Comparing Scheduled Block Time Setting in Europe and China Based on Multiple Linear Regression

DAI Wei, MA Lingling and ZHAO Yifei, Civil Aviation University of China
Rainer KOELLE, EUROCONTROL Performance Review Unit
About CAUC

Where are we

CAUC

1000+ Teachers
20,000+ Undergraduate Students
1200+ Graduate Students
Introduction

Motivation

• Traffic growth have been seen in civil aviation in both Europe and China. In 2016, the raise of number of flights was 2.4% in Europe and 7.89% in China.

• Traffic growth supports thriving economics. Enhancing air traffic performance while not impeding the traffic growth remains a great challenge.

Flight Scheduling

• Flight schedule design is a key function in airline business planning. It also plays a core role in determining air traffic demand.

• Good scheduling improves air traffic performance by influencing capacity-demand-balance.
Scheduled Block Time (SBT)

- **Block time & scheduled block time**

- **SBT** is a lever that affects reliability and profitability. It's a key component of an airline's operational and cost performance. It also points at the efficiency impact of air navigation.
Part 2
Methodology

PART 02
Methodology

Study Approach

Original Database of China

Original Database of Europe

Data Pre-Processing

Chinese Dataset for Modeling

European Dataset for Modeling

Multiple Regression Model

Chinese SBT Model

Hub Airport Analysis

European SBT Model

Coefficients Comparing

Differences in SBT Setting Behavior between CHN and EUR

Previous Results

Industry Practice
Methodology

SBT Modeling

- The SBT model for the study is based on literature and augmented for both, the Chinese and European, context.
- Multiple linear regression has been proved working well in SBT modeling in literature.

Data Preparation

- In Europe, the air traffic network was approximated by all flights between 20 main airports chosen based on the cumulative number of flights connections of OD pairs for 2014 through 2016. Approximately 2.55 million pieces of data were used for the fitting.
- In China, a dataset covering all flights from 2014 to 2016 was employed. After careful data cleaning, records of approximately 7.55 million flights were used in the modeling.
The work of this study was implemented in an open-source “eco-system”:

- **Source data**
- **Analytical data**
- **Analysis, model**
- **Results & visualisations**

**Methodology**

“Joint & Collaborative” Data Analysis – Reproducible Research

- R/Rstudio
- Rmarkdown, knitr
- Git/Bitbucket
- Tidyverse “packages”
Part 3
EUR/CHN System Specifications
China and Europe show similarities in terms of their regional civil aviation systems, while the air traffic in Europe is approximately twice as high as the traffic in China.

<table>
<thead>
<tr>
<th></th>
<th>Year 2016</th>
<th>China</th>
<th>Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airspace Area (10^6 km²)</td>
<td>10.8</td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td>Number of Air Traffic Controllers</td>
<td>8522</td>
<td>17370</td>
<td></td>
</tr>
<tr>
<td>Number of Flights (million)</td>
<td>4.96</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Flight Density (flight hours p. km²)</td>
<td>0.71</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Number of Airports</td>
<td>218</td>
<td>415</td>
<td></td>
</tr>
<tr>
<td>Data provider</td>
<td>ATMB</td>
<td>EUROCONTROL</td>
<td></td>
</tr>
</tbody>
</table>
China and Europe show similarities in terms of their regional civil aviation systems, while the air traffic in Europe is approximately twice as high as the traffic in China.
On the airport level there is a stronger difference between China and Europe in terms of air traffic characteristics and traffic concentration.
Part 4
Scheduled Block Time Modeling
Scheduled Block Time Modeling

Model Description

- Distribution of difference phases of block times

<table>
<thead>
<tr>
<th>ZBAA-ZSPD</th>
<th>Departure Delay</th>
<th>Taxi-Out Time</th>
<th>En-Route Flying Time</th>
<th>Taxi-In Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>32.511</td>
<td>23.051</td>
<td>105.906</td>
<td>7.561</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>79.185</td>
<td>28.905</td>
<td>10.557</td>
<td>6.320</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>2.436</td>
<td>1.254</td>
<td>0.100</td>
<td>0.836</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EGLL-EDDF</th>
<th>Departure Delay</th>
<th>Taxi-Out Time</th>
<th>En-Route Flying Time</th>
<th>Taxi-In Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>10.587</td>
<td>20.146</td>
<td>64.486</td>
<td>10.132</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>22.444</td>
<td>6.655</td>
<td>7.077</td>
<td>3.790</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>2.120</td>
<td>0.330</td>
<td>0.110</td>
<td>0.374</td>
</tr>
</tbody>
</table>
Scheduled Block Time Modeling

Scheduled Block Time Model Components

- Basis: “historic” actual block time
  - taxi-out time
  - non-taxi-out time (elapsed time between actual take-off and actual in-block time)
  - departure delay

- Airport-pair characteristic (i.e. hub, non-hub)
- Aircraft type characteristic (i.e. wake vortex category)
- Temporal characteristic – seasonal variation
- Spacial characteristic → Great-circle distance
Scheduled Block Time Modeling

Model Description

- **SBT Model**

\[
SBT_{f,y}^{T_0} = \alpha_1 \times T_0 + \alpha_2 \times \text{non}T_0 + \alpha_3 \times \text{dep}T_0
\]

\[
+ \sum_{i=1}^{5} \beta_i \times dT_0 + \sum_{i=1}^{5} \gamma_i \times dn\text{on}T_0 + \sum_{i=1}^{5} \lambda_i \times d\text{dep}
\]

\[
+ \epsilon_1 \times HUB_0 + \epsilon_2 \times HUB_D + \delta \times \text{Vortex} + \pi \times \text{GCD} + \mu \times \text{Season}
\]

\[
+ \text{const}
\]

- **XX_{0.5} 50^{th} percentile**

\(dXX_{i+4,i+5}\) difference between adjacent 10^{th}-percentiles (characterisation of distribution)

100^{th} percentile replaced by 99.8^{th} to remove the outliers in actual calculation
Scheduled Block Time Modeling

Hub Airport Analysis

- In the absence of a globally accepted definition of hub airport, this project studied a numerical approach to qualify hub characteristic.

- The ratio of non-regional connections to the total number of air services at an airport has been studied.
### Scheduled Block Time Modeling – Model Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>China coefficient</th>
<th>p value</th>
<th>Europe coefficient</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>13.6613</td>
<td>&lt; 0.0001</td>
<td>6.8292</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>TO(_{0.5})</td>
<td>0.1845</td>
<td>&lt; 0.0001</td>
<td>0.7420</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>nonTO(_{0.5})</td>
<td>0.5421</td>
<td>&lt; 0.0001</td>
<td>0.9258</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dep(_{0.5})</td>
<td>0.0317</td>
<td>&lt; 0.0001</td>
<td>0.2092</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dTO(_{5,6})</td>
<td>0.0823</td>
<td>&lt; 0.0001</td>
<td>0.4424</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dTO(_{6,7})</td>
<td>0.0262</td>
<td>&lt; 0.0001</td>
<td>0.3120</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dTO(_{7,8})</td>
<td>0.0267</td>
<td>&lt; 0.0001</td>
<td>0.1480</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dTO(_{8,9})</td>
<td>-0.0235</td>
<td>&lt; 0.0001</td>
<td>-0.3926</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dTO(_{9,10})</td>
<td>0.0008</td>
<td>&lt; 0.0001</td>
<td>0.0066</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dnonTO(_{5,6})</td>
<td>0.4680</td>
<td>&lt; 0.0001</td>
<td>0.1140</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dnonTO(_{6,7})</td>
<td>0.5017</td>
<td>&lt; 0.0001</td>
<td>0.2968</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dnonTO(_{7,8})</td>
<td>0.4644</td>
<td>&lt; 0.0001</td>
<td>0.0749</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>dnonTO(_{8,9})</td>
<td>0.4805</td>
<td>&lt; 0.0001</td>
<td>0.2398</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>China coefficient</th>
<th>p value</th>
<th>Europe coefficient</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dnonTO(_{9,10})</td>
<td>0.0430</td>
<td>&lt; 0.0001</td>
<td>0.0004</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ddep(_{5,6})</td>
<td>-0.0915</td>
<td>&lt; 0.0001</td>
<td>-0.0722</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ddep(_{6,7})</td>
<td>0.0055</td>
<td>0.0151</td>
<td>-0.0593</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ddep(_{7,8})</td>
<td>-0.0097</td>
<td>&lt; 0.0001</td>
<td>-0.1666</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ddep(_{8,9})</td>
<td>0.0225</td>
<td>&lt; 0.0001</td>
<td>0.0029</td>
<td>0.3900</td>
</tr>
<tr>
<td>ddep(_{9,10})</td>
<td>0.0031</td>
<td>&lt; 0.0001</td>
<td>0.0041</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>HUB(_O)</td>
<td>7.7746</td>
<td>&lt; 0.0001</td>
<td>0.4528</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>HUB(_D)</td>
<td>2.3286</td>
<td>&lt; 0.0001</td>
<td>0.5654</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Vortex</td>
<td>-0.2849</td>
<td>&lt; 0.0001</td>
<td>-1.2851</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>GCD</td>
<td>0.0362</td>
<td>&lt; 0.0001</td>
<td>0.0070</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Season</td>
<td>1.9948</td>
<td>&lt; 0.0001</td>
<td>0.7614</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.8499</td>
<td></td>
<td>0.8268</td>
<td></td>
</tr>
</tbody>
</table>

The ordinary least-square (OLS) method was used for the linear regression to estimate coefficients of fitting.
Scheduled Block Time Modeling

Model Results

Coefficients of taxi-out (TO)

Coefficients of nonTO

Coefficients of departure delay
Part 5
Conclusions and Discussions
Conclusions and Discussions

• The SBT model shows a good fit for the Chinese (~85%) and European (~82%) context to model SBT setting behaviour in both regions.
  ➔ supports comparison and explanation of system similarities and differences

• Conceptually, higher flight phase times or delays should be positively related to the SBT. A variety of variable coefficients with a negative sign suggest that SBT decreases when the respective variable increases!
  ➔ This can lead to misleading interpretation of the results of the model fitting.

• Both the approach (multiple linear regression with OLS) and the identification of variables strongly influence the results.
  ➔ further research is needed concerning
    • expressiveness of the model variables
    • analysis of linear correlation

• Future work: analyzing the influence of the model on traffic demand and system performance by changing the value of coefficients.
Thank You!