JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 2, January 2005

Evaluation of Improved Pushback Forecasts Derived from Airline Ground Operations Data

Francis Carr,[*] Georg Theis,[†] John-Paul Clarke,[‡] and Eric Feron[§]
Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139

Accurate and timely predictions of airline pushbacks can potentially lead to improved management of airport surface traffic, including reductions in the variability and average duration of costly airline delays. One factor which affects the realization of these benefits is the level of uncertainty inherent in the turn processes. Novel analyses of the *minimum* inherent uncertainty yield a significant *lower bound* on the predictability of airline pushbacks under the best possible conditions. These analyses are based on a large and detailed dataset of approximately 10^4 real-world turn operations obtained through collaboration with Deutsche Lufthansa AG. Three techniques are developed for predicting time-to-go until pushback as a function of available ground time; elapsed ground time; and the status (not started, in progress, or completed) of individual sub-processes in the turn such as catering, fueling, etc. This lower bound result shows that airport surface traffic management must incorporate robust mechanisms for coping with pushback demand stochasticity. These results also characterize the forecast horizon over which pushback predictions are accurate.

Nomenclature

$c_f(t, \mathbf{X})$	=	instantaneous cost of forecast error
$e(t)$	=	forecast error: $f(t) - (\mathbf{X} - t)$
$f(t)$	=	forecast of remaining life
$F(t)$	=	cumulative probability distribution of $\mathbf{X}$: $\Pr(\mathbf{X} \leq t)$
$G(t)$	=	complementary distribution of $\mathbf{X}$: $\Pr(\mathbf{X} > t)$
$\hat{G}(t)$	=	histogram estimate of $G(t)$
$J(f)$	=	expected integral of forecast error cost over forecast lifetime
$\mathbf{L}_t$	=	remaining life of $\mathbf{X}$: $[\mathbf{X} - t \,\vert \mathbf{X} > t]$
$r(t)$	=	hazard rate of $\mathbf{X}$: $-\dot{G}(t)/G(t)$
$\mathbf{X}$	=	random variable describing time until a given event occurs
α	=	forecast error cost: scale parameter
β	=	forecast error cost: shift parameter
ς	=	forecast error cost: discount parameter
μ_L	=	$\mathrm{E}[\mathbf{L}_t]$
λ_L	=	$\mathrm{Var}(\mathbf{L}_t)$
$\sigma(t)$	=	average instantaneous forecast accuracy
$\#\{x_i > t\}$	=	set cardinality: $\left\Vert \{x_1,...,x_N \vert x_i > t\} \right\Vert$

Received 18 March 2003; revision received 30 July 2003; accepted for publication 26 August 2003.

[*]Graduate Research Assistant, Department of Electrical Engineering and Computer Science. fcarr@alum.mit.edu.
[†]Graduate Research Assistant, Department of Civil and Environmental Engineering. gtheis@mit.edu.
[‡]Associate Professor, Department of Aeronautics and Astronautics. Associate Fellow AIAA. johnpaul@mit.edu.
[§]Associate Professor, Department of Aeronautics and Astronautics. Associate Fellow AIAA. feron@mit.edu.

I. Introduction

THE typical airline passenger thinks of air traffic in terms of *flights* from a departure airport to their destination. From the perspective of an airline, once a flight is successfully en-route, its delivery from point A to point B is (happily) quite routine. In contrast, during an airline's ground operations at a large hub airport, there can be multiple waves every day of inbound and outbound flights — tens of aircraft will simultaneously park at the gates; thousands of passengers and their baggage, additional cargo, and airline crews must be transferred between aircraft; these aircraft must be prepped and sent off to execute the next wave of flights. It is during these *turn processes*, between each aircraft's *gate-arrival* and its subsequent *pushback* from the gate, that airlines must efficiently re-allocate their resources (aircraft, crews, accommodation for delayed passengers, etc.) to profitably maintain their flight schedule in spite of delays and disruptions. Accurate and timely pushback forecasts can potentially improve the management of these ground operations and thereby reduce both the variability and average duration of costly delays.

Examining this idea from the other direction, the stochasticity of pushback times must also impose upper limits on such potential performance improvements. Several authors have derived worst-or average-case characterizations of pushback stochasticity.[2-7] In this paper, we present novel analyses of the *best-case* (minimum) pushback stochasticity. 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 complexity must be considered to derive a lower bound on the minimum uncertainty. In this paper, 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.

These techniques yield pushback predictions with the least possible uncertainty over the longest possible forecast horizon. In actual practice this best-case performance is rarely achieved. This imposes a significant design requirement on methods for managing airport surface traffic: any viable methodology must be robust to this minimum uncertainty, and can only rely on pushback forecasts over a shorter horizon. It also imposes a limitation on potential improvements due to improved management, which must be considered before investing in modernization efforts. Finally, under the current system, Air Traffic Control (ATC) has very little real-time information about the airline turn processes.[1] Our analyses show the predictive value of different types of process information, which enables further research on the cost-benefit tradeoffs of investing in the infrastructure needed to make such information available for planning purposes.

II. Literature Review

The simplest *raw* pushback forecast is just the scheduled departure time for a flight. Air traffic controllers rely on a mental model of upcoming pushbacks learned from their day-to-day experience; this is essentially an aggregated raw forecast. In addition, the US Collaborative Decision Making (CDM) program computes Ground Delay Programs (GDP) up to 6h in advance which are based on raw forecasts, albeit modified by the airlines' up-to-date plans for cancellations.[8]

Several studies have examined the accuracy of raw forecasts. Idris (Ref. 2, pp. 144-147) obtained 5h of flight-strip data for departure operations at Boston Logan International Airport (BOS) and analyzed how well ATC could predict pushback requests. Raw pushback forecasts were also evaluated at Atlanta Hartsfield International Airport (ATL).[3,4] It was noted in the ATL studies that pushback times at that airport have low uncertainty compared to many other U.S. hub airports. Vanderson[5] analyzed pushback forecasts at horizons of zero to six hours from the U.S. CDM program. These CDM forecasts frequently had errors with mean absolute deviation in excess of 20min even under the best weather and traffic conditions; the mean absolute deviation sometimes exceeded 1.5h during periods of inclement weather.

These statistical analyses have shown that the published departure time is a poor forecast of the actual pushback time. Both structural complexity and inherent natural randomness contribute to this lack of predictability.[6] More advanced *model-based* forecasts attempt to factor in the tactical decisions made by each airline during the turn process. Vanderson[5] augmented U.S. CDM predictions using a heuristic model of airline decision-making related to cancellations, swaps, and intentional pushback delays and hastening. The Aircraft Sequencing Model (ASM)

constructed by Andersson[7] is an optimization-based model which, given a perfect forecast of landing times over a 3.25h horizon, minimizes passenger delay by modifying aircraft pushback times.

Cooper et al.[3,4] analyzed the sensitivity of their prototype decision-support tool with respect to pushback forecast accuracy. Their two studies examined how the performance of the Departure Enhanced Planning And Runway/Taxiway-Assignment System (DEPARTS) was expected to vary according to both the magnitude of pushback prediction error, and the horizon over which predictions are available. The results show how the benefits of such a system can be quickly eroded when pushback predictions are noisy or not available in a timely fashion. To our best knowledge, there is no complementary research which characterizes best-case pushback predictions and the corresponding upper bound on achievable performance.

The accuracy of raw and model-based forecasts are compared in Table 1. The inconclusive comparison illustrates that, while these model-based forecasts have reduced some of the stochasticity due to unmodeled dynamics, other significant sources of forecast error remain. One candidate is unobservable dynamics. Idris et al.[1] developed a Petri-net model of a single turn, and noted that most of the relevant state information is not observable outside each local airline station. In their study of an airline operations center, Pujet and Feron[9] concluded that many of the relevant decision processes are not observable externally. Obtaining statistically significant quantities of accurate operations data is a longstanding problem in this research area.

Table 1 Literature on pushback predictions

Author	Type	σ	MAD
Idris[2]	Raw	17.4min	14min
DEPARTS[3,4]	Raw	12.6min	8.5min
Vanderson[5]	Raw	n/a	20–90min
Vanderson[5]	Model	n/a	> 15min
Andersson[7]	Model	9–19min	n/a

σ – standard deviation
MAD – mean absolute deviation

III. Observing the Turn Process

The analyses in this paper are based on a dataset of 17,344 turn operations obtained from the Lufthansa Airlines OBELISK data-warehouse. These data were collected via the ALLEGRO system, a unique real-time control and analysis system for airline ground operations developed under the Operational Excellence on-time initiative at Lufthansa.[10] These turns were continental operations on Lufthansa which transited Frankfurt International Airport (FRA) between 1 February 2003 and 30 April 2003. For each transit through FRA, several observations were measured by ALLEGRO and recorded 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 standard IATA[11] delay codes.

Ground operations data (from ALLEGRO). Both scheduled and actual ground time intervals were specified, and further 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 already been validated.[10]

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 aircraft, 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 since 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:

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 types of passenger-loading equipment were used. Changes in target-times due to different *available*[*] ground time were accounted for.

Only relevant delay-codes. Turns with delays due to late inbound flights; local or downstream weather; missed takeoff slots; and controls imposed by ATC or the airport authority were excluded. Delays due to abnormal aircraft maintenance requirements were also excluded, since those delays do not directly impact the normal turn processes and hence were not directly observable in the ALLEGRO data. The remaining types of delay were specific to aircraft servicing processes, and hence were *not* filtered out.

During day-to-day operations, some causes of delay cannot be predicted in advance. For example, a routine mechanical inspection may discover an unusual problem which requires unscheduled maintenance. In contrast, by working with historical data, these unpredictable sources of uncertainty can be filtered out by censoring any affected turns from the dataset. Hence the analyses presented in this paper 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). At the time these data were collected, the end of boarding along a jetway was being inferred indirectly from other turn processes.[†] The results for *all* of the simple turns have been computed and appear similar, but to exclude possible hidden biases due to this inferencing step, the results reported herein are based solely on the 3820 bus-bus simple turns (22.0% of the total sample).

IV. Penalty Function for Forecast Error

Pushback forecasts must obviously derive from accurate operations data, but it is also necessary to have an accurate model of the cost of forecast errors. Various stakeholders (airlines, the traveling public, ATC, etc.) incur different costs when a pushback forecast is incorrect. While empirical studies of these costs are currently lacking, in the interim we argue for the use of a square-law penalty, and derive our results under this assumption. To support future work, a framework is also laid out in the appendix for forecasting based on alternative cost functions.

It is commonly known that airlines incur financial costs due to delayed flights, including both additional Direct Operating Costs (DOC) and lost revenue. The additional DOC increases linearly with delay due to additional crew wages and (depending on the situation) extra fuel burn. In high-demand markets where increased delays on one itinerary may drive passengers to substitute alternate itineraries or other forms of transportation, Januszewski[12] derives a marginal price-change of approximately $1USD per fare per minute of expected delay.

A delayed flight also induces cascading network disruptions on later flights which share its resources (crew, passengers, aircraft, etc). Beatty et al.[13] modeled this propagation effect as a *delay multiplier* by which the initial flight delay is scaled to account for the total delay incurred in the cascade. In the Beatty study, two significant and general effects were observed: the additional delay due to the cascade tends to increase linearly with both the initial delay, and with 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 (as supported empirically by Januszewski's work), the total cost to the airline increases no slower than the *square* of the initial delay.

It is obvious that airlines pay a delay penalty when the delay is unforeseen. One can make related arguments regarding unforeseen *haste*.[‡] For example, a hastened flight might be swapped into the schedule, capping the growth of a cascading network delay, but only if such an opportunity is detected in time. This example brings up important issues of forecast horizons, and how the penalty for an incorrect forecast depends on its interaction with the planning process. Finding an empirically supported mathematical model of the forecast penalty which accounts for forecast horizons; different stakeholders; and opportunity costs for unforeseen haste is an open research area.

Based on these arguments, this paper assumes a forecast error cost of the form

$$\text{cost} = \alpha \cdot \text{delay}\,(\text{delay} + \beta) \tag{1}$$

or equivalently,

[*]Available ground time is defined as the time between the actual gate-arrival time of an inbound aircraft, and its subsequent scheduled pushback time.

[†]The ALLEGRO managers report that this measurement is currently being redesigned.

[‡]For lack of a colloquial antonym for "delay," we refer to a process which ends early as h (*hastened*), with a corresponding h (*haste*).

$$\text{cost} = \alpha \cdot \text{haste}\,(\text{haste} + \beta) \tag{2}$$

In this cost-function, "delay" is defined as the actual minus the scheduled time of an event, and "haste" is just the negative of delay. The constant α is a scale-factor (e.g. dollars per minute-squared), and the constant β allows a small *benefit* to accrue due to a small haste. The optimal forecast corresponding to this cost structure is mathematically elegant; in particular, it is invariant with respect to the particular scaling coefficient α of each stakeholder. This cost structure also implies two important concepts: cost does not scale in proportion to average delay or haste, and cost increases as the *dispersion*[*] of delay increases. In concrete terms, twice the delay can actually cost four times as much, and a flight which is usually delayed by 5±1min will always be more expensive than a flight which is usually delayed by 5min. It is natural for airlines, ATC and passengers to schedule buffers and slack time in order to stay on-time and robust in the face of uncertain operating conditions. However, the length of these buffers and all controllable sources of uncertainty should be kept under scrutiny to avoid rapidly ballooning costs.

V. □ Forecasts Using Simple Descriptive Statistics

Consider a raw forecast derived from the ALLEGRO data: predict the actual ground time based on the scheduled ground time. In Figure 1, 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. From the plot, the minimum ground time (50min at FRA when these data were measured) is apparent. The median of actual ground time tends to track the scheduled ground time, indicating that gate-arrival and pushback typically occur on schedule. However the variability is high, leading to poor predictions of actual ground time given scheduled ground time.

Much of this variability can be removed by replacing scheduled ground time with available ground time. In particular the effect of early or late inbounds can be accounted for. The resulting decrease in variability is shown in Figure 2, illustrating that it is possible to predict actual ground time with much reduced uncertainty given the available ground time.

However, it is also important to note that, due to the possible need for cancellations or aircraft 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. Since this dataset only describes turns which actually occurred, this problem does not affect these analyses. It is an open question as to when an airline station can accurately estimate the available ground time, i.e. how far in advance an airline station can confidently decide that a particular outbound flight will not be cancelled or swapped.

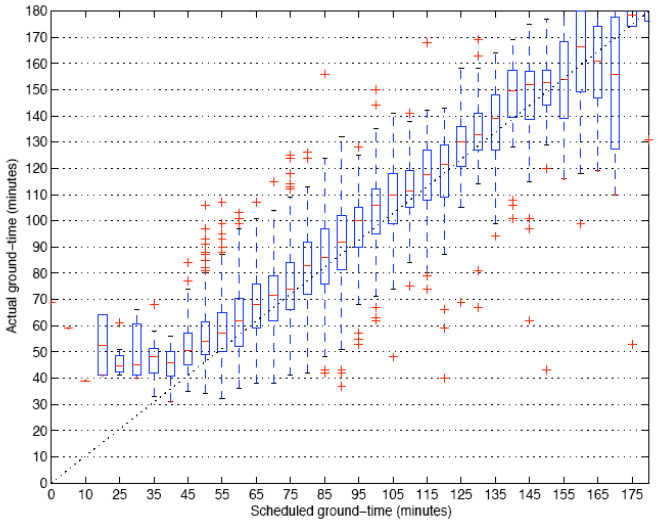

Fig. 1 Actual vs scheduled ground time.

[*]This can be formalized rigorously using the notion of stochastic dispersion order $\leq_{\text{disp}}$ from Ref. 14, p. 40.

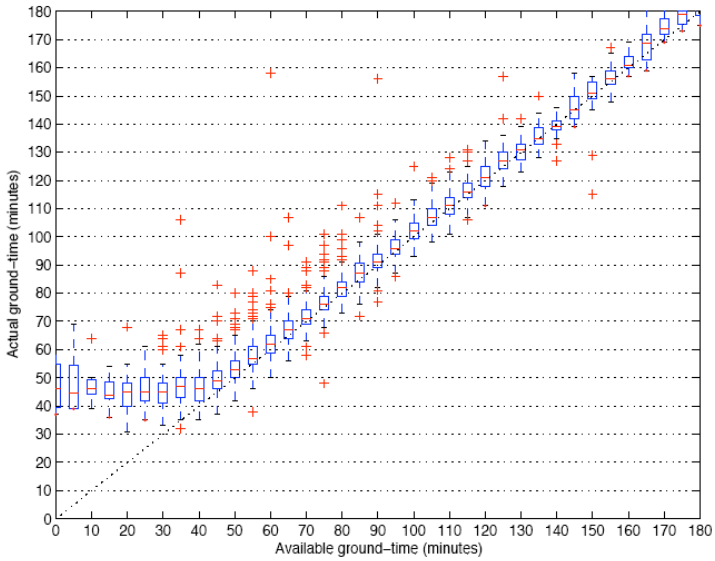

Fig. 2 Actual vs available ground time.

VI. Bayesian Forecasts Using Elapsed Ground time

An initial forecast based on the available ground time can be updated using the currently elapsed ground time. For a turn with some given available ground time, let the random variable **X** denote the actual ground time. Characterize **X** by its *complementary probability distribution G*:

$$G(t) \doteq \Pr(\mathbf{X} > t) \tag{3}$$

(for convenience, the parametrization by the available ground time is elided). At elapsed time t since gate-arrival, the "perfect" forecast of time-to-go is simply **X**-t. To approximate this perfect forecast, consider deterministic functions $\tilde{f}(t)$ with instantaneous forecast error $e(t) = \tilde{f}(t)$ - (**X**-t) and total forecast cost computed as the instantaneous cost integrated over the life of the forecast. Assuming the instantaneous cost is given by Eq. (1), the expected total cost is minimized by the forecast.

$$f(t) = E[\mathbf{X} - t \,|\, \mathbf{X} > t] + \frac{\beta}{2} \tag{4}$$

(see appendix for derivation). Note that while the total cost is defined by an integral over a random-length interval, the forecast depends on a pointwise-optimal Bayesian estimate of the *remaining life* $\mathbf{L}_t \doteq [\mathbf{X} - t \,|\, \mathbf{X} > t]$. The following theorem is useful for characterizing the remaining life:

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

$$\mathrm{E}[L_t^n] = \int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1} dv \,. \tag{5}$$

(See appendix for derivation.) Given this theorem, to calculate the forecast in Eq. (4) it only remains to approximate G from a dataset of samples of **X**. Two approximation techniques are described below.

A. Nonparametric Approximation

One approach is to estimate G nonparametrically. Given N i.i.d. 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 the Bayesian estimate $E[\mathbf{L}_t]$ and its associated mean-square error $\mathrm{Var}(\mathbf{L}_t)$. From Eq. (5),

$$E[\mathbf{L}_t] = \int_t^\infty \frac{G(v)}{G(t)}\,dv \approx \int_t^\infty \frac{\#\{x_i > v\}}{\#\{x_i > t\}}\,dv$$

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

The approximate integrands are piecewise constant and easy to evaluate.

B. Parametric Approximation

An alternative approach is to fit a probability distribution to the data and directly compute the integrals of Theorem 1. If $\mathbf{X}$ has a smooth complementary distribution function, the *hazard rate* $r(t) \doteq -\frac{d}{dt}\ln G(t)$ is well behaved. It is then possible to derive a closed-form recursion on the evolution of $E[\mathbf{L}_t^n]$ as a function of t:

Theorem 2 (Hazard-Rate Recursion)

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

$$\frac{\partial}{\partial t} E[\mathbf{L}_t^n] = -nE[\mathbf{L}_t^{n-1}] + r(t)E[\mathbf{L}_t^n] \tag{6}$$

(see appendix for derivation). The dynamics of the mean μ_L and variance λ_L are of particular interest:

$$\begin{aligned} \frac{d}{dt}\mu_L(t) &= -1 + r(t)\mu_L(t) \\ \frac{d}{dt}\lambda_L(t) &= r(t)[\lambda_L(t) - u_L^2(t)] \end{aligned}$$

These ODEs can be evaluated numerically.

C. Performance on the Simple Turns

Both the nonparametric and parametric approaches were applied to 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 3. These cumulative distributions are then used to derive the pair of forecasts (here β=0) shown in Figure 4. 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 correspondingly rare, the nonparametric forecast must depend on only a handful of datapoints and some deviations are apparent.

For a specific sample of N turns, 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 5. Note that the apparent accuracy of the parametric forecast may be misleading, since it is derived under the assumption that the underlying data are in fact drawn from a Gaussian distribution.

These forecasts are derived from a Bayesian minimum-error formulation. They have been tested against a dataset of turns from which all sources of uncertainty external to the airline turn processes have been removed. Thus any remaining forecast uncertainty must either be attributed to inherent randomness in the turn processes or lack of observability. The observability issue is treated in the following sections by considering the other type of observations which are available, namely the status of various processes in the turn.

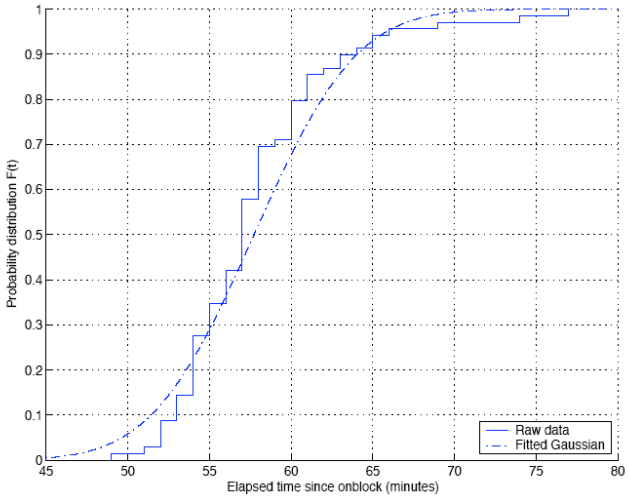

Fig. 3 Probability distribution of 55min simple bus-bus turns.

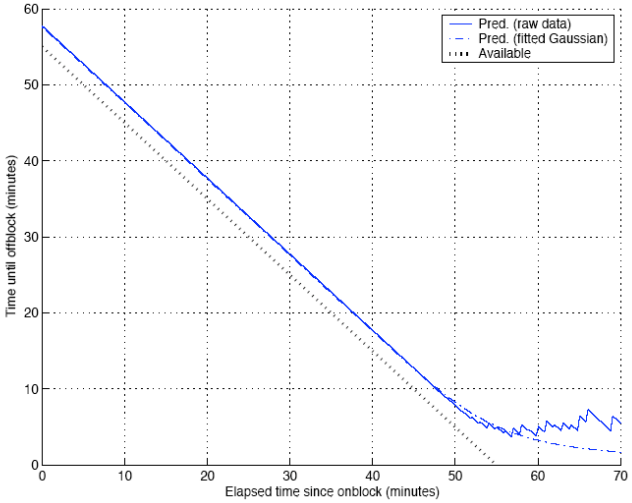

Fig. 4 Age-based forecast for 55min simple bus-bus turns.

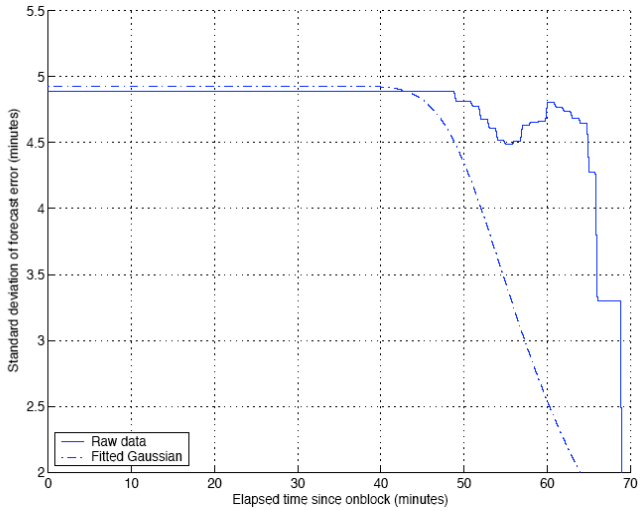

Fig. 5 Age-based forecast: uncertainty as a function of elapsed ground time.

VII. Forecasts Using Coupled Updates of Process Status

A. Derivation of the Model

Another 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 its 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.[1,15-17]

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 6. There are other turn processes including baggage handling, loading of freight, flight crew preparation, etc. but it was indicated during the interviews that these activities only rarely contribute to the critical path.

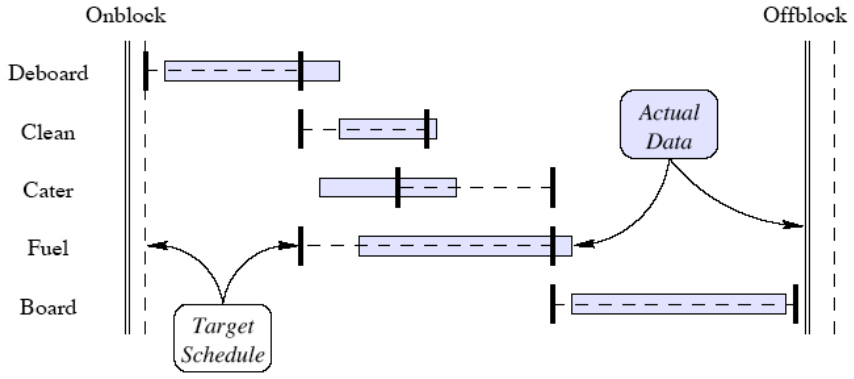

Fig. 6 Illustration of the critical-path events in the turn process.

Because late turns are of much greater importance operationally (early turns are 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 for real-world mechanisms linking the durations. In normal operations, 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. For the sample of 55min bus-bus simple turns, no correlations were observed between the durations of deplaning, servicing and boarding. 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 by Idris et al.[1] By focusing on sequential critical-path processes, it is possible to develop practical forecasting techniques. This model also extends the PERT model of the airline turn process given in Ref. 6, 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 new PERT model is shown in Figure 7.

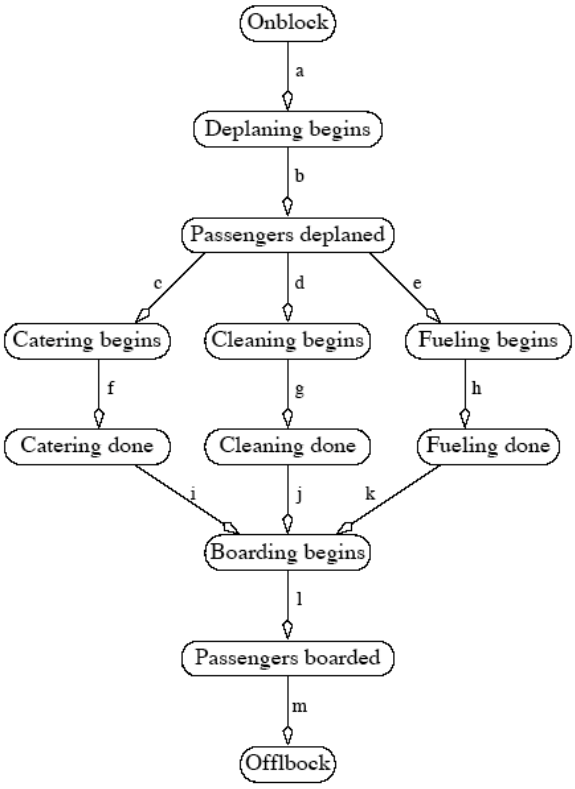

Fig. 7 PERT representation of the critical-path model (55min bus-bus turns).

B. 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 8 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 8 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 measured: since 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.[18] At 95% significance, approximately 11% of the bins were rejected as deviating from a Gaussian distribution. This is acceptable given the expected 5% false-rejection rate.

Based on these mild assumptions, we can derive a 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. The optimal regression of $\mu_k(x)$ is then given by the PAVA algorithm (see Ref. 19) which essentially smoothes the usual estimated averages $\mu_{est}(x)$ 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 8 and 9. 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. Finally, a bootstrap procedure was used to validate the monotonic regressions.

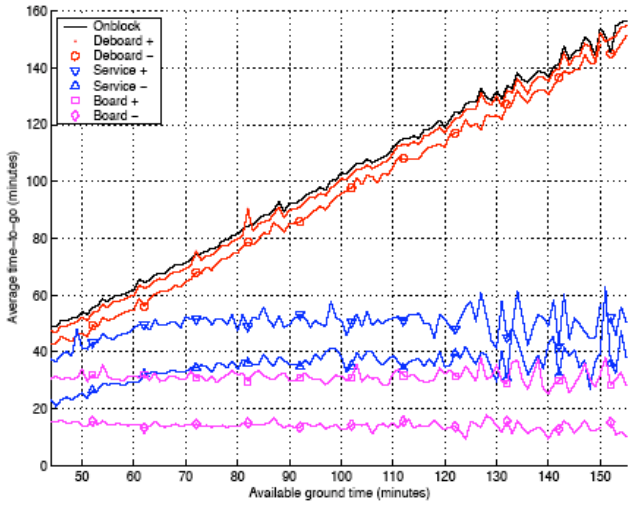

Fig. 8 Average time-to-go as a function of process status.

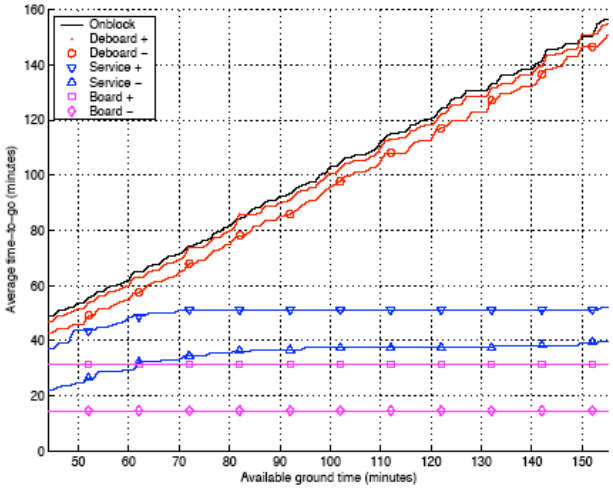

Fig. 9 Average time-to-go, smoothed via PAVA algorithm.

C. Forecast Algorithm

The status-based forecast is constructed from the monotonic regression as follows. When a turn with available ground time x arrives at the gate, the initial forecast of time-to-go until pushback is just the average turn 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 given that status change, and again the forecast counts downwards. This update method is sufficient for most turns in the dataset.

There are three conditions under which this 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. Since the phase which ended out of sequence is obviously not blocking the turn, its status-update can be ignored.

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 ground time. When such a case arises, the forecast is instead updated to the minimum ground time minus the currently elapsed time.

An example is shown in Figure 10. The initial forecast based on the available ground time is successively updated as new status-changes occur. For comparison, the initial forecast based only on available ground time; the "perfect" forecast based on the actual duration; and the scheduled and actual ALLEGRO data are also plotted. The average forecast accuracy $\sigma \doteq \left(\frac{1}{\mathbf{X}} \int_0^{\mathbf{X}} e^2(t)dt \right)^{1/2}$ is also computed.

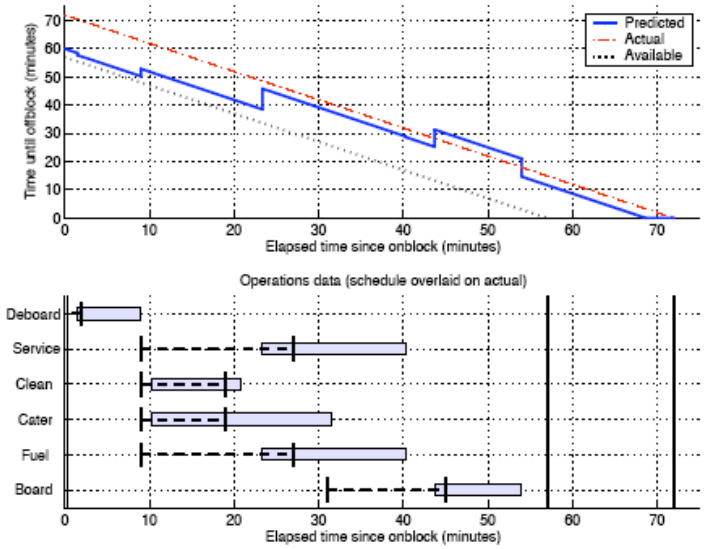

Fig. 10 Example of status-based forecast.

D. Analysis

The average instantaneous forecast accuracy for a subset of the simple bus-bus turns is shown in Figure 11. Again a lower bound on the forecast accuracy throughout the turn is apparent. In Figure 9 the average time-to-go clearly decreases as successive phases start and finish. However the standard deviation of the forecast accuracy in Figure 11 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 Figure 12. This confirms that the status-based forecast is indeed behaving correctly; the lower bound observed from Figure 11 is due to the fact that the actual time-to-go is relatively uncertain and cannot be observed directly.

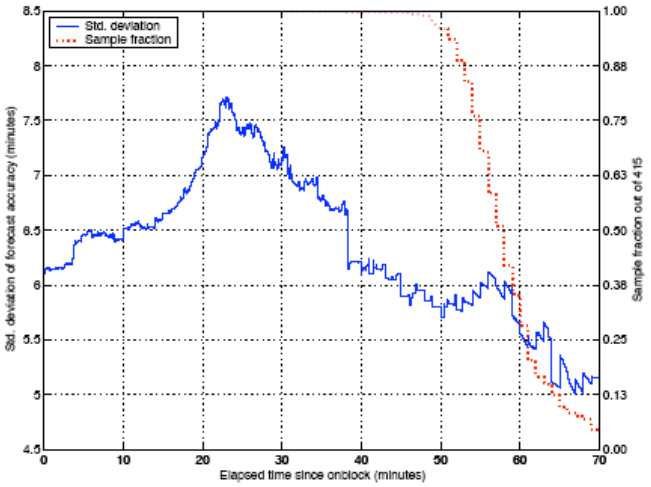

Fig. 11 Status-based forecast: uncertainty as a function of elapsed time.

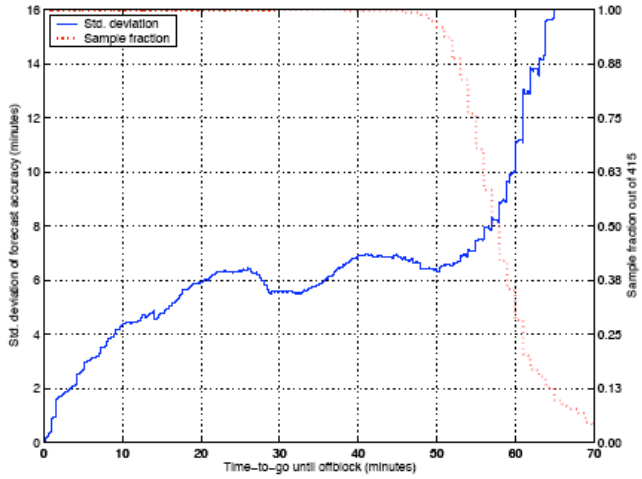

Fig. 12 Status-based forecast: uncertainty as a function of time-to-go.

VIII. Refinement to the Best Status-Based Forecast

One important feature of Figure 11 is that the forecast uncertainty *increases* over certain portions of the turn! At first glance this behavior is counter-intuitive: additional status-changes observed later in the turn should only decrease the forecast uncertainty. Further analysis provides a partial answer. Several of the turn processes have some slack time in the schedule, e.g. the caterers and cleaners may 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 $(L\text{-}S)^2/12$ to the observed variance.

To remedy this problem, a search was conducted for the most informative status changes in the turn. It was assumed that the initial gate-arrival 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 forecast 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 forecast uses only the initial gate-arrival event, the start of deboarding and the end of boarding. The forecast accuracy as a function of time is plotted in Figures 13 (vs elapsed time) and 14 (vs time-to-go). 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 forecast algorithm is required. One possibility for a more complex status-based forecast 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. Since the current schedule does not appear to yield informative thresholds, these critical points must be determined empirically. It is an open problem to solve this thresholding problem; note that finding these thresholds is a dual problem to improving the design of airline schedules.

Note also that this best status-based forecast is evaluated using a symmetric error cost function. It may also be useful to have a forecast which concentrates only on delayed pushbacks. Therefore a similar search was conducted using the asymmetric error cost function $(\max\{0,-e(t)\})^2$. The search revealed an identical best forecast! This 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 strongly penalizes forecasts which are unable to predict large delays, and therefore favors the same forecasts as the asymmetric error cost function. An ongoing project is to optimize and evaluate status-based forecasts 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.

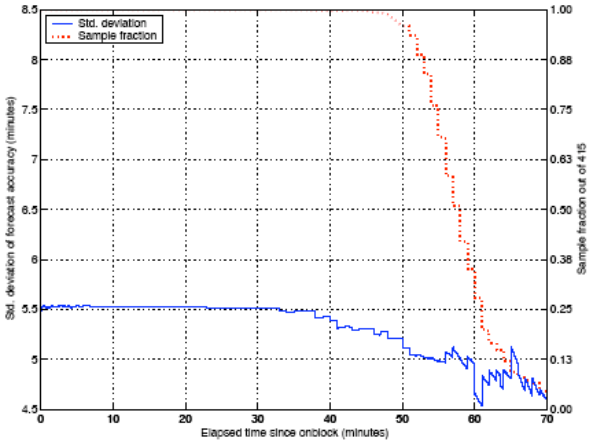

Fig. 13 Best status-based forecast: uncertainty as a function of elapsed time.

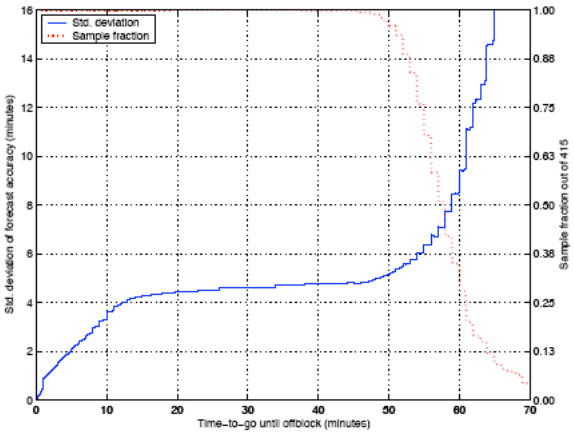

Fig. 14 Best status-based forecast: uncertainty as a function of time-to-go.

IX. 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 Eq. (4). This approach has been successfully applied using Eq. (4) and the three status changes (gate-arrival, begin deboarding, and end boarding) from the best status-based forecast.

A comparison of the three forecasts is shown in Figure 15. 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 Eq. (4) simply decreases the predicted time-to-go at a steady slope of -1sec/sec. Together these observations indicate that the long prediction horizons observed in Figures 5 and 13 are also occurring for turns with much longer available ground times.

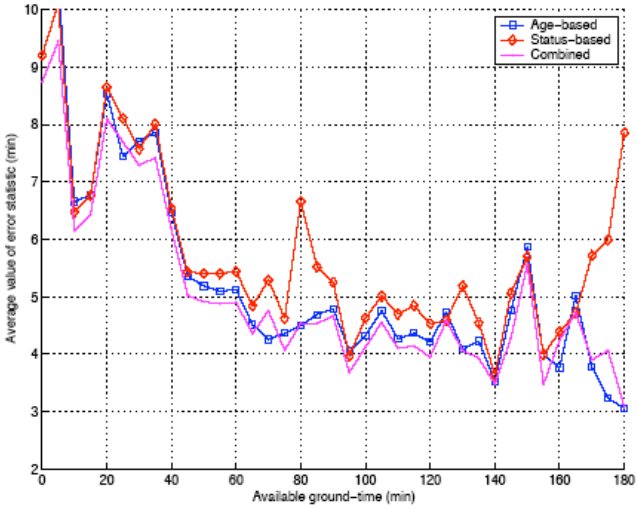

Fig. 15 Comparison of age-based, best status-based and combined forecasts.

X. Conclusions

Uncertainty in the airline turn process imposes limitations on pushback predictability, which in turn limits the performance of airport surface traffic management. The forecast techniques developed in this paper 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. We have concluded that robustness issues in airport surface traffic management will *not* be resolved by future improvements in observing the pushback process. Air traffic management must be designed with 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. Automated decision-support systems (DSS) operating on a long time-horizon, well beyond the 15-30min look-ahead typically practiced by human air traffic controllers, 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 contrast, any DSS operating on a shorter time-scale must adapt to the "sliding delays" observed in Figure 4. As an expected pushback time approaches, the absence of expected status changes (e.g. 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, DSS cannot 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 planning capabilities is an ongoing research topic.

Acknowledgments

The ground event timing data is remarkable, and enables an exceptional opportunity for systematic analyses of the turn process using accurate observations collected over a long period of time. We are sincerely indebted to Deutsche Lufthansa AG for this opportunity. We 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 the NASA/Ames Cooperative Agreement #NCC 2-1147 (Departure Planner Project), the FAA Grant #FAA 95-G-017 (Joint University Program), and the NEXTOR Grant DTFA 01-01-C-00030.

Appendix: Derivations

This appendix contains abbreviated derivations for some of the equations presented in the text; Carr[20] gives additional details.

A.1 Minimizing the Cost of Forecast Error

Consider a nonnegative random variable $\mathbf{X}$. When $\mathbf{X}$ is the lifetime of an object or process, $\mathbf{L}_t \doteq \left[\mathbf{X} - t \middle| \mathbf{X} > t\right]$ 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 $\mathbf{X}>t$. Assume that forecast errors $e(t) \doteq f(t) - (\mathbf{X} - t)$ have an instantaneous cost given by $c_f(t, \mathbf{X})$. Our aim is then to minimize the expected total cost over the life of the forecast:

$$J(f) \doteq E_{\mathbf{X}}\left[\int_0^{\mathbf{X}} c_f(t,\mathbf{X})dt\right] \tag{A1}$$

Because the integration interval depends on $\mathbf{X}$, this integral is nonlinear with respect to $\mathbf{X}$ and cannot be directly interchanged with the expectation operator. Let $F(t) \doteq \Pr(\mathbf{X} \le t)$ be the cumulative distribution of $\mathbf{X}$. Then $J(f)$ can be expanded 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 (Ref. 21, p. 404), the integrals may be exchanged to yield

$$\begin{aligned} J(f) &= \int_0^{\infty} \int_0^{\infty} c_f(t,x)dF(x)dt \\ &= \int_0^{\infty} E\left[c_f(t,\mathbf{X}) \middle| \mathbf{X} > t\right] \Pr(\mathbf{X} > t)dt \end{aligned} \tag{A2}$$

Note that to minimize $J(f)$, it suffices to minimize $\mathrm{E}\left[c_f(t,\mathbf{X})\middle|\mathbf{X}>t\right]$ pointwise, treating $f(t)$ as a free parameter at each instant t.

A.2 Quadratic Error Costs

Consider the instantaneous forecast error cost given by Eq. (1)

$$c_f(t,\mathbf{X}) = \alpha \cdot e(t) \cdot \left(e(t) - \beta\right) \tag{A3}$$

and the simple case when forecast errors which occur over any horizon can equally disrupt planning activities. Substituting into Eq. (A2) and using the shorthand $f(t) \equiv f$, the minimizing value of f may be computed via

$$\begin{aligned} 0 &= \frac{d}{df}\mathrm{E}\left[c_f(t,\mathbf{X})\middle|\mathbf{X}>t\right] \\ &= \alpha\left(2f - 2\mathrm{E}\left[\mathbf{L}_t\right] - \beta\right) \\ \Rightarrow f(t) &= \mathrm{E}\left[\mathbf{X} - t\middle|\mathbf{X}>t\right] + \frac{\beta}{2} \end{aligned} \tag{A4}$$

Using this result, the costs $\mathrm{E}\left[c_f(t,\mathbf{X})\middle|\mathbf{X}>t\right]$ and $J(f)$ can be expressed as quadratic functions of α, β^2, and $\mathrm{Var}(\mathbf{L}_t) = \mathrm{Var}\left(\mathbf{X} - t\middle|\mathbf{X}>t\right)$.

A.3 Discounted quadratic error costs

Forecast errors which occur long before the actual event itself may contribute relatively little to the total error cost. In this case a reasonable approach may be to discount a quadratic error cost by a time-varying factor $e^{-\zeta(\mathbf{X}-t)}$ where $\zeta > 0$ is a fixed parameter:

$$c_f(t,\mathbf{X}) = \boldsymbol{\alpha} \cdot e(t) \cdot \left(e(t) - \boldsymbol{\beta}\right) \cdot e^{-\boldsymbol{\zeta}(\mathbf{X}-t)} \tag{A5}$$

A similar procedure yields the optimal forecast

$$f(t) = \frac{E\left[\mathbf{L}_t e^{-\zeta \mathbf{L}_t}\right]}{E\left[e^{-\zeta \mathbf{L}_t}\right]} + \frac{\boldsymbol{\beta}}{2} \tag{A6}$$

The forecast costs now depend on the *discounted* moments of $\mathbf{L}_t$. For example, $\mathrm{E}\left[c_f(t,\mathbf{X})\middle|\mathbf{X}>t\right]$ simplifies to

$$\alpha\left(\mathrm{E}\left[\mathrm{L}_t^2 e^{-\zeta \mathrm{L}_t}\right] - \frac{\mathrm{E}^2\left[\mathrm{L}_t e^{-\zeta \mathrm{L}_t}\right]}{\mathrm{E}\left[e^{-\zeta \mathrm{L}_t}\right]} - \frac{\beta^2}{4}\mathrm{E}\left[e^{-\zeta \mathrm{L}_t}\right]\right) \tag{A7}$$

Similar results hold for a wide class of discount factors.

A.4 Moments of Remaining Life

To compute these forecasts, 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 (Ref. 22, p. 8)

$$\mathrm{E}[\mathbf{Z}] = \int_0^\infty G_{\mathbf{Z}}(t)\,dt\,. \tag{A8}$$

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

$$\Pr\left(\mathbf{L}_t^n > \tau\right) = \frac{G\left(\tau^{1/n} + t\right)}{G(t)} \tag{A9}$$

Applying Eq. (A8) yields the desired result:

$$E\left[\mathbf{L}_t^n\right] = \int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1}\, dv \tag{A10}$$

A.5 Hazard-Rate Recursion

Recall Leibniz' integral rule (Ref. 23):

$$\frac{\partial}{\partial z}\int_{a(z)}^{b(z)} f(x,z)\, dx = \int_{a(z)}^{b(z)} \frac{\partial f}{\partial z} dx + f\left(b(z),z\right)\frac{\partial b}{\partial z} - f\left(a(z),z\right)\frac{\partial a}{\partial z} \tag{A11}$$

For the case $n = 1$,

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{E}\left[\mathbf{L}_t\right] &= -1 - \frac{\dot{G}(t)}{G(t)}\int_t^\infty \frac{G(v)}{G(t)}\, dv \\ &= -1 + r(t)\,\mathrm{E}\left[\mathbf{L}(t)\right]\end{aligned} \tag{A12}$$

while for $n > 1$,

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{E}\left[L_t^n\right] &= -n\int_t^\infty \frac{G(v)}{G(t)}(n-1)(v-t)^{n-2}\, dv \\ &\quad - \frac{\dot{G}(t)}{G(t)}\int_t^\infty \frac{G(v)}{G(t)} n(v-t)^{n-1}\, dv \\ &= -n\,\mathrm{E}\left[\mathbf{L}_t^{n-1}\right] + r(t)\,\mathrm{E}\left[\mathbf{L}_t^n\right]\end{aligned} \tag{A13}$$

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

$$\frac{\partial}{\partial t}\mathrm{E}\left[\mathbf{L}_t^n(t)\right] = -n\mathrm{E}\left[\mathbf{L}_t^{n-1}\right] + r(t)\mathrm{E}\left[\mathbf{L}_t^n\right] \tag{A14}$$

The simplicity of this telescoping recursion suggests that a more elegant derivation must exist.

References

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

[2]Idris, Husni, "Observations and Analysis of Departure Operations at Boston Logan International Airport," PhD thesis, Massachusetts Inst. of Technology, 2000. Also available via Ref. 24.

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

[4]Cooper, Wayne, Jr., Cormier, Robert H., Foster, J. Glenn, Mills, Michael J., and Mohleji, Satish, "Use of the Departure Enhanced Planning and Runway/Taxiway Assignment System (DEPARTS) for optimal departure scheduling," *Proceedings of the 21st DASC*, IEEE, Oct. 2002.

[5]Vanderson, William W., "Improving aircraft departure time predictability," Master's thesis, Massachusetts Inst. of Technology, Sept. 2000.

[6]Feron, Eric R., Hansman, R. John, Odoni, Amedeo R., Cots, Ruben Barocia, Delcaire, Bertrand, Feng, Xiaodan, Hall, William D., Idris, Husni R., Muharremoglu, Alp, and Pujet, Nicolas, "The Departure Planner: A Conceptual Discussion," white paper, Massachusetts Inst. of Technology, International Center for Air Transportation, Dec. 1997. Available online at http://icat-server.mit.edu/Library/ (cited Nov. 2004).

[7]Andersson, Kari, "Potential benefits of information sharing during the arrival process at hub airports," Master's thesis, Massachusetts Inst. of Technology, 2000. Also available as Technical Report CSDL-T-1374, The Charles Stark Draper Laboratory, Inc., Cambridge MA; and via Ref. 25 in this list.

[8]Ball, Michael O., Hoffman, Robert, Chen, Chien-Yu, and Vossen, Thomas, "Collaborative decision making in air traffic management: Current and future research directions," NEXTOR Technical Report 2000-3, National Center of Excellence in Aviation Operations Research, 2000.

[9]Pujet, Nicolas, and Feron, Eric, "Modeling an Airline Operations Control Center," *Air Traffic Control Quarterly*, Vol. 7, No. 4, 2000.

[10]Theis, Georg, "Telematik Anwendungen im Luftverkehr," ("Telematics in Air Transportation," English translation), *Internationales Verkehrswesen*, Vol. 54, No. 5, 2002, pp. 225–228.

[11]International Air Transport Association, *Airport Handling Manual*, 23rd ed., Dec. 2002.

[12]Januszewski, Silke I., "The Effect of Air Traffic Delays on Airline Prices," MIT Job Market Paper, Economics Department, Massachusetts Inst. of Technology, 2003. Available from the MIT Libraries as part of Januszewski's 2003 PhD thesis.

[13]Beatty, Roger, Hsu, Rose, Berry, Lee, and Rome, James, "Preliminary evaluation of flight delay propagation through an airline schedule," *Proceedings of the Second USA/Europe Air Traffic Management Seminar ATM-1998*, EUROCONTROL and the U.S. Federal Aviation Administration, Dec. 1998.

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

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

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

[17]Anagnostakis, Ioannis, Idris, Husni R., Clarke, John-Paul, Feron, Eric, Hansman, R. John, Odoni, Amedeo R., and Hall, William D., "A conceptual design of a departure planner decision aid," *Proceedings of the Third USA/Europe Air Traffic Management Seminar ATM-2000*, EUROCONTROL and the U.S. Federal Aviation Administration, June 2000.

[18]Baker, R. A., Lilliefors test, implemented as "lillietest" function. Part of source-code for Matlab™ "stats" package v1.4, Aug 1998, based on Conover, W. J., *Practical Nonparametric Statistics*, John Wiley and Sons, Inc., 1980.

[19]Robertson, Tim, Wright, F. T., and Dykstra, R. L., "Order Restricted Statistical Inference," *Wiley Series in Probability and Mathematical Statistics*, John Wiley and Sons Ltd., Great Britain, 1998, Chapter 1, pp. 1–58.

[20]Carr, Francis, "Robust Decision-Support Tools for Airport Surface Traffic," PhD thesis, Massachusetts Inst. of Technology, 2004.

[21]Marsden, Jerrold E., and Tromba, Anthony J., *Vector Calculus*, 3rd ed. W. H. Freeman and Company, New York, 1988.

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

[23]Weisstein, Eric W., "Leibniz Integral Rule," From MathWorld—A Wolfram Web Resource. http://mathworld.wolfram.com/ LeibnizIntegralRule.html.

[24]Idris, Husni, and Hansman, R. John, "Identification of communication and coordination issues in the U.S. air traffic control system," ICAT Report No. ICAT-2000-7, Massachusetts Inst. of Technology, 2000.

[25]Andersson, Kari, Hall, William, Atkins, Stephen, and Feron, Eric, "Optimization-based analysis of collaborative airport arrival planning," *Transportation Science*, Vol. 37, No. 4, 2003, pp. 422-433.