Service Level Expectation Setting for Air Traffic Management: Project Overview

September 12, 2014

Sponsors: FAA Air Traffic Organization – Mission Support Services: Rich Jehlen, Rob Hunt, Yong Li
NEXTOR-II Team: Mike Ball, Alex Estes, Prem Swaroop, U of Maryland
Cindy Barnhart, Chiwei Yan, MIT
Mark Hansen, Lei Kang, Yi Liu, UC Berkeley
Vikrant Vaze, Dartmouth
Current Practice on TMI Planning

Strategic planning telecons

TMI decisions
Operational Challenge

• Flight operators participate in strategic TMI planning by verbal input. Operators can sometimes have a disproportionate influence on decisions that affect a broad range of others who are less vocal.

• Discussion focuses on specific parameters rather than performance goals.

• Different traffic managers may create different plans for the same situation.

• The planning process is ad-hoc and subjective.
SLE Concept

• The Service Level Expectation (SLE) setting project has produced a conceptual approach and prototype software tool designed to address the above deficiencies.

• The SLE concept takes into account the input of all involved flight operators and generates an output that represents a consensus of those flight operators in making Traffic Flow Management Initiative (TMI) decision.
SLE Concept

- The SLE mechanism allows operators to submit quantitative input that represent their preferred system performance goals (capacity, predictability and efficiency).
- It then appropriately weighs and aggregates operators’ inputs to determine consensus performance goals.
- These goals can then used to determine TMI parameters that are expected to best achieve the performance expectations.

Underlying models, analysis and mechanisms are results of SLE project.

“Step 2”: requires additional research – performance based TMI planning.
A NextGen Vision: Performance-Based ATM

Current Practice:

- Expected Operating Environment
- Planned Operational Response
- Response Execution
- Operational Outcome

NextGen Vision:

- Expected Operating Environment
- Service Expectations
- Planned Operational Response
- Response Execution
- Operational Outcome

Philosophy:

- Airlines provide “consensus” service expectations
- FAA develops operational plan to meet those expectations
COuNSEL: CONsensus Service Expectation Level Planning

Information to users:
candidate performance vectors

<table>
<thead>
<tr>
<th>Capacity</th>
<th>Efficiency</th>
<th>Predictability</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1: 0.9</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>V2: 0.7</td>
<td>0.7</td>
<td>0.9</td>
</tr>
<tr>
<td>V3: 0.8</td>
<td>0.6</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Inputs from user 1:
Grades for vectors and candidate vectors

Grades:
100%, 95%, 90%, 85% ...

Consensus vector:
e.g. (.89, .76, .65)
Consensus Vector Chosen using **Majority Judgment**

- Suppose:
  - 6 airlines (voters), voting on 3 candidates: $V_1$, $V_2$, $V_3$
  - grades: 100%, 99%, 98%, 97%, 96%, 95%, 94%, ...

- Grades sorted after voting from worst to best:

<table>
<thead>
<tr>
<th></th>
<th>80%</th>
<th>80%</th>
<th>90%</th>
<th>94%</th>
<th>95%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1$</td>
<td>80%</td>
<td>80%</td>
<td><strong>90%</strong></td>
<td>94%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>$V_2$</td>
<td>75%</td>
<td>83%</td>
<td><strong>85%</strong></td>
<td>87%</td>
<td>88%</td>
<td>90%</td>
</tr>
<tr>
<td>$V_3$</td>
<td>65%</td>
<td>70%</td>
<td><strong>88%</strong></td>
<td>90%</td>
<td>93%</td>
<td>95%</td>
</tr>
</tbody>
</table>

**Majority grades:** majority would give at least that grade.

... in this example 4th grade from right.

**Vector with highest majority grade will be selected.**

There is a tie-breaking rule – not discussed here.
Performance Goals in SLE

• **Capacity**: maximize throughput
  – Avoid underestimating capacity and encourage quick response if weather clears early

• **Efficiency**: minimize delay cost
  – Take delay on the ground instead of in the air

• **Predictability**: provide timely, accurate, information
  – Announce GDPs well ahead of start times
  – Avoid overestimating or underestimating capacity; make program revisions unlikely
Interpretation of Performance Goals

All metrics take on values between 0 and 1

1 $\rightarrow$ perfect performance
0 $\rightarrow$ worst possible performance

The system only allows goal vectors that are “feasible”, e.g. even on a near-perfect day (1,1,1) would not be possible – perfect performance across all dimensions.

The system forces the flight operators to make tradeoffs:

\[(.91, .83, .85) \rightarrow (.86, .89, .85)\]

Reduce capacity goal: .91 $\rightarrow$ .86

... in order to improve efficiency goal: .83 $\rightarrow$ .89
Interpretation of Performance Goals

Capacity:

1 ➔ maximum airport throughput achieved (perfect weather day)

As metric decreases, flights will be delayed, cancellations may be necessary, diversions are a possibility, etc.
Interpretation of Performance Goals

Efficiency:
1 ➔ each flight will be executed in a minimum (user) cost manner: no airborne holding or vectoring, minimum taxi-in/out times, no diversions (note: an assigned ground delay is not counted against user cost as this cost is captured under capacity/throughput)

As metric decreases, airborne delays (and diversions) become more likely, the need to take suboptimal routes becomes more likely, etc.
Interpretation of Performance Goals

Predictability:

1 ➔ each flight’s departure and arrival time known with perfect accuracy well in advance of flight

As metric decreases, flight departure time estimates will vary over course of day, enroute times will become less predictable, there will be less advance warning of FAA actions, TMI parameters will be more likely to change over time, etc.
Design Tradeoffs

• SLE will enable flight operators to influence TMI design tradeoffs

• Predictability vs. Throughput
  – Predictability—assume lower rates and long duration so that initially assigned delays are unlikely to be extended
  – Throughput—assume higher rates and shorter duration in order to increase demand pressure

• Efficiency vs. Throughput
  – Efficiency—minimize airborne delay by imposing more ground delay
  – Throughput—employ higher arrival rates to increase demand pressure but (possibly) at the expense of more airborne delay

• Predictability vs. Efficiency
  – Predictability—make decisions well in advance, even though this increases the risk that they will be based on erroneous forecasts
  – Efficiency—make decisions later when better information is available, reducing the risk of airborne delay
SLE Features

• Airline votes are weighted by number of flights involved in the TMI
• Voting process is iterative—new candidate vectors are determined by ratings of previous candidate vectors
• Only feasible candidate vectors are allowed — set of feasible vectors is based on conditions of the day
• Airlines may develop their own tools to assess how different candidate vectors affect their individual business objectives
• Multiple applications of COuNSEL might be used as conditions change; could be applied nationwide or to regional problem area
Significant Research Components

• Generating candidate vectors, COuNSEL iteration mechanism: must generate promising candidates for infinite space of possible vectors – employs optimization and statistical estimation models.

• Definition of space of feasible candidate vectors: analytic models of TMIs – relationship between parameter setting and performance metrics.

• Understanding user impact and benefit mechanisms, gaining user acceptance: outreach to flight operators; formal flight operator surveys; human-in-the-loop simulation, involving flight operators and FAA.

• Modeling benefit mechanism and flight operator impact: use of historical data analysis and simulation to relate flight operator performance to TMI parameter settings.

• Modeling user voting/grading behavior: game theory and related models to understand user payoff functions and incentives for good (and bad) voting behavior.
Benefits of SLE

• A more fair and inclusive decision-making process where all the flight operators’ voices will be heard

• A goal-oriented decision-making process where performance criteria are clear to the flight operators

• A more consistent decision-making process where decision are less dependent on managers’ experience and personality
Topic 1: Choice of Performance Categories
NextGen: Performance-Based ATM

• “Performance based” ATM for National Air Space (NAS)
  – Support airline operators’ business objectives subject only to system-level objectives like safety and security
• Consistent with global and other regions’ visions of future ATM

1. Capacity  
2. Cost-effectiveness  
3. Efficiency  
4. Flexibility  
5. Predictability  
6. Access & Equity  
7. Environment  
8. Global interoperability  
9. Participation by the ATM Community  
10. Security  
11. Safety
NextGen: Performance-Based ATM

- “Performance based” ATM for National Air Space (NAS)
  - Support airline operators’ business objectives subject only to system-level objectives like safety and security
- Consistent with global and other regions’ visions of future ATM

1. Capacity
2. Cost-effectiveness
3. Efficiency
4. Flexibility
5. Predictability
6. Access & Equity
7. Environment
8. Global interoperability
9. Participation by the ATM Community
10. Security
11. Safety
NextGen: Performance-Based ATM

- “Performance based” ATM for National Air Space (NAS)
  - Support airline operators’ business objectives subject only to system-level objectives like safety and security
- Consistent with global and other regions’ visions of future ATM

1. Capacity
2. Cost-effectiveness
3. Efficiency
4. Flexibility
5. Predictability
6. Access & Equity
7. Environment
8. Global interoperability
9. Participation by the ATM Community
10. Security
11. Safety

These would be set as global/strategic requirements and not manipulated on a day to day basis.
NextGen: Performance-Based ATM

- “Performance based” ATM for National Air Space (NAS)
  - Support airline operators’ business objectives subject only to system-level objectives like safety and security
- Consistent with global and other regions’ visions of future ATM

1. Capacity
2. Cost-effectiveness
3. Efficiency
4. Flexibility
5. Predictability

Refers to ANSP cost effectiveness; not likely that flight operators would have incentive to reduce this.

10. Security
11. Safety
NextGen: Performance-Based ATM

• “Performance based” ATM for National Air Space (NAS)
  – Support airline operators’ business objectives subject only to system-level objectives like safety and security
• Consistent with global and other regions’ visions of future ATM

1. Capacity
2. Cost-effectiveness
3. Efficiency
4. Flexibility
5. Predictability

There could be good arguments for including, e.g. TMI designs that allow flight operators greater ability to substitute and internally optimize would certainly be viewed positively.
NextGen: Performance-Based ATM

- “Performance based” ATM for National Air Space (NAS)
  - Support airline operators’ business objectives subject only to system-level objectives like safety and security

1. Capacity
2. Cost-effectiveness
3. Efficiency
4. Flexibility
5. Predictability
6. Access & Equity
7. Environment
8. Global interoperability
9. Participation by the ATM Community
10. Security
11. Safety

Important category; however, flight operators would vote based on whether they were currently getting good or bad end of inequitable treatment; perhaps ANSP should somehow control equity metric.
Specific Metrics

• The metrics used in each category were chosen for specific reasons related to status of research and prototype development:
  – *We anticipate that these will change based on more research and priorities set by various other groups within the FAA.*
Topic 2: Choice of Underlying Mechanism
Research Problem

• Design a consensus-building mechanism, incorporating airline operators’ preferences, for determining the levels of service expectations at NAS-level, usable by the Air Navigation Service Provider (ANSP), to design Planned Operational Response, for the day-of-operations
Desirable Properties

1. **single winner determination.**
   – Leads to a unique “winner”.

2. **confidentiality.**
   – Minimal private information requirements from the airlines.

3. **practicality.**
   – Easy to administer, not involving time-consuming information gathering and / or processing steps.

4. **consensus-building.**
   – Maximum acceptability among the airlines.

5. **equitability.**
   – Perceived to be fair to all parties involved from the outset.

6. **strategy-proof.**
   – As far as possible, encourage truth-telling behavior.
Mechanisms Considered

• “Investment” / Marketplace / Combinatorial Auction
  – Requires creation of artificial “currency”
  – Metrics are not really goods being split up
    • Strategic behavior unavoidable: free-rider problem

• Multi-player Non-cooperative Game
  – Useful in modeling the strategic behavior
  – Existence of unique Nash equilibrium established
  – Outcomes not “desirable”: extreme solutions, without desired tradeoffs

• Voting
  – Natural way to model the decision making paradigm
  – Challenges exist in modeling
  – Two alternatives considered:
    • Weighted Instant Runoff Voting
    • Majority Judgment
  – Game theory to be used for analysis
Majority Judgment

- Recently proposed procedure (Balinski and Laraki, '10)
- Bypasses Arrow’s Impossibility Theorem (1950)
  - when voters have three or more distinct alternatives, no voting system can convert the ranked preferences of individuals into a community-wide (complete and transitive) ranking while also meeting a certain set of criteria, namely: unrestricted domain, non-dictatorship, Pareto efficiency, and independence of irrelevant alternatives.
- Claimed by authors to be “a better alternative to all other known voting methods, in theory and in practice.”
Majority Judgment – Definition

Majority Judgment is a *social decision function*

- Grading of each candidate by all voters in a common language
  - instead of preference rankings
  - more natural, richer preference elicitation

- Many good properties: highly resistant to strategic voting
Consensus Vector Chosen using **Majority Judgment**

• Suppose:
  – 6 airlines (voters), voting on 3 candidates: $V_1$, $V_2$, $V_3$
  – grades: 100%, 99%, 98%, 97%, 96%, 95%, 94%, ...

• Grades sorted after voting from worst to best:

<table>
<thead>
<tr>
<th></th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80%</td>
<td>75%</td>
<td>65%</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>83%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>85%</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>94%</td>
<td>87%</td>
<td>90%</td>
</tr>
<tr>
<td></td>
<td>95%</td>
<td>88%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>90%</td>
<td>95%</td>
</tr>
</tbody>
</table>

**Majority grades:** majority would give at least that grade.
.... in this example 4\textsuperscript{th} grade from right.

*Vector with highest majority grade will be selected.*
There is a tie-breaking rule – not discussed here.
MJ in Perspective

• The use of the median grade as the majority grade is key to the good properties of MJ, i.e. it greatly reduces the potential gain from “strategic” grading.

.... Yet, in terms of global welfare, one would prefer the average grade. Even in the limited set of examples explored in the HITL, this issue was very notable to participants (and made some participants question the MJ criterion).

Idea worth exploring: use median criterion to identify set of nearly equivalent vectors and allow ANSP to break near-ties using other criteria, e.g. average grade, equity, etc.
Idea worth exploring: use median criterion to identify set of nearly equivalent vectors and allow ANSP to break near-ties using other criteria, e.g. average grade, equity, etc.
Topic 3: Majority Judgment – Adaptation for Use in COuNSEL
Challenge in Application of MJ to Service Level Expectation Setting

• The basic application of MJ allows flight operators to make a consensus choice among possible goal vectors.

• **Challenge 1:** given conditions on a particular day of operations what are appropriate “possible goal vectors” that should be presented to flight operators.
  – Partial Answer: In concept there will be many (an infinite number) of vectors that represent the possible tradeoffs among the performance vectors given the weather and traffic conditions for the scenario of interest. Thus, challenge 1, becomes the problem of representing the space of performance metric tradeoffs for the TMIs under consideration.

• **Challenge 2:** given some representation of the space of possible goal vectors, what is a process for choosing among these the ones that flight operators will grade as part of the MJ process?
Solution to Challenge 1: *Set of constraints that define feasible vectors for particular day in the NAS.*

Bad weather day – sample vectors: (.90, .75, .80), (.85, .80, .83), (.85, .90, .79).

Good weather day – sample vectors: (.98, .95, .90), (.99, .92, .91), (.95, .97, .90).

\[ \mathbf{m} \text{ is possible metric vector:} \]

\[ \mathbf{m} \in \text{FEAS}_{\text{METRIC}} \]
**Majority Judgment**
*with small set of vectors*

- Suppose:
  - 6 airlines (voters), voting on 3 candidates: $V_1$, $V_2$, $V_3$
  - grades: 100%, 99%, 98%, 97%, 96%, 95%, 94%, ...

- Grades sorted after voting from worst to best:

<table>
<thead>
<tr>
<th></th>
<th>80%</th>
<th>80%</th>
<th>90%</th>
<th>94%</th>
<th>95%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1$</td>
<td>80%</td>
<td>80%</td>
<td>90%</td>
<td>94%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>$V_2$</td>
<td>75%</td>
<td>83%</td>
<td>85%</td>
<td>87%</td>
<td>88%</td>
<td>90%</td>
</tr>
<tr>
<td>$V_3$</td>
<td>65%</td>
<td>70%</td>
<td>88%</td>
<td>90%</td>
<td>93%</td>
<td>95%</td>
</tr>
</tbody>
</table>

**Majority grades:** majority would give at least that grade.

.... in this example 4th grade from right.

*Vector with highest majority grade will be selected.*

There is a tie-breaking rule – not discussed here.
COuNSEL Architecture

- Feasibility Constraints
- Airline 1 Input
  - Grades
  - Candidates
- Airline 1 Assessment
- Airline Weights
  (Fn. of num. flts. impacted)
  ANSP Input
- Service Expectations
  Resolution (Majority Judgment)
- Evaluation / Generation of New Candidates
Applying MJ with infinite set of candidates:

1. Define optimization model (MJ-Opt) that finds Majority Judgment winner assuming each airline’s grading function $g^a(m)$ is known.

2. Iteratively generate candidate vectors and based on airline grades use statistical methods to estimate $g^a(m)$
   - Candidate generation employs MJ-Opt to generate candidates likely to be close to MJ winner.

Also:

*Allow flight operators to supply their own candidates.*
Majority Judgment

**Winner**

- All possible candidates
- All possible majority-forming sets (Subset_Opt(b))
- “best” majority-forming set involving each player (Player_Opt(i’))

“Majoritarian Set”: set of players that determine MG

Player with the lowest grade in MS determines MG

Candidate with the highest MG wins
“Best” majority-forming set for a player
“Best” majority-forming set for...
“Best” majority-forming set for [image]
“Best” majority-forming set for...
"Best" majority-forming set for player $i'$

$I_i = 1$ if $i \in M_{i'}$; 0 otherwise

$\tilde{z}_{i'} = \max \ x_{i'}$

s.t. $x_{i'} \leq G_{\text{max}}(1 - I_i) + x_i$

$\sum_{i \in N} w_i I_i \geq \overline{W}'$

$I_i \in \mathbb{B}$

$x_i = g_i(m)$

$m \in \mu$

This model can be solved efficiently with integer programming software.
Majority Judgment Winner

\[ u^* = \max_{i \in N} \tilde{z}_i \]

\[ u^* = \max_{m \in \mu} v(m) \]

Candidate with the highest MG wins

“best” majority-forming set involving each player

(Player_Opt(i'))
New Candidate Vectors

\[ I_i = 1 \text{ if } i \in M_i'; 0 \text{ otherwise} \]

\[ \hat{z}_{i'} = \max x_{i'} \]
\[ \text{s.t. } x_{i'} \leq G^{\max}(1 - I_i) + x_i \]
\[ \sum_{i \in N} w_i I_i \geq \bar{W}' \]
\[ I_i \in \mathbb{B} \]
\[ x_i = \hat{g}_i(m) \]
\[ m \in \mu \]

Estimate grade function
Constrained least-squares regression (for concavity)
Same formulation: Two Uses

1. Majority Judgment Winner determination over continuous candidate space
   Uses knowledge of grade functions
   Theoretical

2. New candidate generation
   Estimates grade functions (constrained least-squares)
   Practical
Chicago area (ORD + MDW)

October 2007

Calendar for October 2007:

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Heterogeneous airline operations

<table>
<thead>
<tr>
<th>Category</th>
<th>Airlines</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hub and Spoke</td>
<td>6</td>
<td>1243</td>
</tr>
<tr>
<td>Point-to-point</td>
<td>8</td>
<td>292</td>
</tr>
<tr>
<td>International</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Charter</td>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td>Cargo</td>
<td>5</td>
<td>11</td>
</tr>
</tbody>
</table>

- **Airlines**: 47
- **Operations**: 1603
Long tail in distribution of operations

- UA: 635
- AA: 500
- WN: 242
- NW: 34
- DL: 28
- US: 27
- CO: 19
- FL: 18

20
19
Airlines’ **best** vectors are **spread out**
Optimal vectors are hard to find

<table>
<thead>
<tr>
<th>Wt Sch</th>
<th>UA’s %age</th>
</tr>
</thead>
<tbody>
<tr>
<td>nops</td>
<td>39.5</td>
</tr>
<tr>
<td>log.2</td>
<td>10.8</td>
</tr>
<tr>
<td>root.10</td>
<td>10</td>
</tr>
<tr>
<td>eqwt</td>
<td>2.1</td>
</tr>
</tbody>
</table>
Winning vectors are close to Optimal

%age Deviation in Majority Grades
Winners vs. Optimal Vectors

Sequence of bars: eqwt, root.10, log.2, nops

Overall accuracy of procedure: \(0.2\%\)
Computing times are manageable.

Dell Inspiron 5520
Intel Core i7-3612 @2.10GHz, 8GB RAM
Windows 7 Ultimate 64-bit

R 2.15.1 32-bit
CPLEX 12.4 via Rcplex 0.3-0
quadprog 1.5-4
Final Thoughts

• Simulation has shown approach to be computationally effective for 2-metric spaces – have not fully tested process for 3-dimensional vectors but looks quite doable.

• Practical Perspective: as was done in the HITL, the system can work quite well with “more modest” ways of generating candidate vectors, e.g. allowing flight operators to submit candidates, creating list ahead of time based on intuition, using various “heuristic” criteria.

... the sophisticated integer programming approach to candidate generation may not be critical in practice (but determining this will require more experimentation).
Topic 4: Definition of Space of Feasible Candidate Vectors
Characteristics of Space of Feasible Performance Goal Vectors:

- A basic assumption of the performance metrics is a higher value of any metric is preferred to a lower value (by any flight operator), e.g. any flight operator would prefer (.91, .88, .85) to (.91, .82, .85) since the first and last metric values are the same but the 2\textsuperscript{nd} is higher in the first vector (we say the 1st point dominates the 2\textsuperscript{nd}).

- Also, it is assumed (somewhat for conceptual and mathematical convenience) that if two vectors are possible/feasible then any vector on the line segment between them is feasible, e.g. if (.91, .88, .87) and (.91, .82, .91) are both feasible then a point in between, e.g. \( \frac{1}{2} (.91, .88, .87) + \frac{1}{2} (.91, .82, .91) = (.91, .85, .89) \) is also feasible.
Characteristics of Space of Feasible Performance Goal Vectors

• Thus we can define the space of feasible vectors by a set of linear constraints with the structure illustrated below.

• Only the points of the efficient frontier are of interest as possible goal vector.

![Diagram showing points on the efficient frontier and dominated points.](Image)
Format of Constraints Defining Space of Performance Goal Vectors

• Based on the previous discussion, if performance vectors are denoted by \((V_1, V_2, V_3)\) then any constraint defining the region of feasible performance goal vectors has the form:

\[
A_1 V_2 + A_2 V_2 + A_3 V_3 \leq B
\]

where \(A_1, A_2, A_3 \geq 0\) and \(B > 0\)

• The COuNSEL software tool accepts a list of constraints in this format.
Generating Constraints

• Approach to generating constraints defining space of feasible performance vectors:
  – Step 1: generate set of possible performance vectors given the weather and demand conditions of the day.
  – Step 2: find set of constraints that encloses the points generated in step 1, in a feasible region with the appropriate properties.
Solution to Step 2

- There are well known methods that find a set of constraints defining the convex hull of a set of given points – such methods can be accessed as functions in various computational toolkits.
- This “almost” provides a solution to Step 2: before applying such a method, it may be necessary to add some points to insure the set of points have the structured described earlier.
- The figure below illustrates the points that may need to be added.

The points added insure that all dominated are feasible and that the interior constraints defining the region contain only non-dominated points.
Solution to Step 1:
Performance Vector Generation for GDPs Based on Analysis of Historical Days

• Research carried out so far assumes a GDP plan is characterized (only) by the planned airport arrival rate vector (PAAR)
• The performance achieved by choosing a particular PAAR is determined by the actual airport arrival capacity profile that occurs (AAAR)
• The conditions on a particular day (weather forecast) will determine an AAAR distribution for that day, i.e. a list of possible AAAR together with associated probabilities
• Performance vectors can be enumerated by enumerating possible PAARs and computing an associated performance vector for each PAAR by applying the AAAR distribution
The Logic

• Identify a set of possible capacity profiles for the given day-of-operation

• Each possible capacity profile may be selected as the planned capacity profile
The Logic (II)

• For each planned capacity profile, the feasible candidate vector (SLE metric) is estimated as an average of the realized system performances over all the possible capacity profiles that may realize:

\[
\bar{M}_i^k = \frac{\sum_{j=1}^{J} M_{i,j}^k}{J}
\]

where, \(\bar{M}_i^k\) is SLE metric for performance goal \(k\) with planned capacity profile \(i\);

\(M_{i,j}^k\) is the realized performance for performance goal \(k\) if capacity profile \(i\) is planned and capacity profile \(j\) is the actual capacity profile.
Currently, all the profiles are assumed to be equally likely.
Performance Goals

• Currently, we are considering the following performance criteria:
  – Capacity utilization
  – Efficiency
  – Predictability

• More criteria could be considered upon users’ request
Capacity Utilization

This metric is defined to measure how much capacity is planned when the GDP is first implemented against the capacity under VMC condition:

\[ M^{1}_{i,j} = \alpha_{cu,i,j} = \frac{N_{R,i,j}}{N_{VMC,i,j}} \]

where,

\( \alpha_{cu,i,j} \) is the capacity utilization metric with planned capacity profile \( i \) and actual capacity profile \( j \);

\( N_{R,i,j} \) is the count of realized arrivals between GDP start time and end time when capacity profile \( i \) is planned and profile \( j \) is realized;

\( N_{VMC,i,j} \) is the count of arrivals that could have been landed assuming VMC capacity and infinite demand during the same period for the same pair of profiles.
Efficiency

Efficiency is defined referring to the motivation of GDP: transforming airborne delay to cheaper ground delay:

\[ M_{i,j}^2 = \alpha_{e,i,j} = \frac{\sum_k GD_{i,j,k}}{\sum_k TD_{i,j,k}} \]

where,

- \( \alpha_{e,i,j} \) is the efficiency metric with planned capacity profile \( i \) and actual capacity profile \( j \);
- \( GD_{i,j,k} \) is the ground delay incurred by flight \( k \) for the same pair of capacity profiles;
- \( TD_{i,j,k} \) is the total delay incurred by flight \( k \), equal to realized ground delay plus realized airborne delay.
Predictability

• Predictability is defined to capture the accuracy in estimating capacity rates. In the strategic planning telecons, most of the debate is on setting capacity rates.

• On one hand, we want to make sure available capacity will be effectively utilized. On the other hand, we also appreciate the accuracy of the guess on capacity rates. The former is considered in the capacity utilization and the latter is considered by predictability metric.
Predictability (II)

\[ M_{i,j}^3 = \alpha_{p,i,j} = \frac{1}{T} \sum_{t=1}^{T} \frac{\min(PAAR_{i,t},AAAR_{j,t})}{\max(PAAR_{i,t},AAAR_{j,t})} \]

where,

- \( \alpha_{p,i,j} \) is the predictability metric with planned capacity profile \( i \) and actual capacity profile \( j \);
- \( t \) is the index for the 15-minute interval and \( T \) is the total number of intervals;
- \( PAAR_{i,t} \) is the planned airport acceptance rate for interval \( t \) given plan capacity profile as \( i \);
- \( AAAR_{j,t} \) is the actual airport acceptance rate for interval \( t \) when the actual capacity profile is \( j \).
How to Generate the Set of Possible Capacity Profiles?

Weather forecast For a given day:
- Ceiling
- Visibility
- Wind
- Thunderstorm
- Snow

(Ref: Liu et al., Icrat 2014)
Methodology: learn from history
## Logic in the Method

<table>
<thead>
<tr>
<th>Given day</th>
<th>Day G</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical days</strong></td>
<td>Day H₁  Day H₂  ...  Day H_m</td>
</tr>
<tr>
<td><strong>Capacity profiles</strong></td>
<td>Profile 1  Profile 2  ...  Profile m</td>
</tr>
<tr>
<td><strong>Total distances</strong></td>
<td>$TD_{G,H₁} &lt; TD_{G,H₂} &lt; ... &lt; TD_{G,H_m}$</td>
</tr>
<tr>
<td><strong>Similarity</strong></td>
<td>Highest  →  Lowest</td>
</tr>
</tbody>
</table>

$TD_{G,H_i}$ : Total distance in weather forecast between Day G and Day $H_i$
Total distance between Days G and H

\[ TD_{G,H} = \sum_{i=1}^{n} (d_i)^2 \]

Day G:

Day H:

hourly distances in weather forecasts
Hourly Distance between Hours $j$ and $k$

$$d_{j,k}(A) = \sqrt{(WF_j - WF_k)^T \cdot A \cdot (WF_j - WF_k)}$$

Weather Forecast vector

$$[x_1, x_2, x_3]$$

Matrix of distance coefficients

$$\begin{bmatrix}
\Delta x_1 & \Delta x_2 & \Delta x_3 \\
\Delta x_1 & a_{1,1} & a_{1,2} & a_{1,3} \\
\Delta x_2 & a_{2,1} & a_{2,2} & a_{2,3} \\
\Delta x_3 & a_{3,1} & a_{3,2} & a_{3,3}
\end{bmatrix}$$

$\Delta$’s: difference between the weather variables from hour $i$ and hour $j$

$$d_{j,k}(A) = \sqrt{a_{11} \cdot \Delta x_1^2 + a_{12} \cdot \Delta x_1 \cdot \Delta x_2 + a_{13} \cdot \Delta x_1 \cdot \Delta x_3 + \cdots}$$
Weather forecast distance between two hours depends on difference in capacity between these two hours.
Similarity/Dissimilarity Sets

• A pair of hourly weather forecasts, \((WF_j, WF_k)\)
  – belongs to the similarity set, S, if difference in realized capacity rates is small
  – belongs to the dissimilarity set, D, if difference in realized capacity rates is large

The objective here is to predict hourly capacity
Matrix of Distance Coefficients, A

Objective: \( \min_A \sum_{(WF_j, WF_k) \in S} [d_{j,k}(A)]^2 \)

Minimize the weather forecast distances for the hour pairs in the similarity set

Constraints:

\( \sum_{(WF_j, WF_k) \in D} \|WF_j - WF_k\|_A \geq 1 \)

So \( A \neq 0 \)

and

\( A \geq 0 \)

A is positive and semi-definite, so \( d_{j,k}(A) \) is satisfying non-negativity

(Eric et al., 2012)
Distance Matrix, $A$

In the literature

\[
\begin{bmatrix}
\Delta x_1 & \Delta x_2 & \Delta x_3 \\
\Delta x_1 & 1 & 0 & 0 \\
\Delta x_2 & 0 & 1 & 0 \\
\Delta x_3 & 0 & 0 & 1 \\
\end{bmatrix}
\]

\[d_{j,k} = \sqrt{\Delta^2 x_1 + \Delta^2 x_2 + \Delta^2 x_3}\]

In the proposed work

\[
\begin{bmatrix}
\Delta x_1 & \Delta x_2 & \Delta x_3 \\
\Delta x_1 & a_{1,1} & a_{1,2} & a_{1,3} \\
\Delta x_2 & a_{2,1} & a_{2,2} & a_{2,3} \\
\Delta x_3 & a_{3,1} & a_{3,2} & a_{3,3} \\
\end{bmatrix}
\]

- Different weights for different weather variables
- Weights for the interactions between weather variables
Recipe

Given day

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>start</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Historical day

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>start</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hourly distances between WFs, $d_i$

Total distance between two days, $\sum_{i=1}^{n} (d_i)^2$

Extract capacity profiles from similar days with short total distances. These profiles are taken as planned capacity profiles.
Topic 5: Benefit Mechanisms and User Grading Models
Benefits of SLE

• A goal-oriented decision-making process where performance criteria are clear to the flight operators
• A more consistent decision-making process where decisions are less dependent on managers’ experience and personality
• Reduction in NAS-wide operating (delay and disruption) cost via better support of airlines’ business objectives
• A more fair and inclusive decision-making process where all the flight operators’ voices will be heard

This set of slides focus on the last two
Assessment Methods

• **CoUNSEL Design**
  – COuNSEL design is informed by assuming airlines vote according to the value functions computed by our modeling approach. (Aside from modeling approach, we also conducted a Human-In-The-Loop (HITL) experiment to get airline inputs)

• **Benchmarks**, compare COuNSEL design to
  – *Centralized (state-of-research) design*: the design which has the least total aircraft delay cost (sum of ground delay and airborne delay cost) for all GDP-impacted incoming flights
  – *System-optimal design*: the design which has the least total delay and disruption cost (both aircraft and passenger delay/disruption) by summing over the delay cost of each airline. This approach accounts for airline recovery actions.

• **Notes**
  – FAA traffic managers make decisions in designing GDPS and these decisions impact airlines’ operating bottom lines. COuNSEL design most likely will not necessarily lead to an improvement in traditional system performance metrics, e.g. overall throughput or delay. Rather it will lead to a better economic performance for the airlines and fairer distribution of outcomes among different airlines.
Core Modeling Approaches

• In order to assess airline’s value function of different GDP designs, we built...
  
  – An integrated simulation platform
    • Generate different GDP designs (rate, duration and scope).
    • An integrated recovery module for each airline to simulate airline response to GDP programs.
    • Evaluate under capacity uncertainty.

  – An airline recovery module
    • Given disruptions, how to swap fleet, cancel flights, re-accommodate passengers to minimize total delay cost.
Assessment Flowchart

**GDP design:**
- *Rate*
- *Duration*
- *Scope*

**Airline Recovery Module**

- Airline planned cost
- Airline unplanned cost

**Airline Cost:**
- Aircraft delay
- Passenger disruption
- Crew disruption

**Capacity is uncertain...**
GDP rate may be under/over-estimated
GDP duration may be too long/short
...leads to early cancellation and late extension

**Unplanned cost:** Additional airborne delay, passengers disruptions, fleet disruptions due to inaccurate delay information provided by FAA
Experimental Setup

- **mm/dd/yy:** 6/16/2007
- **Airport:** SFO
- **Actual duration:** Uniform[3 hrs, 9 hrs]
- **14 candidate designs for evaluation:** planned duration: 3-9 hours, with an increment of 0.5 hour
- **Program arrival rate:**
  - outside GDP duration: VFR rate
  - inside GDP duration: IFR rate
- **Airline Itinerary data source:**
  Generated by Barnhart et al., 2011.

- **Delay cost coefficient estimation:**
  BTS Form 41 financial data. Estimated separately for different airlines and different fleet types.
- **Carriers involved:**

<table>
<thead>
<tr>
<th>Carriers</th>
<th># Impacted Operations</th>
<th># Fleet Types (# Aircraft in Each Category)</th>
<th># Impacted Passengers</th>
<th>% Connecting Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>United &amp; SkyWest</td>
<td>359</td>
<td>10 (17,4,8,3,9,1,3,7,5,27)</td>
<td>24236</td>
<td>32.33%</td>
</tr>
<tr>
<td>American &amp; American Eagle</td>
<td>70</td>
<td>5 (4,2,4,3,9)</td>
<td>7678</td>
<td>27.39%</td>
</tr>
<tr>
<td>US Airways</td>
<td>40</td>
<td>4 (1,4,1,4)</td>
<td>4007</td>
<td>31.57%</td>
</tr>
<tr>
<td>Continental &amp; ExpressJet</td>
<td>30</td>
<td>5 (1,1,3,1,2)</td>
<td>3244</td>
<td>20.43%</td>
</tr>
<tr>
<td>Delta Airlines</td>
<td>26</td>
<td>4 (1,1,2,2)</td>
<td>3750</td>
<td>30.29%</td>
</tr>
<tr>
<td>Alaska Airlines</td>
<td>25</td>
<td>2 (4,3)</td>
<td>2461</td>
<td>9.47%</td>
</tr>
<tr>
<td>Northwest Airlines</td>
<td>23</td>
<td>4 (2,2,2,1)</td>
<td>3232</td>
<td>25.46%</td>
</tr>
<tr>
<td>Frontier Airlines</td>
<td>15</td>
<td>2 (2,2)</td>
<td>1351</td>
<td>31.68%</td>
</tr>
<tr>
<td>JetBlue Airways</td>
<td>9</td>
<td>1 (2)</td>
<td>1180</td>
<td>8.05%</td>
</tr>
<tr>
<td>AirTran Airways</td>
<td>8</td>
<td>1 (4)</td>
<td>973</td>
<td>32.58%</td>
</tr>
</tbody>
</table>
Revealing Airline’s Preference

<table>
<thead>
<tr>
<th>Airline</th>
<th># Impacted Operations</th>
<th># Fleet Types (# Aircraft in Each Category)</th>
<th># Impacted Passengers</th>
<th>% Connecting Passengers</th>
<th>Average Load Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Airways</td>
<td>40</td>
<td>4 (1,4,1,4)</td>
<td>4007</td>
<td>31.57%</td>
<td>80.43%</td>
</tr>
</tbody>
</table>

small number of total operations, multiple different fleet type

little flexibility for recovery
(reduces 6.6% cost through recovery at most)

total cost trend: preference for aggressive design (shorter planned duration)
Revealing Airline’s Preference

<table>
<thead>
<tr>
<th>Airline</th>
<th># Impacted Operations</th>
<th># Fleet Types (# Aircraft in Each Category)</th>
<th># Impacted Passengers</th>
<th>% Connecting Passengers</th>
<th>Average Load Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>American &amp; American Eagle</td>
<td>70</td>
<td>5 (4,2,4,3,9)</td>
<td>7678</td>
<td>27.39%</td>
<td>75.53%</td>
</tr>
</tbody>
</table>

medium number of total operations, multiple different fleet type

medium flexibility for recovery
(reduces 32.4% cost through recovery at most)

total cost trend: preference for moderate design (intermediate planned duration)
Revealing Airline’s Preference

<table>
<thead>
<tr>
<th>Airline</th>
<th># Impacted Operations</th>
<th># Fleet Types (# Aircraft in Each Category)</th>
<th># Impacted Passengers</th>
<th>% Connecting Passengers</th>
<th>Average Load Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>United &amp; SkyWest</td>
<td>359</td>
<td>10 (17,4,8,3,9,1,3,7,5,27)</td>
<td>24236</td>
<td>32.33%</td>
<td>75.29%</td>
</tr>
</tbody>
</table>

extremely large number of total operations, multiple different fleet type

great flexibility for recovery
(reduces 62.3% cost through recovery at most)

total cost trend: preference for conservative design (longer planned duration)
Revealing Airline’s Preference

<table>
<thead>
<tr>
<th>Airline</th>
<th># Impacted Operations</th>
<th># Fleet Types (# Aircraft in Each Category)</th>
<th># Impacted Passengers</th>
<th>% Connecting Passengers</th>
<th>Average Load Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>AirTran Airways</td>
<td>8</td>
<td>1 (4)</td>
<td>973</td>
<td>32.58%</td>
<td>82.32%</td>
</tr>
</tbody>
</table>

small number of total operations, single fleet type

great flexibility for recovery
(reduces 47.6% cost through recovery at most)

total cost trend: preference for conservative design (longer planned duration)
## NAS-wide Benefits Assessment

<table>
<thead>
<tr>
<th>Preference Category</th>
<th>Airline - GDP Cost Matrix</th>
<th>Aggressive Design</th>
<th>GDP Planned Duration (hours)</th>
<th>Conservative Design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>Moderate</td>
<td>American &amp; American Eagle</td>
<td>141340</td>
<td>124389</td>
<td>123490</td>
</tr>
<tr>
<td></td>
<td></td>
<td>117142</td>
<td>125420</td>
<td>115407</td>
</tr>
<tr>
<td>Aggressive</td>
<td>Frontier</td>
<td>60362</td>
<td>76148</td>
<td>69646</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82396</td>
<td>83363</td>
<td>85580</td>
</tr>
<tr>
<td>Aggressive</td>
<td>US Airways</td>
<td>83058</td>
<td>84186</td>
<td>85994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>87735</td>
<td>90418</td>
<td>96115</td>
</tr>
<tr>
<td>Aggressive</td>
<td>Continental &amp; ExpressJet</td>
<td>34152</td>
<td>37247</td>
<td>37844</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33511</td>
<td>36526</td>
<td>33968</td>
</tr>
<tr>
<td>Moderate</td>
<td>JetBlue</td>
<td>9705</td>
<td>9849</td>
<td>10766</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8939</td>
<td>8252</td>
<td>7983</td>
</tr>
<tr>
<td>Moderate</td>
<td>Delta</td>
<td>36256</td>
<td>35408</td>
<td>34897</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34846</td>
<td>34860</td>
<td>34132</td>
</tr>
<tr>
<td>Conservative</td>
<td>AirTran</td>
<td>16600</td>
<td>15049</td>
<td>15050</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13499</td>
<td>13363</td>
<td>11954</td>
</tr>
<tr>
<td>Aggressive</td>
<td>Northwest</td>
<td>22247</td>
<td>36705</td>
<td>32657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>31738</td>
<td>31265</td>
<td>34185</td>
</tr>
<tr>
<td>Conservative</td>
<td>United &amp; SkyWest</td>
<td>489250</td>
<td>448340</td>
<td>426198</td>
</tr>
<tr>
<td></td>
<td></td>
<td>408230</td>
<td>402122</td>
<td>386515</td>
</tr>
<tr>
<td>Moderate</td>
<td>Alaska</td>
<td>41167</td>
<td>35758</td>
<td>35713</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35713</td>
<td>37002</td>
<td>34810</td>
</tr>
<tr>
<td>Moderate</td>
<td>NAS wide</td>
<td>934137</td>
<td>903079</td>
<td>869554</td>
</tr>
<tr>
<td></td>
<td></td>
<td>850262</td>
<td>847815</td>
<td>844941</td>
</tr>
<tr>
<td>Conservative</td>
<td>Centralized Objective</td>
<td>244986</td>
<td>235343</td>
<td>226604</td>
</tr>
<tr>
<td></td>
<td></td>
<td>221638</td>
<td>214614</td>
<td>210056</td>
</tr>
</tbody>
</table>

Note: GDP values are shown in bold.
NAS-wide Benefits Assessment

Linearity transform costs into 100-scale grades...

<table>
<thead>
<tr>
<th>Airline - GDP Grade Matrix</th>
<th># impacted operation</th>
<th>weights</th>
<th>Aggressive Design</th>
<th>GDP Planned Duration (hours)</th>
<th>Conservative Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>American &amp; American Eagle</td>
<td>70</td>
<td>17.97</td>
<td>80</td>
<td>91</td>
<td>92</td>
</tr>
<tr>
<td>Frontier</td>
<td>15</td>
<td>6.30</td>
<td>100</td>
<td>79</td>
<td>90</td>
</tr>
<tr>
<td>US Airways</td>
<td>40</td>
<td>12.28</td>
<td>100</td>
<td>99</td>
<td>97</td>
</tr>
<tr>
<td>Continental &amp; ExpressJet</td>
<td>30</td>
<td>10.1</td>
<td>98</td>
<td>90</td>
<td>89</td>
</tr>
<tr>
<td>JetBlue</td>
<td>9</td>
<td>4.45</td>
<td>78</td>
<td><strong>77</strong></td>
<td><strong>70</strong></td>
</tr>
<tr>
<td>Delta</td>
<td>26</td>
<td>9.16</td>
<td>94</td>
<td>96</td>
<td>98</td>
</tr>
<tr>
<td>AirTran</td>
<td>8</td>
<td>4.11</td>
<td>58</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>Northwest</td>
<td>23</td>
<td>8.43</td>
<td>100</td>
<td>61</td>
<td>68</td>
</tr>
<tr>
<td>United &amp; SkyWest</td>
<td>359</td>
<td>54.63</td>
<td>61</td>
<td>67</td>
<td>70</td>
</tr>
<tr>
<td>Alaska</td>
<td>25</td>
<td>8.92</td>
<td><strong>79</strong></td>
<td>92</td>
<td>92</td>
</tr>
</tbody>
</table>

Majority judgement winner coincides with system-optimal design!!
NAS-wide Benefits Assessment

Centralized Design, System Optimal Design, COuNSEL Design and Historical Design

To centralized design:
COuNSEL reduces 2.0% in NAS-wide total cost

To historical design:
COuNSEL reduces 4.2% in NAS-wide total cost
NAS-wide Benefits Assessment

To centralized design: COuNSEl reduces total ground delay by 9.8%, total passenger delay by 3.3%. It increases total airborne delay from 214 minutes to 318 minutes. On a per flight basis, from 0.95 minutes/flight to 1.46 minutes/flight.

To historical design: COuNSEl reduces 22.8% total ground delay, 13.7% total passenger delay, while only inducing an airborne delay of 1.46 minutes/flight.

NAS performance is improved by being operated slightly aggressive!

<table>
<thead>
<tr>
<th>GDP Planned Duration (hours)</th>
<th>NAS wide Total Cost ($)</th>
<th>Total Ground Delays (minutes)</th>
<th>Total Airborne Delay (minutes)</th>
<th>Average Ground Delay per Flight (minutes)</th>
<th>Average Airborne Delay per Flight (minutes)</th>
<th># Disrupted Passengers</th>
<th>Total Passenger Delay (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>934137</td>
<td>2337</td>
<td>422</td>
<td>10.40</td>
<td>1.87</td>
<td>674</td>
<td>480618</td>
</tr>
<tr>
<td>3.5</td>
<td>903079</td>
<td>2835</td>
<td>432</td>
<td>12.62</td>
<td>1.92</td>
<td>725</td>
<td>469141</td>
</tr>
<tr>
<td>4</td>
<td>869554</td>
<td>3012</td>
<td>403</td>
<td>13.41</td>
<td>1.79</td>
<td>679</td>
<td>467044</td>
</tr>
<tr>
<td>4.5</td>
<td>850262</td>
<td>3269</td>
<td>386</td>
<td>14.56</td>
<td>1.71</td>
<td>681</td>
<td>456964</td>
</tr>
<tr>
<td>5</td>
<td>847815</td>
<td>3589</td>
<td>382</td>
<td>15.98</td>
<td>1.70</td>
<td>685</td>
<td>453177</td>
</tr>
<tr>
<td>5.5</td>
<td>844941</td>
<td>3857</td>
<td>376</td>
<td>17.18</td>
<td>1.67</td>
<td>674</td>
<td>458074</td>
</tr>
<tr>
<td>6</td>
<td>818413</td>
<td>4090</td>
<td>371</td>
<td>18.21</td>
<td>1.65</td>
<td>671</td>
<td>459613</td>
</tr>
<tr>
<td>6.5</td>
<td>822104</td>
<td>4429</td>
<td>358</td>
<td>19.72</td>
<td>1.59</td>
<td>673</td>
<td>472420</td>
</tr>
<tr>
<td>7</td>
<td>809297</td>
<td>4607</td>
<td>328</td>
<td>20.52</td>
<td>1.46</td>
<td>668</td>
<td>480232</td>
</tr>
<tr>
<td>Centralized design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>811352</td>
<td>4748</td>
<td>257</td>
<td>21.14</td>
<td>1.14</td>
<td>678</td>
<td>485759</td>
</tr>
<tr>
<td>8</td>
<td>825808</td>
<td>5105</td>
<td>214</td>
<td>22.73</td>
<td>0.95</td>
<td>673</td>
<td>496769</td>
</tr>
<tr>
<td>8.5</td>
<td>820144</td>
<td>5546</td>
<td>123</td>
<td>24.70</td>
<td>0.54</td>
<td>646</td>
<td>520816</td>
</tr>
<tr>
<td>Historical design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>834271</td>
<td>5690</td>
<td>85</td>
<td>25.34</td>
<td>0.37</td>
<td>667</td>
<td>531846</td>
</tr>
<tr>
<td>9.5</td>
<td>845128</td>
<td>5970</td>
<td>0</td>
<td>26.59</td>
<td>0</td>
<td>662</td>
<td>556366</td>
</tr>
</tbody>
</table>
Assessment Results

- Compared to centralized design COuNSEL produces more equitable GDP design.
- Compared to system optimal design COuNSEL produces the same level of equity.

<table>
<thead>
<tr>
<th>Airline</th>
<th>Cost under Preferred Design</th>
<th>Cost under COuNSEL Design</th>
<th>Percentage Increment (COuNSEL)</th>
<th>Cost under Centralized Design</th>
<th>Percentage Increment (Centralized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>American &amp; American Eagle</td>
<td>112998</td>
<td>127462</td>
<td>11.35%</td>
<td>128174</td>
<td>11.84%</td>
</tr>
<tr>
<td>Frontier</td>
<td>60362</td>
<td>91783</td>
<td>34.23%</td>
<td>101507</td>
<td>40.53%</td>
</tr>
<tr>
<td>US Airways</td>
<td>83058</td>
<td>95663</td>
<td>13.18%</td>
<td>101711</td>
<td>18.34%</td>
</tr>
<tr>
<td>Continental &amp; ExpressJet</td>
<td>33511</td>
<td>39935</td>
<td>16.09%</td>
<td>41300</td>
<td>18.86%</td>
</tr>
<tr>
<td>JetBlue</td>
<td>7577</td>
<td>7707</td>
<td>1.69%</td>
<td>9446</td>
<td>19.79%</td>
</tr>
<tr>
<td>Delta</td>
<td>34132</td>
<td>35531</td>
<td>3.94%</td>
<td>38467</td>
<td>11.27%</td>
</tr>
<tr>
<td>AirTran</td>
<td>9592</td>
<td>10338</td>
<td>7.22%</td>
<td>9592</td>
<td>0.00%</td>
</tr>
<tr>
<td>Northwest</td>
<td>22247</td>
<td>36074</td>
<td>38.33%</td>
<td>36690</td>
<td>39.36%</td>
</tr>
<tr>
<td>United &amp; SkyWest</td>
<td>300218</td>
<td>330232</td>
<td>9.09%</td>
<td>322038</td>
<td>6.78%</td>
</tr>
<tr>
<td>Alaska</td>
<td>32724</td>
<td>34573</td>
<td>5.35%</td>
<td>36882</td>
<td>11.27%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Standard Deviation</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11.88%</td>
<td>12.44%</td>
</tr>
</tbody>
</table>
• Various user support tools are developed to help airlines and FAA make corresponding decisions under SLE framework.
  – SLE metrics tradeoff curves
  – SLE metrics to TMI parameters mapping
  – SLE metrics to airline performance mapping
User Support Tool #1: SLE Metric Tradeoff Curves

- Each slide gives four tradeoff curves showing the tradeoff between two SLE metrics for four values of the third SLE metric.
Efficiency vs Capacity Utilization Tradeoff

Eff vs CapUtil for Pred =
0.768
0.803
0.839
0.874
Capacity Utilization vs Predictability Tradeoff

For $\text{Eff} = \{0.366, 0.576, 0.785, 0.994\}$
Efficiency vs Predictability Tradeoff

Eff vs Pred for CapUtil =

- 0.56
- 0.618
- 0.676
- 0.735
User Support Tool #2: SLE Vectors to TMI Parameters Mapping

• For the FAA traffic managers and for each scenario a mapping is given from a set of SLE metric vectors to corresponding TMI plans/parameters.
Goal Vectors to TMI Parameters Mapping
(Use GDP as an example)

- Goals: capacity utilization, efficiency, predictability
- GDP parameters: start time, end time, planned called rates

<table>
<thead>
<tr>
<th>Goal vectors</th>
<th>TMI parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacity utilization</td>
</tr>
<tr>
<td></td>
<td>0.651</td>
</tr>
<tr>
<td></td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td>0.668</td>
</tr>
<tr>
<td></td>
<td>0.617</td>
</tr>
<tr>
<td></td>
<td>0.615</td>
</tr>
<tr>
<td></td>
<td>0.582</td>
</tr>
<tr>
<td></td>
<td>0.647</td>
</tr>
<tr>
<td></td>
<td>0.560</td>
</tr>
<tr>
<td></td>
<td>0.735</td>
</tr>
<tr>
<td></td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td>0.718</td>
</tr>
<tr>
<td></td>
<td>0.671</td>
</tr>
<tr>
<td></td>
<td>0.617</td>
</tr>
<tr>
<td></td>
<td>0.658</td>
</tr>
</tbody>
</table>
User Support Tool #3:
SLE Vectors to User Performance Indicator Mapping

- For each flight operator and for each scenario a mapping is given from a sample of SLE metric vectors to user performance indicators.
SLE Vectors to User Performance Indicator Mapping: American Airlines

- American Airline: with 89 total impacted flights, 9863 impacted passengers
- With great recovery capability, it prefers low capacity-high efficiency GDP design
- Total operating cost includes: 1) flight delay cost (fuel and other aircraft operating cost) 2) passenger delay cost

<table>
<thead>
<tr>
<th>Cap.</th>
<th>Eff.</th>
<th>Pred.</th>
<th>Expected Total Operating Cost ($1,000)</th>
<th>Num of Cancellations</th>
<th>Expected Ground Delay Minute</th>
<th>Expected Airborne Delay Minute</th>
<th>Expected Passenger Total Delay Minute</th>
<th>Expected Delay Minute per Nondisrupted Passenger</th>
<th>Expected Num of Disrupted Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.996 0.484 0.791</td>
<td>727 0 2,852 320</td>
<td>451,228 47.3</td>
<td>261</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.992 0.529 0.796</td>
<td>748 0 3,073 337</td>
<td>462,632 48.5</td>
<td>267</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.989 0.571 0.798</td>
<td>754 0 3,011 345</td>
<td>485,702 50.8</td>
<td>245</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.984 0.610 0.798</td>
<td>730 0 3,013 217</td>
<td>471,783 49.4</td>
<td>254</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.970 0.705 0.789</td>
<td>726 0 3,708 295</td>
<td>430,993 45.2</td>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.964 0.731 0.784</td>
<td>744 0 3,794 288</td>
<td>466,915 48.9</td>
<td>275</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.959 0.753 0.777</td>
<td>669 2 3,395 221</td>
<td>382,214 40.0</td>
<td>276</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.942 0.798 0.752</td>
<td>695 2 3,780 261</td>
<td>418,667 43.8</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.931 0.812 0.731</td>
<td>621 0 4,869 265</td>
<td>487,811 50.6</td>
<td>209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.926 0.814 0.720</td>
<td>603 0 4,810 255</td>
<td>479,422 49.6</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.921 0.813 0.708</td>
<td>600 0 4,820 239</td>
<td>463,497 48.1</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.917 0.814 0.696</td>
<td>571 0 4,970 120</td>
<td>461,740 47.8</td>
<td>187</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.912 0.814 0.683</td>
<td>611 2 3,972 155</td>
<td>379,071 39.5</td>
<td>264</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.908 0.815 0.670</td>
<td>576 2 4,273 128</td>
<td>386,789 40.1</td>
<td>226</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.904 0.815 0.656</td>
<td>566 2 4,250 91</td>
<td>382,581 39.8</td>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.900 0.816 0.643</td>
<td>528 0 5,102 50</td>
<td>440,946 45.5</td>
<td>174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.899 0.816 0.639</td>
<td>570 2 4,809 20</td>
<td>397,665 41.3</td>
<td>231</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

• In most of the cases, COuNSEL has the capability to reduce system-wide total delay cost, and produce more equitable design.

• COuNSEL leads to a better economic performance for the airlines and fairer distribution of outcomes among different airlines.
Topic 6: COuNSEL Software Tool
Software Tool

• Users are divided into administrators and participants
• Administrators create polls, approve submissions and can view detailed submission results
• Participants submit candidates, rank candidates and can view only the winning vector
Process

1. Administrator creates poll
2. Participants submit candidates
3. Administrator approves candidates and opens grading
4. Participants grade candidates
5. Results are shown
Necessary Inputs

• The following inputs will be required for each poll:
  – User Accounts: each participants must have an account
  – Group: participants are organized into groups
  – Metric table: a table of constraints defining the feasible set of candidates
  – Weight set: an assignment of weights to the participants
Groups

• Individuals are organized into groups
• When you make a poll, you need to create a group for that poll which contains the users that will vote in that poll
• Individuals may belong to more than one group
Metric Table

• The metric table is a list of constraints which describe the feasible set of candidates.

• These constraints take the form
  \[ A_1 \times \text{capacity} + A_2 \times \text{efficiency} + A_3 \times \text{predictability} \leq B \]
  where \( A_1, A_2, A_3 \) and \( B \) are all positive numbers.

• This tool requires that these numbers be at least 0.0001
Weight Sets

• Weight sets describe how much weight is given to each user during voting
• Weights can be any positive number with at most two decimal place
Candidate Format

• In the software, candidates are represented as a three dimensional vector:

  \((\text{capacity}, \text{efficiency}, \text{predictability})\)

• Each element of a candidate is usually represented as an integer percentage from 0 to 100

• Example: the candidate which achieves 50% capacity, 70% efficiency and 70% predictability is represented as

  \((50,70,70)\)
Candidate Format

• However, metric constraints are written in terms of decimal values instead of percentages
• Example: the constraint that the sum of the three metrics is no more than 200% for any candidate would be:

\[ 1 \times \text{capacity} + 1 \times \text{efficiency} + 1 \times \text{predictability} \leq 2 \]
Administrator Home Page

Create new polls

Active polls – participants are submitting candidates

Active polls - waiting for administrator to approve candidates

Active polls – participants are grading candidates
Completed polls – can view results, or can create new polls based on the results of the completed poll.
Participant home page

Active polls – participant may submit candidates

Active polls – participant may grade candidates

Results from completed polls are shown
Poll Creation - Administrator

The poll is given a name and description.

We can create a table of metric constraints or choose an existing table.
The metric table is given a name. After we create a table for one poll, we will be able to select it in later polls.

Each line in this table is a constraint of the form

\[ A \times \text{cap.} + B \times \text{eff.} + C \times \text{pred.} = D \]

Every entry must have a value of at least 0.0001.

<table>
<thead>
<tr>
<th>Capacity Constraint</th>
<th>Efficiency Constraint</th>
<th>Predictability Constraint</th>
<th>RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0001</td>
<td>0.0001</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>3.5</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Poll Creation - Administrator

Add Voters

Select an existing set of grader weights from the drop-down or create a new one

Create New Weight Set

Select airline groups that are eligible to vote

Create New Group

Add Candidates

Values for capacity, efficiency, and predictability should be integers between 0 and 100

The set of weights of voters is specified
Poll Creation - Administrator

The set of weights is given a name. We will be able to select this set in future polls.

Weights should be numbers greater than zero with up to two decimal places.

<table>
<thead>
<tr>
<th>Grader</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Canada</td>
<td>10</td>
</tr>
<tr>
<td>Alaska</td>
<td>20</td>
</tr>
<tr>
<td>American</td>
<td>45.5</td>
</tr>
<tr>
<td>Cubby</td>
<td></td>
</tr>
<tr>
<td>Delta</td>
<td>15</td>
</tr>
<tr>
<td>DJ</td>
<td></td>
</tr>
</tbody>
</table>

Weights can be any positive number with up to two decimal places.
If a user is not voting in our poll, we leave the entry blank.
Poll Creation - Administrator

Add Voters

Select an existing set of grader weights from the drop-down or create a new one

Select airline groups that are eligible to vote

Add Candidates

Values for capacity, efficiency, and predictability should be integers between 0 and 100

We create a group of users for the poll or select an existing group
Poll Creation - Administrator

Group Creation

Enter a name for your new group and select the users you wish to add.

Group Name (Required): EWR_Example

The group is given a name. We will be able to select this group in future polls.

We select which users will be able to participate in our poll.

<table>
<thead>
<tr>
<th>User</th>
<th>□</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Canada</td>
<td>✔</td>
</tr>
<tr>
<td>Alaska</td>
<td>✔</td>
</tr>
<tr>
<td>American</td>
<td>✔</td>
</tr>
<tr>
<td>Cubby Brendle</td>
<td>□</td>
</tr>
<tr>
<td>Delta</td>
<td>✔</td>
</tr>
<tr>
<td>DJ Funk</td>
<td>□</td>
</tr>
<tr>
<td>ExpressJet</td>
<td>✔</td>
</tr>
</tbody>
</table>
Poll Creation - Administrator

We click on a group to add the users in that group to our poll.

Eligible users will appear in the “Eligible Groups” section.

Note: if an individual is in several groups, then they might be appear under a group other than the one we selected. This does not affect the functionality of the poll.

We can also select individuals one at a time.

Add Voters

Select an existing set of grader weights from the drop-down or create a new one

Select airline groups that are eligible to vote

Add Candidates

Values for capacity, efficiency, and predictability should be integers between 0 and 100
Poll Creation - Administrator

Add Candidates
Values for capacity, efficiency, and predictability should be integers between 0 and 100

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Capacity</th>
<th>Efficiency</th>
<th>Predictability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Candidate submitted for EWR example</td>
<td>74</td>
<td>60</td>
<td>90</td>
</tr>
</tbody>
</table>

Add Another

Voting Procedure
Decide whether or not to allow custom candidate submissions

- Allow Candidate Submission
- Open Poll to Voting

Candidate Submission Restrictions
Specify acceptable ranges that airlines should conform to when creating custom candidates

- Maximum number of candidate submissions per grader: 2

Solicit Candidates

- Duration to Accept Candidates: 10 minutes
- This poll contains 1 candidates.

Start Accepting Candidates  Save as Draft

The administrator can choose to submit some candidates for grading. These should be feasible, but the software allows infeasible ones. Choose whether to also allow participants to submit candidates. Choose how many candidates each user may submit. Choose a time duration and then click the button to start accepting candidates.
Candidate Submission - Participant

User uses sliders to set values of two parameters

Third slider set automatically to highest feasible value

Red color indicates that this slider was automatically set

Description
Candidate approval - Administrator

Administrator may approve or reject any submitted candidate.
Candidate approval - Administrator

Once the administrator has finished approving candidates, then the poll is opened to voting.
Grading Candidates - Participants

Participants assign each candidate a grade using a slider. The rank that the user has given each candidate is shown.

The rank that the user has given each candidate is shown.
**Viewing Results - Administrator**

**Description:**
This is an example at Newark.

**Winner:** (20% | 42% | 99%)

- Order results by candidate number
- Order results by grade

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Description</th>
<th>Capacity Efficiency/Predictability</th>
<th>Include in next round?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate 1</td>
<td>(No Description)</td>
<td><img src="image" alt="Graph" /></td>
<td>No</td>
</tr>
<tr>
<td>Majority Grade: 85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candidate 2</td>
<td>High capacity</td>
<td><img src="image" alt="Graph" /></td>
<td>No</td>
</tr>
<tr>
<td>Majority Grade: 20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The winning candidate is shown
- All candidates and their majority grades are shown
- The administrator may select candidates for inclusion in another poll
Viewing Results - Administrator

The administrator may also view each user’s grade of each candidate.

The administrator may create another poll based on the results of this poll.
Viewing Results - Participant

Participants may see the winner, but cannot see more detailed results.