Modeling and Detecting Anomalous Safety Events in Approach Flights Using ADS-B Data

Bonifazi, A.; Sun, J.; van Baren, Gerben; Hoekstra, J.M.

Publication date
2021

Document Version
Final published version

Citation (APA)
Bonifazi, A., Sun, J., van Baren, G., & Hoekstra, J. M. (2021). *Modeling and Detecting Anomalous Safety Events in Approach Flights Using ADS-B Data*. Paper presented at 14th USA/Europe Air Traffic Management Seminar.

Fourteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2021)

Modeling and Detecting Anomalous Safety Events in Approach Flights Using ADS-B Data

Alberto Bonifazi*, Junzi Sun*, Gerben van Baren†, Jacco Hoekstra*

*Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands

†Inspectie Leefomgeving en Transport Minister van Infrastructuur en Waterstaat Den Haag, the Netherlands

***Abstract*—Not all flight data anomalies correspond to operational safety concerns. But anomalous safety events can be linked to anomalies in flight data. During the final phases of a flight, two significant safety events are unstable approach and go-around. In this paper, using Automatic Dependent Surveillance-Broadcast (ADS-B) data, we develop several exceedance and anomaly detection techniques to identify these events. Rule-based algorithms and data-driven Gaussian Mixture Models (GMM) are proposed to identify unstable approaches. A fuzzy logic approach is developed to model and to identify go-arounds. We extend our analysis combining runway information and meteorological reports to provide deeper insights on flight safety during the approach. These identification models are also applied to the ADS-B data from the Schiphol Airport area in Amsterdam in 2018. By using a reference report provided by the Dutch transportation regulatory agency, the chosen GMM model can identify 25% to 30% of reported unstable approaches, and the go-around detection model can identify 98% of go-arounds.**

***Index Terms*—flight safety, anomaly detection, safety monitoring, ADS-B, Schiphol Airport, data mining**

I. Introduction

One of the goals of air traffic management is to ensure safe aircraft operations. To this end, aviation authorities around the world have promoted safety information sharing reporting mechanisms. The most notable ones are the European Co-ordination Centre for Accident and Incident Reporting Systems (ECCAIRS) and the Aviation Safety Information Analysis and Sharing (ASIAS) from the FAA. Although the majority of stakeholders join these initiatives, much of the safety knowledge they generate remains within the boundaries of the organizations. Under the current framework, only serious occurrences are communicated. It is also often challenging for researchers to use safety related data due to confidentiality.

Automatic Dependent Surveillance-Broadcast (ADS-B) technology can complement this information, and it is compulsory for most aircraft as of 2020. With ADS-B, aircraft continuously broadcast their position, velocity, track angle, and other related flight information. In addition to flight monitoring, combining ADS-B data with data-mining techniques can also give us extra safety knowledge. Researchers and regulators can rely on this independent data source to monitor and verify safety reports.

The approach and landing phase is the most critical phase of a flight. In this phase, 65% of accidents occurred between 2011 and 2015 [1], particularly unstable approaches and go-arounds. The first event corresponds to 14% of the accidents occurring during approach and landing [1]. The second one is a standard procedure that can be initiated by the pilot or ATC for different reasons, such as an unstable approach, conflicting traffic, or adverse weather. For this reason, there is a strong link between go-arounds and anomalous safety events.

Extracting safety knowledge from aircraft data is an active field of research. Commonly, safety knowledge is extracted using exceedance detection algorithms [2], which are dependent on certain thresholds. The main pitfall of this method is that it fails to detect unknown events. The focus of recent research is on anomaly detection techniques, which have shown better performance but may demonstrate a false-positive rate as high as 70% [3].

The most well-recognized technique in anomaly detection is the Once Class Support Vector Machine (OC-SVM) variant called the Multiple Kernel Anomaly Detection (MKAD) algorithm [4]. This algorithm has been tested extensively for approximately 10 years for various phases of flights, and it is able to detect unstable approaches and go-arounds [5], [6]. However, its focus is on discovering flight level anomalies, and it has shown the best results when used with a single type of aircraft and FOQA data. To overcome these limitations, energy metrics [7], [8] have shown potential in identifying instantaneous anomalies in general aviation operations.

This paper aims to contribute to the field of safety monitoring during the approach phase of commercial flights. The goal is to produce insights based on the safety knowledge extracted from ADS-B data and to develop an independent safety monitoring mechanism. We combine both rule-based and data-driven unstable approach detection. We also develop a go-around detection model based on fuzzy logic extending our previous research [9].

In the rest of this paper, section II and section III discuss the detection of anomalous safety events using a rule-based algorithm, GMM, and fuzzy logic. Section II addresses the detection of unstable approaches and section III the detection of go-arounds. Section IV covers aspects regarding the data in use and the preprocessing strategy. The findings are presented in section V and section VI. The result consists of a set of

safety indicators that can be aggregated in different manners to produce insights.

II. Detecting Unstable Approaches

Several definitions of an unstable approach have been used in air traffic management depending on the operator and the entity [10]. All definitions consider trajectories that are not aligned with the correct flight path, being too fast or too slow.

When analyse unstable approaches, it is common to relate them to corresponding meteorological conditions, which are 1000 feet above airport elevation in Instrument Meteorological Condition (IMC) or 500 feet above airport elevation in Visual Meteorological Condition (VMC). When these conditions can not be met, go-arounds should be executed.

We propose two strategies in this paper to detect unstable approaches. The first one checks the horizontal boundaries of stable approach operations with the flight data. The second one models the normal specific energy boundaries of flight operations and checks the specific energy profile of particular approach with this model.

A. Horizontal Compliance

This method assumes that approaches in the airport of interest (EHAM) are ILS approaches. Hence flights must flow within one-dot of the localizer [10], which represents a dot on the Course Deviation Indicator. The definition of dot depends on the instrument. For instrument landing system localizer (LOC), this corresponds to one degree, while for a very high frequency omni-directional range (VOR), it corresponds to two degrees.

The ideal ILS-intercept trajectory is defined for each runway as the line connecting the final approach fix (FAF), runway threshold, and localizer. Information is collected from the instrument approach chart. A horizontal compliance region is constructed as the area comprised within 1 degree of the ILS-intercept. It results in areas with the same dimension and shape but positioned differently depending on the runway.

In Figure 1, we visualize the horizontal compliance area for each runway in gray. As defined in regulations, an aircraft is considered unstable if it only stabilizes after 1000 ft in IMC, or 500ft in VMC.

B. Energy Compliance

An unstable approach can commonly display abnormal energy levels during the flight [11]–[13]. This strategy makes use of various specific energy parameters during the approach. Specific energy (e) refers to energy (E) per unit mass (m):

$$e = \frac{E}{m} \tag{1}$$

The detection strategy is composed of two steps: energy features generation and anomaly detection. The energy features are derived from ADS-B data. Subsequently, anomaly detection is performed using a Gaussian Mixture Model (GMM). GMM is an unsupervised learning model, with the advantage of not requiring a priori knowledge of anomalous data. The GMM algorithm clusters normal operations based on a weighted sum of Gaussian component densities, and

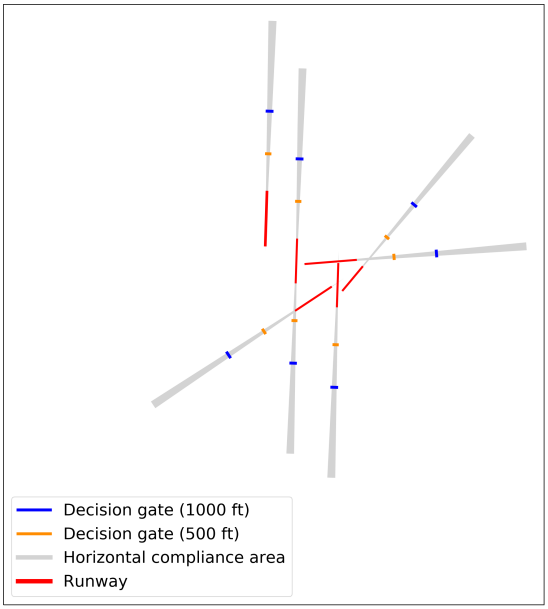

Figure 1: Example of Schiphol airport compliance areas. The horizontal compliance area (in gray) corresponds to the area in which an aircraft needs to be flying once it intercepts the ILS.

anomalous operations can be found where energy features are beyond certain confidence intervals. The main challenges in this approach are to identify the number of Gaussian mixtures and the appropriate confidence interval to determine anomalies.

1) Energy Features Generation: The energy features we chose for unstable approach detection are: specific total energy (e), specific kinetic energy (e_k), specific potential energy (e_u), energy angle (γ_p), and the specific energy rates ($\dot{e}$, $\dot{e}_k$, $\dot{e}_u$). In addition, the energy angle is a measure of how the flight-path angle can change given the current energy level [14].

These features are similar to the ones used by Puranik and Marvis in their work [11], and they are computed as follows:

$$e = e_k + e_u \tag{2}$$
$$e_k = 0.5 \cdot v^2 \tag{3}$$
$$e_u = h \cdot g \tag{4}$$
$$\gamma_p = \arcsin\left(\frac{VS}{v}\right) + \frac{\dot{v}}{g} \tag{5}$$
$$\dot{e} = \dot{e}_k + \dot{e}_u \tag{6}$$
$$\dot{e}_k = v \cdot \dot{v} \tag{7}$$
$$\dot{e}_u = VS \cdot g \tag{8}$$

where v is the true airspeed, and h is the altitude above the ground. They require additional information than ADS-B to be computed, as explained in section IV-B. g is the gravitational acceleration, and VS is the vertical rate.

Once the energy features are calculated, the enhanced trajectory data are re-sampled to for training the GMM. We sample the data in space using a median window of 0.5 NM, which corresponds roughly to 12 seconds. The time

interval varies greatly depending on the flight stage. The analysis considers aircraft flying between 0.5 to 10NM from the runway threshold. This is the area where final approach procedures occur, as shown in the Instrument Approach Charts. The last portion, between 0 and 0.5, is not used due to lower ADS-B data quality closer to surface.

As an example, Figure 2 shows the variation of the specific potential energy, the specific kinetic energy, and the specific total energy during the approach phase for all the flights we analyze in this paper.

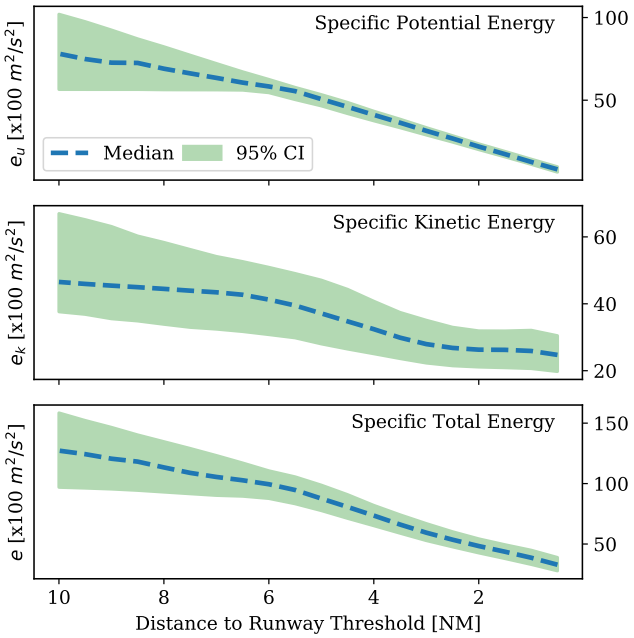

Figure 2: Variation of the specific potential energy, the specific kinetic energy and the specific total energy depending on the distance to runway threshold

2) Anomaly Detection: The anomaly detection step makes use of GMM. A GMM is fully defined by the following equation:

$$p(\mathbf{x}|\lambda) = \sum_{i=1}^{k} w_i g(\mathbf{x}|\mu_i, \Sigma_i) \tag{9}$$

Equation (9) represents a parametric probability density function that is a weighted (w_i) sum of Gaussian components. $g(\mathbf{x}|\mu_i, \Sigma_i)$ indicates a single component, where Σ is the covariance matrix that captures the relation between the different features. k is the number of components in the mixture. The majority of these parameters can be obtained using the expectation-maximization (EM) algorithm. Only the type of covariance matrix and the number of components need to be specified beforehand. This approach is very common in the literature [7], [15].

The type of covariance matrix has direct consequences on the shape of the Gaussian components. [7] and [15] use a diagonal covariance matrix because of the lower computational cost, which restricts the shape of the Gaussian components forcing it to be oriented along the coordinate axes. During the experimentation process, the computational time of using any kind of covariance matrix is reasonable.

Because aircraft's behavior changes while performing an approach, a single GMM would over-generalize behavior. Hence, we propose a multi-GMM approach comprised of three separate Gaussian mixtures, referred to as *3-GMM model* henceforth. The 3-GMM approach can be described as follows:

1) The first GMM learns what it means to be stable during the final approach. This GMM is trained on aircraft flying at a distance from the runway threshold between 0.5 NM and 4 NM. This covers the most interesting aircraft operations because it corresponds to the decision area described by the Flight Safety Foundation as it includes the gates of 500ft and 1000ft.
2) The second GMM learns how aircraft descend and intercept with the ILS. It comprises the final approach area that goes at a distance from the runway threshold between 4 NM to 7 NM.
3) The last GMM learns a broader spectrum of aircraft behaviors as the area expands 3NM starting at a distance of 7NM from the runway threshold. At this stage, an aircraft could be descending, turning, or at flying level.

In addition to the energy features we have discussed earlier, 3-GMM includes two more features, which are time and distance to runway threshold. These two new features are significant due to the fact that each energy level is closely related to a particular position and time before landing.

We use a full covariance matrix for the 3-GMM model so that the shape of GMM components are not constrained. The number of components is chosen using the Calinski–Harabascz (C-H) index, as suggested by Puranik in [7]. This is an internal evaluation criterion that measures how compact and how separated components are.

Based on this information, the 3-GMM model can be trained after normalizing the data. We perform the training with data that are within 95% confidence interval of its parameter values. It is worth noting that this does not automatically classify data beyond 95% confidence interval as anomalies. The final anomaly detection still relies on GMM thresholds, which need to be carefully chosen.

In Figure 3, an example GMM model is shown. The GMM is trained based on the first segment from 10 NM to 7 NM to runway threshold. The figure contains around 50 k flights. The shade of green represents the log-probability of a trajectory during normal operation.

After the 3-GMM model is trained, it is possible to obtain the probability density of every point, as shown in Figure 3.

Finally, to detect unstable approaches, a threshold is selected such that it contains a percentage of points with the lowest probability density. Puranik in [7] suggests to use 0.05%, or 0.1% of points. In this paper, the amount of points is selected experimentally, as will be detailed in section V-D2. The anomaly detection goes as follows:

1) Each trajectory is fed into to the 3-GMM model, which will return the probability density for each of its points.
2) If the probability density of a point is above the anomaly threshold, the point is labeled as anomalous.

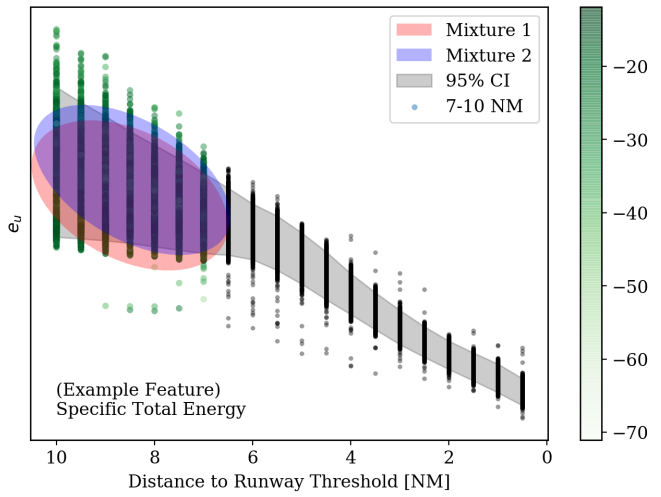

Figure 3: The GMM model visualized for the first segment. Model is trained on all features, while only specific total energy and distance to runway threshold are illustrated. The color bar represents the log-probability of a flight point being a normal operation.

3) A trajectory is considered anomalous if at least two points are anomalous in the last 7 NM.

III. Detecting Go-arounds

A go-around is an operation that occurs when landing is aborted. Go-around deviates from the intended operation, which increases the risk that can be caused by increased workload of pilots. Accommodating a go-around in busy airspace also increases the workload for air traffic controllers. In some cases, pilots can refrain from performing this maneuver, where the approach is continued and may remain unstable. Thus, it May lead to loss of control, runway excursion, or controlled flight into terrain.

On top of the inherent risks of go-arounds, it is important to understand what are the circumstances surrounding a go-around because they are potentially anomalous safety events themselves. Generally, these are unstable approach, conflicting traffic, or adverse weather. The detection of go-arounds from ADS-B data follows a two-steps approach: identification of a possible go-around and evaluation of go-around score.

In some part of the world, go-around may be considered as a normal operation to mitigate capacity constraints. However, it is worth noting that, for the airport in this paper (EHAM), go-rounds are only issued if there is a safety risk.

1) Identification of a Possible Go-Around: A go-around consists of climbing to a predetermined altitude prescribed in the Instrument Approach Chart and then turning 360 degrees around the runway. Thus, an intuitive strategy to detect a go-around is identifying when an airplane starts climbing and then changes its course.

This idea forms the basis of the detection model presented in this paper. This step employs the technique developed by Sun et al. in [9]. Applying this algorithm allows for the identification of rapid changes in aircraft behavior. The flight phase detector can distinguish between five different phases: climb (CL), ground(GND), descent (DE), level (LVL), and cruise (CR). As shown by Proud [16], this is particularly useful in the case of a go-around as the algorithm detects changes from DE to LVL/CL.

The change in phase is the indication for the beginning of a possible go-around; this will be referred to as the *starting position* in the text. Four ADS-B variables are analyzed: vertical rate, altitude, ground speed, and track angle. During a go-around, these variables are expected to change in a very specific way and in a specific time interval. A moving average with a window of 15 seconds is used to reduce the susceptibility of the model to outliers.

The vertical rate and the altitude indicate that the airplane is climbing and gaining altitude. It is expected that these two variables will be changing immediately after the *starting position* and up to the next 2 minutes. In particular, the model analyzes the *delta*, namely the change in altitude from the *starting position*.

In this go-around detection approach, the ground speed and the track angle are expected to change some time up to 10 minutes from the *starting position*.[1] These two states will start varying after the climb. The track angle keeps changing until the aircraft is aligned with a runway for landing. The algorithm constantly computes the change in ground speed and track angle from the *starting position*.

2) Evaluation of Go-Around Score: In this step, we want to evaluate how much the data resemble a go-around. We describe the expected behavior of these variables using four S-functions, which output one for highest possibility. Every variable has its specific S-shape function with the following general definition:

$$S(x;a,b) = \begin{cases} 0, & \text{if } x \le a \\ 2\left(\frac{x-a}{b-a}\right)^2, & \text{if } a \le x \le \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2, & \text{if } \frac{a+b}{2} \le x \le b \\ 1, & \text{if } x \ge b \end{cases} \tag{10}$$

where a and b are values where the score is 0 and 1, as shown in Figure 4. The functions for each variable are defined as follows, where delta represents the difference between the current position and the *starting position*:

$$\begin{aligned} \Delta X(m) &= \mathrm{S}(m;30,300) && [deg] \\ \Delta H(n) &= \mathrm{S}(n;100,1000) && [ft] \\ \Delta V(p) &= \mathrm{S}(p;5,80) && [kn] \\ VS(q) &= \mathrm{S}(q;10,1000) && [fpm] \end{aligned} \tag{11}$$

These 4 S-functions in (11) represent a simplified version of the behavior of an aircraft during a go-around, where X corresponds to the track angle, H to the altitude, V to the ground speed, and VS to the vertical rate. By applying to actual flight data, we obtain go-around scores. Ideally, there should be points in the trajectory considered where all these functions have scores of 1. This is not always the case since

[1]This 10 minute value is chosen because a go-around typically adds a flight delay of this amount. https://www.casa.gov.au/safety-management/advice-air-travellers/go-arounds

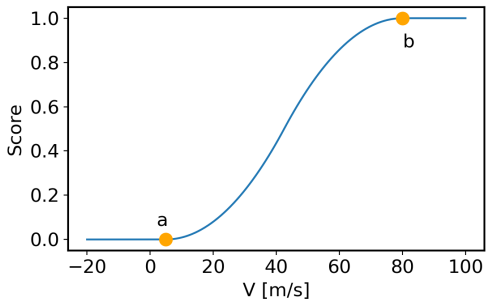

Figure 4: S-function of the ground speed

flight data are prone to errors. For this reason, the model identifies the trajectory of a go-around if the average of the maximum score obtained from all four S-functions is higher than 0.5. Figure 5 provides direct insights into the working mechanism of this strategy.

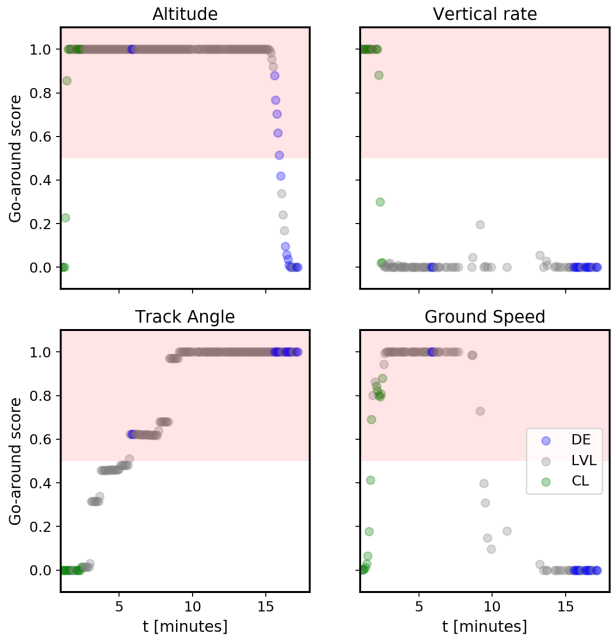

Figure 5: Evolution of the scores of the various ADS-B variables during a go-around. The red areas indicate scores above the 0.5 go-around detection threshold.

The figure shows how the score is obtained using the S-functions and the ADS-B data during a go-around procedure with the red areas corresponding to the 0.5 detection threshold. The detected flight phases are labelled using different colors. The change in color represents a change in phase corresponding to different stages of a go-around. In particular, time zero corresponds to the starting position that is the point in which the phase changes from descent (DE) to climb (CL)/ level (LVL).

In this example, a go-around is clearly detected because all variables have at least a point above the 0.5 detection threshold. As expected, the altitude and the vertical rate vary rapidly. In the vertical rate plot, there is a sudden jump. This is not surprising as this variable represents the instantaneous change in altitude. Furthermore, once the aircraft has reached the go-around altitude, this variable drops as the altitude of aircraft does not increase anymore. The altitude score remains high during the whole go-around and then it starts decreasing gradually during the final descent towards the airport.

IV. Data processing

A. Data sources

Data used in this research consist of flight data and weather data. Flight data include ADS-B data, an aircraft database from OpenSky[2], and Schiphol's Aeronautical Information Services (AIS) publications[3]. The weather data are METeorological Aerodrome Report (METAR) reports and Global Forecast System (GFS) data from the National Oceanic and Atmospheric Administration (NOAA)[4].

The ADS-B data are collected through the antenna positioned at the Aerospace Faculty of the Delft University of Technology. ADS-B data provide aircraft identification, positions, and velocities. The aircraft database from OpenSky contains additional information on the aircraft, including manufacturer, aircraft model and typecode, registration, and operator. Among Schiphol's data from the AIS publication, the Instrument Approach Charts and runway data are used.

METAR reports indicate the weather perceived at the airport on any particular day. In this case, METAR reports from Schiphol are downloaded from IOWA ASOS network[5]. These reports are generated every 30 minutes, and they include many weather variables. The ones that are used in this research are timestamp, temperature, dew point temperature, wind direction, wind speed, pressure, visibility, wind gust, cloud coverage, and weather codes. GFS data is used for higher altitudes, which offers wind information at intervals of 700 ft with updates every 6 hours.

B. Preprocessing

In this step, the raw data are cleaned and combined to create a clean dataset for further analysis. A preliminary procedure consists of removing general aviation aircraft, helicopters, and ground vehicles using the OpenSky aircraft database.

It is possible to know exactly which aircraft communicated the ADS-B data point because the message includes the ICAO's identifier. However, the same aircraft might land and take-off multiple times on the same day. For this reason, the data points of a particular aircraft are further divided into trajectories. This is a crucial step as we will be using trajectories later in the analysis. Given the fact that we only use the area around Schiphol for the analysis, it is easy to detect inbound and outbound traffic, which gets labeled respectively as approaching and taking-off traffic. In this way, we can further refine the trajectories used for the analysis by selecting only approaching traffic. Throughout this process, incomplete trajectories are removed. Finally, using data from Schiphol's AIS publication, it is possible to determine the landing runway for each approaching trajectory.

[2]opensky-network.org/aircraft-database
[3]en.lvnl.nl/information-for-airmen/publications-for-airmen
[4]ncdc.noaa.gov/data-access/model-data/model-datasets
[5]https://mesonet.agron.iastate.edu/request/download.phtml

The track angle included in ADS-B is a valuable source of information. However, they are not always accurate. Identifying the causes of this behavior is beyond the scope of this work, but a track angle fix is proposed. It is observed that sometimes when an aircraft performs a go-around or after landing when it moves along the taxiways, its track angle indicator doesn't follow the aircraft movements. The track angle communicated via ADS-B data doesn't change whereas it is clear it should have. To solve this issue, the track angle information communicated from the ADS-B data is compared to a bearing estimated using its position. If the difference between the two is higher than 60 deg, the estimate is used as track angle data. To limit the influence of poor measurements in the track angle estimation, a window of 40 seconds is considered and a minimum amount of 5 points.

Furthermore, ADS-B data provide ground speed and barometric altitude. These are dependent on the weather, and thus limit the comparison of aircraft in different meteorological conditions. Nonetheless, it is possible to remove their dependency using METAR reports and GFS NOAA data. Since barometric altitude assumes standard temperature and pressure, these two values are adjusted based on METAR reports as follows:

$$
\begin{aligned}
T_A &= T + a \cdot h_A \\
P_A &= P \cdot \left(\frac{T_A}{T}\right)^{\frac{-g}{a \cdot R}} \\
P_{A,M} &= P_A \\
T_{A,M} &= T_M \cdot \left(\frac{P_{A,M}}{P_M}\right)^{\frac{a \cdot R}{-g}} \\
h_{A,M} &= \frac{T_{A,M} - T_M}{a}
\end{aligned}
\tag{12}
$$

where P_A, T_A, and h_A refer to pressure, temperature, and altitude of the aircraft assuming standard atmospheric conditions. The variables P, T, g, a, and R are International Standard Atmosphere constants, while, $P_{A,M}$, $T_{A,M}$ and $h_{A,M}$ are the airplane's pressure, temperature, and altitude calculated using P_M and T_M from the METAR report.

Ground speed is dependent on the wind speed, and it can be corrected to obtain true airspeed by subtracting ADS-B ground speed from the wind speed. METAR reports are used to correct speed up to an altitude of 100 m, where the difference in wind speed is estimated to be approximately 1.5 m/s. After this point, wind information is extracted from the GFS NOAA data.

It is interesting to understand the relationship between safety anomalous events and the weather. For this reason, we establish a way to assess the severity of the weather with a meaningful score using EUROCONTROL's ATMAP weather algorithm [17]. In this metric, a high score corresponds to poor weather with four being the threshold for a weather condition that disrupts airport operations. The algorithm also takes consideration if an aircraft is flying in VMC or IMC [18].

V. Experiments and Results

The results are produced using the data collected in 2018. This section shows examples of the detected events, the validation for the three detection models, and insights drawn from monitoring safety indicators.

A. Example of Horizontal Compliance

Figure 6 shows an example of the type of trajectory that the horizontal compliance strategy is able to detect. In this case, it shows a trajectory, purple line, that stabilizes at 500 ft, this gate is represented in orange.

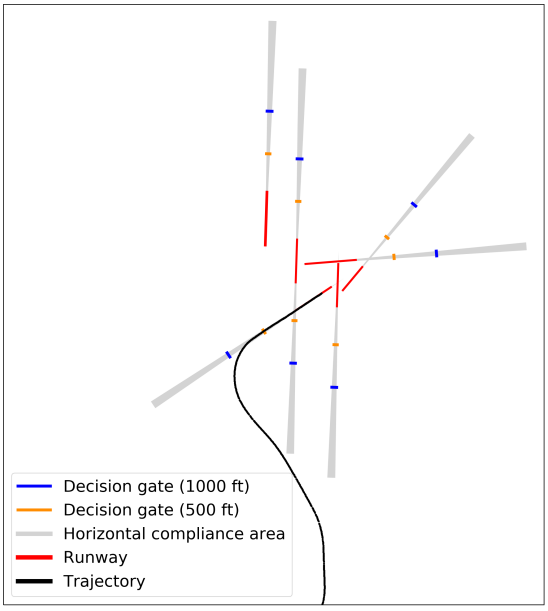

Figure 6: A flight that stabilizes after the 1000ft light blue gate and before the 500ft green gate.

B. Example of Energy Compliance

The flight shown in Figure 7 represents an unstable approach. The red dots are points belonging to this particular aircraft, while the green area comprises 95% of data, with the dashed line as its median.

It is considered anomalous even though some of the parameters vary within normal energy bounds. This is where the multivariate nature of GMM anomaly detection comes into play because it discovers anomalies based on a combination of different features.

Looking at the specific potential energy plot in Figure 7, it is clear that the aircraft is approaching higher than usual. The reason might be that a high speed is maintained until the beginning of the final approach phase, which is shown in the specific kinetic energy plot. This situation leads to an overall higher than usual specific total energy level. It is possible to infer that the aircraft tried to dissipate its excess energy, which is visible in the rate plots. When we inspect the trajectory from the time perspective, the aircraft is advancing faster than usual towards the threshold.

This flight is also present in the validation list, which is described in detail in section V-D. This occurrence is accompanied by the following explanation as: *after accepting*

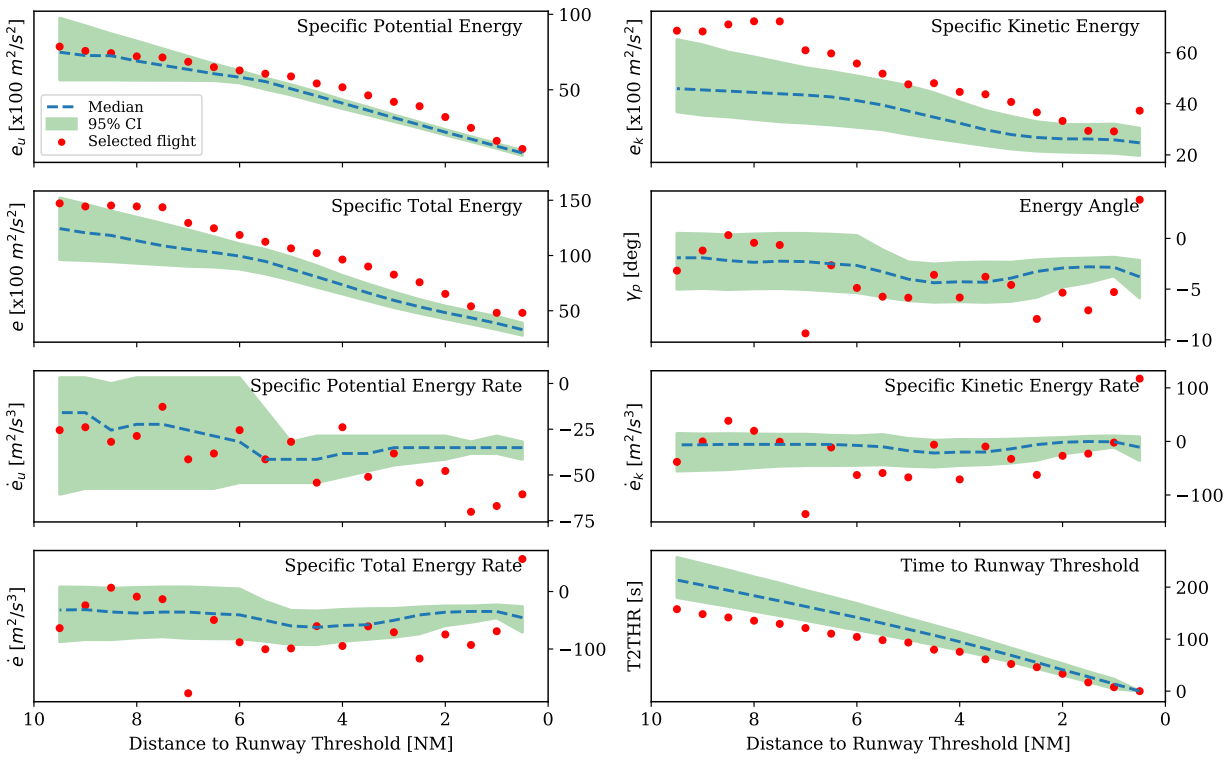

Figure 7: The energy metrics of an aircraft performing an unstable approach

a short line-up, the approach is unstable because the aircraft is high during the approach, which visible in the energy metrics plot.

C. Detection of Go-Around

An example of a go-around detected on runway 06 by the algorithm is presented in Figure 8. In this figure, the flight trajectory is illustrated in black, and we can clearly see that a go-around is performed.

D. Validation

Validating the results of the detection algorithms is challenging because data is unlabeled. However, we could rely on a list of known go-arounds and unstable approaches provided by ILT (Environment and Transport Inspection), a part of the Ministry of Infrastructure that in charge of the airports' oversight in the Netherlands. The list, which will be called the validation list, comprises 65 go-arounds and 48 unstable approaches.

There is an important methodological difference that needs to be made. Since this report data is generated by manual reporting process, it is possible that a flight present in the validation list is not present in the ADS-B data. For go-around detection, this is less likely to happen, since go-around can be visually inspected. However, for unstable approaches, this is not always the case since energy compliance information can not be inspected directly.

On the other hand, energy metrics do not always corresponds to unstable approaches. This dilemma is presented by LI et al. in [15] where four experts could not agree on which situation poses safety concerns.

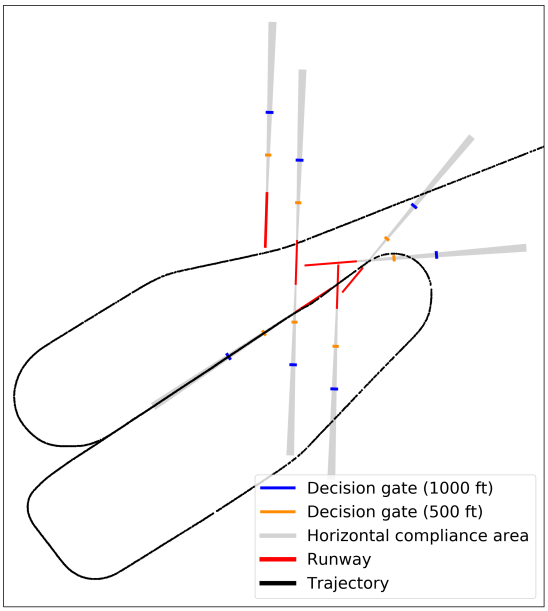

Figure 8: An airplane performs a go-around and then lands on runway 06.

1) Horizontal Compliance: There are only two unstable approaches detected for 2018 using this strategy, none of which is present in the validation list. Furthermore, by inspecting the detected cases, one detection is a false-positive as it is caused by poor data quality. The analysis shows that landing aircraft are at most times within horizontal stability limits.

2) Energy compliance: Comparing the flights identified in the ADS-B data with the ones of the validation data, we find that 31 are present in both.

As mentioned in section II-B, a threshold is used to determine if a point is anomalous. The goal is to select a threshold that includes all anomalous safety occurrences while limiting as much as possible the number of false-positives. Table I shows the relation between a particular GMM threshold, number of detected events, detection accuracy, and ratio of positives. They are defined as:

1) The number of detected events refers to the occurrences that are also present in the validation list.
2) The detection accuracy is the ratio of the number of detected events over the total number of events found in the validation list.
3) The ratio of positives provides an overview of the amount of approaches that are considered unstable in an year relative to the total number of approaches occuring that year.

TABLE I. Comparing the number of unstable approaches detected from the validation list depending on a GMM threshold

GMM threshold [%]	**Detected**[a]	**Detection accuracy**[b] **[%]**	**Ratio of positives**[c] **[%]**
0.01	3	9.68	0.50
0.05	6	19.36	1.08
0.1	8	25.81	1.60
0.5	9	29.03	4.78
1	11	35.48	7.55
2	15	48.39	12.70
3	17	54.84	17.55
5	21	67.74	26.23
10	29	93.55	42.80

[a]Number of detected unstable approaches that are present in the validation list

[b]Ratio between detected unstable approaches also present in the validation list and the total number of elements in the validation list

[c]Ratio between number of detected unstable approaches and overall number of detected landing aircraft

As expected, when increasing the threshold the number of detected events increases. In the last row, when the threshold equals 10 %, the detection accuracy rises to roughly 94%. However, 43% of landing trajectories are labeled as unstable. It is likely that the higher percentage of positives results underlies a high percentage of false positives. According to [19], approximately 3% of the approaches are unstable. For this reason, a reasonable choice of threshold is between 0.1% and 0.5%.

In our study, we recommend a lower threshold, 0.1%, to limit the number of false positives. This corresponds to a detection accuracy of 26%, and it reveals an overall number of unstable approaches equal to 3000 for the year 2018.

3) Detection of Go-Around: In this paragraph, two validation tests are performed: 1) comparing the results with the validation date set, and 2) manual inspection for the detection of false-positives.

The first test reveals that 18 trajectories are not present in the ADS-B data, approximately 28% of all 65 go-arounds present in the validation list. The missing data is likely due to the availability of ADS-B data for low altitude. Among the present trajectories, 46 go-arounds are detected and 1 is undetected, which yields a accuracy of 98%.

Another way to evaluate this algorithm is by manually inspecting its output. This means logging when the result is a go-around. The results are that 285 are true positives and 7 false positives out of 292 detected go-arounds in 2018. Thus, false positives occur only 2% of the time.

E. Monitoring the Safety Indicators

In this section, safety indicators are constructed by aggregating the results of the energy compliance model and the go-around detection model. With the knowledge extracted from the ADS-B data, it is possible to gain insights into operations by analyzing the relationship between different variables.

1) Energy Compliance: Table IIa shows the relation between unstable approaches, months of the year, and weather conditions. The weather column in the table represents unstable approaches happening in adverse weather situations.

This table shows that the months with the highest number of unstable approaches are March, January, and July, with approximately 300 unstable approaches per month. February is the month with the least unstable approaches, around 180. We can see a clear relationship between weather conditions and unstable approaches. As expected, the months with the largest portion of unstable approaches happening during poor weather are January and December. In these months, approximately 25% of the unstable approaches are linked with the weather. During June and July, this occurs only 3% of the time.

Furthermore, it is possible to see how the unstable approaches vary depending on the runway and the weather conditions. This is shown in Table IIb. The runway with the most unstable approaches is 36R with 809, almost 30% of all. The runway where there seems to exist a strong link between adverse weather and unstable approaches is runway 27. Almost 40% throughout the year, with peaks in January and December, where poor weather is concurrent to unstable approaches 60% and 50% of the time respectively.

2) Detection of Go-Around: The similar analysis is applied to go-arounds. Table IIIa and Table IIIb show the relationship between the number of go-arounds, weather conditions, unstable approach, and separation to closest aircraft. Compared to the analysis of unstable approaches, there are two extra columns. The unstable column contains information on whether the go-around is preceded by an unstable approach. The separation column indicates if the closest aircraft to the one performing the go-around is at a distance between 1.5NM and 3NM. These values are chosen because there is no aircraft closer than 1.5 NM, while 3NM is the minimal for terminal airspace operations.

From Table IIIa, we can see that 25% of go-arounds happen in December and January. The cause can be linked to poor weather conditions since 60% of go-arounds in this period are linked with adverse weather conditions. In September, approximately 60% of go-arounds are also linked with bad weather. April is the month where go-arounds are preceded by

TABLE II. Overview of detected approaches and unstable approaches by month, runway, and weather

(a) By month and weather

Month	Unstable[a]	Weather[b]	Total[c]
Jan	297	77	15269
Feb	182	33	14370
Mar	304	30	16794
Apr	215	8	17645
May	256	36	17212
Jun	259	8	18276
Jul	298	7	19409
Aug	265	48	19554
Sep	213	48	18948
Oct	226	26	19292
Nov	214	14	16227
Dec	270	72	16249

(b) By runway and weather

Runway	Unstable[a]	Weather[b]	Total[c]
06	528	57	48394
18C	409	20	35393
18R	662	86	66637
22	107	21	4018
27	351	131	17304
36C	133	27	9742
36R	809	65	27757

[a]Number of unstable approaches
[b]Number of unstable approaches with a weather score higher than 4
[c]Number of approaches detected in the data

TABLE III. Overview of detected go-arounds by month, runway, weather, unstable approach, and separation

(a) By month, weather, unstable approach, and separation

Month	Total[a]	Weather[b]	Unstable[c]	Separation[d]
Jan	44	29	2	1
Feb	12	1	3	1
Mar	29	6	4	3
Apr	22	0	9	1
May	32	5	9	6
Jun	26	2	8	1
Jul	28	2	7	1
Aug	22	4	3	1
Sep	23	13	3	2
Oct	11	1	4	1
Nov	10	1	1	0
Dec	26	12	7	1

(b) By runway, weather, unstable approach, and separation

Runway	Total[a]	Weather[b]	Unstable[c]	Separation[d]
06	49	2	14	2
18C	33	3	5	5
18R	68	19	14	2
22	11	6	2	2
27	65	36	5	1
36C	12	3	4	1
36R	47	7	16	6

[a]Number of go-arounds
[b]Number of go-arounds with a weather score higher than 4
[c]Number of go-arounds linked with an unstable approach
[d]Number of go-arounds where the separation to the closest aircraft is between 1.5NM and 3NM

an unstable approach most often, at 40% of the time, followed by October, at 36%. When we look at the separation between aircraft, it seems to exert very little influence on the number of go-arounds.

Table IIIb shows how the number of go-arounds varies depending on the runway. The majority happens on runway 18R (68 go-arounds) and 27 (65 go-arounds). On this last runway, 55% of go-arounds are linked with poor weather conditions. In particular, 22 of 25 total go-arounds in January are concurrent with a difficult weather situation. Also, runway 18R has 28% of go-arounds associated with adverse weather. The peaks happen in January and September where go-arounds are associated with poor weather 73% and 87% of the time respectively. Unstable approaches precede 34% of go-arounds on runway 36R and 29% on runway 06. In particular, in April 80% of go-arounds on runway 06 are preceded by an unstable approach.

VI. Discussion

Different models are tested to monitor unstable approaches and go-arounds using ADS-B data. Overall all methods would benefit by having better data availability, as 28% of the go-arounds in the validation list are not present in the data, and 35% for unstable approaches. In particular, there are fewer data available at a lower altitude because this study relies on an antenna 40 km away. This impacts the detection of unstable approaches more than the detection of go-arounds because this last maneuver is performed at higher altitudes.

The horizontal compliance algorithm is the one that shows the most limited applicability. The analysis shows that landing aircraft are always within horizontal stability limits, thus the method has shown low efficacy in detecting unstable approaches.

The energy compliance algorithm, instead, shows promising results with a detection accuracy of 26% on the validation list. After improving the availability of data especially at lower altitudes, the threshold used for anomaly detection could be reiterated by a safety expert to improve the detection accuracy. GMM is chosen as the anomaly detection strategy due to its interpretability and its ability to handle multivariate data. The ability to detect anomalies at discrete flight point level makes the model more resilient in case of missing data.

A limitation of this model is that it is linked with the true airspeed calculation because the vertical wind velocity is not taken into account and it relies on the GFS NOAA wind data that is updated every 6 hours at intervals of 700 ft.

Go-around detection is the best performing model with a false positive rate of 2% and over 98% detection accuracy on the validation list. Minor variations of this model can potentially detect other aircraft operations, such as holding patterns, which provide a measure of the ATC workload. This paper attempts to understand the most likely circumstances of go-arounds. Another aspect that can be analyzed is the wake-category of the preceding aircraft. Nevertheless, some factors remain difficult to investigate such as an ATC instruction

because of an occupied runway.

The thresholds used for the S-functions of the go-around detection strategy are chosen empirically. Further studies could be performed to improve the choice. However, given the low false-positive rate and high detection accuracy, the final difference might be marginal.

It is also interesting to see that between December and January 25% of the unstable approaches are linked to severe weather conditions and 60% of go-arounds. This could indicate that pilots are more careful in severe weather conditions and tend to not perform landing in unstable approaches.

VII. Conclusions

This paper shows a data-driven approach to construct the basis of a safety monitoring system for approaching flights at an airport. The methods are able to extract safety knowledge from flight trajectory data. Of all three methods, two focus on detecting unstable approaches. The first one relies on the idea that a stable approach is constrained within certain horizontal bounds. The second one assumes that an unstable approach is characterized by an anomalous energy level. Thus, energy features are derived from ADS-B data, and anomalies are revealed using a GMM. The third method detects go-arounds using fuzzy logic and s-functions that dynamically scores go-arounds in trjactory data.

These models can detect anomalous safety events using ADS-B data from 2018 collected around Schiphol Airport. These results are aggregated to derive useful insights. For example, December and January are the months in which adverse weather conditions have the highest impact on the presence of unstable approaches and go-arounds. Also, on runway 27 unstable approaches and go-arounds are often linked with an adverse weather situation.

From the results, it can be concluded that the data-driven methodology proposed in this paper has the potential to enable independent monitoring of aircraft operations using aviation and meteorological open data.

Future research could focus on detecting other events that may impact the safety of operations at different flight phases, such as during take-off and ground operations. Another possible development is make use the detection model for the prediction of go-arounds and unstable approaches at a earlier stage of flights.

References

[1] IATA, "Unstable Approaches Risk Mitigation Policies, Procedures and Best Practices 2nd Edition," Tech. Rep., 2016.

[2] Z. Wang, L. Sherry, and J. Shortle, "Airspace risk management using surveillance track data: Stabilized approaches," in *ICNS 2015 - Innovation in Operations, Implementation Benefits and Integration of the CNS Infrastructure, Conference Proceedings*. Institute of Electrical and Electronics Engineers Inc., jun 2015, pp. W31–W314.

[3] A. Dahleni Kraemer, E. Villani, A. D. Kraemer, and E. Villani, "On the gap between aircraft FDI methods in industry and academy: challenges and directions," in *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics Inc, AIAA, 2019, p. 506.

[4] S. Das, B. L. Matthews, A. N. Srivastava, and N. C. Oza, "Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 47–56.

[5] S. Das, B. L. Matthews, and R. Lawrence, "Fleet level anomaly detection of aviation safety data," in *2011 IEEE International Conference on Prognostics and Health Management, PHM 2011 - Conference Proceedings*, 2011, pp. 1–10.

[6] B. Matthews, D. Nielsen, J. Schade, K. Chan, and M. Kiniry, "Comparative study of metroplex airspace and procedures using machine learning to discover flight track anomalies," in *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*, IEEE. Institute of Electrical and Electronics Engineers Inc., oct 2015, pp. 2G41–2G415.

[7] T. G. Puranik and D. N. Mavris, "Identification of Instantaneous Anomalies in General Aviation Operations Using Energy Metrics," *Journal of Aerospace Information Systems*, vol. 17, no. 1, pp. 51–65, jan 2020. [Online]. Available: http://dx.doi.org/10.2514/1.I010772

[8] T. Puranik, H. Jimenez, and D. Mavris, "Energy-based metrics for safety analysis of general aviation operations," *Journal of Aircraft*, vol. 54, no. 6, pp. 2285–2297, 2017.

[9] J. Sun, J. Ellerbroek, and J. Hoekstra, "Flight extraction and phase identification for large automatic dependent surveillance-broadcast datasets," *Journal of Aerospace Information Systems*, vol. 14, no. 10, pp. 566–571, aug 2017.

[10] Flight Safety Foundation, "ALAR BRIEFING NOTE 7.1: Stabilized Approach," Tech. Rep., 2009.

[11] T. G. Puranik and D. N. Mavris, "Identification of Instantaneous Anomalies in General Aviation Operations Using Energy Metrics," 2019.

[12] G. Jarry, D. Delahaye, F. Nicol, and E. Feron, "Aircraft atypical approach detection using functional principal component analysis," *Journal of Air Transport Management*, vol. 84, p. 101787, may 2020.

[13] R. Deshmukh and I. Hwang, "Anomaly Detection Using Temporal Logic Based Learning for Terminal Airspace Operations," in *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics Inc, AIAA, 2019, p. 682.

[14] M. C. L. Van Den Hoven, P. M. A. De Jong, C. Borst, M. Mulder, and M. M. Van Paassen, "Investigation of Energy Management during Approach-Evaluating the Total Energy-Based Perspective Flight-Path Display," 2010.

[15] L. Li, R. J. Hansman, R. Palacios, and R. Welsch, "Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 45–57, mar 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X16000188

[16] S. R. Proud, "Go-Around Detection Using Crowd-Sourced ADS-B Position Data," *Aerospace*, vol. 7(2), pp. 16–, feb 2020.

[17] Performance Review Unit and ATMAP MET working group, "Algorithm to describe weather conditions at European airports," Tech. Rep., 2011.

[18] ICAO, "ICAO Annex 2: Rules of the Air, Chapter 4," ICAO, Edition 42.0, nov 2009.

[19] The Boeing Edge, "Boeing Aeromagazine Issue 54 Quarter 02 2014." [Online]. Available: http://www.boeing.com/commercial/aeromagazine/articles/2014_q2/pdf/AERO_2014q2.pdf