An Operational Evaluation of the Ground Delay Program (GDP) Parameters Selection Model (GPSM)

Lara Shisler, Christopher Provan – Mosaic ATM
Dave Clark – MIT Lincoln Lab
William Chan, Shon Grabbe – NASA Ames
Ken Venzke, Christine Riley – NWS
Dan Gilani, Ed Corcoran – FAA

ATM Seminar, Chicago, IL
June 11, 2013
GDP Parameters Selection Model (GPSM) Background

- NASA funded Mosaic ATM to explore how the use of the probabilistic forecast of stratus clearing could be integrated with ATM decision making to achieve benefits at SFO.
- GPSM is the resulting software prototype designed to provide guidance to decision makers in selecting GDP parameters for SFO airport during summer stratus events.
- Provides recommended GDP parameters for SFO based on probabilistic forecast of the clearing of stratus, bridging the gap between the forecast product and the Flight Schedule Monitor (FSM) used to issue GDPs.
- Provides relative indication of risk and benefit of the recommended GDP parameters vs. alternative options given the uncertainty in the forecast.
GDP Parameters Selection Model (GPSM) Overview

- Historical forecasts are used to create the forecast error distributions by forecast type (run time and forecast confidence)
- The forecast errors are applied to the latest forecast to create a distribution of all possible clearing times
- All clearing times and their probabilities are then evaluated against each GDP scenario

Min Cost = Optimal GDP End Time
Cost Function to Select Optimal GDP Parameters

Unnecessary Ground Delay
Airborne Holding Delay
Holding Queue Size
Cumulative Cost

Weighted Objective Cost

GDP End Time (GMT)

Min Cost @ 19:15

Increasing risk of over-delivery (Airborne holds and diversions)

Increasing risk of unnecessary ground holds
### Initial GDP Guidance

**GPSM: High Confidence**

<table>
<thead>
<tr>
<th>11Z GDP RECOMMENDATIONS</th>
<th>-&gt; Clear at 17:19 GMT</th>
<th>[GOOD]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Data</td>
<td>12:30 GMT</td>
<td></td>
</tr>
<tr>
<td>No GDP</td>
<td>Alt-1</td>
<td>Primary</td>
</tr>
<tr>
<td>Start Time</td>
<td>14:15</td>
<td>14:15</td>
</tr>
<tr>
<td>End Time</td>
<td>19:14</td>
<td>19:44</td>
</tr>
<tr>
<td>Scope</td>
<td>n/a</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>+CZV_AP</td>
<td>+CZV_AP</td>
</tr>
<tr>
<td>AAR</td>
<td>n/a</td>
<td>45+n/a</td>
</tr>
<tr>
<td></td>
<td>60@17:45</td>
<td>60@18:15</td>
</tr>
<tr>
<td>Risk Exceed Max Queue</td>
<td>*****</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td>97%</td>
<td>22%</td>
</tr>
<tr>
<td>Benefit Delay Reduction</td>
<td>$$$$$$$</td>
<td>$$$</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>41%</td>
</tr>
</tbody>
</table>

**Expanded statistics**

- GPSM Questions/Feedback
- ATC/SCC Operational Support

### Continuous Monitoring

**GPSM: High Confidence**

<table>
<thead>
<tr>
<th>15Z GDP RECOMMENDATIONS</th>
<th>-&gt; Clear at 16:34 GMT</th>
<th>[GOOD]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Data</td>
<td>15:52 GMT</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>Alt-1</td>
<td>Primary</td>
</tr>
<tr>
<td>Start Time</td>
<td>14:15</td>
<td>15:52</td>
</tr>
<tr>
<td>End Time</td>
<td>19:44</td>
<td>19:44</td>
</tr>
<tr>
<td>Scope</td>
<td>1000 +CZV_AP</td>
<td>1000 +CZV_AP</td>
</tr>
<tr>
<td></td>
<td>+CZV_AP</td>
<td>+CZV_AP</td>
</tr>
<tr>
<td>AAR</td>
<td>7-8-7/8-7-8-7 /8-7-8-7 /8-7-8-7 /8-15-15-15 /15-15-15</td>
<td>45@16:45</td>
</tr>
<tr>
<td></td>
<td>60@17:45</td>
<td>60@18:15</td>
</tr>
<tr>
<td>Risk Exceed Max Queue</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>3%</td>
<td>13%</td>
</tr>
<tr>
<td>Benefit Delay Reduction</td>
<td>$</td>
<td>$$$$$</td>
</tr>
<tr>
<td></td>
<td>11%</td>
<td>70%</td>
</tr>
</tbody>
</table>

**Expanded statistics**

- GPSM Questions/Feedback
- ATC/SCC Operational Support
ATM Weather Integration Plan: Levels of Weather Integration

- **Level 0:** Stand-Alone Displays
- **Level 1:** On-the-Glass Weather Integration
  - Corridor Integrated Weather System (CIWS) added to the Traffic Situational Display (TSD)
  - Weather and Radar Processor (WARP) displays on controllers’ Display System Replacement (DSR).
- **Level 2:** Translated Weather Integration
  - Weather Avoidance Field (WAF)
  - Wind shear function of Integrated Terminal Weather System (ITWS)
  - Route Availability Planning Tool (RAPT)
- **Level 3:** Impact Integration
  - Integrated Departure Route Planning (IDRP)
- **Level 4:** Machine-to-Machine Integration – Constraints (Level 2) and impacts (Level 3) are used by DSTs through M2M integration. This level of integration provides automated recommendations for ATM decisions without the need of human interpretation or translation.

**GPSM is the first Level 4 ATM/Weather Integrated DST to be operationally evaluated**
GPSM and the NextGen Vision of Weather Integration

GPSM 2011 Shadow Evaluation

- **Goal:** Evaluation of GPSM’s recommendations for SFO GDPs by traffic managers in a real-time environment independent from the actual decision making in order to:
  - Gain better familiarity and understanding of the product
  - Assess the robustness and readiness of the product for operational decision making
  - Identify potential enhancements
  - Assess potential benefits

- **Approach:** Staff a new position with all tools used in running SFO GDPs, including the GPSM prototype, to monitor and evaluate GPSM’s recommendations while observing actual decision making. Actual decision makers do not have access to GPSM, providing an environment where GPSM recommendations can be compared to decisions made without GPSM, and benefits can be captured.
Shadow Evaluation Lessons Learned #1

Expect the unexpected in the weather

- Higher number of non-typical stratus events in 2011, increasing even more in 2012
- 2011 due to high amount of rain fall/non-stratus ceilings in June through mid-July
- 2012 due in part to “stagnant omega block” through mid/late July

SFO Stratus GDPs by Year

Not a typical pattern:
- 10 less GPSM days by 7/15 in 2011, and 7 less again in 2012
- Decreasing GPSM days from 2009 through 2012, with significant drops in 2011 and 2012
Shadow Evaluation Lessons Learned #2

The importance of the meteorologist-in-the-loop

- CWSU forecasters have expertise in recognizing weaknesses and biases in automated forecast system
- Meteorologists can identify days to disregard the automated forecasts, and GSPM, entirely
- Can recommend more “aggressive” or more “conservative” GDP alternatives based on their familiarity with system biases
- This important role was not realized in the shadow mode evaluation

Ken Venzke, Meteorologist In Charge, FAA Oakland ARTCC
NWS Preparation for 2012 Operational Evaluation

- GPSM Training
- New Universal Language (get everyone speaking the same language)
- Not a forecast for the Terminal

Classification:
- GPSM, High Confidence
- GPSM, Low Confidence
- Not GPSM
2012 Operational Evaluation

- All FAA facilities and NAS users had access to GPSM’s recommendations for SFO GDP parameters.
- The CWSU provided guidance on whether or not the use of the MSFS and GPSM was appropriate at the start of each morning.
- On “GPSM days” the recommendations were discussed on the planning telcon, and users were encouraged to follow the recommendations for the GDP.
2012 Challenges

• Unusual weather patterns
  – Only 2 GPSM days by July 15 (usually 17 at this point in time)
  – Huge impact on initial user impression of GPSM’s usefulness

• New CWSU procedures for classifying the days
  – Took until mid/late July before use of procedures across meteorologists stabilized

• Extraordinary challenges with sensor equipment due to construction at SFO
  – Key sensors were not available to the forecast models in the MSFS
  – True impact to forecast accuracy is not known, but forecast bias shifted from an average of -9 minutes (clearing 9 minutes later than forecasted) in 2009-2011, to 20 minutes in 2012, an approximate **30 minute shift in forecast bias**
On GPSM days where GPSM recommendations were followed for the initial GDP, delays were on average **1,630 minutes less per GDP**.

Unnecessary delay (actual – ideal) has gone up in 2012 due to:

- Traffic increase of ~20%
- 11Z forecasts errors went from clearing 10 minutes earlier than forecasted in 2011 to 20 minutes later on GPSM days in 2012
- Stratus clearing times 5 minutes earlier on average on 2012 GPSM days
Missed Planned Slots
Based on Initial GDP Parameters

- Actual missed planned slots in previous years since 2009 were consistent, with very slight improvements year-over-year.
- Significant drop is observed in 2012 when GPSM recommendations were followed, a 26% reduction from 2011. This drop increases to 60% when forecast errors are adjusted to be consistent with previous years.
- No change on days when GPSM was not followed, but a significant impact is observed when forecast errors are adjusted to be consistent with previous years, showing GPSM’s impact on decision making even when recommendations are not strictly followed.
Unnecessary Delay by Year
Based on Final GDP Parameters

- Absolute unnecessary delay nearly 1000 minutes greater per GDP in 2012 than previous years.
- When forecast errors are adjusted to be consistent with previous years, unnecessary delay drops to its lowest level since 2007, despite 20% traffic increase between 2011 and 2012, and a 30% rise since 2007.
- When GPSM recommendation was not strictly followed, the adjusted unnecessary delay is approximately 500 minutes higher per GDP than on days that GPSM was followed.
2012 GPSM Assessment Summary

- GPSM utilization reduced delay
  - 1,630 less minutes of delay per initial GDP when GPSM used, a 20% reduction
- Planned use of arrival slots post-clearing improved
  - 26% less wasted planned slots than 2011, an increase in arrival slot utilization by 10 slots per GDP
  - Benefits even larger at 60% when adjusted for changes in forecast errors in 2012
- Unnecessary delay was at its lowest levels since 2007 (when adjusted for changes in forecast errors) on days where GPSM recommendations were followed, despite increase in traffic of ~30%
- Benefits achieved with negligible increase in risk
2012 GPSM Assessment Summary

• Further study/enhancements indicated
  – How to better plan for the transition from a 30 rate to a 60 rate
  – How to minimize the changes in recommendations as traffic data changes
  – Improving the forecasters’ understanding of when GPSM’s use is appropriate
  – Value of improvements to the SFO Marine Stratus Forecast Guidance models and evaluation of the forecast model run times
How Does Forecast Accuracy Impact GPSM Potential Delay Savings?

- There is value in improving forecast accuracy
- There is little value in improving forecast accuracy past ~60%
- A statistical guidance model like GPSM can still provide benefits even when the forecast quality is degraded.
Next Steps in Research

• Can the GPSM concept be applied at other airports?

**GPSM at SFO**

SFO Marine Stratus Forecast System → Historical Forecast Performance → Probabilistic Capacity Estimate → GPSM → SFO GDP Parameters Recommendation

**GPSM at Other Airports**

Weather Forecast → Weather Translation Model → Probabilistic Capacity Estimate → GPSM → GDP Parameters Recommendation

• GPSM is separate from the translation of weather into capacity estimates, allowing for use at other airports with customized WTMs

• Research is required to understand the weather product requirements at other airports that can result in GPSM providing benefits
Challenges for the Future in the Deployment of Weather-Integrated DSTs

- Controller perception of benefits
- Use of probabilistic information
- One “bad day” can ruin the reputation of a statically-based tool
- Time lag between testing prototypes and operational deployment
Questions

Lara Shisler  
Mosaic ATM, Inc.  
lshisler@mosaicatm.com

Dave Clark  
MIT Lincoln Lab  
davec@ll.mit.edu