“Friendlier Flying”: Stochastically Modeling Airport Arrival Capacities During Inclement Weather

Tasha R. Inniss,
Clare Boothe Luce Professor of Mathematics,
Trinity College

Minorities and Applied Mathematics-Connections to Industry and Government Laboratories
May 4, 2001
Goal/Motivation

- Goal: Estimate airport arrival capacity distributions during inclement weather conditions

- Why?
  - Bad weather reduces capacity below demand
  - Implicit relationship between weather and capacity
  - Stochastic nature of weather makes it difficult to deterministically predict capacity
  - Required input to a class of stochastic ground holding models
Outline

- Background
  - Discussion of a Ground Delay Program
  - Hoffman-Rifkin Static Stochastic Ground Holding (H-R) Model
- Capacity Scenarios (Arrival Capacity Distributions)
  - Conceptual Representation of ACDs
  - Generating overall distribution of ACDs
- Deriving Seasonal Distributions via “Seasonal Clustering”
- Current and Future Work
Method is needed to efficiently address capacity-demand imbalances.

To address and manage these imbalances, ATCSCC may institute ground delay programs (GDPs).

Determining the amount of delay to assign in a GDP is known as the ground holding problem (GHP).

GDP planning has become more efficient under a new collaborative process known as Collaborative Decision Making (CDM).
What is a Ground Delay Program (GDP)?

- Delayed departures
- Delayed arrivals/no airborne holding
- Delayed departures
- Delayed departures
Ground vs. Airborne Delay

- In a GDP, determining the optimal amount of ground delay to assign is known as the **Ground Holding Problem (GHP)**.

- **Conservative vs. Liberal Policies**: more ground holding vs. less ground holding (more airborne holding)

  Flights scheduled to arrive = 30; Capacity (AAR) = 20
Hoffman-Rifkin (H-R) Static Stochastic Ground Holding Model

- Determines number of flights to delay on the ground and number expected to be air delayed per unit time
- Explicitly takes into account the uncertainty of weather
- **Formulation:**
  \[
  \text{Min} \quad \sum_{t=1} \ c_g G_t + \sum_{t=1}^{T+1} c_a p_q W_{q,t} \\
  \text{subject to} \quad A_t - G_{t-1} + G_t = D_t \quad t = 1, \ldots, T+1 \\
  G_0 = G_{T+1} = 0 \\
  -W_{q,t-1} + W_{q,t} - A_t \geq -M_{q,t} \quad t = 1, \ldots, T+1 \\
  q = 1, \ldots, Q \\
  W_{q,0} = W_{q,T+1} = 0 \\
  A_t \in \mathbb{Z}_+, \ W_{q,t} \in \mathbb{Z}_+, \ G_t \in \mathbb{Z}_+ \\
  \]

- **Inputs:** aggregate demand for each time period \((D_t)\), ground delay cost factor \((c_g)\), airborne delay cost factor \((c_a)\), capacity scenarios \((Q)\) and associated probabilities \((p_q)\)
Is a capacity-demand imbalance likely?

YES

GDP Planner

GDP Parameters

Hoffman-Rifkin Model (Stochastic GH Model)

# of arrival slots per unit time

CDM Procedures (RBS, Substitution, Compression)

Slot assignments of individual flights

Arrival Capacity Distribution Generator

Capacity Scenarios and Probabilities

Weather Forecast (conditions & start time)

Month (Season)

Projected Demand (build-up)

Airborne/ Ground Cost Ratio

Month (Season)
Representative Structures of Capacity Scenarios

General Arrival Capacity Distribution

2-Level ACD

2-Parameter ACD

1-Parameter ACD
Empirical (Historical) Data Sets

- Ground Delay Programs’ Data
  - Logged at ATCSCC and archived by Metron, Inc.
  - Contains GDP parameters such as duration of GDP, scope of GDP and Airport Acceptance Rate (AAR-capacity)
  - Includes 1995, 1996, 1997 GDPs at SFO
  - Can be used for performance analysis
  - Can be used to generate Capacity Probabilistic Distributions Functions (CPDFs) when weather data not available
Data (continued)

- Weather Data
  - Contained in “Surface Airways Hourly” collected by National Climatic Data Center (NCDC)
  - Contains data such as cloud ceiling height, visibility, wind direction and wind speed
  - Can be used to estimate distribution of inclement weather conditions (Instrument Flight Rules-IFR)
- Want combination of GDP data and weather data to get distribution of IFR conditions given a GDP is planned (conditional distribution)
Overall Capacity PDF with 1-Parameter ACDs

"Conditional" Distribution of Duration of IFR Conditions

\[ P(S_j) = \frac{\text{frequency of } j}{\text{total sum of frequencies}} \]

- \( P(S_0) = 0.25 \)
- \( P(S_3) = 0.12 \)
- \( P(S_6) = 0.03 \)

Relative frequency histogram created by binning historical weather data for San Francisco.
Time Series Plot of Average GDP Length
Choosing Seasons of Least Cost

- Want to “cover” an entire year by a finite number of covers or seasons
- Determine which season to assign a month in a least costly fashion
- Cost ($C_j$) of a season $j$ can be the sum of squared deviations between a season’s value and the values of the months contained within that season:

$$\sum_{i=1}^{n} (\text{SeasonAvg} - \text{MonthAvg}_i)^2$$

- Formulation:

$$\text{Minimize } \sum_{j=1}^{n} C_j x_j$$
$$\text{subject to } \sum_{i=1}^{n} x_j = N$$
$$\sum_{j=1}^{n} a_{ij} x_j = 1 \text{ for each month } i$$
$$x_j \in \{0,1\}$$
### Cost Functions for Set Partitioning (Differences in Means)

<table>
<thead>
<tr>
<th>Sum of Squared Deviations (SoSqs)</th>
<th>( \sum_{j=1}^{m} (X_{.j} - \overline{X}.)^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized SoSqs</td>
<td>( \frac{1}{m} \sum_{j=1}^{m} (X_{.j} - \overline{X}.)^2 )</td>
</tr>
<tr>
<td>Seasonal Variances</td>
<td>( \frac{1}{m-1} \sum_{j} \sum_{i} (X_{ij} - \overline{X}.)^2 )</td>
</tr>
</tbody>
</table>

\( \overline{X}_{.j} \) is the average over all days \( i \) in month \( j \);  
\( \overline{X}. \) is the (overall) seasonal average over all days \( i \) and all months \( j \);  
\( X_{ij} \) is the GDP length on day \( i \) in month \( j \).
Cost Functions Based on Differences in EDFs

- Calculate an EDF for each month $j$ ($F_j$) in a given season.
- Calculate a seasonal EDF (pooled EDF):
  \[ F = \frac{1}{n} \sum_j (n_j F_j) \]
- Compute the cost of a given season by calculating a Kolmogorov-Smirnov (KS) statistic for the season:
  \[ KS = \max_x \sqrt{\sum_j \left( \frac{n_j}{n} \right) \left[ F_j(x) - F(x) \right]^2} \]
Post Analysis for Evaluating Sets of Seasons

- Single-Factor ANOVA with multiple comparisons

\[ Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i=1,\ldots,12 \text{ and } j=1,2,3, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \]

- Single-factor ANOVA used to test if there exist statistically significant differences in means of seasons.

- Multiple comparisons used to test for equality between two seasonal means.

\[ \frac{\sum_s n_s (\bar{Y}_{s.} - \bar{Y}_{..})^2}{k - 1} \]

- Mean Square Ratio:

\[ \frac{\sum_s \sum_j (Y_{js} - \bar{Y}_{s.})^2}{n - k} \]
Results of Post Analysis

- ANOVA multiple comparisons’ results (for seasons resulting from set partitioning of GDP data)

<table>
<thead>
<tr>
<th>Contiguous Seasons</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr-Jun vs Jul/Aug</td>
<td>.0288</td>
</tr>
<tr>
<td>Jul/Aug vs Sep/Oct</td>
<td>.0388</td>
</tr>
<tr>
<td>Sep/Oct vs Nov-Mar</td>
<td>.0053</td>
</tr>
</tbody>
</table>

- Mean Square ratios (for seasons resulting from set partitioning of weather data)

<table>
<thead>
<tr>
<th>Contiguous Seasons</th>
<th>Mean Square Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar-Jun vs Jul-Sep</td>
<td>14.06</td>
</tr>
<tr>
<td>Jul-Sep vs Oct-Feb</td>
<td>24.39</td>
</tr>
</tbody>
</table>
Perspectives on Seasonal “Clustering”

Intra-season homogeneity
(From Set Partitioning Integer Program)

Non-contiguous seasons can be similar

Transition Period

Inter-Season Variability
(From Post Analysis using Mean Square Ratio)

Developed in dissertation and used to find seasonal distributions
Relative Frequency Histograms for Weather Seasons

Histogram for "Heavy Fog" Season (Oct-Feb)

Histogram for "Rainy" Season (Mar-Jun)

Histogram for "Summer Weather" Season (Jul-Sep)
Current and Future Work

Current Work:
- Input capacity scenarios (ACDs) and associated probabilities into H-R model
- Developed methodology to adjust ground delay appropriately in dynamic GDPs
- Created algorithm to compare model results to current operational procedures

Future Work:
- Determine seasonal distributions with arbitrary start/end days
- Model airports using 2-Parameter ACD