYEL: open
  COATE: open
  ELIOT: open
  BIGGY: 40MIT
- Time: 11:15
  BIGGY: 15MIT
- Time: 11:30
  PARKE: open
- Time: 11:45
  BIGGY: open
- Time: 12:30
  IAD: closed
- Time: 12:45
  GAYEL: 10MIT
  COATE: 10MIT
- Time: 13:10
  LANNA: 40MIT
  BIGGY: 40MIT
  WHITE: closed
- Time: 13:15
  ATL traffic on LANNA to PARKE
  DAB traffic on BIGGY to SHIPP
  FLL traffic on BIGGY to SHIPP
  JAX traffic on BIGGY to SHIPP
  MIA traffic on BIGGY to SHIPP
  PBI traffic on BIGGY to SHIPP
  DAB traffic on LANNA to LINND
  FLL traffic on LANNA to LINND
  JAX traffic on LANNA to LINND
  MIA traffic on LANNA to LINND
  PBI traffic on LANNA to LINND
  CLT traffic on BIGGY to WHITE
  ORD: 30MIT
- Time: 14:10
  ELIOT: 20MIT
- Time: 15:15
  COATE: closed
  ORF traffic on WHITE to alternate fix
  CLT traffic on WHITE to alternate fix
- Time: 15:45
  ELIOT: closed
  PARKE: closed
- Time: 15:50
  LANNA: closed
  BIGGY: closed
- Time: 16:30
  ELIOT: 20MIT
  PARKE: 20MIT

- Time: 17:45
  LANNA: 20MIT
  BIGGY: 20MIT

- Time: 18:08
  ELIOT: closed
  PARKE: closed
  LANNA: closed
  BIGGY: closed
  IAD: 20MIT

Bibliography

[AB02] Stephen Atkins and Christopher Brinton. Concept description and development plan for the Surface Management System. *Journal of Air Traffic Control*, 44(1), January-March 2002.

[ABW02] Stephen Atkins, Christopher Brinton, and Deborah Walton. Functionalities, displays and concept of use for the Surface Management System. In *Proceedings of the 21st DASC*, Irvine CA, October 2002. IEEE. IEEE Catalog Number 02CH37325C.

[AC02] Ioannis Anagnostakis and John-Paul Clarke. Runway operations planning: A two-stage heuristic algorithm. In *Proceedings of the AIAA Aircraft, Technology, Integration and Operations Forum*, Los Angeles, CA, October 1–3 2002.

[AC03] Ioannis Anagnostakis and John-Paul Clarke. Runway operations planning: A two-stage solution methodology. In *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS-36)*, Big Island, HI, January 6–9 2003.

[ACBV01] Ioannis Anagnostakis, John-Paul Clarke, Dietmar Böhme, and Uwe Völckers. Runway operations planning and control: Sequencing and scheduling. *AIAA Journal of Aircraft*, 38(6), November-December 2001. Also presented at the 34th Hawaii International Conference on System Sciences (HICSS-34), Maui, HI, January 3–6, 2001.

[ACFH00] K. Andersson, F. Carr, E. Feron, and W. Hall. Analysis and modeling of ground operations at hub airports. In *Proceedings of the Third USA/Europe Air Traffic Management Seminar ATM-2000*, Napoli Italy, June 2000. EUROCONTROL and the US Federal Aviation Administration.

[AHAF02] Kari Andersson, William Hall, Stephen Atkins, and Eric Feron. Optimization-based analysis of collaborative airport arrival planning. *Transportation Science*, Accepted July 2002.

[AIC+00] Ioannis Anagnostakis, Husni R. Idris, John-Paul Clarke, Eric Feron, R. John Hansman, Amedeo R. Odoni, and William D. Hall. A conceptual design of a departure planner

decision aid. In *Proceedings of the Third USA/Europe Air Traffic Management Seminar ATM-2000*, Napoli Italy, June 2000. EUROCONTROL and the US Federal Aviation Administration.

[Ana03] Ioannis Anagnostakis. *A Multi-Objective, Decomposition-Based Algorithm Design Methodology and its Application to Runway Operations Planning*. PhD thesis, Massachusetts Institute of Technology, 2003.

[And00] Kari Andersson. Potential benefits of information sharing during the arrival process at hub airports. Master's thesis, Massachusetts Institute of Technology, 2000. Also available as Technical Report CSDL-T-1374, The Charles Stark Draper Laboratory, Inc., Cambridge MA; and via reference [AHAF02].

[Atk02] Stephen C. Atkins. Estimating departure queues to study runway efficiency. *Journal of Guidance, Control, and Dynamics*, 25(4):651–657, July-August 2002.

[AW02] Stephen Atkins and Deborah Walton. Prediction and control of departure runway balancing at Dallas/Fort Worth Airport. In *Proceedings of the 2002 American Control Conference*, Anchorage AK, May 2002. IEEE. IEEE Catalog Number 02CH37301C.

[B̈94] Dietmar Böhme. Improved airport surface traffic management by planning: Problems, concepts and a solution—TARMAC. In Heinz Winter, editor, *Advanced Technologies for Air Traffic Flow Management: Proceedings of an International Seminar Organized by DLR, Bonn, Germany, April 1994*, Lecture Notes in Control and Information Sciences. Springer-Verlag, 1994. ISBN 3540198954.

[BCOQ92] François Baccelli, Guy Cohen, Geert Jan Olsder, and Jean-Pierre Quadrat. *Synchronization and Linearity: An Algebra for Discrete Event Systems*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 1992.

[Bet99] Gordon Bethune. *From Worst to First: Behind the Scenes of Continental's Remarkable Comeback*. John Wiley & Sons, 1999. ISBN 0471356522.

[BHBR98] Roger Beatty, Rose Hsu, Lee Berry, and James Rome. Preliminary evaluation of flight delay propagation through an airline schedule. In *Proceedings of the Second USA/Europe Air Traffic Management Seminar ATM-1998*, Orlando FL, December 1998. EUROCONTROL and the US Federal Aviation Administration.

[BKSA00] J. E. Beasley, M. Krishnamoorthy, Y. M. Sharaiha, and D. Abramson. Scheduling aircraft landings—the static case. *Transportation Science*, 34(2):180–197, May 2000.

[BKSA03] J. E. Beasley, M. Krishnamoorthy, Y. M. Sharaiha, and D. Abramson. The displacement problem and dynamically scheduling aircraft landings. White paper, Imperial College, London, January 2003. Submitted to Transportation Science in 1995, 1998.

[BS00a] Michael A. Bolender and G. L. Slater. Analysis and optimization of departure sequences. In *Proceedings of 2000 AIAA Guidance, Navigation and Control Conference and Exhibit*. American Institute of Aeronautics & Astronautics, August 14–17 2000. AIAA 2000-4475.

[BS00b] Michael A. Bolender and G. L. Slater. Cost analysis of the departure-en route merge problem. *Journal of Aircraft*, 37(1):23–29, January-February 2000.

[BS00c] Michael A. Bolender and G. L. Slater. Evaluation of scheduling methods for multiple runways. *Journal of Aircraft*, 37(3):410–416, May-June 2000.

[Car01] Francis Carr. Stochastic modeling and control of airport surface traffic. Master's thesis, Massachusetts Institute of Technology, 2001. Note: these results are also documented (more accessibly) in [ACFH00].

[CCD+01] Wayne Cooper, Jr., Dr. Ellen A. Cherniavsky, James S. DeArmon, J. Glenn Foster, Dr. Michael J. Mills, Dr. Satish C. Mohleji, and Frank Z. Zhu. Determination of minimum push-back time predictability needed for near-term departure scheduling using DEPARTS. In *Proceedings of the Fourth USA/Europe Air Traffic Management Seminar ATM-2001*, Santa Fe NM, December 2001. EUROCONTROL and the US Federal Aviation Administration.

[CCF+02] Wayne W. Cooper, Jr., Robert H. Cormier, J. Glenn Foster, Dr. Michael J. Mills, and Dr. Satish Mohleji. Use of the Departure Enhanced Planning and Runway/Taxiway Assignment System (DEPARTS) for optimal departure scheduling. In *Proceedings of the 21st DASC*, Irvine CA, October 2002. IEEE. IEEE Catalog Number 02CH37325C.

[CECF02] Francis Carr, Antony Evans, John-Paul Clarke, and Eric Feron. Modeling and control of airport queueing dynamics under severe flow restrictions. In *Proceedings of the 2002 American Control Conference*, Anchorage AK, May 2002. IEEE. IEEE Catalog Number 02CH37301C.

[CEFC02] Francis Carr, Antony Evans, Eric Feron, and John-Paul Clarke. Software tools to support research on airport departure planning. In *Proceedings of the 21st DASC*, Irvine CA, October 2002. IEEE. IEEE Catalog Number 02CH37325C.

[Che02] Victor H. L. Cheng. Collaborative automation systems for enhancing airport surface traffic efficiency and safety. In *Proceedings of the 21st DASC*, Irvine CA, October 2002. IEEE. IEEE Catalog Number 02CH37325C.

[Chi93] Ming-Cheng Chiang. A study of errors in predicting arrival fix times in air traffic control. Master's thesis, Massachusetts Institute of Technology, 1993.

[Dea76] Roger G. Dear. *The Dynamic Scheduling of Aircraft in the Near Terminal Area*. PhD thesis, Massachusetts Institute of Technology, 1976. See also Flight Transportation Lab Report R76-9, Massachusetts Institute of Technology.

[DFB02] David Dugail, Eric Feron, and Karl D. Bilimoria. Conflict-free conformance to en route flow-rate constraints. In *Proceedings of 2002 AIAA Guidance, Navigation and Control Conference and Exhibit*. American Institute of Aeronautics & Astronautics, August 2002. AIAA 2002-5013.

[DGH99] Persi Diaconis, Ronald Graham, and Susan P. Holmes. Statistical problems involving permutations with restricted positions. Technical report, Stanford University, 1999. Available from http://www-stat.stanford.edu/ susan/papers/perm8.ps.

[DH01] Hayley J. Davison and R. John Hansman. Identification of communication and coordination issues in the U.S. air traffic control system. ICAT Report No. ICAT-2001-2, Massachusetts Institute of Technology, June 2001.

[DH02] Stephen W. Dareing and Debra Hoitomt. Traffic management and airline operations. In *Proceedings of the 2002 American Control Conference*, Anchorage AK, May 2002. IEEE. IEEE Catalog Number 02CH37301C.

[DS89] Roger G. Dear and Yosef. S. Sherif. The dynamic scheduling of aircraft in high density terminal areas. *Microelectronics Reliability*, 29(5):743–749, 1989.

[DS91] Roger G. Dear and Yosef. S. Sherif. An algorithm for computer assisted sequencing and scheduling of terminal area operations. *Transportation Research Part A, Policy and Practice*, 25:129–139, 1991.

[EC02] Antony D. Evans and John-Paul B. Clarke. Responses to airport delays—a system study of newark international airport. ICAT Report No. ICAT-2002-5, Massachusetts Institute of Technology, June 2002.

[EUR95] EUROCONTROL. *Operational Concept Document for the EATCHIP Phase III System Generation*, June 1995. OPR-ET1-ST02-1000-OCD-01-00.

[FAA95] FAA, Course 50115. *Traffic management coordinator specialized training student manual*. National Air Traffic Training Program, U.S. DOT, 1995.

[FAA02] FAA. *Order 7110.65N: Air Traffic Control*, February 21 2002.

[FHO+97] Eric Feron, R. John Hansman, Amedeo R. Odoni, Ruben Barocia Cots, Bertrand Delcaire, Xiaodan Feng, William D. Hall, Husni R. Idris, Alp Muharremoglu, and Nicolas Pujet. The departure planner: A conceptual discussion. White paper, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, International Center for Air Transportation, December 1997.

[Gal98] Robert G. Gallager. *Discrete Stochastic Processes*. The Kluwer International Series in Engineering and Computer Science. Communications and information theory. Kluwer Academic Publishers, Boston, 1998.

[GG97] B. J. Glass and Y. Gawdiak. Integrated navigation applications in airport surface traffic management. In *4th St. Petersburg International Conference on Integrated Navigation Systems*, pages 105–113, St. Petersburg, Russia, May 26–28 1997. AIAA A97-30869 07-35.

[Gor93] Jesse Goranson. Looking for trouble: How well the FAA's Enhanced Traffic Management System predicts aircraft congestion. Master's thesis, Massachusetts Institute of Technology, 1993.

[HB98] Henk Hesselink and Niels Basjes. Mantea departure sequencer: Increasing airport capacity by planning optimal sequences. In *Proceedings of the Second USA/Europe Air Traffic Management Seminar ATM-1998*, Orlando FL, December 1998. EUROCONTROL and the US Federal Aviation Administration.

[ICBK02] Husni Idris, John-Paul Clarke, Rani Bhuva, and Laura Kang. Queueing model for taxi-out time estimation. *Air Traffic Control Quarterly*, 10(1):1–22, 2002.

[IDA+98] Husni R. Idris, Bertrand Delcaire, Ioannis Anagnostakis, William D. Hall, John-Paul Clarke, R. John Hansman, Eric Feron, and Amedeo R. Odoni. Observations of departure processes at Logan Airport to support the development of departure planning tools. In *Proceedings of the Second USA/Europe Air Traffic Management Seminar ATM-1998*, Orlando FL, December 1998. EUROCONTROL and the US Federal Aviation Administration.

[IDA+00] Husni R. Idris, Bertrand Delcaire, Ioannis Anagnostakis, William D. 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.

[Idr00] Husni Idris. *Observations and Analysis of Departure Operations at Boston Logan International Airport*. PhD thesis, Massachusetts Institute of Technology, 2000.

[Int] International Air Transport Association. *Airport Handling Manual*, 23rd edition.

[Jan03] Silke I. Januszewski. The effect of air traffic delays on airline prices. MIT job market paper, Economics Department, Massachusetts Institute of Technology, 2003. After Sept. 2003, available from MIT Libraries as part of Januszewski's thesis.

[Leh70] D. H. Lehmer. Permutations with strongly restricted displacements. In P. Erdös, A. Renyi, and V. Sós, editors, *Combinatorial Theory and its Applications II*, pages 755–770. North Holland, Amsterdam, 1970. Originally published in the Proceedings of the Colloquium at Balatonfured, 1969.

[MHV+99] P. Martin, A. Hudgell, S. Vial, N. Bouge, N. Dubois, H. de Jonge, and O. Delain. Potential applications of Collaborative Decision Making. EEC Note No. 9/99, Flight Data Research, Eurocontrol Experimental Centre, July 1999.

[MS02] Alfred Müller and Dietrich Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2002.

[OBD+98] Amedeo R. Odoni, Jeremy Bowman, Daniel Delahaye, John J. Deyst, Eric M. Feron, R. John Hansman, Kashif A. Khan, James K. Kuchar, Nicolas Pujet, and Robert W. Simpson. Existing and required modeling capabilities for evaluating ATM systems and concepts. ICAT Report No. ICAT-98-2, Massachusetts Institute of Technology, 1998.

[PDF00] 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.

[PF99] Nicolas Pujet and Eric Feron. Input-output modeling and control of the departure process of busy airports. In *Proceedings of 1999 Guidance, Navigation and Control Conference and Exhibit*, Portland, OR, August 1999. AIAA.

[PF00] Nicolas Pujet and Eric Feron. Modeling an airline operations control center. *Air Traffic Control Quarterly*, 7(4), 2000.

[Psa79] Harilaos N. Psaraftis. *Dynamic Programming Algorithms for Specially Structured Sequencing and Routing Problems in Transportation*. PhD thesis, Massachusetts Institute of Technology, 1979. See also Flight Transportation Lab Report R78-4, *A Dynamic Programming Approach to the Aircraft Sequencing Problem*, Massachusetts Institute of Technology.

[Puj99] 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 [PDF00] and [PF99].

[RSG+02] Jay M. Rosenberger, Andrew J. Schaefer, David Goldsman, Ellis L. Johnson, Anton J. Kleywegt, and George L. Nemhauser. A stochastic model of airline operations. *Transportation Science*, 36(4):357–377, November 2002.

[RTC00] RTCA Select Committee for Free Flight Implementation. *National Airspace System Concept of Operations Addendum 4: Free Flight Phase 2*, 2000.

[RWD98] Tim Robertson, F. T. Wright, and R. L. Dykstra. *Order Restricted Statistical Inference*, chapter 1, pages 1–58. Wiley series in probability and mathematical statistics. John Wiley and Sons Ltd., GB, 1998.

[The02] Georg Theis. Telematik Anwendungen im Luftverkehr. *Internationales Verkehrswesen*, 54(5):225–8, 2002.

[Tri98] Dionyssios A. Trivizas. Optimal scheduling with Maximum Position Shift (MPS) constraints: A runway scheduling application. *Journal of Navigation*, 51(2):250–266, May 1998.

[Van00] William W. Vanderson. Improving aircraft departure time predictability. Master's thesis, Massachusetts Institute of Technology, September 2000.

[VB97] Uwe Völckers and Dietmar Böhme. Dynamic control of ground movements: State-of-the-art review and perspectives. AGARD R-825, DLR, 1997.

[WB00] Jerry D. Welch and Steven R. Bussolari. Initial Surface Management System preliminary operational concept. ATC Project Memorandum 92PM-AATT-0008, Massachusetts Institute of Technology Lincoln Laboratory, Lexington MA, Aug 31 2000.

[Wei80] William Edmond Weiss. A flexible model for runway capacity analysis. Master's thesis, Massachusetts Institute of Technology, 1980.

[YAS+03] Gang Yu, Michael Argüello, Gao Song, Sandra M. McCowan, and Anna White. A new era for crew recovery at Continental Airlines. *Interfaces*, 33(1):5–22, January-February 2003.