How robust are robust schedules in reality?

N. Eggenberg, M. Salani, M. Bierlaire
Transport and Mobility Laboratory, EPFL, Switzerland

Funded by:
SNSF - Project 200021-118547
Introduction
Some numbers

- Huge economical impact
  - Air France-KLM 35 Mio € / day
  - Lufthansa 48 Mio € / day
  - IATA: $200 Mio / day to air sector

- Spill out due to disrupted / blocked passengers
Why robustness appeals for airline scheduling

- **Airlines have low profitability**
  - < 2% profit margin (US, 2007)

- **High delays and implied delay costs**
  - 4.3 Billion hours delay (US, 2008)
  - $41 Billion delay costs (US, 2008)
Worse is still to come

- **Growth:**
  - 2.5% more flights annually
  - Every 1% additional flights incur an additional 5% delays  
    \((\text{Schaefer et al., 2005})\)
  - => Yearly increase of delays of 12.5%

- Europe: 50% of flights in 2030 depart or land at congested airports

- Airlines must react – we try to help
  - Improve operations in a congested network
Outline

- Optimization under uncertainty
  - In general
  - In airline scheduling

- Robust Maintenance Routing Problem
  - Definitions
  - “Robust” and “Recoverable” models

- Simulation – preliminary results
  - Methodology to evaluate and compare robust solutions
  - Preliminary a priori and a posteriori results
General Optimization Problems

- Planning
- Observing
- Adapting
Robustness: plan for stability and reliability

- Optimized solutions have
  - Highest “expected” revenue/yield/profit
  - Known to be sensitive to noise

- Robust solutions have
  - Lower expected revenue/yield/profit
  - Higher reliability
Definition of robustness

- Unclear in literature
  - For more “stable” solutions (that remain feasible)
  - For more “flexible” solutions
  - For solutions with lower “operational costs”

- How to determine what “more robust” means?
  - What metric to use?
  - Should it be a priori or a posteriori?
Parallel to Stochastic Programming

- What is the equivalent to robustness
  - Stochastic optimization
  - Stochastic optimization with recourse
  - Risk management / chance constraint programming?

- Or are these robust methods themselves?
Airline Scheduling: An iterative Process

-60 to -6 months
- Route Choice

-6 months
- Fleet Assignment

-6 to -2 months
- Maintenance Routing

-6 to -2 months
- Crew Pairing

-2 to -1 months
- Crew Rostering

Day of Operations (Disruption Management)

-6 months to day D
- Revenue Management (passenger booking)
Robustness in airline scheduling

Robust airline schedules are

- Operationally more efficient
- Less sensitive to delay
  - i.e. with reduced delay propagation
Delay Propagation

- 2 types of delays for each flight
  - **Independent** delay: generated during a flight
    - At any stage (taxi, runway, landing,...)
  - **Propagated** delay
    - Delay due to previously delayed flight
    - Propagation is downstream (possibly to several flights)

- $Del(f) = ID(f) + PD(f)$

- Robustness proxy = expected PD
  - To be minimized
Other meanings of robustness

- Robustness is also used as a “flexibility” measure
  - Facilitates recovery
  - Reduces recovery costs

- We differentiate
  - ROBUSTNESS vs RECOVERABILITY
Robust Maintenance Routing Problem (MRP)

- **Deterministically known**
  - Original schedule (1 maintenance route/aircraft)

- **To determine**
  - New routes for each aircraft
  - And/or new departure times for each flight

- **Constraints**
  - Maintenance routes are feasible for each aircraft
  - All flights are covered exactly once
  - Each flight is retimed by at most $\pm 15$
  - Total retiming of all flights of at most $C$ minutes (500 or 1000)

- **Objective**
  - Optimize robustness metric
Used Uncertainty Feature Optimization (UFO) Models

- Use different UFs:
  - IT: maximize total idle time
  - MIT: maximize sum of minimal idle time of each route
  - CROSS: maximize nbr plane crossings
  - PCON: maximize passenger idle connection time
  - MinPCON: maximize minimal PCON

- Solved with CG algorithm (COIN-OR – BCP package)
Benchmark

- Models from literature
  - EPD: minimize expected propagated delay (Lan et al., 2006)
    - No retiming
    - Allow only plane swaps
  - EPD2: minimize expected propagated delay (AhmadBeygi et al., 2008)
    - No plane swaps
    - Allow for retiming by ± 15 minutes
    - Total retiming bounded (500 or 1000 minutes)

- Solved with same CG algorithm (COIN-OR – BCP package)
Measuring Recoverability: Methodology

- Solve Robust MRP using different models

- Apply some disruption scenarios
  - Differentiate *independent* and *propagated* delay
  - Update propagated delay according to schedule

- Solve the recovery problem
  - Using same recovery algorithm

- Evaluation with external recovery cost evaluator
  - Data and cost-evaluator provided by the *ROADEF Challenge 2009*
Scenario Generation

- EPD and EPD2 require expected delay for each flight
  - Generate two distributions using historical data from similar airline (scenarios 1 and 2)
  - Generate several scenarios drawing from each scenario
    - No variability (perfect information)
    - Low variability ($\sigma = 0.1 \hat{\mu}$)
    - High variability ($\sigma = 0.5 \hat{\mu}$)
  - Evaluate solutions on all scenarios and apply recovery algorithm
Generated schedules

- UFO solutions are the same for both scenarios
  - UFs are non-predictive models

- EPD solutions are different
  - Solution depends on estimated delay distribution

- Use two “realities” to simulate erroneous predictive models
### Simulation Overview – UFO solutions

<table>
<thead>
<tr>
<th>Scenario/Solution</th>
<th>Solutions Sc. 1</th>
<th>Solutions Sc. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>NEUTRAL</td>
<td>NEUTRAL</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>NEUTRAL</td>
<td>NEUTRAL</td>
</tr>
</tbody>
</table>
# Simulation Overview – EPD and EPD2

<table>
<thead>
<tr>
<th>Scenario/Solution</th>
<th>Solutions Sc. 1</th>
<th>Solutions Sc. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>OK</td>
<td>WRONG DISTRIBUTION</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>WRONG DISTRIBUTION</td>
<td>OK</td>
</tr>
</tbody>
</table>
Comparison Criteria

- Compare a priori AND recovery statistics
- A priori
  - UF values
  - EPD
- Recovery statistics
  - Recovery costs
  - Aircraft statistics
    - Total aircraft delay
    - Canceled flights
  - Passenger statistics
    - Total passenger delay
    - Rerouted passengers
    - Canceled passengers
Used Instance

- 608 flights
- 85 aircraft
- 36010 passengers
- 1 day
A priori robustness statistics
(max retiming = 500 minutes)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Original</th>
<th>IT</th>
<th>MIT</th>
<th>PCON</th>
<th>EPD</th>
<th>EPD2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPD [min]</td>
<td>8453</td>
<td>8265</td>
<td>8431</td>
<td>8496</td>
<td>8411</td>
<td><strong>7953</strong></td>
</tr>
<tr>
<td>IT [min]</td>
<td>12000</td>
<td><strong>12185</strong></td>
<td>12010</td>
<td>12135</td>
<td>12010</td>
<td>12060</td>
</tr>
<tr>
<td>PCON [min]</td>
<td>10815</td>
<td>10950</td>
<td>10860</td>
<td><strong>11815</strong></td>
<td>10815</td>
<td>10795</td>
</tr>
<tr>
<td><strong>Scenario 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPD [min]</td>
<td>7282</td>
<td>7185</td>
<td>7221</td>
<td>7221</td>
<td>7251</td>
<td><strong>6732</strong></td>
</tr>
<tr>
<td>IT [min]</td>
<td>12000</td>
<td><strong>12185</strong></td>
<td>12010</td>
<td>12135</td>
<td>12065</td>
<td>12110</td>
</tr>
<tr>
<td>PCON [min]</td>
<td>10815</td>
<td>10950</td>
<td>10860</td>
<td><strong>11815</strong></td>
<td>10815</td>
<td>10855</td>
</tr>
</tbody>
</table>
# Simulation Overview – EPD and EPD2

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>OK</td>
<td>WRONG DISTRIBUTION</td>
</tr>
<tr>
<td>S2</td>
<td>WRONG DISTRIBUTION</td>
<td>OK</td>
</tr>
</tbody>
</table>
Average Results (25 scenarios in each “reality”)

<table>
<thead>
<tr>
<th>Scenario</th>
<th># canc. Flts</th>
<th>Original</th>
<th>IT</th>
<th>MIT</th>
<th>PCON</th>
<th>EPD</th>
<th>EPD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>13.2</td>
<td>13.2</td>
<td>12.3</td>
<td>11.8</td>
<td>8.5</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>P.D. [min]</td>
<td>17,738</td>
<td>17,352</td>
<td>17,692</td>
<td>17,843</td>
<td>17,827</td>
<td>16,866</td>
<td></td>
</tr>
<tr>
<td>Rec Cost [€]</td>
<td>872,942</td>
<td>#</td>
<td>#</td>
<td>714,236</td>
<td>676,273</td>
<td>866,298</td>
<td></td>
</tr>
<tr>
<td>Scenario 2</td>
<td>9.9</td>
<td>9.8</td>
<td>9.4</td>
<td>8.1</td>
<td>6.5</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>P.D. [min]</td>
<td>14,115</td>
<td>13,973</td>
<td>14,029</td>
<td>14,052</td>
<td>13,967</td>
<td>13,310</td>
<td></td>
</tr>
<tr>
<td>Rec Cost [€]</td>
<td>548,194</td>
<td>#</td>
<td>#</td>
<td>422,551</td>
<td>423,997</td>
<td>449,128</td>
<td></td>
</tr>
</tbody>
</table>
Simulation Overview – EPD and EPD2

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>OK</td>
<td>WRONG DISTRIBUTION</td>
</tr>
<tr>
<td>S2</td>
<td>WRONG DISTRIBUTION</td>
<td>OK</td>
</tr>
</tbody>
</table>
Average Results (25 scenarios in each “reality”)

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>EPD_S1</th>
<th>EPD_S2</th>
<th>EPD2_S1</th>
<th>EPD2_S2</th>
<th>PCON</th>
</tr>
</thead>
<tbody>
<tr>
<td># canc. Flts</td>
<td>8.5</td>
<td>8.6</td>
<td>11.2</td>
<td>11.6</td>
<td>11.8</td>
</tr>
<tr>
<td>P.D. [min]</td>
<td>17,827</td>
<td>17,697</td>
<td>16,866</td>
<td>17,186</td>
<td>17,843</td>
</tr>
<tr>
<td>Rec Cost [€]</td>
<td>676,273</td>
<td>684,246</td>
<td>866.298</td>
<td>915,433</td>
<td>714,236</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 2</th>
<th>EPD_S1</th>
<th>EPD_S2</th>
<th>EPD2_S1</th>
<th>EPD2_S2</th>
<th>PCON</th>
</tr>
</thead>
<tbody>
<tr>
<td># canc. Flts</td>
<td>6.5</td>
<td>6.5</td>
<td>7.9</td>
<td>7.7</td>
<td>8.1</td>
</tr>
<tr>
<td>P.D. [min]</td>
<td>13,971</td>
<td>13,967</td>
<td>13,624</td>
<td>13,310</td>
<td>14,052</td>
</tr>
<tr>
<td>Rec Cost [€]</td>
<td>428,885</td>
<td>423,997</td>
<td>461,774</td>
<td>449,128</td>
<td>422,551</td>
</tr>
</tbody>
</table>
Conclusions

- No absolute meaning of robustness
  - How to measure?
  - How to evaluate?

- Methodology to compare solutions
  - A priori using pre-defined proxies
  - A posteriori using recovery statistics

- Preliminary results show that
  - Proxies are inter-correlated
  - Using evaluation approach allows better understanding of these inter-correlations and their implications
Open Research Directions

- Extend simulations and perform deeper analysis to
  - Better understand relations between proxies
  - Understand correlations between
    - a priori proxies
    - a posteriori proxies (recovery statistics)
    - Structure of the recovery algorithm

- Will this analysis allow to define robustness...
  - ... with respect to a given recovery algorithm?
  - ... with respect to a chosen proxy?
The End

Thank you for your attention!