Novel terminal arrival airspace robustness metrics via topological density clustering
A case study of the Chicago O’Hare International Airport

Max Z. Li, Megan S. Ryerson, Ph.D.
University of Pennsylvania

June 26, 2018
Presentation Overview

1. Introduction and motivation
2. Case study airport (Chicago O’Hare International Airport)
3. Data overview
   - Good vs. bad weather
   - Using FlightAware ADS-B trajectory data
   - Trajectory re-sampling
4. Density-based spatial clustering of applications with noise (DBSCAN) clustering within the terminal arrival airspace
5. Results + discussion + implications
6. Future work and conclusions
The Terminal Airspace

Complex and dynamic airspace within the NAS
Identified need for increased resiliency – how tolerant, or robust, airport-airspace operations are to operational perturbations – and predictability – reduction in variability and uncertainty – within the NAS (FAA, 2016).
Identified need for increased resiliency – how tolerant, or robust, airport-airspace operations are to operational perturbations – and predictability – reduction in variability and uncertainty – within the NAS (FAA, 2016).

Motivation: Effects of convective weather on different terminal arrival airspaces are not well understood

⇒ Difficult to compare terminal airspace resiliency across airports
⇒ Difficult for effective airport peer-to-peer learning
New metrics + Augmentation + Better Identification + Flexibility

- New terminal airspace resiliency metrics
- Potential to augment existing metrics to be spatially-sensitive, e.g. AARs, probabilistic capacity profiles (*Buxi & Hansen, 2013*)
- Identification of shifting terminal airspace routing structures
- Applicable to other terminal airspaces with potential to expand to metroplexes
Why Chicago-O’Hare (ORD) and the surrounding terminal airspace (C90 TRACON)?
Case study airport/airspace choice

- **Major domestic and international airport**
  - 75+ million passengers served
  - 850,000+ aircraft movements
  - 200+ direct-flight destinations

- **Dual hub** for two legacy US carriers (AAL, UAL)

- Previous research shows terminal airspace delays at ORD have **internal and external impacts** (Le et al., 2003), (Laskey et al., 2006), (Churchill, 2007), (Diana, 2009)
Ord Instrument Approach Procedures (IAPs)
7 in-use runway (as of June 21, 2018); all parallel east-west runways with exception of 4L/22R and 4R/22L
8 active runways in final post-O’Hare Modernization Plan (OMP) airfield
ORD *west-flow* configuration

Arrive 27R/27L/28C; depart 22L/28R; AAR/ADR at 114/100 aircraft per hour

*Primary preferred configuration* at ORD (~70% of ops)
Arrive 9L/10C/10R; depart 10L/9R; AAR/ADR at 114/100 aircraft per hour

Secondary preferred configuration at ORD (~30% of ops)
Data Overview

Gather → Process → Cluster

**Time frame:** April → June, 2017

1. **Gather weather and trajectory data**
   - Historical ORD weather reports via METAR archive
   - Terminal airspace trajectory data via FlightAware
   - *Hourly configuration + AAR/ADR/SAER* via ASPM

2. **Subdivide trajectory data into “good” versus “bad” weather subsets**

3. **Standardize all trajectories via linear interpolation-based re-sampling**

4. **Cluster terminal airspace trajectories via DBSCAN**

5. **Calculate resiliency metrics** based off of clustering results
Bad weather criteria

Time period within ORD historical METAR report considered **bad weather** if **ANY** of the following criterion is satisfied:

- Visibility ($< 3$ miles)
- Lowest ceiling altitude ($< 1,000$ feet)
- Wind speed ($> 20$ knots)
- Wind gust ($> 20$ knots)
- Present weather code (**is one of the following ...**)
  - VCTS, VCTS -RA
  - TS, TSRA BR
  - FG, BR, RA BR

If none of the above criteria is satisfied $\Rightarrow$ time period considered **good weather period**
**NB: Collection periods**

<table>
<thead>
<tr>
<th>Weather condition</th>
<th>Weather condition criteria</th>
<th>Start time (hh:mm CDT, M/D/Y)</th>
<th>End time (hh:mm CDT, M/D/Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOD</td>
<td>- No bad weather criteria satisfied</td>
<td>16:55, 4/12/2017 → 00:43, 4/16/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>13:51, 4/16/2017 → 23:51, 4/16/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>07:55, 4/21/2017 → 23:56, 4/24/2017</td>
<td></td>
</tr>
<tr>
<td>BAD</td>
<td>- Visibility &lt; 3 miles OR</td>
<td>14:51, 4/9/2017 → 19:32, 4/10/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Lowest ceiling altitude &lt; 1,000 feet OR</td>
<td>10:20, 4/20/2017 → 23:00, 4/20/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Wind speed &gt; 20 knots OR</td>
<td>04:19, 4/29/2017 → 07:40, 4/30/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Wind gust &gt; 20 knots OR</td>
<td>10:05, 4/30/2017 → 12:05, 4/30/2017</td>
<td></td>
</tr>
</tbody>
</table>

Want approximately same number of flights in **good** ($n = 6,895$) and **bad** ($n = 7,817$) weather. Note the differences in the collection intervals between **good** and **bad** weather.
Walk-through with a **bad** weather example (1/2)

**Time frame:** 16:00 CDT, May 17 → 05:00 CDT, May 18

Wind speed + gust average $\sim$25 knots at 200° $\Rightarrow$ disruption to preferred east-west flow configuration at ORD
Wind speed + gust average $\sim 25$ knots at $200^\circ \Rightarrow$ disruption to preferred east-west flow configuration at ORD

Arrival runway configuration change $27L/27R/28C \rightarrow 22L/22R$ with corresponding AAR drop from $114 \rightarrow 56$ aircraft per hour.
The trajectory irregularity problem

We cannot easily apply DBSCAN cluster to the FlightAware trajectory data due to:

1. Sampling rate (i.e. interval between radar observations) not identical between flights
2. Length of trajectory not equal between flights

What we’re given: $M$ aircraft trajectories sampled irregularly, with lengths $N_1, N_2, ..., N_M$.

Goal: Re-sampled trajectory data set of $M$ aircraft in $\mathbb{R}^{N' \times 2 \times M}$ with common time index and common trajectory length $N'$. 

(NB: We only cluster on lat-long $(\varphi, \lambda)$, resulting in $\mathbb{R}^{N' \times 2 \times M}$ as desired output space)
The trajectory irregularity problem - SOLUTION

What we’re given: \( M \) aircraft trajectories sampled irregularly, with lengths \( N_1, N_2, \ldots, N_M \).

Trajectory re-sampling via linear interpolation

Goal: Re-sampled trajectory data set of \( M \) aircraft in \( \mathbb{R}^{N' \times 2 \times M} \) with common time index and common trajectory length \( N' \).
Data re-shaping

Entire trajectory data set is re-shaped so that:

- Each observation is one aircraft
- Each observation has $2N'$ variables containing properly-ordered $N'$ latitudes and $N'$ longitudes

Our data space is now $\mathbb{R}^{M \times 2N'}$, and we can now perform DBSCAN clustering.
Clustering via DBSCAN

DBSCAN clustering (process + parameter selection)

Parameter selection: $\varepsilon$ radius
Parameter selection: $m$ min. points

Construct $\mathcal{B}(\cdot, \varepsilon)$ around each point in $\mathbb{R}^{2N'}$ data space

Compute reachability between points via the usual norm on $\mathbb{R}^{2N'}$

$\mathbb{R}^{2N'}$

Cluster 1

$\mathcal{B}(\bullet, \varepsilon)$

Outlier

Cluster 2

= core points
= edge points
= outliers
= reachable

Assign trajectory cluster labels to individual aircraft

Longitude

Latitude

ORD
Results and Discussion

STAR cases and resiliency metrics

1. BENKY FOUR STAR (Southwest arrival gate)

2. VEECK THREE STAR (Southeast arrival gate)

3. Derived resiliency metrics
   - Number/scale of irregular clusters
   - Number/scale of flights within irregular clusters
Major southwest arrival gate for ORD
Major **southwest** arrival gate for ORD
Important numbers to keep in mind:

- 452 flights observed in good weather
- 831 flights observed in bad weather
- Same $\varepsilon$ (0.6) and $m$ ($\geq 6$) for both weather conditions

Now onto clustering results ...
Closer look: BENKY FOUR STAR (3/5)

Good weather clustering results
Bad weather clustering results
### Closer look: BENKY FOUR STAR (5/5)

<table>
<thead>
<tr>
<th>STAR (Corner-post) Weather Condition</th>
<th>( n )</th>
<th>( \varepsilon )</th>
<th>( m )</th>
<th>Cluster Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOD</td>
<td>452</td>
<td>0.6</td>
<td>( \geq 6 )</td>
<td>1  Outlier 16</td>
</tr>
<tr>
<td>BAD</td>
<td>831</td>
<td>0.6</td>
<td>( \geq 6 )</td>
<td>2  East Flow; East Flow straight-in 258</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3  West Flow 172</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4  Flight Plan Change (FYTTE4) West Flow 6</td>
</tr>
<tr>
<td>BENKY FOUR (Southwest)</td>
<td></td>
<td></td>
<td></td>
<td>1  Outlier 38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2  Flight Plan Change (FYTTE4) West Flow; Flight Plan Change (FYTTE4) 22L/22R 21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3  East Flow; East Flow straight-in 132</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4  West Flow; 22L/22R 625</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5  Flight Plan Change (FYTTE4) East Flow 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6  Flight Plan Change (FYTTE4) East Flow with tromboning 7</td>
</tr>
</tbody>
</table>
Major southeast arrival gate for ORD
Major **southeast** arrival gate for ORD
Important numbers to keep in mind:

- **1592 flights** observed in **good weather**
- **1478 flights** observed in **bad weather**
- **Same $\varepsilon (0.8)$ and $m (\geq 7)$** for both weather conditions

Now onto clustering results ...
Closer look: VEECK THREE STAR (3/5)

Good weather clustering results
Bad weather clustering results
### Closer look: VEECK THREE STAR (5/5)

<table>
<thead>
<tr>
<th>VEECK THREE (Southeast)</th>
<th>GOOD</th>
<th>1592</th>
<th>0.8</th>
<th>≥ 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BAD</td>
<td>1478</td>
<td>0.8</td>
<td>≥ 7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>Event Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Outlier</td>
<td>38</td>
</tr>
<tr>
<td>2</td>
<td>East Flow</td>
<td>1129</td>
</tr>
<tr>
<td>3</td>
<td>West Flow</td>
<td>335</td>
</tr>
<tr>
<td>4</td>
<td>West Flow shortened IAP</td>
<td>90</td>
</tr>
<tr>
<td>1</td>
<td>Outlier</td>
<td>115</td>
</tr>
<tr>
<td>2</td>
<td>East Flow</td>
<td>359</td>
</tr>
<tr>
<td>3</td>
<td>West Flow</td>
<td>707</td>
</tr>
<tr>
<td>4</td>
<td>West Flow shortened IAP; 22L/22R</td>
<td>228</td>
</tr>
<tr>
<td>5</td>
<td>Flight Plan Change (TRTLL4/BENKY4) West Flow</td>
<td>41</td>
</tr>
<tr>
<td>6</td>
<td>East Flow Lake Michigan vectors</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>Flight Plan Change (TRTLL4/BENKY4) East Flow straight-in</td>
<td>16</td>
</tr>
</tbody>
</table>
FYTTE FOUR STAR clustering results for good (left) and bad (right) weather
TRTLL FOUR STAR clustering results for good (left) and bad (right) weather
WATSN THREE STAR clustering results for *good* (*left*) and *bad* (*right*) weather
WYNDE EIGHT STAR clustering results for good (left) and bad (right) weather
Spatially-aware resiliency metrics

Visual inspection of clustering results ⇒ significant spatial differences in terminal airspace operations

Common airspace metrics such as:

- AARs
- ADRs
- SAERs

do not take into account spatial differences, i.e. two periods may have the same AAR and SAER but using different arrival trajectories.
Incorporating spatial differences

**Goal:** Address insensitivity to spatial differences by rating STAR resiliency based off of clustering results

⇒ Label cluster as an IROP cluster if that operation modality does not match nominal ORD flow operations (west- or east-flow)

⇒ Compute % of arrivals utilizing IROP cluster for each STAR as:

\[
\% \text{ STAR arrivals } \in \text{IROP} = 100\% \times \frac{\sum \text{IROP} n}{n_{\text{STAR}}} \quad (1)
\]
Four metrics to consider for each STAR

1. Number of IROP clusters during bad weather
2. Difference between number of IROP clusters during good versus bad weather
3. Percent of arrivals in an IROP cluster
4. Change in percent of arrivals in an IROP cluster between good versus bad weather

Each STAR scored in all four spatially-aware resiliency metrics; score of 1 for best-performing STAR → score of 6 for worst-performing STAR
### Results and Discussion

#### Derived resiliency metrics

**And the winner is ...**

<table>
<thead>
<tr>
<th>STAR (Corner-post) Weather Condition</th>
<th>IROP Cluster ID</th>
<th>% arrivals in IROP cluster</th>
<th>Performance Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td># IROP clusters</td>
</tr>
<tr>
<td>BENKY FOUR (Southwest)</td>
<td>GOOD</td>
<td>1, 4</td>
<td>4.9%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 2, 5, 6</td>
<td>8.9%</td>
</tr>
<tr>
<td>FYTTE FOUR (Northwest)</td>
<td>GOOD</td>
<td>1</td>
<td>3.1%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 4, 5, 6</td>
<td>9.2%</td>
</tr>
<tr>
<td>TRILL FOUR (Southwest)</td>
<td>GOOD</td>
<td>1</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 4, 5, 6, 7, 8</td>
<td>9.7%</td>
</tr>
<tr>
<td>VEECK THREE (Southeast)</td>
<td>GOOD</td>
<td>1</td>
<td>2.4%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 5, 6, 7</td>
<td>12.4%</td>
</tr>
<tr>
<td>WATSN THREE [ESSPO THREE] (Southeast)</td>
<td>GOOD</td>
<td>1</td>
<td>2.8%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 4, 5, 6, 7</td>
<td>20.1%</td>
</tr>
<tr>
<td>WYNDE EIGHT (Northwest)</td>
<td>GOOD</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 5</td>
<td>3.9%</td>
</tr>
</tbody>
</table>
Northern airspace sectors scored better than southern sectors
  • WYNDE EIGHT + FYTTE FOUR average resiliency score of 7.5 (out of max. 24)
  • Southern arrival gates/STARs average resiliency score of 16.25 (out of max. 24)

WYNDE EIGHT rated as particularly robust
  • Little difference between good & bad weather clusters
  • Low number of arrivals in IROP clusters that do appear
Possible reasons for this difference:

1. Larger amount of incoming traffic from major southeastern airports
2. Operations at Chicago-Midway (MDW), located 15 mi. (24 km.) southeast of ORD

C90 terminal arrival airspace flows; ORD arrivals (blue) and MDW arrivals (red) depicted \((Li & Ryerson, 2018)\).
Terminal controller preference? SW vs. SE arrival gates

SW gates show high cluster fragmentation but low usage levels

SE gates behaviors are in general reversed from SW gates
Summary of work

• Application of trajectory clustering in characterizing complex terminal arrival airspaces

• Scored ORD arrival airspace resources according to performance in convective weather
  • Four proposed + calculated resiliency metric

• Obtained insights into spatially-differing airspace robustness within C90 TRACON
  • Northern vs. southern airspace sectors
  • Insights into terminal controller preference
The work’s not done!

- More robust parameter selection as resultant metrics depend on cluster parameters
- Choice of good and bad weather criteria
- More sophisticated trajectory processing techniques
- Scheme for incorporating spatially-varying resiliency into existing metrics (AAR, ADR, SAER, etc.)
- Expand resiliency metric calculation to other airports/metroplexes
Thank you for your attention!

Questions? Comments?