10.2 AN OPERATIONAL EVALUATION OF A GROUND DELAY PROGRAM (GDP) DECISION SUPPORT TOOL USING AN INTEGRATED FORECAST OF STRATUS CLEARING AT SAN FRANCISCO AIRPORT

Lara Shisler Cook*, Christopher Provan
Mosaic ATM, Leesburg, VA
Christine Riley, Austin Cross
National Weather Service, Monterey, CA

1. INTRODUCTION

A major challenge for air traffic managers in the San Francisco area is the summertime low altitude cloud layer that develops overnight in the San Francisco Bay. This layer, called marine stratus, has a tremendous impact on San Francisco Airport (SFO) arrivals since it precludes simultaneous arrival operations on closely spaced parallel runways, which reduces the arrival capacity from 60 to 30 flights per hour. Frequently, the stratus layer is anticipated to burn off after the first bank of scheduled arrivals, and a strategic Ground Delay Program (GDP) must be implemented. A GDP sets target arrival rates that keep traffic at or below predicted capacity by assigning ground delays to flights at their departure airports in order to defer excess demand to later time periods with available capacity. If a GDP is issued with rates that are too conservative, unnecessary delay is absorbed on the ground at the departure airports, and arrival capacity is wasted. On the other hand, if the GDP rate rises above 30 flights per hour before the stratus burn-off time then airborne holding and diversions may be necessary. Air traffic controllers prefer to avoid the high cost and risk of holding and diversions, and thus historic practices tended to result in overly conservative programs.

The frequency of marine stratus occurrence at SFO and its tremendous impact on aviation operations motivated the Federal Aviation Administration (FAA) Aviation Weather Research Program (AWRP) to sponsor development of the prototype SFO Marine Stratus Forecast System (MSFS) [1], an automated forecast product designed specifically to predict the time of stratus clearing in the SFO approach zone. Technical development of the prototype was led by MIT Lincoln Laboratory in collaboration with the National Weather Service (NWS), San Jose State University (SJSU), and the University of Quebec at Montreal (UQAM). The topology and meteorology of the San Francisco Bay area allowed for the development of a model that would predict stratus clearing with reasonable accuracy. National Aeronautics and Space Administration (NASA) Ames Research Center first recognized the opportunity presented by the MSFS to investigate the integration of a probabilistic weather forecast with air traffic management (ATM) decision making. Analysis showed that despite the deployment of the MSFS in 2004 and the improved accuracy in forecast stratus clearing times provided by the system, there was no measurable improvement in the GDP planning process or the efficiency of the resulting GDPs. The conclusion was that the probabilistic nature of the forecast product was difficult to interpret for traffic managers, and that in order to improve GDP efficiency, a model would need to be developed to translate the probabilistic forecast into traffic flow management (TFM) decisions [2].

NASA-funded research conducted by Mosaic ATM toward this goal led to the development of the GDP Parameters Selection Model (GPSM). GPSM integrates the forecast of stratus clearing from the MSFS into the current process of modeling and issuing GDPs at SFO. Utilizing historical forecast performance to build a probabilistic error distribution, the model selects GDP parameters that best balance the objectives of minimizing delay and managing risk. GPSM represents one of the first fully developed tools to achieve the major Next Generation Air Transportation System (NextGen) goal of

**Corresponding author address:* Lara Shisler Cook, Mosaic ATM, Inc., lshisler@mosaicatm.com, 801 Sycolin Road SE, Leesburg, VA, 20175.

integrating probabilistic weather forecasts into TFM decision making [3].

The GPSM model and preliminary benefits assessment were described in a paper at the 2009 ATM Seminar [4]. This led to interest by the FAA, who then sponsored the development of a prototype of the GPSM that could be tested by the operational community. GPSM was designed to fit within today's operational environment with no required changes to the tools used to issue GDPs and only minor procedural changes. This allowed for an operational evaluation of GPSM prior to any NextGen changes to the United States National Airspace System (NAS) infrastructure.

In this paper we report on the conduct of the GPSM operational evaluation over 2011 and 2012. We first provide the NextGen vision for ATM-Weather integration in the NAS. This is followed by a brief overview of the design and implementation of GPSM, followed by an overview of the operational evaluation and lessons learned. After an overview of GPSM's benefits to the operational community, a discussion of the impact of forecast accuracy follows prior to concluding.

2. ATM-WEATHER INTEGRATION IN THE NAS

Aviation operations are significantly impacted by weather. Weather delays account for 70 percent of the $41 billion annual cost of air traffic delays within the NAS [5]. Approximately two thirds ($19 billion) of weather delays are considered to be avoidable, i.e., unnecessary if weather was forecast with 100% accuracy [6].

The NextGen Concept of Operations [3] defines eight new capabilities to be developed, one of which focuses on assimilating weather into decision making. Because of the profound impact adverse weather has on air transportation, there is a major focus on developing new aviation weather information capabilities that will help stakeholders at all levels make better weather-related TFM decisions [7]. Those capabilities will be developed based on three major tenets:

- A common weather picture for all air transportation decision makers and aviation system users;
- Weather directly integrated into sophisticated decision support capabilities to assist decision makers;
- Use of internet-like information dissemination capabilities to realize flexible and cost-efficient access to all necessary weather information.

NextGen decision support tools (DSTs) will directly incorporate probabilistic weather data and aid in the human interpretation of probabilistic weather. This will allow decision makers to determine the best response to mitigate the potential operational impact of weather on both a tactical and strategic time horizon while minimizing delays and restrictions. Using automation to better manage uncertainties associated with weather minimizes capacity limitations and reduces the likelihood of overly conservative actions while maintaining acceptable levels of risk across the NAS.

The NextGen Joint Planning and Development Office (JPDO) sponsored the development of a plan for the integration of ATM and Weather in 2010 [8]. As noted in this plan, the Weather–ATM Integration Working Group (WAIWG) of the NAS Operations Subcommittee of the FAA's Research, Engineering and Development Advisory Committee (REDAC) conducted a 12-month study to examine the potential benefits of integrating weather and ATM. The report of this committee made several recommendations regarding the potential for weather integration to help reduce delays by improving the quality and method of use of weather information and integrating weather support in the NAS.

This plan includes a conceptual flow of weather integration, shown in Fig. 1, which has been accepted by the weather and ATM communities. It serves as an overview of the envisioned NextGen weather concept for enhancing ATM decision making in the face of adverse weather.

Five levels of weather integration were also defined, each moving closer towards the conceptual vision depicted in Fig. 1.

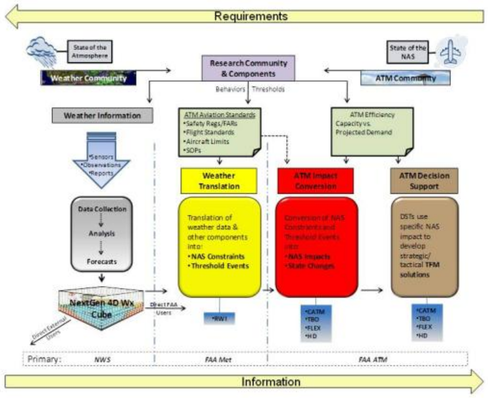

Figure 1. Conceptual flow of weather integration

- Level 0: Stand-Alone Displays – Weather data are displayed on dedicated interfaces separate from any ATM data
- Level 1: On-the-Glass Weather Integration – Weather overlays are added to ATM tools. Examples include the Corridor Integrated Weather System (CIWS) added to the Traffic Situational Display (TSD) and the Weather and Radar Processor (WARP) displays on controllers' Display System Replacement (DSR).
- Level 2: Translated Weather Integration – Automation translates weather data into a constraint, such as a Weather Avoidance Field (WAF). Other examples include the wind shear function of the Integrated Terminal Weather System (ITWS) and the Route Availability Planning Tool (RAPT).
- Level 3: Impact Integration – User-in-the-Loop Tools – Built upon Level 2 technologies, they ingest NAS traffic and other data to determine impact.
- Level 4: Machine-to-Machine (M2M) Integration – Constraints (Level 2) and impacts (Level 3) are used by DSTs through M2M integration. Tools provide automated recommendations for ATM decisions without the need of human interpretation or translation.

There are currently no operational Level 4 integration tools in use, or even Level 3. The ATM-Weather Integration Plan identified three maturing capabilities that are not yet in operational use but are the most mature new concepts in ATM-Weather integration. These included Integrated Departure Route Planning (IDRP), Collaborative Trajectory Options Program (CTOP), and GPSM. Of these three capabilities, only GPSM meets the criteria for a Level 4 DST by moving beyond impact assessment and providing actual automated recommendations for Traffic Management Initiatives (TMIs). Thus, during the summer of 2012, GPSM became the first Level 4 ATM-Weather integration DST to begin operational trials.

3. GPSM OVERVIEW

GPSM is built around a core optimization model that evaluates large sets of GDP parameters and selects the parameters that minimize a weighted sum of ground delay and expected airborne holding subject to certain risk mitigation constraints. The parameters that are optimized as part of the model include start time, end time, airport arrival rate (AAR), and, optionally, geographic scope. For each candidate set of GDP parameters, metrics such as excess ground delay (which occurs when GDPs are too conservative), airborne holding (which occurs when GDPs are too aggressive), and a variety of additional risk metrics are calculated for each possible stratus clearing time based on an error distribution built around the MSFS clearing time forecast. The probabilities of each clearing time are used to calculate an expected value for each metric under each GDP scenario. Then the metrics are combined into an objective function that calculates a cost for a particular GDP scenario, and the GDP scenario with the lowest cost is selected as the recommended program to implement. References [3, 9] provide a more detailed description of the model.

The implementation of GPSM for SFO is actually a combination of two underlying component models: a weather translation model that translates the forecast of stratus clearing into a probabilistic estimate of capacity, and the optimization model that uses that probabilistic capacity estimate and traffic data to determine the optimal GDP parameters.

The weather translation component in the SFO implementation uses a historical database of MSFS forecast performance since 1997 (a 15-year archive) to build an error distribution in real-time around any newly generated forecast. This error distribution is dependent on forecast run time and the automated MSFS confidence rating assigned to that forecast. A single outcome from the probabilistic clearing time distribution is translated into an arrival capacity scenario by assuming an arrival rate of 30 flights per hour prior to stratus clearing and 60 flights per hour thereafter.

This weather translation component is specific to the SFO implementation of GPSM. However, because it is independent from the core GPSM optimization model, it can be replaced when operating at another airport by a Weather Translation Model (WTM) that uses the appropriate weather forecast products that best capture the weather factors influencing capacity at that airport and that can translate these weather forecasts into probabilistic estimates of airport arrival capacities.

The GDP parameters that are selected by the optimization component model are displayed in a GPSM table integrated into the web interface for the MSFS forecast tool (Fig. 2). All FAA and collaborative users can access this web page. The GPSM table also provides two alternative sets of parameters to show users the impact of issuing a more aggressive or conservative GDP than recommended. A variety of delay and risk metrics are displayed for the recommended and alternative parameters as well as for any SFO GDP currently in place.

Figure 2. GPSM recommendations integrated with the MSFS webpage.

4. LESSONS LEARNED FROM THE OPERATIONAL EVALUATION

The GPSM evaluation was conducted during 2011 and 2012 during the stratus season, which runs from May 15th to October 15th of each year. The evaluation was split into two periods: a shadow evaluation during the 2011 stratus season followed by a full operational evaluation in 2012.

During the 2011 shadow evaluation, GPSM was not used as part of the operational decision making process for issuing SFO GDPs. Instead, a shadow position was designated at the Air Traffic Control System Command Center (ATCSCC) for the personnel conducting the GPSM evaluation. This position neighbored the position responsible for monitoring and issuing west coast airport TMIs, which allowed the staff at the shadow position to monitor the same data sources and to listen to any planning conference calls or discussions. The operational west coast position was not given access to the GPSM recommendations, and GDPs were planned and issued under the existing procedures. The shadow position had all of the same ATM tools at their disposal and additionally had access to GPSM. The recommendations provided by GPSM were monitored from the initial GDP planning time frame until the stratus cleared and the GDP was cancelled. Observations made by the shadow position were captured in the National Traffic Management Log (NTML).

The shadow position personnel were tasked to monitor GPSM to determine answers to the following questions:

- Are GPSM's recommendations for the initial GDP sound?
- How stable are GPSM's recommendations over time as traffic and weather forecasts evolve?
- Do the improved forecasts in the 15Z hour, when visible satellite imagery is available, result in GPSM recommendations that are better and that can guide revisions?

- Are there user interface and human factors considerations that should be addressed before operational use?
- Is it clear to the traffic managers on which days GPSM's use is appropriate (i.e., when ceilings are due to "typical" stratus, for which the forecast system is designed)?

One of the most important outcomes of the shadow evaluation was related to the procedures for communication between meteorologists and traffic managers. As mentioned previously, the MSFS is designed to forecast stratus clearing during typical summer weather patterns, wherein the daily cloud dissipation mechanism is dominated by local physical processes, as is common during the warm season. When larger scale transient weather systems impact the region, as is more common in winter months, there can be a significant degradation of forecast accuracy. The system is designed to automatically recognize some of the conditions under which degraded performance might be expected and issue an indication of lower forecast confidence. However, there are additional days on which the human forecaster can recognize other unusual conditions that may impact the automated forecast quality. Forecast performance in turn impacts the quality of GPSM recommendations at SFO, which rely on the MSFS forecasts to generate probabilistic capacity scenarios. Since the MSFS is fully automated and will generate forecasts of clearing as long as the system's sensors detect that low ceilings are in place in the Bay Area, this provides a challenge for traffic managers who need to recognize the days for which the GPSM recommendation is not intended. Though it may be straightforward for a meteorologist to determine the applicability of the MSFS forecast to GPSM on a given low ceiling day, this is far from obvious for traffic managers. During the shadow evaluation, the staff at the shadow position frequently needed to call the Oakland Center Weather Service Unit (CWSU) to better understand the weather situation on a given day in order to determine whether or not the evaluation of GPSM on that day was appropriate. Though those conversations were helpful, they were time consuming and even sometimes confusing. The meteorologists providing the synopsis of the current conditions and their judgment regarding the forecast often used meteorological terminology, and it was not always clear to the specialists at the GPSM shadow position what the implication was for the use of GPSM in the planning process. At the end of the 2011 stratus season, it was clear that a better procedure was needed in order to more clearly communicate to the traffic management community whether or not GPSM should be used.

In preparation for the 2012 operational evaluation, procedures for communicating whether or not GPSM's use was appropriate on a given day were developed. The work conducted to develop these procedures was an excellent example of inter-agency collaboration between the FAA and the NWS. NWS staff at both Monterey and at the CWSU and the MSFS technical team at MIT Lincoln Laboratory worked together to develop a concept of classifying a low ceiling day as one of the following three categories:

- GPSM: High Confidence – Typical stratus day for which the MSFS was designed. The forecast system and GPSM can be used for guidance with high confidence.
- GPSM: Low Confidence – Typical stratus day, but with some other weather influences contributing to lower confidence in the forecast system. GPSM can still be used for guidance, but with caution.
- Not GPSM – Low ceilings are caused primarily by weather factors other than typical marine stratus, or no ceilings are present. GPSM use is not appropriate.

The meteorologists worked together to define the key weather factors and their threshold parameters that would classify a day into each of these categories. This process facilitated a common understanding of the automation between the developers of the forecast system (MIT Lincoln Lab) and the meteorologist using that system, and it also resulted in clear procedures for the CWSU for doing classification such that individual human judgment or subjective interpretation could be minimized. By making the classification specific to GPSM, it largely removed meteorological interpretation from the role of the traffic manager. A joint training session was conducted between the FAA and NWS on GPSM (Fig. 3), and the NWS conducted their own internal training sessions for all of their meteorologists on the classification procedures developed.

Figure 3. GPSM Training at the NWS Monterey.

These new procedures were put in use at the start of the 2012 operational evaluation. During the 2012 status season, unlike 2011, GPSM's recommendations were available to all stakeholders at SFO and were used to guide GDP decision making. The CWSU was responsible each morning for evaluating the weather conditions and manually entering into the MSFS display the day's classification (GPSM High Confidence, GPSM Low Confidence, or Not GPSM) and their reasoning.

At first, the CWSU meteorologists were overly conservative in classifying a day as a "GPSM Day", but by July, they were more comfortable with the new procedures and had a better understanding of the weather conditions for which the MSFS was designed to provide reliable forecasts. The new procedures worked well for the remainder of the season and proved to be an excellent of example of the importance of the "meteorologist-in-the-loop" even when utilizing Level 4 ATM/Weather Integrated DSTs. Additional observations from the operational evaluation related to "meteorologist-in-the-loop" and meteorologist-traffic manager communications included:

- The CWSU forecasters have expertise in recognizing weaknesses and biases in the automated forecasts generated by MSFS.
- Meteorologists can recommend more aggressive or conservative GDP alternatives based on their familiarity with system biases.
- The procedures provided a way for the forecasters to communicate to the traffic managers in a way that did not require an understanding of "meteorologist-speak".

Figure 4. Ken Venzke, Meteorologist in Charge, FAA Oakland ARTCC.

Another lesson learned during both the 2011 and 2012 evaluations was to expect the unexpected when it comes to weather. Despite years of historical observations and data that show a fairly consistent number of typical stratus days per year, 2011, and even more so, 2012, proved to be outliers.

In 2011, the beginning of the stratus season was plagued with a much higher rate of rain fall and non-stratus ceilings in June through mid-July than is typical at that time of year. In fact, a record was broken for the wettest June on record in 2011.

2012 again resulted in a slow start to the stratus season, due in part to a stagnant omega block over North America which persisted through July. By July 15th, there had only been two GPSM days versus an average of nearly 17 typical stratus days up to that point over the previous 5 seasons. The small number of opportunities to use GPSM slowed user acceptance of the tool and made progress at the CWSU towards adjusting to the new procedures surrounding the classification of days difficult. The initial impression of many users was that GPSM could not be relied on to help plan GDP parameters at SFO.

The pattern quickly changed in mid-July 2012 as the weather became more typical and the NWS refined their procedures for designating GPSM days. The second half of the month added an additional eight GPSM days, and after August and September, the total count of GPSM days was 30. Even with this increase in activity, the total number of GPSM days in 2012 was well below the number of typical stratus days in previous years, as shown in Fig. 5.

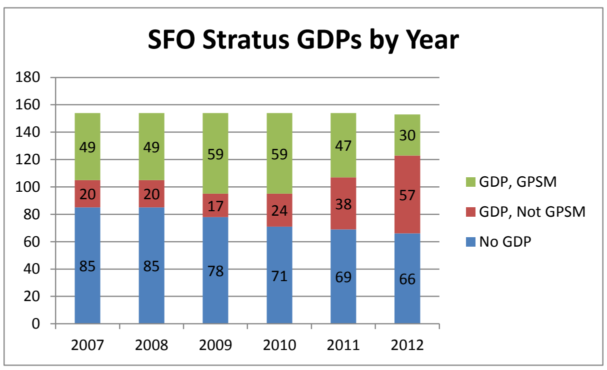

Figure 5. 2011 and 2012 weather patterns resulted in fewer opportunities for evaluating GPSM.

This experience further emphasized that any ATM-Weather Integrated DST must be prepared to account for uncertainty with regards to the weather, and sometimes this uncertainty occurs in ways completely unanticipated, as it did with the unusual weather patterns the past two summers.

5. BENEFITS

The benefits measured during the 2012 operational evaluation of GPSM are summarized as follows:

- There were over 1,600 fewer minutes of initial delay per GDP (a 20% reduction) when GPSM recommendations were followed than when they were not.
- Planned use of arrival slots post-clearing improved by 29% over 2009-2011 levels when GPSM recommendations were followed.
- When adjusted for changes in forecast errors, planned use of arrival slots was improved by 62% relative to 2009-2011, and unnecessary delay was at its lowest levels since 2007 on days where GPSM recommendations were followed.
- Even when GPSM recommendations were not strictly followed, after adjusting for changes in forecast errors, planned use of arrival slots was 42% more efficient than the 2009-2011 average, and unnecessary delay was comparable to recent years even though 2012 traffic levels were substantially higher.
- GPSM benefits were achieved with negligible increase in risk.

These results indicate that GPSM can provide benefits in terms of reduced delays and increased capacity utilization at SFO during typical summer stratus events.

6. THE IMPACT OF FORECAST ACCURACY

NextGen DSTs like GPSM do not negate the importance of improving the accuracy of weather forecasts. One outcome of the GPSM operational evaluation was the question from the user community on whether or not improvements to the MSFS could provide additional benefits when integrated with GPSM. Any improvements to the forecast models in the MSFS would potentially require significant investment, so it is important to understand and quantify the benefits of any improvements before making any investment decision.

Quantifying the benefits to the NAS users due to new or improved weather forecasts has historically proven to be difficult. With a model like GPSM, one that uses forecast performance data to provide ATM recommendations, analyses can now be conducted to fully quantify the impact of improvements to forecast quality on ATM decision making.

Fig. 6 shows how the expected GPSM benefits vary based on forecast accuracy using the data from the 2011 stratus GDPs. The x-axis is the percent improvement to the forecast accuracy while the y-axis is the percent reduction in delay relative to actual 2011 GDPs that could be achieved by utilizing GPSM for that specific forecast accuracy. The 0% point shows the benefits being achieved today with the given forecast quality, a delay reduction of approximately 25%.

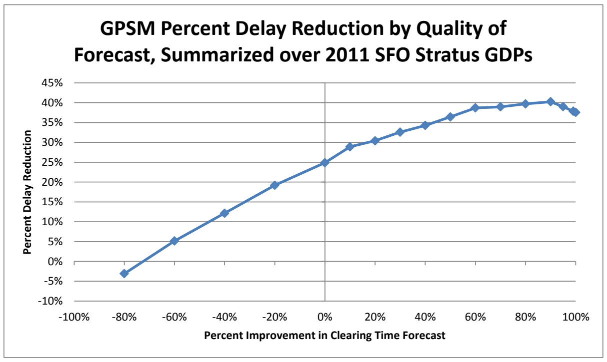

Figure 6. The impact of forecast quality improvements on delays at SFO can be directly measured using GPSM.

There is definite value in improving the quality of the forecasts generated by the MSFS, even by small amounts. A larger improvement of 50% would have reduced overall delay by 36% in 2011, compared to a baseline value of 25%. But even just a 10% improvement in forecast quality would have reduced delay by over 11,000 minutes, increasing the delay reduction from 25% to 29%.

The peak amount of delay that can be reduced from the actual GDPs is at 40% (some amount of delay is needed in order to maintain a 30 rate prior to stratus clearing, so this can never reach 100%). This maximum point is reached when the forecast quality is improved by 90%. But an improvement of 60% results in delay reduction benefits of 39%, just one point below the value when forecast improvement is at 90%. Thus, there is little value in improving the forecast past a 60% reduction in forecast errors, at least where the use of GPSM is concerned.

After reaching a certain threshold in forecast quality improvement (at some point greater than 90%), there are no additional benefits by improving the quality of the forecast further.

A statistical guidance model like GPSM can still provide benefits even when the forecast quality is low. Since the model accounts for the forecast accuracy in its recommendations, the science will work regardless of the forecast quality, though with diminishing returns as forecast quality degrades. The wider the distribution in historical forecast errors, the more conservative the recommendation provided by GPSM, thus increasing the likelihood of more unnecessary delay. Assuming that the GDPs issued by human decision makers remain at the current level of efficiency, there is some point reached in degrading forecast quality at which GPSM can no longer provide benefits, though this level of quality is fairly low.

7. CONCLUSIONS AND NEXT STEPS

Though there were clear measurable benefits to the use of GPSM during the operational evaluation, the evaluation also brought to light the human factors issues surrounding the deployment of an ATM-weather integrated DST. Tools that use weather forecasts will have to address uncertainty, and this uncertainty makes users slow to understand and accept these kinds of DSTs. No matter how well-designed the tool, there will still be days where the weather evolves in a completely unanticipated way, and a decision made on an earlier forecast will not be optimal. Users have to understand that these types of tools have to be evaluated over a large sampling of days in order to truly capture the expected outcomes and benefits. If an ATM-weather integrated DST provides significant benefits 99 out of 100 times it is used, should it not be deployed because in one out of 100 cases the DST's recommendations were not acceptable? What if the first time a user is exposed to the tool is the one in 100 times that the recommendations were unacceptable? How do you gain user acceptance? The GPSM evaluation brought these issues to light, but there are no clear answers on how to successfully address the challenge of user acceptance of tools that are based on probabilistic data.

The other human factors-related discovery was the importance of the role of the meteorologist-in-the-loop. Though a level 4 ATM-Weather integrated tool fully automates all steps of using weather forecasts to guide ATM decisions, the role of the forecaster in the process is most likely still going to be critical to the successful use of an integrated tool, as discovered with GPSM. Without the meteorologists' guidance on whether or not the use of the MSFS and GPSM was appropriate, ATM users would either have to become experts themselves in interpreting the weather conditions driving the low ceilings or would be trying to use the tool under conditions it was not designed to support.

An area of further research is the application of GPSM to other airports using available or newly designed airport-specific WTMs. SFO provided a unique opportunity for testing the concept of integrating a probabilistic weather product with TFM decision making due to the fact that there was already an operational automated forecast product available (MSFS) with an established forecast error profile, and that the translation to capacity predictions is straightforward due to the dependence of SFO capacity primarily on the single forecast dimension of stratus clearing time. WTMs for other airports will likely require more complex models that consider a range of different weather factors, such as wind speed, wind direction, ceilings, visibility, and convection. Probabilistic predictions will be made more difficult by the interdependence of these various forecast dimensions.

Though it is unclear whether or not GPSM will transition from a prototype to a future operational tool in the Traffic Flow Management System (TFMS), the conduct of the operational evaluation was an important step towards better understanding both the challenges and potential benefits of a Level 4 ATM-weather integrated DST.

ACKNOWLEDGEMENTS

This work was funded by NASA Ames Research Center (ARC) under contract number NNA11AA18C. The authors would like to thank Shon Grabbe and William Chan at NASA ARC for their support of this effort. The GPSM operational evaluation required the involvement of many individuals across multiple organizations. We appreciate the collaborative spirit that existed amongst many personnel at NASA, FAA, NWS, and the technical team of Mosaic ATM and MIT Lincoln Laboratory. Lastly, we acknowledge the significant work done by many individuals at the FAA ATCSCC to support and conduct the evaluation, led by Ed Corcoran and Daniel Horton.

Figure 7. Ed Corcoran, FAA ATCSCC; Lara Shisler Cook, Mosaic ATM; Dan Horton, FAA ATCSCC; Christine Riley and Austin Cross, NWS Monterey

REFERENCES

[1] Clark, D. A. and Wilson, F. W., "The San Francisco marine stratus initiative," 7th Conference on Aviation, Range and Aerospace Meteorology, Long Beach, CA, pp. 384-389, 2003.

[2] Clark, D., "Investigating a new ground delay program strategy for coping with SFO stratus," Aviation, Range, and Aerospace Meteorology Special Symposium on Weather - Air Traffic Management Integration, 89th AMS Annual Meeting, Phoenix, AZ, 11-15 January, 2009.

[3] Joint Planning and Development Office, "Concept of operations for the next generation air transportation system v 2.0," June 2007.

[4] Cook, L. and Wood, B., "A model for determining ground delay program

parameters using a probabilistic forecast of stratus clearing," Eighth USA/Europe Air Traffic Management Research and Development Seminar, Napa, CA, June 2009.

[5] Joint Economic Committee Majority Staff, "Your flight has been delayed again; flight delays cost passengers, airlines, and the U.S. economy billions," May 2008.

[6] Weather-ATM Integration Working Group, "Research, engineering and development advisory committee," October 2007.

[7] Weather Integrated Product Team, Joint Planning and Development Office, "Weather concept of operations, v 1.0," May 2006.

[8] Joint Planning and Development Office: ATM-Weather integration plan, "Where we are and where we are going, v2.0," September 2010.

[9] Cook, L. and Wood, B., "A model for determining ground delay program parameters using a probabilistic forecast of stratus clearing," Air Traffic Control Quarterly, Volume 18 (1), 2010.