A novel machine learning model to predict abnormal Runway Occupancy Times and observe related precursors

Seattle ATM R&D seminar

Floris Herrema
27th of June 2017
Overview

- Introduction and Objective
- Approach
- Charles de Gaulle and Vienna data
- Methodology
- Conclusions and Recommendations
- Prototype live trial
### Arrival mixed mode per single runway operation

<table>
<thead>
<tr>
<th><strong>Arrival Runway Occupancy Times (AROT)</strong></th>
<th><strong>Next day Post processing (next day)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence earlier</td>
<td>Multiple Runway Aiming Points (MRAP)</td>
</tr>
</tbody>
</table>

**1 min - 5 minutes**
- Tactical tool alert controller on extending AROT due to:
  - Wind or bad visibility
  - Predicted missed runway exit
  - Runway excursion
  - Cut-off taxi ways leads to hold line up clearance for departure

**+- 15 min**
- Planning tool for departure only
- Sequence algorithm could change for arrival and flow departure

**1 h – 3h**
- Strategic for ATCO supervisor (traffic manager)
  - Coordination of the runway configuration
  - Selection criteria landing on which runway
  - Sequence algorithm could change
  - Input for AMAN/DMAN and flow manager
Can we better predict Arrival Runway Occupancy Time?

- We are working on solutions impacting the Runway for maintaining runway Arrival and Departure throughput.

- As a first outcome of this process we have looked at “better predicting the True airspeed profile, Time to fly and Taxi-Out Times”.

- This produced interesting first descriptive results.

- This is the objective of my study.

- To develop a real time model that forecasts the abnormal Arrival Runway Occupancy Time (AROT) for different aircraft types, operational parameters and weather conditions.
Runway will focus on the following sub scenarios since these are the most limiting factors:

- Unstable Approach
- ATC Procedures
- RECAT-EU
- Reduction of Separation
- Go around
- Taxi-Out Times
- Runway excursiion
- Runway Occupancy Time

Introduction - the problem
Introduction - the problem

In this study we focus on the contributing factor Arrival Runway Occupancy Times

- Unstable Approach
- ATC Procedures
- RECAT-EU
- Reduction of Separation
- Go around
- Taxi-Out Times
- Runway Occupancy Time
- Runway excursion
Introduction – What is AROT?

Currently, there is no support system that ensures controller awareness on abnormal Arrival Runway Occupancy Times during High Intensity Runway Operations (HIRO).
Prior studies reported in the literature have attempted to predict AROT, however these predictions are lacking on;

1. Comparing key AROT features using statistical methods and ML feature techniques.

2. Applying and combining;
   - feature selection techniques before training the model.
   - feasible state of the art ML algorithms during training of the model.

3. Providing decision tree distributions associated with the individual AROT predictions.

4. Developing algorithm that is capable of real-time computation
• Consider real data from Charles de Gaulle (CDG) airport
• 2 year of data (2014/2015)
• Retrain the model every day with 2 hours of data
• Assess AROT for each possible Runway and Aircraft type
• For instance:
  • Air France A320 from London
  • Runway 08L
  • Start prediction 08:00
  • Update frequency 5 min
  • Number of simulations 1500
## Data collection CDG

350,000 Arrival flights from CDG

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>01.01.2015</td>
<td>1</td>
<td>2015</td>
<td>AF3007</td>
<td>AF007</td>
<td>F80JU</td>
<td>AFR</td>
<td>AF</td>
<td>AIR FRANKF</td>
<td>JKF</td>
<td>01-01-2015 07:49</td>
<td>01-01-2015 07:35</td>
<td>01-01-2015 08:30</td>
<td>01-01-2015 08:30</td>
<td>01-01-2015 07:54</td>
<td>01-01-2015 07:49</td>
</tr>
<tr>
<td>01.01.2015</td>
<td>1</td>
<td>2015</td>
<td>AF359</td>
<td>AF093</td>
<td>F80SU</td>
<td>AFR</td>
<td>AF</td>
<td>AIR FRANKF</td>
<td>HND</td>
<td>01-01-2015 05:14</td>
<td>01-01-2015 05:05</td>
<td>01-01-2015 05:30</td>
<td>01-01-2015 05:30</td>
<td>01-01-2015 05:23</td>
<td>01-01-2015 05:14</td>
</tr>
<tr>
<td>01.01.2015</td>
<td>1</td>
<td>2015</td>
<td>AF345</td>
<td>AF078</td>
<td>F80LG</td>
<td>AFR</td>
<td>AF</td>
<td>AIR FRANKF</td>
<td>BZU</td>
<td>01-01-2015 05:26</td>
<td>01-01-2015 05:17</td>
<td>01-01-2015 05:40</td>
<td>01-01-2015 05:40</td>
<td>01-01-2015 05:23</td>
<td>01-01-2015 05:26</td>
</tr>
<tr>
<td>01.01.2015</td>
<td>1</td>
<td>2015</td>
<td>AF341</td>
<td>AF080</td>
<td>F80LG</td>
<td>AFR</td>
<td>AF</td>
<td>AIR FRANKF</td>
<td>PMR</td>
<td>01-01-2015 05:34</td>
<td>01-01-2015 05:26</td>
<td>01-01-2015 05:48</td>
<td>01-01-2015 05:48</td>
<td>01-01-2015 05:54</td>
<td>01-01-2015 05:54</td>
</tr>
<tr>
<td>01.01.2015</td>
<td>1</td>
<td>2015</td>
<td>AF342</td>
<td>AF081</td>
<td>F80ML</td>
<td>AFR</td>
<td>AF</td>
<td>AIR FRANKF</td>
<td>CMN</td>
<td>01-01-2015 07:26</td>
<td>01-01-2015 07:15</td>
<td>01-01-2015 08:00</td>
<td>01-01-2015 08:00</td>
<td>01-01-2015 06:00</td>
<td>01-01-2015 06:00</td>
</tr>
</tbody>
</table>

**Notes:**
- CDG: Charles de Gaulle Airport
- SDOT: Scheduled Departure Time
- EDOBT: Estimated Departure Time
- TOBT: Time of Arrival
- ASOT: Actual Departure Time
- CTOT: Current Time
Prior studies reported in the literature have attempted to predict AROT, however these predictions are lacking on:

1. **Comparing key AROT features using statistical methods and ML feature techniques.**

2. Applying and combining;
   - feature selection techniques before training the model.
   - feasible state of the art ML algorithms during training of the model.

3. Providing decision tree distributions associated with the individual AROT predictions.

4. Developing algorithm that is capable of real-time computation
Before feasible Machine Learning techniques can be applied the AROT response is extracted by calculating the time between the ALDT and the RWEP.
Introduction – What is AROT?

- AROT versus different CDG runway exits for ICAO categories
### TXOT Prediction variables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Anne Year</td>
</tr>
<tr>
<td>2.</td>
<td>Caractredevol Commercial or private flight</td>
</tr>
<tr>
<td>3.</td>
<td>CodeIATA IATA code company</td>
</tr>
<tr>
<td>4.</td>
<td>CodeAeroportOACI Airport origin ICAO code</td>
</tr>
<tr>
<td>5.</td>
<td>CodeAeroportIATA Airport origin IATA code</td>
</tr>
<tr>
<td>6.</td>
<td>Compagnie Airline</td>
</tr>
<tr>
<td>7.</td>
<td>Crosswind Crosswind vector</td>
</tr>
<tr>
<td>8.</td>
<td>DateReal Actual date</td>
</tr>
<tr>
<td>9.</td>
<td>Deep landing The runway length available beyond the touchdown point</td>
</tr>
<tr>
<td>10.</td>
<td>IdentifiantvolATC ATC call sign</td>
</tr>
<tr>
<td>11.</td>
<td>Long flare Estimate the start of the flare until touchdown</td>
</tr>
<tr>
<td>12.</td>
<td>Mois Month</td>
</tr>
<tr>
<td>13.</td>
<td>NumFlight</td>
</tr>
<tr>
<td>14.</td>
<td>Postedestationnement</td>
</tr>
<tr>
<td>15.</td>
<td>QFU</td>
</tr>
<tr>
<td>16.</td>
<td>Semaine Week</td>
</tr>
<tr>
<td>17.</td>
<td>Tailwind</td>
</tr>
<tr>
<td>18.</td>
<td>Temp Temperature</td>
</tr>
<tr>
<td>19.</td>
<td>TimeReal Actual time of the day</td>
</tr>
<tr>
<td>20.</td>
<td>Typeavion Aircraft type</td>
</tr>
<tr>
<td>21.</td>
<td>Visibility METAR visibility conditions</td>
</tr>
<tr>
<td>22.</td>
<td>ACSpeedPoint Speed of the aircraft at 2NM, and 1NM out, Threshold and the Runway Exit Point (RWEP)</td>
</tr>
<tr>
<td>23.</td>
<td>ALDT Actual Landing Time</td>
</tr>
</tbody>
</table>

### Additional TXOT Prediction variables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.</td>
<td>Arrival runway throughput The amount of landings that is performed on the runway during the last 30 min</td>
</tr>
</tbody>
</table>

- **23 variables are used to train the model**
- **ALDT and RWEP are only used to extract the AROT response**
- **1 variable is added to the model**
Prior studies reported in the literature have attempted to predict AROT, however these predictions are lacking on;

- 1. Comparing key AROT features using statistical methods and ML feature techniques.

- 2. Applying and combining;
  - feature selection techniques before training the model.
  - feasible state of the art ML algorithms during training of the model.

- 3. Providing decision tree distributions associated with the individual AROT predictions.

- 4. Developing algorithm that is capable of real-time computation
Methodology – Data preparation

We had 22 variables

Top 10 Features for AROT

Consistent results leading to top 10 explanatory variables

Two feature selection techniques were applied

Top 3 variables identical to statistical approach:
1. QFU
2. Aircraft type
3. Arrival RWY throughput
First we learn the AROT for 350,000 flights with the top 10 features. Four models will be learned using the Lasso, Multi-Layer perception, Neural Networks and the Regression forests technique.

Second we merge the predicted AROT results of all 4 models to one final matrix for 350,000 flights.

Finally we apply the regression tree technique to the final matrix obtained in the previous step.

Using 10 features instead of all features gives us a model that:
- Has a faster computational time,
- Is more robust to changes
- Has a similar MSE
Methodology – MSE metric

- Before applying the regression tree technique, the following performances are obtained:

  - Max Squared Error = 18.7 sec
  - Mean Squared Error = 2.6 sec
An abnormal AROT is considered when it is $+2\sigma$ deviated from the normal standard deviation mean.

- Outliers example of abnormal AROT flights for training the model with 10 features.
Prior studies reported in the literature have attempted to predict AROT, however these predictions are lacking on:

- 1. Comparing key AROT features using statistical methods and ML feature techniques.
- 2. Applying and combining;
  - feature selection techniques before training the model.
  - feasible state of the art ML algorithms during training of the model.
- 3. Providing decision tree distributions associated with the individual AROT predictions.
- 4. Developing algorithm that is capable of real-time computation
Methodology – Regression tree

- MSE versus Tree Depth for different leaf size and amount of features

![Graph showing MSE versus Tree Depth for different leaf size and amount of features. The graph displays two lines: one for 600 leaf size with all features and another for 600 leaf size with top 10 features. The x-axis represents the Min Leaf Size ranging from 100 to 900, and the y-axis represents the cross-validated error ranging from 0 to 8. The graph also shows a decrease in MSE with increasing tree depth.](image-url)
During regression tree modelling we observe AROT precursors which is done by extracting ‘what-if’ scenarios based on the 10 important features.

Air France A320 from London Runway 08L
Start prediction 08:00
Update frequency 5 min
Number of simulations 1500

1. ArrRwyThroughput < 30
2. TypeAvion = Medium
3. Visibility < 935 m
4. Crosswind > 14 kts
5. Time between 07:00 and 09:00
AROT = 62 sec
MSE = 2.8 sec
The probability distributions we considered for the terminal leave nodes include the Normal, Gumbel, Gamma, and F distributions. The following equation shows the Gumbel distribution which fits best:

\[ f(x) = \frac{1}{\beta} e^{-\left(\frac{x-\mu}{\beta} + e^{\frac{x-\mu}{\beta}}\right)} \]
What do we do with this?

- We know what influence what…

- Remember the question:
  - Can we better predict the *abnormal* AROT profile?
    - for better predicting the go-arounds
      - for better predicting the number of landings
    - for better predicting the arrival runway throughput.

YES!
Prior studies reported in the literature have attempted to predict AROT, however these predictions are lacking on:

1. Comparing key AROT features using statistical methods and ML feature techniques.

2. Applying and combining;
   - feature selection techniques before training the model.
   - feasible state of the art ML algorithms during training of the model.

3. Providing decision tree distributions associated with the individual AROT predictions.

4. Developing algorithm that is capable of real-time computation
Methodology – Real time modelling

- Based on what we have learned in the previous steps and the data availability we develop a prototype model to forecast abnormal AROT at CDG and VIE.

```
>> CDGRwyOccupancyTime
ML technique (Combine Lasso, Multi-Layer perception and Neural Networks (yes or no)):yes
Forecast window (min):120
Number of simulations:1500
Runway:08R
Forecast resolutions (separated with commas):1,5,15,60
Update frequency (min):5
Starting time to predict (YYYY-MM-DD HH:MM:SS):2015-07-20 08-00-00
Ending time to predict (YYYY-MM-DD HH:MM:SS):2015-07-20 14-00-00
```

- The operational data sets can be accessed in real time with exception of the actual variables; ALDT and RWEP
Connect real time modelling to CAST

Real Time Airport Database (AODB)

Interactive Dashboard
Methodology – Prototype model

Date and time when prediction is made: 20-Jul-2015 08:00:00 from all stands to runway 08R

Runway 08L
Start prediction 08:00, 29th of February 2016
Update frequency 5 min
Number of simulations 1500

For this case study the first 30 minutes have a significantly lower error compared to the remaining 90 minutes prediction.
Prototype model

- Average error differences per trial for the time prediction window 0-30 minutes and 30-90 minutes.

<table>
<thead>
<tr>
<th></th>
<th>0 - 30 minutes</th>
<th>30 - 90 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>6%</td>
<td>10%</td>
</tr>
<tr>
<td>Trial 2</td>
<td>4%</td>
<td>8%</td>
</tr>
<tr>
<td>Trial 3</td>
<td>4%</td>
<td>9%</td>
</tr>
<tr>
<td>Trial 4</td>
<td>2%</td>
<td>6%</td>
</tr>
</tbody>
</table>
Conclusion

- Machine learning techniques seem to be very effective in terms of abnormal AROT performance prediction.

- Predictions can be computed in just a few seconds after the tree has been extracted.

- The tree help make quick and intuitive decisions.

- Such predictions are useful for:
  - Tactical tool to alert ATC on extending ROT
  - Strategic tool to support ATC supervisor on;
    - Coordination of the runway configuration and changing of the sequence algorithm
    - Input AMAN/DMAN
Recommendations

- The methodology can easily be transposed to any other airport processes such as the prediction of the runway exit taken.

- Include runway conditions and FDM data.

- Presenting Vienna AROT results

- Live trial at CDG
  - To validate the accuracy of the predicted AROT per flight.
  - To validate the predicted number of landings per hour.
  - To test the prediction speed in real time.
Thank you

Any questions?

- With this study a better prediction will be established of the AROT patterns and precursors, this research will stimulate further Dynamic Pair-wise Separation studies....