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

Max Z. Li
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, Pennsylvania
lizhaoyu@seas.upenn.edu

Megan S. Ryerson, Ph.D.
Department of City and Regional Planning
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, Pennsylvania
mryerson@design.upenn.edu

Abstract— We leverage topological density clustering to analyze radar trajectory data, resulting in novel airspace robustness metrics with respect to convective weather in the terminal arrival airspace. Our method extracts trajectory clusters that occur during varying weather conditions and derives metrics based off of the quantity and quality of these clusters. We study these clusters for Chicago O’Hare International Airport (ORD), and obtain insights into the spatially-varying robustness of ORD arrival resources in convective weather conditions. We found that the southern arrival airspace resources at ORD fared worse in convective weather, indicated by a higher prevalence of and population within anomalous clusters. Our airspace robustness metrics via trajectory density clustering allow for cross-airport comparisons of resiliency, enhancing the ability of airport operators to learn and adapt traffic management strategies utilized at peer airports with similar airfield and airspace attributes.

Keywords: Air traffic flow management; Terminal arrival airspace; Topological density clustering; Arrival resource identification; Spatially-varying airspace robustness

I. INTRODUCTION

The terminal airspace surrounding a major airport is a complex and dynamic environment. To ensure efficiency and reduce delays and congestion, airport operators, airlines, and air traffic controllers seek to enhance system robustness: how tolerant airport-airspace operations are to operational perturbations. Within the terminal airspace, a key gap in knowledge exists that impedes increased system robustness: the effects of convective weather on different terminal arrival airspaces are not well understood. This renders cross-airport comparisons of resiliency difficult, complicating the ability of airport operators to learn and adapt traffic management strategies utilized at peer airports with similar airfield and airspace attributes. We use topological density clustering to analyze aviation trajectory data sets in order to address this gap. The actionable insights obtained from our analysis will lead to an increase in system robustness on the arrival end, cascading into overall improvements in terminal airspace efficiency, the benefits of which are felt throughout the National Airspace (NAS).

The rest of the paper is organized as follows: In Section 2 we present some current research literature in three pertinent subjects: (1) airspace complexity, (2) terminal airspace modeling with convective weather effects, and (3) current network and topological approaches in aviation. Section 3 provides an overview of the terminal arrival airspace at our case study airport, as well as our data collection and processing methods. We present our methodology in Section 4, and summarize the implementation of our method on the trajectory data set for our case study airport. We discuss our results and distill insights into the air traffic situation at the case study airspace in Section 5. Section 6 concludes with a summary of the motivation, methods, results, and insights presented in this paper, as well as future directions for our work.

II. LITERATURE REVIEW

We chose to characterize the terminal arrival airspace via topological density clustering motivated by the inherently topological – aircraft trajectories are paths embedded in \( \mathbb{R}^2 \) – and high-dimensional nature of aircraft radar trajectory data. This section provides an overview of the complex system we are working with (Section 2.A), current efforts in terminal airspace-level modeling (Section 2.B), and network-topology research in air transportation (Section 2.C).

A. The many facets of airspace complexity

The terminal, transition, and en route airspaces are highly complex environments with multiple deterministic and stochastic, linear and non-linear components, all of which can introduce uncertainty and compromise robustness within the practice of air traffic flow management (ATFM). In a comprehensive look at applying complexity science to air
transportation, [1] explores how high dimensional and sparse data sets, along with operational, equipment, and weather uncertainty lead to tremendous challenges in modeling, predicting, and improving the performance of the air transportation system. Examples of complexity factors throughout the air transportation system include number of aircraft, aircraft proximity measures, and density indicators [2]. However, these factors are largely limited to the en route airspace and do not provide a characterization of terminal airspace robustness and complexity, particularly when convective weather is a compounding variable.

Previous research focuses on a more macroscopic view of the air transportation system, choosing to examine the entirety of a given air route network (ARN) using network science [3, 4, 5]. Through such a wide scope of analysis, insights such as the varying levels of dynamic complexity and how that variability plays into delay and fuel inefficiencies have been uncovered [6, 7]. These insights help address another facet of airspace complexity: the impacts of different stakeholders – airport operators, airlines, air navigation service providers – on the ARN as a whole [8]. Each stakeholder has a different perspective of the ARN, resulting in a different evaluation of system utility and performance efficiency [9]. While these insights are critical quantitative foundations upon which routing decisions and airspace utilization should be made at the stakeholder level, they cannot be substituted as metrics that can be controlled or optimized in a tactical ATFM sense.

B. Terminal airspace modeling with convective weather effects

Whereas we provided a purview of research literature scoped at the ARN level in the preceding subsection, we now zoom in to focus specifically on tactical ATFM terminal airspace models incorporating convective weather. When severe weather conditions develop within the terminal airspace, the pilots and the terminal controllers evaluate the situation and decide on a course of action from their perspective. [10] found that weather location and intensity, as well as flight range and total time spent in the terminal area are the critical factors that determine whether the flight crew attempts to fly through convective weather. Models from the controller perspective examined aspects such as the optimal runway configuration switching times via a hybrid optimal control formulation [11]; the evolution of Airport Arrival Rates (AARs) throughout the day, conditioning on Terminal Area Forecasts (TAFs) [12]; the identification of terminal area flow patterns with respect to arrival fixes [13] and robust arrival routes [14, 15].

C. Current network and topological approaches to characterizing air transportation networks

In the scope of ARNs, graph-based approaches have been used to characterize the complexity of airspace sectors [16], to quantify the connectivity of the route network structure [17], and to find correlations between stakeholder metrics and the network topology of the ARN [18]. [19] applied methods in fluid flow dynamics to characterizing en route airspace corridors, resulting in a quantitative benefit assessment of new RNAV-enabled “Q” jet routes. [20] recognized the need for a multi-layered network approach in order to account for the added complexity of air traffic controller interactions.

Density-based clustering has been used in the en route, transition [21], and terminal airspace [22], but with proprietary radar data sets. We show that the same level of density-based clustering can be performed on publicly available data sets from FlightAware, given that appropriate re-sampling and interpolation is performed prior to clustering.

III. Data and case study airport preliminaries

We selected the terminal arrival airspace at Chicago O’Hare International Airport (ORD) to apply trajectory density clustering in order to characterize the air traffic situation during periods of good weather as well as inclement weather conditions. As a major domestic and international hub, terminal airspace arrival and departure delays at ORD cascade throughout the NAS, and have been shown to be significant factors in gate arrival delays for ORD departures [23, 24, 25, 26]. We provide insights on how the ORD terminal arrival airspace behaves in differing weather conditions through topological density clustering. Such insights can lead to the development of new control strategies during inclement weather at ORD, and reduce overall terminal airspace delay.

We provide an overview of the Chicago Terminal Radar Control Airspace, referred to by the Federal Aviation Administration (FAA) as the C90 TRACON. We will also define “good weather” and “bad weather” periods, and identify time periods for trajectory data collection. Given these time periods, we collect historical trajectory data for ORD arrivals via FlightAware, and provide a summary of the data. We also summarize our re-sampling and interpolation heuristic.

A. The C90 TRACON

The C90 TRACON is roughly centered on the eastern suburbs of Chicago, and encompasses the terminal arrival and departure airspaces of ORD as well as Chicago Midway International Airport (MDW). Several smaller airports primarily handling general aviation operations are also included within the purview of C90. An arriving aircraft bound for ORD would be passed on at an appropriate altitude from a neighboring en route sector to a C90 terminal controller. The terminal controller ensures that the arriving aircraft is following a filed Standard Terminal Arrival Route (STAR), which are standardized procedures published by the FAA to help funnel arriving traffic into busier terminal airspaces. We consider the airspace fix that designates an entrance into the terminal arrival airspace as the *arrival fix* particular to a specific STAR. For example, the TRTLL arrival fix accompanies the TRTLL FOUR STAR, and aircraft filing this
STAR procedure must be at a specific altitude when crossing the TRTLL arrival fix. The remainder of the STAR procedure guides the aircraft in a fix-to-fix manner from the entrance of the terminal arrival airspace towards the airport. We provide a diagram of the ORD terminal arrival airspace with appropriate overlaid STARs in Figure 1. For the purposes of this paper, we consider the seven most often used STARs as depicted in Figure 1. The STARs were obtained from the FAA digital publications archive [27].

Figure 1. The ORD terminal arrival airspace with pertinent STARs overlaid

B. Convective weather criteria and collection periods

In order to assemble a representative data set containing approximately the same amount of arrival operations conducted under good weather and convective weather conditions, we parsed through historical Meteorological Terminal Aviation Routine Weather Reports (METARs) archived for ORD. We record in Table 1 the criteria for a period of time during which ORD experienced inclement weather, as well as the associated start and end times for each time period. The five variables of interest contained within the METAR reports that were visibility, altitude of the lowest ceiling, present weather code, wind speed, and wind gusts. All of these variables indicate conditions that could impact regular operations within C90, necessitating control actions such as runway configuration changes and implementation of traffic management initiatives (TMIs).

Within the combined trajectory data set, a total of 6,895 arrival aircraft trajectories were collected during all good weather time periods, and a total of 7,817 arrival aircraft trajectories were collected during all bad weather time periods. When the trajectory data set is decomposed according to the filed STAR procedure, each STAR procedure depicted in Figure 1 has a representative subset. Operationally, the WATSN THREE STAR was often filed and used in lieu of ESSPO THREE. Thus, the trajectory density clustering analysis combines the two STARs under the WATSN THREE STAR as it is the dominant east-southeast corner-post within the ORD terminal arrival airspace. A planar trajectory plot of a sample subset of the bad weather trajectory data set colorized by the utilized STAR is given in Figure 2. The accompanying subsection provides motivation and precedence for the utility of publicly available radar trajectory data as well as a description of the variables contained within the data set prior to re-sampling and interpolation.

<table>
<thead>
<tr>
<th>Weather condition</th>
<th>Weather condition criteria</th>
<th>Start time (hmm CDT, MDY)</th>
<th>End time (hmm CDT, MDY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOD</td>
<td>No bad weather criteria satisfied</td>
<td>16:55, 4/12/2017 → 06:43, 4/16/2017</td>
<td></td>
</tr>
<tr>
<td>BAD</td>
<td>Visibility &lt; 3 miles OR Lowest ceiling altitude &lt; 1,000 feet OR Present weather code ∈ {VCTS, VCTS-RA, TS, TSRA BR, FG, BR, RA BR} OR Wind speed &gt; 20 knots OR Wind gust &gt; 20 knots</td>
<td>14:51, 4/9/2017 → 04:49, 4/30/2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10:20, 4/20/2017 → 04:19, 4/29/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>22:39, 4/26/2017 → 04:05, 4/30/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14:35, 4/30/2017 → 09:51, 5/01/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17:50, 5/10/2017 → 09:51, 5/17/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12:51, 5/16/2017 → 10:15, 5/20/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11:40, 6/10/2017 → 20:51, 6/10/2017</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Standard METAR weather codes: VCTS = thunderstorms in the vicinity; VCTS-RA = VCTS with light rain; TS = thunderstorms; TSRA BR = thunderstorm, rain, with mist; FG = fog; BR = mist; RA BR = rain with mist

Figure 2. Subset of ORD trajectory data set; arrival fixes given in yellow

C. Trajectory data summary

Researchers have successfully leveraged data obtained from publicly available sources such as FlightAware and Flightradar24 in aviation research. The scope of the data collection effort in the literature ranges from the trajectory of a single aircraft [28], to the speed distribution for a particular aircraft type [29], to large datasets containing many flights of interest [30, 31]. These web-based flight tracking service providers draw on their large networks of private and commercial Automatic Dependent Surveillance-Broadcast (ADS-B) receivers to supplement private airport radar datalinks, gathering accurate positional and movement data from flights operating over serviced regions. For the aviation research community, these providers are a convenient and
flexible alternative to obtain time-indexed positional and movement data.

Within the assembled trajectory data set of ORD arrivals, the call-sign and flight number is recorded, along with the filed route beginning with the Standard Instrument Departure (SID) and ending with the STAR at the terminal airspace of the destination airport. The critical information we extract from the route code is the filed STAR of the ORD arrival, as this is how we condition and segment our trajectory data sets with respect to ORD STARs. Positional information (latitude, longitude, altitude) and aircraft dynamics (course, direction, speed, climb or descent rate) are retrieved at 20 second to 1-minute intervals, depending on the location of the aircraft and the ADS-B coverage. Each observation is timestamped, and the origin of the observation is recorded as well.

D. Pre-clustering trajectory resampling

We first introduce some notation that will be used throughout the rest of the paper in order to concisely refer to the various components of our trajectory data set. Given an arriving aircraft within our trajectory data set with \( N + 1 \) recorded observations for the duration of its trajectory, each observation is indexed in time by vector \( \tilde{t} = [t_{\varphi_1, \lambda_1, z_1}, t_{\varphi_2, \lambda_2, z_2}, \ldots, t_{\varphi_{N+1}, \lambda_{N+1}, z_{N+1}}]' \) in \( \mathbb{R}^{N+1} \). The position of the aircraft is indexed by the matrix \( \mathcal{P} \) in \( \mathbb{R}^{(N+1) \times 3} \) given by (1), where each row corresponds to an entry in \( \tilde{t} \) and the columns contain latitude, longitude, and altitude information, respectively.

\[
\mathcal{P} = \begin{bmatrix}
\varphi_1 & \lambda_1 & z_1 \\
\varphi_2 & \lambda_2 & z_2 \\
\vdots & \vdots & \vdots \\
\varphi_{N+1} & \lambda_{N+1} & z_{N+1}
\end{bmatrix} \in \mathbb{R}^{(N+1) \times 3} \tag{1}
\]

Due to irregularity in the sampling intervals between two arrivals (i.e. one aircraft has a 30 second interval between two consecutive \( t_{[\varphi_i, \lambda_i, z_i]} \) and \( t_{[\varphi_{i+1}, \lambda_{i+1}, z_{i+1}]} \) whereas another aircraft has a 1 minute interval), even if the time vectors for all aircraft are normalized with respect to the initial time index \( t_{[\varphi_1, \lambda_1, z_1]} \), the normalized time vectors will not be equal across all aircraft within the trajectory data set. Furthermore, the length of the observation set is not invariant between aircraft, since it is not guaranteed that each aircraft will have \( N + 1 \) observations, even if the sampling intervals were identical (e.g. due to differing trajectories arising from scenarios such as tromboning, holding, and missed approaches). We will dedicate the rest of this subsection to discussing our methodology for standardizing the time sampling interval, as well as the trajectory lengths. The goal is to obtain a re-sampled trajectory data set of \( M \) aircraft in \( \mathbb{R}^{N' \times 3 \times 2} \) with a common time index and observation length \( N' \) so that trajectory density clustering can be done in \( \mathbb{R}^{M \times 2N'} \) using latitude and longitude as features.

To simplify bookkeeping, we restrict ourselves to the latitude versus normalized time profile for a single ORD arrival. We examine the first column \( \mathcal{P}_{[\varphi_1, \lambda_1, z_1]} \in \mathbb{R}^{N+1} \) of the \( \mathcal{P} \) matrix for a single flight and the time vector \( \tilde{t} \in \mathbb{R}^{N+1} \) for this flight. The lower bound of the entire re-sampling interval is the first time index \( t_{\varphi_1} \) in \( \tilde{t} \) (the subscript on elements of \( \tilde{t} \) is simplified since we are only looking at latitudes \( \varphi_1, \varphi_2, \ldots, \varphi_{N+1} \)). The upper bound of the entire re-sampling interval is the last time index \( t_{\varphi_{N+1}} \). We specify a parameter \( \Delta t \) that represents the re-sampling period, and partition the interval \([t_{\varphi_1}, t_{\varphi_{N+1}}]\) by (2):

\[
[t_{\varphi_1}, t_{\varphi_{N+1}} + N\Delta t] = \bigcup_{n=0}^{N-1} [t_{\varphi_1} + n\Delta t, t_{\varphi_1} + (n + 1)\Delta t] \tag{2}
\]

Note that the end of the re-sampling interval is fixed to be \( t_{\varphi_1} + N\Delta t \leq t_{\varphi_{N+1}} \). The entirety of the trajectory for this aircraft will be re-sampled such that for each time \( t_{\varphi_1} + n\Delta t, \forall n \in \{0, N - 1\} \) a corresponding latitude is given, obtained from piecewise pairwise linear interpolation on the original latitude data points. A linear function \( \phi_k(t) \) is constructed via (3) for each pairwise set of latitude-normalized times \( \left\{t_{\varphi_k, \lambda_k}, t_{\varphi_{k+1}, \lambda_{k+1}}\right\} \in \tilde{t} \times \mathcal{P}[\lambda_k,1], \forall k \in [1, N] \).

At each re-sample time index \( t_j, t_j + \Delta t, \ldots, t_j + N\Delta t \) the corresponding re-sampled latitude is given by the evaluation of the appropriate \( \phi_k(t) \).

\[
\phi_k(t) = \frac{t - t_{\varphi_k}}{t_{\varphi_{k+1}} - t_{\varphi_k}} \phi_{k+1} + \frac{t_{\varphi_{k+1}} - t}{t_{\varphi_{k+1}} - t_{\varphi_k}} \phi_k, \quad \forall k \in [1, N] \tag{3}
\]

We apply the same methodology to all flights within the trajectory data set, re-sampling and interpolating both latitude and longitude. By fixing the parameter \( \Delta t \), we fix the observation length \( N' \) across the entire trajectory data set. Each aircraft within the total \( M \) aircraft has \( N' \) re-sampled and normalized time indices, each with an associated latitude and longitude. The entire data set is re-shaped so that each observation is one aircraft, and the \( 2N' \) columns contain \( N' \) latitude and \( N' \) longitude observations, ordered by the re-sampled and normalized time indices. We perform trajectory density clustering in the next section, where each of the \( M \) data points within the data space is a row vector with \( 2N' \) features extracted from the re-shaped trajectory data set in \( \mathbb{R}^{M \times 2N'} \).

IV. TOPOLOGICAL DENSITY CLUSTERING IN C90

Given our trajectory data set of 14,712 ORD arrivals, we seek to characterize the behavior and robustness of the terminal arrival airspace under different weather conditions. Towards this goal, we apply trajectory density clustering to identify major terminal arrival airspace flow patterns during differing weather conditions. The visualization of a subset of the trajectory data set as seen in Figure 2 shows the existence of
trajectory “clusters,” or spatially similar trajectories where it can be assumed that the arriving aircraft flew the same procedures and/or were given similar radar vectors. After passing the trajectory data set through the re-sampling, interpolation, and transformation procedure, we can envision each aircraft as a point in $\mathbb{R}^{2N'}$ where $N'$ is the number of re-sampled and interpolated time-indexed latitude-longitude observations per aircraft. Suppose aircraft $x_i \in \mathbb{R}^{2N'}$ and $x_j \in \mathbb{R}^{2N'}$ flew in the same trajectory cluster, whereas aircraft $x_j \in \mathbb{R}^{2N'}$ flew in another spatially separated and distinct trajectory cluster. We expect that:

$$\|x_i - x_j\|_{2N'} > \|x_i - x_j\|_{2N'}$$ (4)

In (4), the operator $\|\cdot\|_{2N'}$ is the standard Euclidean norm in $\mathbb{R}^{2N'}$. We utilized a density-based clustering algorithm – Density-based spatial clustering of applications with noise (DBSCAN) [32] to cluster a data set containing such flight observations $x_i \in \mathbb{R}^{2N'}$. Given an aircraft data point $x \in \mathbb{R}^{2N'}$, a $2N' + 1$ dimensional ball is constructed such that:

$$x \in B(x_i, \varepsilon) \subset \mathbb{R}^{2N'+1}$$ (5)

The ball $B(x_i, \varepsilon)$ is centered at $x_i$ and has radius $\varepsilon$. Given two flights $x_i$ and $x_j$ again in the same trajectory cluster, depending on the choice of $\varepsilon$, if the relation given in (6) holds, then we consider $x_i$ to be reachable by $x_j$.

$$\|x_i - x_j\| < \varepsilon \Rightarrow x_i \in B(x_i, \varepsilon) \wedge x_j \in B(x_j, \varepsilon)$$ (6)

Consider a set of aircraft $A = \{x_1, x_2, ..., x_i, ..., x_n\} \subset \mathbb{R}^{2N'}$. We now pick a parameter $m \in \mathbb{N}$ such that if (7) is satisfied, then we consider a cluster to have been formed, where all aircraft in $A$ is part of this newly formed cluster. Any remaining aircraft that do not satisfy relation (6) and (7) are classified as outliers or noise.

$$m \leq \sum_{\{x_i, x_j\} \in A} \mathbb{I}(\|x_i - x_j\| < \varepsilon), \quad \forall i, j \in [1, n]$$ (7)

$I(\cdot)$ is the indicator function that evaluates to 1 if any of the equivalent relations in (6) are satisfied, and 0 otherwise. The two parameters $\varepsilon$ and $m$ are the two necessary inputs into DBSCAN, along with the trajectory data set to cluster. The parameter choice already provides DBSCAN with an inherent advantage over other partition-based clustering methods such as k-means: there is no bias with regards to the number of clusters the algorithm produces.

We apply DBSCAN with chosen parameter pair $\varepsilon$ and $m$ on a trajectory data subset of all flights filing the same STAR procedure. Suppose the clustering algorithm produces $k$ trajectory clusters $A_1, A_2, ..., A_i, ..., A_k$; the question of how well the set of clusters $A_1$ through $A_k$ represent the trajectory clusters and outliers for this data subset remains. Following guidance given in [21] regarding parameter selection and cluster evaluation, we perform pairwise evaluations on resultant clusters $A_i$ and $A_j$ via the Davies-Bouldin index $D_{A_i, A_j}$. The Davies-Bouldin index provides a measurement of how separated two clusters $A_i$ and $A_j$ are by computing the distance in $\mathbb{R}^{2N'}$ between the centroids of clusters $A_i$ and $A_j$, then dividing the summed scatter within the two cluster by the separation [33]. We use the set of Davies-Bouldin index along with comparisons of the resultant clusters against common ORD runway configurations to select an optimal $\varepsilon$ for each trajectory data subset per STAR.

We chose the $m$ parameter dependent on the amount of data that is present within each trajectory data subset for a given STAR. The overarching goal of choosing optimal $\varepsilon$ and $m$ parameters is to achieve a balance between too many clusters with too few aircraft observations per cluster and too few clusters with little to no differentiation amongst the clusters. We perform DBSCAN clustering on all six STARs, and summarized the results in Table 2.

**TABLE II. ORD TRAJECTORY DENSITY CLUSTER ANALYSIS SUMMARY**

<table>
<thead>
<tr>
<th>STAR (Corner-post)</th>
<th>Weather Condition</th>
<th>$n$</th>
<th>$\varepsilon$</th>
<th>$m$</th>
<th>Cluster Information</th>
<th>Cluster Populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BENSKY FOUR</td>
<td>GOOD</td>
<td>452</td>
<td>0.6</td>
<td>6</td>
<td>1. East Flow</td>
<td>16</td>
</tr>
<tr>
<td>(Southwest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. East Flow straight-in</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. West Flow</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. Flight Plan Change (FYTTE4)</td>
<td>West Flow</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5. Flight Plan Change (FYTTE4) West Flow</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6. Flight Plan Change (FYTTE4) 22L/22R</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7. Flight Plan Change (FYTTE4) East Flow</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8. West Flow;</td>
<td>625</td>
</tr>
<tr>
<td>FYTTE FOUR</td>
<td>GOOD</td>
<td>1486</td>
<td>0.7</td>
<td>8</td>
<td>1. East Flow</td>
<td>1072</td>
</tr>
<tr>
<td>(Northwest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. East Flow</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Outlier</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. East Flow</td>
<td>545</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5. West Flow</td>
<td>975</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6. Flight Plan Change (BENK4) East Flow</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7. Flight Plan Change (BENK4) West Flow</td>
<td>9</td>
</tr>
<tr>
<td>TRILL FOUR</td>
<td>GOOD</td>
<td>628</td>
<td>0.8</td>
<td>6</td>
<td>1. Outlier</td>
<td>11</td>
</tr>
<tr>
<td>(Southwest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. East Flow</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. East Flow</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. Flight Plan Change (FYTTE4) 22L/22R</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5. Flight Plan Change (FYTTE4) East Flow</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6. Flight Plan Change (FYTTE4) West Flow</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7. Flight Plan Change (FYTTE4) West Flow</td>
<td>8</td>
</tr>
</tbody>
</table>

(table continues on next page ...)
After selecting the optimal parameters and performing the DBSCAN clustering, another metric that we computed and used to evaluate the quality of clustering was the percentage of observations per trajectory data subset that were classified as outliers. A percentage exceedingly close to 0% triggers concerns of over-clustering, in that anomalous flights that should have been classified as outliers were labeled as part of a cluster due to poor parameter selection. If the percentage of observations classified as outliers is too high, then that indicates over-conservative parameter selections. \[21\] reports percentages between 6.5% to 11.0%, with no conditioning on weather conditions. The percentage we computed obtained from Table 2 is on average 2.4% for good weather conditions across all STARs, and on average 6.4% for bad weather. It should be noted that the scope examined by \[21\] focused on the en route and transition airspace, whereas our scope is only within the terminal airspace.

Figures 3 and 4 plots the trajectory data subset for the VEECK THREE STAR, for both good and bad weather periods. A specific arrival aircraft trajectory is colored according to the cluster that the aircraft belongs to. The description for each trajectory cluster is given and annotated on the figures according to their entry in Table 2. We provide an in-depth discussion regarding the clusters from the trajectory density clustering analysis in the next section.

![Figure 3. Annotated clustering results for VEECK THREE STAR during good weather conditions. Nominal trajectory clusters are labeled in white; anomalous trajectory clusters are labeled in yellow](image)

V. ANALYSIS OF RESULTS AND DISCUSSION

From visual inspection of Figures 3 and 4, there are significant spatial differences between good weather and bad weather terminal arrival airspace operations at ORD. To distinguish between the two modalities of operations, we refer to the trajectory clusters within the arrival airspace configurations during bad weather as Irregular Operation (IROP) clusters. While most common airspace metrics such as the AAR, Airport Departure Rate (ADR), and the System Airport Efficiency Rate (SAER) provide a measure of capacity and operational efficiency within the terminal airspace, they do not take into account the spatial differences seen between good weather and bad weather time periods. For example, two time periods may have the same AAR and associated SAER, but are landing aircraft using vastly different arrival trajectories.

We address the insensitivity to spatial trajectory differences by analyzing the results of our trajectory density clustering and comparing derived performance metrics across the ORD STARs examined in our case study. Table 3 documents the relevant cluster information obtained from Section 4. Any cluster that was not indicated to be a nominal arrival trajectory is considered to be an IROP cluster, including the “cluster” of outliers. The percent of ORD arrivals utilizing an IROP cluster is computed as per (8):

\[
\% \text{ STAR arrivals} \in \text{IROP} = 100% \times \frac{\sum n_{\text{IROP}}}{n_{\text{STAR}}}
\] (8)
Annotated clustering results for VEECK THREE STAR during bad weather conditions. Nominal trajectory clusters are labeled in white; anomalous trajectory clusters are labeled in yellow.

We formulate four performance metrics, and present them in Table 3. The best performing STAR within each of the four categories was given a score of 1, progressing to the worst performing STAR with a score of 6. The rightmost column in Table 3 gives the sum of the scores across the four metrics, for each of the STAR.

There are several key insights into the ORD terminal arrival airspace that can be extracted from the trajectory density cluster analysis. The northern sectors of the airspace tend to fare better than the southern sectors of the airspace. The two northern corner-post arrival fixes and associated STAR procedures, WYNDE EIGHT and FYTTE FOUR, both scored relatively low in all four categories of performance metrics. WYNDE EIGHT in particular stands out as a very robust STAR resource within the ORD terminal arrival airspace, with little difference between IROP and good weather clusters, as well as a very low number of arrivals within the IROP clusters that do appear. This stands in contrast with the southern corner-post arrival fixes and STARs, scoring an average of 16.25 out of a maximum 24 points. There are several reasons as to why the southern sector of the ORD terminal arrival airspace may not be as robust: (1) larger amount of incoming traffic from major southeastern hubs (Miami, Atlanta, Charlotte, etc.); (2) operations and runway configurations at MDW, located 15 miles southeast of ORD, will impact these southern STARs.

Interesting differences between the behavior of the southwest versus southeast corner-post STARs may stem from differences in terminal controller preference. The TRTLL FOUR STAR servicing arrivals from the southwest tend to split into multiple IROP clusters during bad weather, scoring a 6 for the number of IROP clusters as well as the change in IROP cluster numbers between good and bad weather. The scores for TRTLL FOUR regarding the actual percent usage of these IROP clusters were not as severe, scoring 4’s for both categories. On the other hand, the WATSN THREE STAR was worst off in IROP cluster usage while also scoring poorly (5’s) in the number of IROP clusters and change in IROP clusters category. This indicates that terminal controllers responsible for arrivals flying TRTLL FOUR tended to use IROP trajectories but are able to rebound and switch back to the nominal trajectories. On the other hand, WATSN THREE arrivals during bad weather were often assigned to IROP trajectories that remained in use for a longer period of time. This again highlights the impact of MDW and larger southeast-originating flows on the ORD terminal arrival airspace.

### Table III. ORD Terminal Airspace Robustness Scores

<table>
<thead>
<tr>
<th>STAR (Corner-post)</th>
<th>Weather Condition</th>
<th>IROP Cluster ID</th>
<th>% arrivals in IROP cluster</th>
<th>% IROP clusters</th>
<th>% change in IROP clusters</th>
<th>Performance Scores</th>
<th>% worst IROP</th>
<th>% worst IROP</th>
</tr>
</thead>
<tbody>
<tr>
<td>BENKY FOUR</td>
<td>GOOD</td>
<td>1, 4</td>
<td>4.9%</td>
<td>2 (4 - 4)</td>
<td>2 (4 - 4)</td>
<td>2 (4 - 4)</td>
<td>2 (4 - 4)</td>
<td>2 (4 - 4)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 2, 5, 6</td>
<td>8.9%</td>
<td>2 (5 - 4)</td>
<td>2 (5 - 4)</td>
<td>2 (5 - 4)</td>
<td>2 (5 - 4)</td>
<td>2 (5 - 4)</td>
</tr>
<tr>
<td>FYTTE FOUR</td>
<td>GOOD</td>
<td>1, 3</td>
<td>3.1%</td>
<td>2 (3 - 2)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 4, 5, 6, 7, 8</td>
<td>9.2%</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
</tr>
<tr>
<td>NTLLE FOUR</td>
<td>GOOD</td>
<td>1, 1</td>
<td>1.0%</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 4, 5, 6, 7, 8</td>
<td>9.7%</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
<td>6 (8 - 6)</td>
</tr>
<tr>
<td>VEECK THREE</td>
<td>GOOD</td>
<td>1, 2</td>
<td>2.0%</td>
<td>2 (2 - 1)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
<td>3 (3 - 2)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 5, 6, 7</td>
<td>12.1%</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
</tr>
<tr>
<td>WATSN THREE</td>
<td>GOOD</td>
<td>1, 1</td>
<td>2.0%</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
<td>6 (6 - 6)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 5, 6, 7</td>
<td>12.1%</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
<td>5 (6 - 4)</td>
</tr>
<tr>
<td>EIDORTH</td>
<td>GOOD</td>
<td>1, 1</td>
<td>0.8%</td>
<td>1 (2 - 1)</td>
<td>2 (2 - 1)</td>
<td>2 (2 - 1)</td>
<td>2 (2 - 1)</td>
<td>2 (2 - 1)</td>
</tr>
<tr>
<td></td>
<td>BAD</td>
<td>1, 5, 6, 7</td>
<td>3.9%</td>
<td>1 (2 - 1)</td>
<td>1 (2 - 1)</td>
<td>1 (2 - 1)</td>
<td>1 (2 - 1)</td>
<td>1 (2 - 1)</td>
</tr>
</tbody>
</table>

### VI. Conclusion and Future Work

We have demonstrated the applicability of trajectory density clustering to aid in the characterization of complex terminal arrival airspaces. With regards to our case study airport, we rated the various arrival airspace resources at ORD according to their performance in good and bad weather scenarios. Arrival airspace resources in the southern sectors of the ORD terminal arrival airspace tended to be less robust during convective weather, compared to the northern sector. This is the result of factors such as sharing the airspace with MDW, and heavier arrival flows filing for and flying through the southwest and southeast corner-post arrival fixes.

We hope to extract the insights obtained from trajectory density clustering and use it to augment current AAR metrics to produce spatio-temporal sensitive capacity metrics that take into account differences in arrival airspace resources. Instead of relying on one AAR metric providing the number of arrivals
acceptable by an airport per hour, we can provide a trajectory-specific AAR metric based on the commonly-utilized clusters discovered via trajectory density clustering. We also plan to apply this trajectory density clustering method towards other terminal airspaces, particularly in metropolex environments where airspace resources are shared amongst different airports.

ACKNOWLEDGMENT

The authors would like to thank M. C. R. Murça for her insightful input regarding various clustering algorithms, and S. Badrinath for providing data from ASPM.

REFERENCES
