4.6

INVESTIGATING A NEW GROUND DELAY PROGRAM STRATEGY FOR COPING WITH SFO STRATUS*

David A. Clark†
MIT Lincoln Laboratory, Lexington, Massachusetts

1. INTRODUCTION

Integration of weather information into the air traffic flow control decision process is a key component of the planned Next Generation Air Transportation System (NexGen). The NexGen Joint Program Development Office (JPDO) has designated a Weather Integration sub-team to specifically address the issue of translating weather information into air traffic decisions (Andrews et al., 2007). The presumed source of weather information is an authoritative database which contains a distilled representation of weather parameters derived from a variety of external sources. The future concept is that this information will be represented probabilistically, recognizing that this is most suitable for an objective optimization of air traffic flow decisions on a system-wide scale. Many weather products emerging from the research community include a probabilistic representation of weather information, most notably those associated with convective weather which occurs on time and space scales that are particularly challenging for providing reliable deterministic forecasts beyond very short time frames relative to what is required for strategic traffic flow planning.

Translating probabilistic information into decisions that represent optimized system wide benefits has long been a subject of operations research, but is a rare practice in the National Airspace System (NAS) where operational decisions are primarily made based on deterministic information, including forecasts of expected weather conditions. Changes to operational procedures in recent years that have attempted to integrate probabilistic weather information, particularly location and intensity of convective weather, have typically resulted in overly-conservative decisions to restrict traffic flow, failing to exploit improved forecasts provided by the currently emerging suite of convective weather products.

One of the difficulties in applying the probabilistic information in the operational setting is the complex multi-dimensionality associated with both the impact and the prediction of convective weather. In contrast, the prototype San Francisco (SFO) Marine Stratus Forecast System (Clark and Wilson, 1997; Ivaldi et al. 2007). provides a probabilistic representation of weather forecast information with greatly reduced dimensionality, namely the forecast of a single weather parameter (cloud base) at a fixed geographical location (SFO approach zone). This system is currently being used within the NAS decision-making environment, and presents an opportunity to serve as an experimental prototype for this type of integration, i.e. how to modify traffic flow management strategy to best exploit the availability of probabilistic weather information. NASA Ames is sponsoring the National Weather Service (NWS) Forecast Office in Monterey to investigate the use of this prototype for this specific application. Described here are the objectives and groundwork of this investigation, and plans for providing a specific recommendation for modifying the SFO Ground Delay Program (GDP) strategy based on probabilistic weather information.

2. THE SFO MARINE STRATUS FORECAST SYSTEM

2.1 Stratus impact and forecast system development

A stratus cloud deck below 3500 feet in the approach zone prevents dual approaches to SFO's closely spaced parallel runways. This effectively cuts the airport's arrival capacity in half, from 60 to 30 planes per hour. During the warm season (May through October), stratus forms and dissipates on a daily cycle in response to marine air advection and radiative cooling and heating, posing a threat to operations on approximately 75-100 days each year. The stratus typically dissipates from the approach zone sometime between mid-morning and early afternoon, roughly coinciding with the morning arrival push of aircraft into SFO. When stratus is present in the approach zone during the early morning and expected to persist, traffic managers implement a Ground Delay Program (GDP) by holding a portion of upstream aircraft on the ground to reduce the flow of incoming traffic during the period of reduced capacity. This prevents the risk of excessive airborne holding and diversions that would result from an extended period of demand exceeding capacity. The operational cost of such a coping mechanism is that upon stratus clearing, there is a period of wasted arrival capacity while the upstream pipeline of aircraft is filled following release of ground-held planes. Figure 1 shows that a typical range of three to as many as twenty arrival slots are not used following clearing of stratus from the approach zone (MCR Federal, 2004). These wasted slots represent unutilized capacity which, when demand exceeds capacity, implies unnecessary delay.

---

*This work was sponsored by the National Oceanic and Atmospheric Administration (NOAA) under Air Force Contract No. FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

†*Corresponding author address*: David A. Clark, MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02420-9185; e-mail: davec@ll.mit.edu

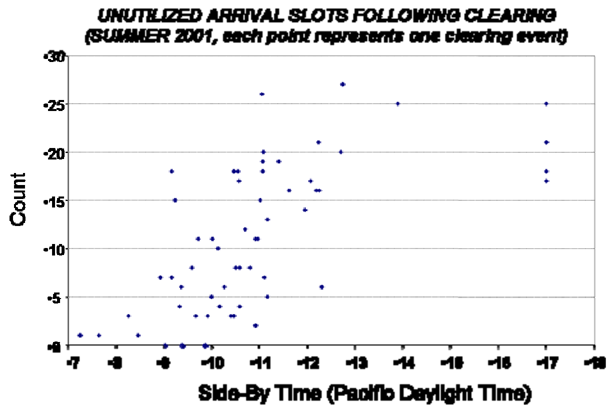

Figure 1. Unutilized arrival slots following stratus clearing, SFO summer 2001.

Primary forecasting responsibility for anticipating the time of stratus clearing is shared by the Center Weather Service Unit (CWSU) at the Oakland Air Route Traffic Control Center (ARTCC), the aviation forecasting desk of the National Weather Service (NWS) Forecast Office in Monterey, and the operations centers of major commercial airlines with significant market share in SFO. Their forecasts are used by traffic managers to determine the duration and scope (number of planes impacted) by a proposed GDP (Figure 2). In 1995, the FAA's Aviation Weather Research Program (AWRP) sponsored an effort to improve these forecast supporting this traffic management decision. This led to the development of the SFO Marine Stratus Forecast System which provides automated forecast guidance specifically for this purpose. Output from the system is made available to forecasters and decision makers via a web-based display.

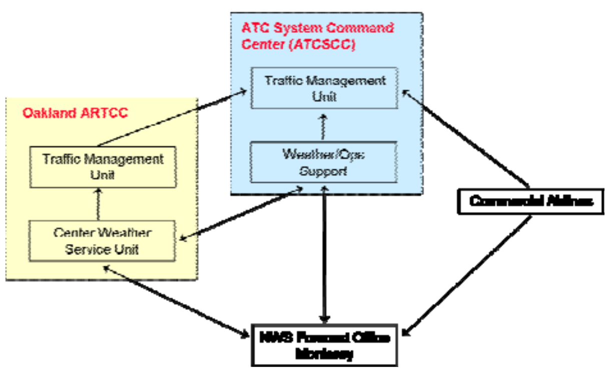

Figure 2. Information sharing for GDP implementation strategy decision [adapted from Strach, 1991].

The system provides a continually updated forecast of the time of stratus clearing each morning. Forecasts are initialized at 9, 11, 13, 15, 16, 17, and 18 GMT (i.e. 2 AM through 11 AM Pacific Daylight Time), with forecasts available approximately 20 minutes after the top of the hour. Initially, the forecasts were presented as a deterministic time of clearing. Following a recommendation by the Traffic Management Unit (TMU) at Oakland Center who were interested in understanding the "risk" associated with the predictions, the forecast presentation was modified to include a probabilistic representation of clearing likelihood by various target times during the high traffic volume period, namely 17, 18, 19, and 20 GMT. The graphic representation of forecasts provided to users is shown in Figure 3.

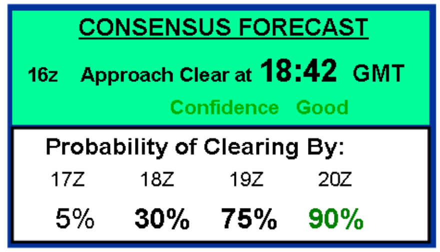

Figure 3. Deterministic and probabilistic display of automated SFO marine stratus forecast.

The experimental automated stratus forecast guidance product was made available to operational users during the summer stratus seasons from 2001-2004. The product became a shared resource for the daily 13 GMT (6 AM PDT) planning conference call for SFO traffic flow strategic planning. It was often cited by users as guidance for decisions to implement GDPs, and it was referenced in the "rule of thumb" guidelines for SFO traffic planners that is provided by the FAA Air Traffic Control System Command Center (ATCSCC).

Following the 2004 stratus season, responsibility for operation and maintenance of the prototype system was transferred to the NWS Forecast Office in Monterey, CA, who continued to provide the forecast guidance product to operational users in an evaluation mode.

2.2 Performance and effectiveness

After operating the prototype for a few years, NWS Monterey undertook an evaluation of system performance and effectiveness during the period 2004-2007 (Delman et al., 2008). The objective of the study was to assess the system impact in improving stratus forecasting performance and, in turn, determine the impact that the system had on traffic flow decisions and decreasing air traffic delay.

The study showed mixed results. It was determined that the forecast guidance provided by the automated system continued to confirm the favorable skill level that was exhibited during the original demonstration period. Specifically, stratus conditions that were predicted to clear with a very high probability by specific target times did, in fact, reliably verify at the expected rate. Furthermore, the forecasts generated by the CWSU were also found to improve after several years experience using the guidance. The study reported that, based solely on forecasts of high clearing probability (i.e. minimal exposure to risk of an incorrect optimistic forecast), there was the opportunity to reduce aircraft delays on approximately 15% of the GDP days throughout the study period, translating to a monetary savings of at least $1.6 million to the airline industry. For perspective, the annual cost to the NWS for maintaining the system is less than $70,000.

In spite of these favorable system performance findings, the study showed that there was almost no measurable positive impact in reduction of GDPs or their associated aircraft delay. Virtually none of the opportunities to proactively release aircraft in anticipation of a high likelihood of clearing was exploited. This strongly suggests that continuing efforts to reduce delay should shift primary focus from improving forecasts to improving strategy for translating forecasts into operational decisions. As stated previously, this finding is extremely germane to the NextGen vision of using improved weather information to improve NAS efficiency.

Findings of the NWS study led to investigation of the hindrances in exploiting the improved forecasting information. It was learned that potential benefits were precluded by the existing strategy for deployment of GDPS, and underlying motivation for adopting changes.

As stated earlier, the decision to implement a GDP based on stratus expectations is initially made as part of a conference call involving stakeholders, specifically ATCSCC traffic managers (who have the ultimate authority for setting GDP parameters), airline operational managers and forecasters, and NWS forecasters, both at the CWSU and the NWS Forecast Office. This decision is made at approximately 13 GMT (6 AM PDT), and consists of three primary parameters: GDP start time, GDP end time, and scope. Planes that are scheduled to arrive during the period defined by the GDP start/stop times are subject to be held on the ground. Scope refers to the number and geographic location (proximity to SFO) of the held planes. In general, the longer the stratus is expected to persist in the approach zone, the longer the duration of the planned GDP, and the broader the scope in terms of impacted aircraft. Historical analysis of GDPs and forecasts shows that typically, *the initial end time of the GDP is planned at roughly two hours after the expected time of stratus clearing.* This represents a very conservative initial strategy to avoid a circumstance wherein held planes are launched prematurely, which would result in demand exceeding capacity, excessive airborne holding, and possible diversions to alternate airports. Excessive airborne holding increases stress and workload for controllers and creates potential safety risk, while diversions are costly to airlines.

An assumption during the product development phase was that there is a fair amount of traffic flow management flexibility to modify established GDPs to exploit an increase in newly anticipated capacity in the near term time frame (less than two hours). Since there are typically many aircraft less than 90 minutes flight time upstream from SFO held on the ground during GDPs, two hours provides sufficient lead time for tactical action to exploit newly anticipated capacity. However, historical analysis shows that changes to GDPs based on an evolving outlook in the stratus forecast are infrequent, and modifications to the GDP occur almost exclusively in a more conservative direction, i.e. when an updated forecast indicates a later clearing than originally thought at the early morning teleconference. It was very rare to observe instances where a GDP was shortened or reduced in scope owing to a stratus forecast modification. This is particularly disappointing in that the skill of the automated forecasts characteristically improves throughout the morning hours following sunrise, peaking in the 15-16 GMT time frame. Delman et al. 2008 report a 20% reduction in mean absolute error in the 2004-2007 forecasts made at 15 GMT compared to those made at 11 GMT.

3. THE RISK AND REWARD OF STRATEGIES TO REDUCE DELAY

Throughout the prototype development process, leadership at the Traffic Management Unit (TMU) at Oakland Center was supportive of a new forecast tool that could aid the stratus forecast, and enthusiastic about its promise for traffic flow efficiency. In practice, however, it was difficult to isolate the impact of the stratus forecast product. During periodic discussions with the TMU, it became apparent that the reliability of the new forecast guidance was an issue in terms of using it to aggressively modify GDPs. The TMU indicated that it would be willing to release ground held planes proactively in anticipation of clearing, provided that the likelihood of clearing prior to the peak demand period was very high, say on the order of 90%. This led to the modification of the system output and display to provide a probabilistic representation of clearing by key target times during the high demand period.

This system modification failed to yield the desired results in terms of reduced ground delay, as evidenced by the NWS study, as traffic flow managers remained reluctant to deviate from the existing strategy of managing with the reduced capacity associated with stratus. What prevents the improved forecast guidance from translating into decisions that reduce delay?

The result seems primarily attributable to the risk/reward structure of increased aggressiveness in reducing delay. Benefits associated with capacity improvement projects are typically estimated in terms of reduced aircraft delay, which is in turn converted to a monetary savings that can be realized by airlines. This conversion is based on items such as crew time, gate fees, fuel loading and fuel burn, etc. These potential benefits, along with less tangible benefits such as passenger goodwill associated with improved on-time performance, are all industry-directed gains of which the airlines (and their passengers) are prime beneficiaries. Of course, the airlines are also exposed to risk penalties in the event of a failed aggressive decision (i.e. early cancellation of GDP program), which are realized as additional fuel costs associated with airborne holding, and diversions to alternate airports.

The airlines, however, are only one voice in the GDP strategy decision. Ultimately, the final authority regarding duration and scope lies with ATCSCC. In addition to risk/reward that can be quantified in monetary terms, there are the overriding considerations of safety risk, a responsibility borne by both the FAA and industry, and controller workload, a burden borne entirely by traffic managers. These are real risk exposures for controllers, but it is difficult to make a relative assessment of the offsetting reward on the traffic management side. There is little question that controllers are motivated to strive for optimum efficiency in management traffic, but their "reward" for squeezing additional efficiency from the system is vague compared to the very real consequences of workload and stress associated with safely managing an unexpected and significant surplus of arrivals. It seems that this is a prime contributor toward reluctance to adopt more aggressive strategies.

4. TOWARD A NEW GDP STRATEGY

In a cooperative effort to support NextGen initiatives, NASA Ames is sponsoring a modest effort to investigate an improved GDP strategy for SFO that would exploit the probabilistic forecast information. This presents an opportunity to implement some aspects of the NextGen vision within the existing operational framework, which could serve as example for developing future methodologies for effective weather data integration. This effort is part of a proposal made to NASA by NWS Monterey, who in turn has sought participation from MIT Lincoln Laboratory to serve as technical lead. Mosaic ATM Inc., UC–Santa Cruz, UC-Berkeley, and San Jose State University are also collaborators in this effort. A kickoff meeting was held in October 2008 to establish objectives and roles.

High level objectives of this initiative have been identified as:

1) Develop an objective methodology for establishing GDP parameters that addresses the risk-reward tradeoff

2) Provide an objective quantification of benefits

3) Quantify risk exposure (frequency and severity) for failed forecasts, and prescribed mitigation plans

4) Formulate a strategy to equitably allot risk and reward amongst impacted airlines

This is an approximately 18-month effort that will culminate with recommendations to FAA for a trial GDP strategy to demonstrate these concepts, with changes made to the operational display to support these recommendations. Outlined here are the key considerations for each objective.

4.1 Analytic methodology for setting GDP parameters

Optimization of ground holding strategy has been the topic of considerable operational research (Ball et al., 2003; Mukherjee et al., 2007). At the outset of this effort, we draw on the work of Mosaic ATM, Inc. (Cook and Wood, 2008), which has done considerable recent work specifically in investigating the translation of weather information into traffic flow strategy using the prototype SFO stratus forecast system as the centerpiece of their analysis. They have made extensive use of the archived deterministic and probabilistic forecast guidance generated by the system, in conjunction with forecasts issued by the CWSU, GDP implementation parameters, delay statistics, and information derived from discussion with traffic flow managers. Their effort establishes an excellent groundwork for moving forward with proposing a new paradigm for GDP implementation.

Their analysis uses a simulation-optimization technique for setting the most critical GDP parameters, i.e. end time and scope. The empirical forecast error is combined with the forecast clearing time to create a probability distribution for clearing time. This distribution is then used to simulate the outcomes of the possible GDP parameter choices for a variety of performance metrics. The chosen GDP parameters are those which minimizes an objective function based on the output probability distributions of the performance metrics and some additional constraints that ensure future extensibility of the program. This model represents a significant improvement over previous such efforts in that it explicitly considers the risk exposure to Air Traffic Control (ATC), and appropriately assigns an exponentially increasing penalty that increases as the actual required GDP exceeds the planned GDP end time.

As an example, the model was used to estimate potential reduction in delay by using an optimized GDP end time, using the 2006 season as a sample data set. The model recommended a reduced GDP for all 20 of the sample GDP events, for an average 59% reduction in GDP duration. Importantly, the recommended end time was too early in only 1 of the 20 events, which corresponds to a similar risk exposure to the current methodology of establishing a GDP end time by simply adding two hours to the expected clearing time.

The model developed by Mosaic ATM to date provides a firm foundation for establishing an objective method of determining GDP parameters that trades off benefit with cost, including risk exposure to ATC. A refined version of this methodology is expected to be central to a recommended GDP strategy.

4.2 Quantitative measure of benefits

The Mosaic ATM report also includes an analysis of estimated benefits on the 20 test days in 2006. Table 1 shows the actual amount of ground delay that occurred on each day, compared to the delay that would have been incurred using the recommended GDP parameters, as well as the delay associated with an "ideal" GDP scenario, which quantifies unavoidable delay. The recommended GDP provided a delay reduction on 18 of the 20 days corresponding to a delay reduction of 16%, which is 32% of the "avoidable" delay. This was converted to a cost savings of $3.01M, using the ATA-provided cost value of $65.80 per minute.

Table 1. Preliminary estimate of SFO delay reduction benefits (minutes) using model-derived GDP parameters [Cook and Wood, 2008]

Date	Actual	Model	Ideal	Actual - Model	Model - Ideal	Actual - Ideal
6/10/06	1339	686	248	653	438	1091
6/28/06	2748	1411	27	1337	1384	2721
6/29/06	4463	3166	1118	1297	2048	3345
7/5/06	3806	3217	926	589	2291	2880
7/12/06	3209	4538	2100	-1329	2438	1109
7/27/06	2729	2260	1718	469	542	1011
7/28/06	4082	1962	1858	2120	104	2224
8/4/06	3243	2545	0	698	2545	3243
8/7/06	6770	5347	6708	1423	-1361	62
8/8/06	3112	2222	1302	890	920	1810
8/14/06	3759	2948	1299	811	1649	2460
8/18/06	3844	3212	1737	632	1475	2107
8/19/06	2346	1406	972	940	434	1374
8/20/06	3172	2967	2425	205	542	747
8/21/06	4654	3221	1232	1433	1989	3422
8/25/06	2798	4903	1708	-2105	3195	1090
8/26/06	2185	2267	1384	-82	883	801
8/27/06	4114	3647	2170	467	1477	1944
8/28/06	3372	2761	1683	611	1078	1689
8/29/06	3145	3011	2900	134	111	245
TOTAL	68890	57697	33515	11193	24182	35375

Traffic managers are not currently provided with this type of objective quantitative estimate of potential benefits to be derived from a successful shortening of GDP duration. This is likely a contributing factor in the reluctance to adopt a more aggressive GDP strategy. An additional project objective, therefore, is to generate this type of benefits estimate on a running basis as associated with each updated forecast. Thus, at each decision point, the operational decision makers will have a more complete information set on which to gauge the reward associated with GDP parameter settings.

4.3 Risk and risk mitigation plans

As described, traffic flow managers have an acute awareness of the consequences associated with premature cancellation of a GDP, namely increased airborne holding causing a substantial increase in controller workload, with the possibility of airborne holding capacity being exceeded and requiring diversions. This represents a substantial operational burden. It was evident in discussing a more aggressive GDP based on new forecast information that there was considerable concern regarding this added risk, and the importance of understanding the potential frequency and severity. The first attempt to address these concerns by converting the deterministic forecast to a probabilistic representation was clearly not sufficient. A new objective will be to provide a more concise quantitative assessment of risk associated with each newly updated forecast. This metric will need to take into account both the expected range of forecast error, and the operational impact in terms of the expected arrival demand profile. Specifically, in addition to providing the probability of clearing by specific target times, the forecast would also include the probability of forecast error exceeding the target time by accumulating increments of 30 minutes, and the expected surplus of demand (number of aircraft) that would exceed capacity during that period of time.

Another possible enhancement would be a prescribed "game plan" for dealing with the premature cancellation of a GDP. For example, in addition to quantifying risk, each risk scenario could be accompanied by a mitigation plan. This could include a predetermination of aircraft (or perhaps a subset of aircraft) to be assigned available slots, and perhaps more importantly, specific identification of aircraft slated for diversion (including the alternate airport). This would need be done consistent with the equitable risk and reward allotment considerations.

4.4 Equitable risk and reward allotment

Once an objective strategy for determining GDP parameters is established, the question remains of how to equitably allot the associated risk and reward. A paradigm shift in GDP implementation strategy would not be considered practicable if it did not address the possibility that individual airlines may be required to incur a level of risk/cost that is not commensurate with their potential reward. This consideration is the subject of parallel work in the field of operations research. For example, Hoffman et al. (2007) presents a "ration-by-distance" methodology for available slot assignment. They present a stochastic model to maximize the throughput into an airport in the presence of weather uncertainty. As its name applies, the approach uses each aircraft's distance (or flight time length) from the destination airport as the primary consideration for assigning an available slot. Recognizing the potential inter-airline inequities that may arrive from the ration-by-distance slot assignment, the methodology is extended to include "assignment exchanges" to constrain the benefits advantages of individual airlines. This type of consideration will need to be included in the proposed strategy in order to gain industry acceptance.

4. SUMMARY

Through sponsorship of NASA Ames, a group of collaborators have begun to investigate the use of probabilistic weather information from the prototype SFO Marine Stratus Forecast System to propose a new strategy for GDP implementation. This effort will provide a simple yet real world example to support development of future NextGen methodologies for integrating weather and air traffic information to optimize NAS efficiency. An objective model which trades off operational risks and costs with potential benefits will be developed to establish GDP parameters based on the probabilistic forecast information. The model will be run repeatedly in conjunction with each stratus forecast update. Objective estimates of both operational risk and benefits will be provided to users with each update to assist in the decision process. The collaboration team intends to coordinate with ATC management to run a real time trial to demonstrate effectiveness and expose deficiencies. Since aggressive strategies for reducing delay currently add individual workload risk with no commensurate personal reward, an appropriate level of support and visibility from ATC management will be required for effectual implementation.

5. REFERENCES

Andrews, M, R. Heuwinkel, J. McCarthy, and S. Brown, 1971: "An update on activities of the Joint Planning and Development Office (JPDO) Weather Integrated Product Team (IOPT) supporting the Next Generation Air Transportation System (NGATS or NexGen)", 23rd Conference on IIPS, American Meteorological Society.

Ball, M.O., R. Hoffman, A. Odoni, and R. Rifkin, 2003: "A stochastic integer program with dual network structure and its application to the ground-holding problem", *Operations Research*, 51, pp. 167-171.

Clark, D.A, and F.W. Wilson, 1997: "The San Francisco Marine Stratus Initiative", 7th Conference on Aviation, Range and Aerospace Meteorology, Long Beach, CA, pp. 384-389.

Cook, L. and B. Wood, 2008: "Translating weather information (convective and non-convective) into TFM constraints – SFO Stratus Forecast System benefits analysis report", NASA NRA Contract No. NNA07BC57C.

Delman, M.J., N.M. Ray, and D.W. Reynolds, 2008: "The Marine Stratus Initiative and effect of stratus on arrival rates at SFO", 13th Conf. on Aviation, Range, & Aerospace Meteorology, New Orleans.

Hoffman, R., M.O. Ball, and A. Mukherjee, 2007: "Ration-by-distance with equity guarantees: A new approach to ground delay program planning and control", Proceedings of the 7th USA/Europe Air Traffic Management Research and Development Seminar, Paper no. 122, Barcelona, Spain.

Ivaldi, C, D.A. Clark, and D. Reynolds, 2006: "Upgrade and technology transfer of the San Francisco Marine Stratus Forecast System to the National Weather Service", 12th Conf. on Aviation, Range, & Aerospace Meteorology, Atlanta.

MCR Federal, Inc., 2004: Ceiling & Visibility at SFO – Marine Stratus Forecast Benefits Estimate. MCR Briefing BR-M021/080-4.

Mukherjee, A. and M. Hansen. 2007: "A dynamic stochastic model for the single airport ground holding problem", *Transportation Science*, vol. 41, no. 4, pp. 444-456.

Strach, W.J., 1991: "The effects of summer time stratus at San Francisco International Airport on the nationwide flow of commercial airline traffic", 1st Aviation Weather Workshop, Kansas City, MO.