Application of Machine Learning for ATM Performance Assessment
Identification of Sources of En-Route Flight Inefficiency

Rodrigo Marcos
Nommon Solutions and Technologies
SESAR Innovation Days
Salzburg, 5th December 2018
Challenges: KPA interdependencies, target setting, ANSP performance benchmarking, performance drivers → need for effective performance modelling approaches

Opportunities: increasing data availability, big data technologies, data science

Goal: explore the potential of data science to improve our understanding of KPA trade-offs, identify cause-effect relationships between performance drivers and KPIs, and develop new decision support tools for ATM performance monitoring and management

- Visual analytics and machine learning algorithms for pattern extraction
- New data-driven modelling techniques
- Prototype DST integrating the new analytical and visualisation functionalities
Sources of Flight Inefficiency
Problem Statement

The problem of target setting and performance evaluation:

- How to isolate the effects of different performance drivers?
- How to link decisions at ANSP/ACC level with overall performance?
  - Diagnosis of low performance episodes
  - Prediction of potential performance gains
Sources of Flight Inefficiency
Case Study and Approach

Approach:

• Case study: February 2017, Bordeaux ACC
• Flight efficiency indicator: interface HFE
• Trajectory data (DDR2, CPR)
• Visual exploration of influence factors and feature selection
• Machine learning model for efficiency prediction trained with historical data
  • Comparison of data sources
  • Analysis of influence factors

\[ HFE_j = \frac{\sum L_{fjp} - \sum H_{fjp}}{\sum H_{fjp}} \]
Sources of Flight Inefficiency

Data Exploration and Feature Selection

Features:

- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
• Flight plan efficiency
• Reference FL
• Take-off time
• Flights per sector
• Distance between ideal and planned entry/exit point
• Route length
• Heading
• Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
• Flight plan efficiency
• Reference FL
• Take-off time
• Flights per sector
• Distance between ideal and planned entry/exit point
• Route length
• Heading
• Airspace crossed
Sources of Flight Inefficiency
Data Exploration and Feature Selection

Features:
- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- **Airspace crossed**
Sources of Flight Inefficiency

Comparison of data sources

<table>
<thead>
<tr>
<th>Data</th>
<th>DDR</th>
<th>CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>0.973</td>
<td>0.980</td>
</tr>
<tr>
<td>Validation</td>
<td>0.804</td>
<td>0.836</td>
</tr>
<tr>
<td>Testing</td>
<td>0.770</td>
<td>0.757</td>
</tr>
</tbody>
</table>

Random forest regressor $R^2$ score

<table>
<thead>
<tr>
<th>Data</th>
<th>DDR</th>
<th>CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>0.973</td>
<td>0.980</td>
</tr>
<tr>
<td>Validation</td>
<td>0.804</td>
<td>0.836</td>
</tr>
<tr>
<td>Testing</td>
<td>0.770</td>
<td>0.757</td>
</tr>
</tbody>
</table>

**Training:**
- petal length (cm) $\leq 2.4500$
  - entropy = 1.58496250072
  - samples = 150

**Validation:**
- petal width (cm) $\leq 1.7500$
  - entropy = 0.0000
  - samples = 50
  - value = [ 50, 0, 0 ]

**Testing:**
- petal width (cm) $\leq 1.7500$
  - entropy = 0.1511
  - samples = 46
  - value = [ 0, 1, 45 ]

**Entropy Calculation:**
- $H(X) = -\sum p(x) \log p(x)$
- $H(Y|X=x) = -\sum p(y|x) \log p(y|x)$
### Sources of Flight Inefficiency

**Comparison of data sources**

<table>
<thead>
<tr>
<th>Data Source</th>
<th>DDR</th>
<th>CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned interface HFE in LFBBCTA</td>
<td>42.08%</td>
<td>43.55%</td>
</tr>
<tr>
<td>Distance between planned and ideal exit point from LFBBCTA</td>
<td>17.57%</td>
<td>17.06%</td>
</tr>
<tr>
<td>Distance between planned and ideal entry point to LFBBCTA</td>
<td>6.02%</td>
<td>7.56%</td>
</tr>
<tr>
<td>Ideal distance in LFFFCTA</td>
<td>4.42%</td>
<td>4.14%</td>
</tr>
<tr>
<td>Ideal distance in LFBBCTA</td>
<td>3.94%</td>
<td>3.52%</td>
</tr>
<tr>
<td>Ideal distance</td>
<td>3.79%</td>
<td>4.20%</td>
</tr>
<tr>
<td>Ideal distance in LECMCTA</td>
<td>2.52%</td>
<td>2.18%</td>
</tr>
<tr>
<td>Ideal distance in LFRRCTA</td>
<td>2.42%</td>
<td>1.50%</td>
</tr>
<tr>
<td>Ideal distance in LFMMCTA</td>
<td>1.65%</td>
<td>1.63%</td>
</tr>
<tr>
<td>Ideal distance in LECBCTA</td>
<td>1.37%</td>
<td>1.48%</td>
</tr>
<tr>
<td>Reference FL in LFBBCTA</td>
<td>1.37%</td>
<td>1.52%</td>
</tr>
<tr>
<td>Flights per sector in LECBCTA</td>
<td>1.12%</td>
<td>0.80%</td>
</tr>
<tr>
<td>Flights per sector in LECMCTA</td>
<td>1.00%</td>
<td>0.75%</td>
</tr>
<tr>
<td>Flights per sector in LFMMCTA</td>
<td>0.85%</td>
<td>0.82%</td>
</tr>
<tr>
<td>Take-off time - cosine</td>
<td>0.83%</td>
<td>0.84%</td>
</tr>
<tr>
<td>Take-off time - sine</td>
<td>0.81%</td>
<td>0.81%</td>
</tr>
<tr>
<td>Flights per sector in LFFFCTA</td>
<td>0.74%</td>
<td>0.67%</td>
</tr>
<tr>
<td>Flights per sector in LFBBCTA</td>
<td>0.73%</td>
<td>0.59%</td>
</tr>
<tr>
<td>Flights per sector in LFRRCTA</td>
<td>0.62%</td>
<td>0.55%</td>
</tr>
</tbody>
</table>

**Random forest regressor R^2 score**

<table>
<thead>
<tr>
<th>Data</th>
<th>DDR</th>
<th>CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>0.973</td>
<td>0.980</td>
</tr>
<tr>
<td>Validation</td>
<td>0.804</td>
<td>0.836</td>
</tr>
<tr>
<td>Testing</td>
<td>0.770</td>
<td>0.757</td>
</tr>
</tbody>
</table>
Sources of Flight Inefficiency

Influence Factors

AIRAC 1702
Testing NRMSE: 4.3 %
Features:
- Flight plan efficiency
- Reference FL
- Take-off time
- Flights per sector
- Distance between ideal and planned entry/exit point
- Route length
- Heading
- Airspace crossed
- Flights crossing LFBBCTA
- Weekday
- Number of the day

<table>
<thead>
<tr>
<th>Feature</th>
<th>Relative importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned interface HFE in LFBBCTA</td>
<td>51.16%</td>
</tr>
<tr>
<td>Distance between planned and ideal exit point from LFBBCTA</td>
<td>14.22%</td>
</tr>
<tr>
<td>Distance between planned and ideal entry point to LFBBCTA</td>
<td>6.41%</td>
</tr>
<tr>
<td>Ideal distance in LFBBCTA</td>
<td>4.86%</td>
</tr>
<tr>
<td>Ideal distance in LFFFCTA</td>
<td>4.36%</td>
</tr>
<tr>
<td>Average heading</td>
<td>3.42%</td>
</tr>
<tr>
<td>Ideal distance</td>
<td>3.17%</td>
</tr>
<tr>
<td>Ideal distance in LFRRCTA</td>
<td>2.34%</td>
</tr>
<tr>
<td>Planned local HFE in LFBBCTA</td>
<td>1.89%</td>
</tr>
<tr>
<td>Ideal distance in LECMCTA</td>
<td>1.87%</td>
</tr>
<tr>
<td>Ideal distance in LFMMCTA</td>
<td>1.84%</td>
</tr>
<tr>
<td>Ideal distance in LECBCTA</td>
<td>1.75%</td>
</tr>
<tr>
<td>Reference FL in LFBBCTA</td>
<td>0.69%</td>
</tr>
<tr>
<td>Take-off time - cosine</td>
<td>0.43%</td>
</tr>
<tr>
<td>Flights per ATCO in LECMCTA</td>
<td>0.22%</td>
</tr>
<tr>
<td>Take-off time - sine</td>
<td>0.20%</td>
</tr>
<tr>
<td>Flights per ATCO in LFBBCTA</td>
<td>0.17%</td>
</tr>
<tr>
<td>Flights per ATCO in LFFFCTA</td>
<td>0.16%</td>
</tr>
<tr>
<td>Flights per ATCO in LFMMCTA</td>
<td>0.16%</td>
</tr>
<tr>
<td>Flights per ATCO in LFRRCTA</td>
<td>0.16%</td>
</tr>
<tr>
<td>Flights per ATCO in LECBCTA</td>
<td>0.15%</td>
</tr>
<tr>
<td>Number of the day</td>
<td>0.11%</td>
</tr>
<tr>
<td>Flights crossing LFBBCTA</td>
<td>0.10%</td>
</tr>
<tr>
<td>Weekday - cosine</td>
<td>0.07%</td>
</tr>
<tr>
<td>Weekday - sine</td>
<td>0.07%</td>
</tr>
</tbody>
</table>
Sources of Flight Inefficiency
Applicability and Future Research

Applications:

- Performance assessment
- Performance optimisation

Future developments:

- Include further factors: weather, military...
- Include further KPIs: fuel, delay...
- Study seasonality effects
- Upscale to the Network, study specificities and commonalities
Application of Machine Learning for ATM Performance Assessment
Identification of Sources of En-Route Flight Inefficiency
Contact: aviation@nommon.es

Thank you very much for your attention!

This project has received funding from the SESAR Joint Undertaking under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 699303

The opinions expressed herein reflect the author’s view only. Under no circumstances shall the SESAR Joint Undertaking be responsible for any use that may be made of the information contained herein.