

ANOMALY DETECTION IN ONBOARD-RECORDED FLIGHT DATA USING CLUSTER ANALYSIS

Lishuai Li, Maxime Gariel, R John Hansman, MIT, Cambridge, MA
Rafael Palacios, Comillas Pontifical University, Madrid, Spain

Abstract

A method has been developed to support Flight Operations Quality Assurance (FOQA) by identifying anomalous flights based on onboard-recorded flight data using cluster analysis techniques. Unlike current techniques, the method does not require pre-defined thresholds of particular parameters, but detects data patterns which differ from the majority of flights by considering all the available flight parameters. The method converts time series data from multiple flight parameters into a high dimensional data vector. Each vector captures all the available information for a single flight. Cluster analysis of the vectors is performed to identify nominal flights which are associated with large clusters and anomalous flights that do not belong to a specific cluster.

The method was applied to a representative Digital Flight Data Recorder (DFDR) dataset from an international airline. Detailed analysis was performed on takeoff and approach for 365 B777 flights. Abnormal flights were detected using the cluster technique which was able to identify anomalous behaviors including: high and low energy states, unusual pitch excursions, abnormal flap settings, high wind conditions. In addition, data clusters representing nominal conditions were also detected. Three distinct takeoff clusters were identified in the B777 data: one represented a majority of the takeoff cases, one correlated with a specific high altitude airport, one correlated with reduced power takeoffs. This initial evaluation indicates that cluster analysis is a promising approach for the identification of anomalous flights from onboard-recorded flight data.

Introduction

In order to improve the current high level of safety in air carrier operations there is increasing emphasis on proactive safety management systems which attempt to identify and mitigate risk areas before they manifest in aircraft accidents or incidents. One approach to proactive safety management is the

the analysis of operational data archived in modern Digital Flight Data Recorder (DFDR). This is often done through Flight Operations Quality Assurance (FOQA) programs which monitor daily flight records to improve airline operations and safety. Traditional FOQA data analysis rely on predefined criteria to identify anomalous flights. While this is effective at tracking known operational issues it does not allow the identification of emergent operational or safety issues.

This paper presents a method that evaluates the flight data and identifies anomalous flight without pre-specifying the criteria which defines the anomaly. Data mining techniques of Cluster analysis are used to group flights sharing common data patterns and to identify those anomalous flights which do not correspond to standard patterns. The anomalous flights can then be further reviewed by domain experts to determine if the anomaly is significant.

Background

FOQA Programs

Flight Operations Quality Assurance (FOQA) programs have been implemented at many US airlines and is mandatory in some European and Asian countries. The objective of FOQA is to improve airline operations and safety by analyzing the detailed flight data recorded during daily flights. FOQA data are recorded by an aircrafts DFDR or Quick Access Recorder (QAR). There is a large set of data for each flight which includes multiple flight parameters sampled at different rates depending on the parameter. The minimum requirements for DFDR data for U.S. carriers have been set by the FAA however many carriers record more parameters than the minimum standard [1] and the actual data set varies by airplane type, recorder type and airline configuration. Moreover, the recording capability on modern airplanes is much larger than the minimum requirements. For example, the DFDR on Boeing 787 will be able to record approximately 2000 flight

parameters for 50 hours, compared to the minimum requirement of 88 flight parameters for 25 hours [1]

FOQA Data Analysis

Currently, FOQA data is evaluated through either exceedance analysis or statistical analysis [2]. The exceedance analysis checks if particular flight parameter exceed a predefined limit under certain conditions. The list of flight parameters to watch and the limits of those parameters need to be specified by safety specialists a priori. The watch list is always chosen to coincide with the airline's standard operating procedures, such as the pitch at takeoff, the speed at takeoff climb, the time of flap retraction, etc [2]. The statistical analysis builds distributions of certain flight parameters during particular operations. An airline can gain a more complete picture of the operations based on the distribution of all flights using the statistical analysis than using the exceedance analysis.

Nevertheless, current FOQA data analysis only examines known operational or safety issues. Both techniques need a pre-defined watch list of key parameters under certain operational conditions. In addition, in the exceedance analysis, the thresholds of the key parameters need to be precisely defined in advance. As such, because current FOQA methodology cannot examine unknown safety threats, the potential for evaluating emerging risks is not being fully achieved.

Related Work

A number of methods have been developed to detect outliers using onboard-recorded flight data. The SequenceMiner algorithm focuses on pilot operations monitored through discrete flight parameters, such as cockpit switch flips [3, 4]. The Morning Report software package was developed to identify atypical flights using continuous flight parameters as part of NASA's Aviation Performance Measurement System (APMS) [5-7]. This work assumes typical flights are centered on the origin in a feature space, and therefore classifies points far away from the origin as atypical flights. The complexity of multiple nominal data patterns was not considered. The Inductive Monitoring System is a more recent method focusing on continuous flight parameters [8]. The method uses cluster analysis, but does not consider the temporal patterns in the data. Multiple

Kernel Anomaly Detection (MKAD) introduces an approach based on kernel functions to incorporate both discrete and continuous data streams [9]. In this approach, one-class Support Vector Machine (SVM), a classification method that requires labeled training data and assumes only one class of nominal flights, was used for anomaly detection.

In this study, we aim to develop an unsupervised method to detect anomalous flights considering temporal patterns in the data based on continuous flight parameters, with also consideration of multiple nominal patterns.

Methodology

Similar flights have common data patterns in the onboard-recorded flight data. Thus, if the data of each flight are mapped into a vector in a high dimensional space, similar flights should be "close" to each other in the hyperspace. Nominal flights would form clusters of proximate vectors and anomalous flights would be far away from any cluster. In order to identify clusters of proximate vectors in the high dimensional space, a type of data mining technique, cluster analysis, is used. The main advantages of cluster analysis include no longer requiring labeled training data and the capability to identify multiple nominal data patterns at the same time as identifying the anomalous outliers.

A three-step process for anomaly detection was developed which include: mapping the data to a high dimensional space, identify nominal data clouds, and detect outliers using cluster analysis. Similar approach has been used and shown to be effective in trajectory clustering for airspace monitoring [10]. Figure 1 provides an overview of this process. The first step transforms the multivariate time series data into high dimensional vectors. Each vector captures the information of a single flight. Then, in the second step, the dimensionality of the vectors is reduced for computational viability while maintaining the most important information. In the last step, cluster analysis is performed to detect outliers and clusters of normal flights in the feature space of reduced dimensions. Each step is described in detail in the following paragraphs.

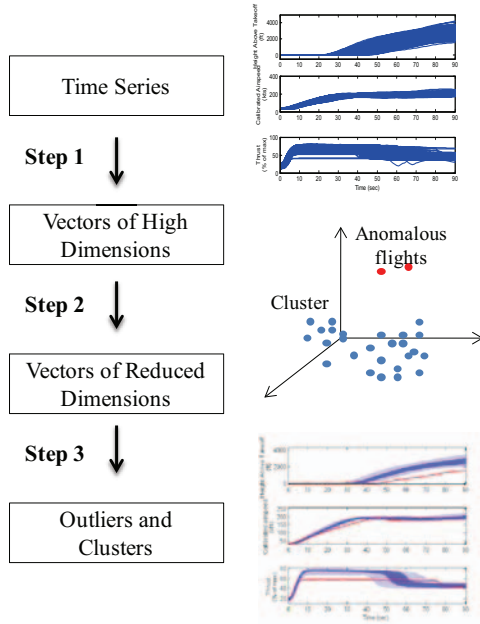


Figure 1. Detection Method for Anomalous Flights Based on Cluster Analysis

Transformation Time-series to Vectors

In order to map the raw data into comparable vectors in the high dimensional space, the time series data from different flights are anchored by a specific event to make the temporal patterns comparable. Then, every flight parameter is sampled at fixed intervals from the time of the reference event. All the sampled values are arranged to form a vector for each flight:

$$[P_{1t_1}, P_{1t_2}, \dots, P_{1t_n}, P_{2t_1}, P_{2t_2}, \dots, P_{2t_n}, P_{3t_1}, P_{3t_2}, \dots, P_{3t_n}, \dots, P_{mt_1}, P_{mt_2}, \dots, P_{mt_n}]$$

where P_{it_k} is the value of the i^{th} flight parameter at sample time t_k ; m is the number of flight parameters; n is the number of samples for every flight parameter. The total dimensionality of every vector is $m \times n$. Each dimension represents the value of a flight parameter at a particular time. The similarity between flights can be measured by the Euclidian distance between the vectors.

The current method is limited to flight phases that start or end with a specific event, such as takeoff or final approach. For the takeoff phase, the time of takeoff power application is used as the reference time and a number of samples are obtained at fixed time intervals, as shown in Figure 2. For the approach phase, the time series are first transformed

into a “distance-series” and then a number of samples are obtained backtracking from the touchdown point (Figure 3). Distance is used as the reference rather than time in the approach phase as procedures during approach are often specified based on distance or height above ground.

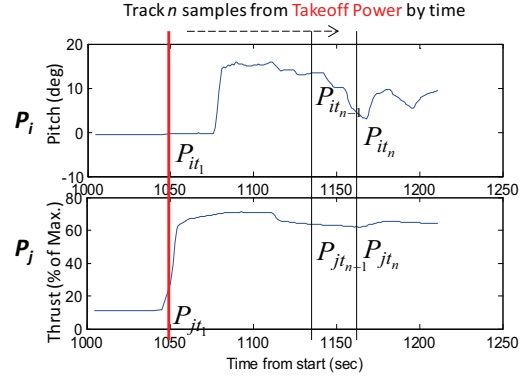


Figure 2. Sampling Time Series in Takeoff Phase

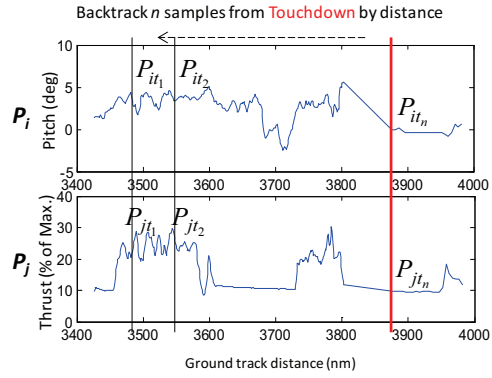


Figure 3. Sampling Time Series in Approach Phase

Dimension Reduction

Because of the temporal aspect, the vectors formed in the first step will normally have thousands of dimensions. For example if 80 parameters are evaluated over 100 time steps, this will result in an 8000 dimension analysis space. In comparison, the number of daily flights of a type of aircraft at a large airline is on the order of 1000 flights. This implies the typical daily dataset will have more dimensions than data points. It is difficult to identify data clouds in such sparse distribution. Therefore, Principal Component Analysis (PCA) was used to reduce the number of dimensions before performing cluster

analysis. PCA is a common procedure to transform data into an orthogonal coordinate system based on the variance in the data [11-13]. The greatest variance by any projection of the data comes to lie on the first principal component, the second greatest variance on the 2nd, and so on. As a consequence, the highest orders of principal components could be dropped to reduce the hyperspace dimension without losing significant information. In this study, the first K principal components that capture 90% of the variance in the data are kept.

$$\sum_{i=1}^K \lambda_i / \sum_{i=1}^N \lambda_i > 90\%$$

where λ_i is the variance explained by principal component i . N is the total number of principal components, which equals to the original number of dimensions. K is the number of principal components kept.

The magnitude of dimensional reduction will vary with the dataset but can be significant. As an example, in the analysis discussed later in this paper, the dimensions were typically reduced from 6188 to 77 for the takeoff data and from 6279 to 95 for the landing data using this criterion.

Cluster Analysis

The clustering analysis aims to identify clouds of data and outliers in the feature space. Several common clustering algorithms can be used to perform the cluster analysis. The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm was chosen because: 1) it can automatically determine the number of clusters; 2) it can handle data with noise/outliers; 3) it can detect outliers while identifying clusters (see Figure 4).

DBSCAN is a common density-based clustering algorithm [9]. It progressively finds clusters based on a density criterion. A cluster forms if at least $MinPts$ points are within ε radius of a circle. The cluster grows by finding the neighbors of the cluster, which also satisfy the same density criterion until no other point can be added into the existing cluster. At this point, it starts to search for a new cluster. Outliers are the points that do not belong to any cluster. Other than the two parameters $MinPts$ and ε to set the density criterion, no other parameters are required in DBSCAN.

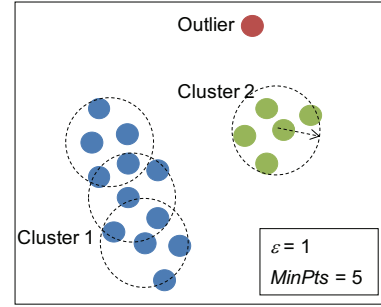


Figure 4. Example of DBSCAN Clustering Process

The selection of the two parameters was based on sensitivity analysis. For a fixed value of $MinPts$, DBSCAN was performed multiple times using a series of ε values ranging from the minimum pairwise distance to the maximum pairwise distance in the data. $MinPts$ was chosen to be the minimum number of similar flights that can be considered as a separate nominal group. While the number of outliers is sensitive to the value of ε , we set the value of ε to match user's preferences: finding the top $x\%$ outliers.

After the cluster analysis, outliers and clusters are identified in the space of reduced dimensions. Behaviors of the normal flights can be summarized by inspecting flights belonging to clusters. Outliers are the anomalous flights of interest. Potential risks might be discovered from the abnormal behaviors of anomalous flights.

Experimental Results

A representative DFDR dataset was obtained from an international airline to test the method. The dataset consisted of 2881 flights including 7 aircraft types with 13 model variants, e.g. A319, A320, and A321. To obtain relatively homogeneous data, the dataset was filtered by model variant. Among the 13 model variants, the set of B777 was the largest, which had 365 flights occurring over one month with various origins and destinations. This paper presents results of testing the method on this dataset.

Outlier detection was conducted separately for the approach phase and the takeoff phase. Three sets of DBSCAN parameters were tested to identify the top 1%, 3% and 5% outliers identified by the method. All the identified anomalous flights were further analyzed to determine the flights were abnormal and if so to characterize the abnormal behaviors. In

addition, when more than one cluster was present the differences in the nominal cluster data patterns were investigated to understand the cause of the different nominal behaviors.

Data Preparation and Parameter Selection

All 365 flights had valid data recordings. Every flight included 69 flight parameters including engine parameters, aircraft position, speeds, accelerations, attitudes, control surface positions, winds, and environmental pressures and temperatures. Radio height was only available during approach phase.

To allow flights at different airports to be compared, the position related flight parameters were first converted to values relative to the airport. For instance, the original recorded altitude values (e.g. pressure altitude, density altitude) were transformed to relative altitudes (e.g. height above takeoff, height above touchdown).

For the time-series to vector transformation, 91 observations were obtained at 1-sec intervals from takeoff power up to 90 seconds after takeoff for the takeoff phase, while for the approach phase the same number of observations was obtained from 6 nm before touchdown to touchdown. After performing the PCA, the number of dimensions was reduced from 6188 (68 flight parameters * 91 samples) to 77 for the takeoff phase and from 6279 (69 flight parameters * 91 samples) to 95 for the approach phase.

This sensitivity to cluster selection criteria (ϵ and $MinPts$) is shown for the Approach and Takeoff data in Figures 5 and 6. It was observed that the selection was insensitive to $MinPts$ (between 3 and 15) but that fewer flights are identified as outliers when ϵ increases. Therefore $MinPts$ was set at a value of 5 and the value of ϵ was selected to find the top 1%, 3% and 5% outliers.

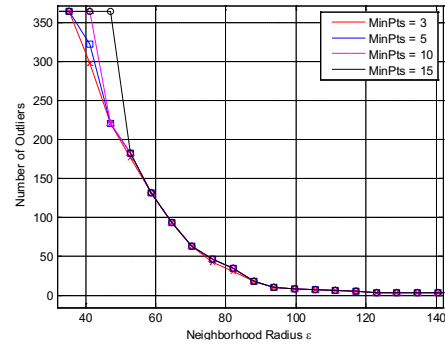


Figure 5. Sensitivity to ϵ and $MinPts$ (Approach)

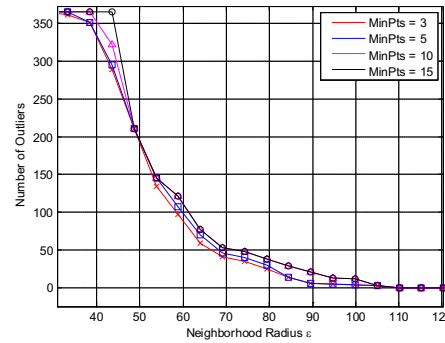


Figure 6. Sensitivity to ϵ and $MinPts$ (Takeoff)

Results Overview

Three sets of anomalous flights were identified using different parameter settings to match the top 1%, 3% and 5% outlier criterion. The results are summarized in Table 1. Further examination confirmed that anomalous flights found using a smaller outlier criterion were always included in the results obtained using a larger outlier criterion. The results showed that the cluster analysis using different setting of DBSCAN was path-independent.

Table 1. Number of Anomalous Flights Identified

Find Outlier	Top x%	DBSCAN Setting ($MinPts = 5$)	Number of Anomalous Flights
Approach Phase	1%	$\epsilon = 122.5$	3
	3%	$\epsilon = 93.9$	10
	5%	$\epsilon = 89.7$	16
Takeoff Phase	1%	$\epsilon = 100.0$	4
	3%	$\epsilon = 85.8$	9
	5%	$\epsilon = 83.4$	22

All of the identified anomalous flights were further analyzed to determine if they exhibited abnormal behaviors by comparing the anomalous flight parameters with distribution of flight parameters from all the flights. All identified flights exhibited some identifiable degree of anomaly.

For the approach phase, Table 2, the most frequent abnormal behaviors were high energy approaches and low energy approaches. Some flights were found to have unusual operations, such as abnormally high pitch, unusual flap settings, and lining up with localizer relatively late. In addition, environmental anomalies, such as strong crosswind and high atmospheric temperature, were found in some of the anomalous flights.

Table 2. Abnormal Behaviors in Approach Phase

Flight ID: Abnormal behaviors
High energy approaches
371040: Fast
373547: Fast
377860: Fast
378688: Fast, unstable airspeed
377844: High, line up late
377288: Initially fast, then normal
379685: Initially fast, then slow
Low energy approaches
383780: Low, slow
375698: Low, high power
383270: Low, unusual yaw trim
Other Unusual operations
383285: Unusual flap setting
384110: Unusual flap setting
371044: Abnormal high pitch
371045: Line up late
Environmental anomalies
372235: High atmosphere temperature
379665: Strong crosswind

The analysis of the takeoff phase was performed in a similar way and is summarized in Table 3. The most frequent abnormal behaviors were high and low power takeoffs which often include other notable factors. Also observed were: excessive reduction of power after takeoff, double rotation, and high pitch attitude during takeoff. It should be noted that not all anomalous flights identified indicate safety concerns. Some flights were identified as anomalous but were

benign cases, such as the takeoff in strong wind and the flight that turned soon after takeoff.

Table 3. Abnormal Behaviors in Takeoff Phase

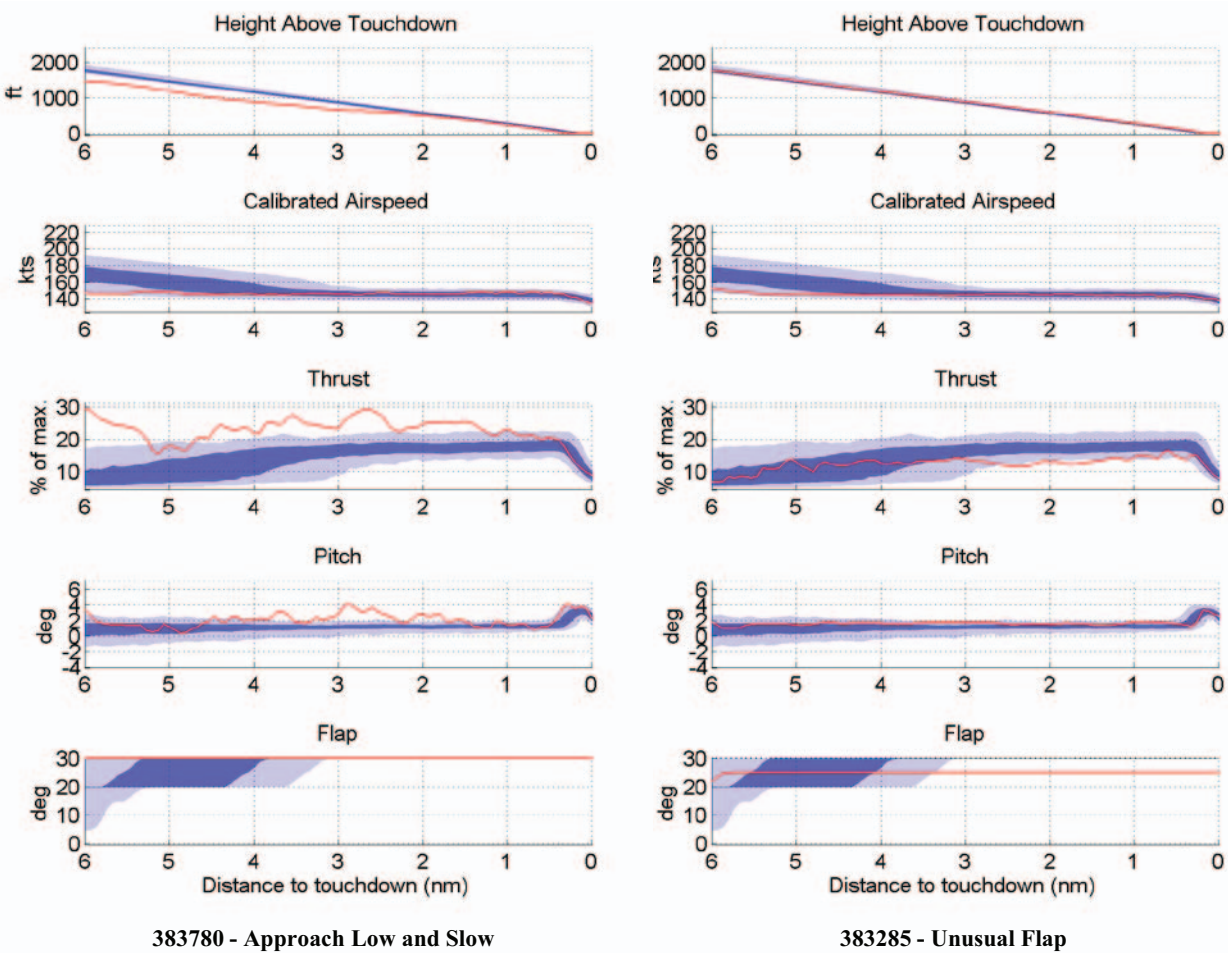
Flight ID: Abnormal behaviors
High power takeoffs
377862: Early rotation, high & fast climb out
380217: Early rotation, early turn
380219: Early rotation, crosswind
370715: Light, accelerate fast, climb fast
383285: Light, early rotation, early turn
384110: Light, climb out high, early turn
385160: High climb out, high pitch rotation
Low power takeoffs
368486: Reduced power, low & slow climb out
368487: Reduced power, low & slow climb out
369755: Reduced power, low & slow climb out
370019: Reduced power, low & slow climb out, extended period of high pitch
371045: Reduced power, low climb out
371046: Reduced power
379636: Reduced power, low climb out
385444: Reduced power, low & slow climb out
Abnormal power settings
369204: Excessive power reduction after takeoff
372209: Start with reduced takeoff power then switch to normal takeoff power, low & slow climb out
378692: Extended period of takeoff power
Other Unusual operations
373921: Double rotation
385702: High pitch rotation, climb out high
386369: Early turn after takeoff
Environmental anomalies
370723: Rise of spoiler, strong wind

Examples of Anomalous Flights

Two example anomalous flights with operational significance are presented in detail in this section for both the approach phase and the takeoff phase. The most distinctive flight parameters for each example are presented in graphs that use the same format. The anomalous flights are shown by red lines. The patterns of most flights are depicted by blue bands. The dark blue bands indicate the 25th to the 75th percentile of all flights' data; the light blue bands encompass the 5th to the 95th percentile. Respectively, the dark blue region contains 50% of the data, while the light blue region covers 90%.

Flight 383780 is a low and slow approach (see Figure 7 left column). The vertical profile is always below the common glide slop until 2 nm before touchdown and the calibrated airspeed is lower than most other flights until 3 nm before touchdown. Moreover, the flap is set to the landing configuration, 30, from at least 6 nm before touchdown. Therefore, this flight has to use a much higher thrust than most others until touchdown. It is also noted that a higher than normal pitch attitude is used to catch the glide slop between 3 nm and 2 nm before touchdown.

Flight 383285 uses Flap 25 all the way up from 6 nm before touchdown until landing, while most other flights are using Flap 30 as the landing configuration, as shown in Figure 7 right column. So less thrust is needed for the final part of the approach than most flights. Meanwhile, major indicators of the approach performance, the altitude, the airspeed and the pitch, are within the 90% normal range.

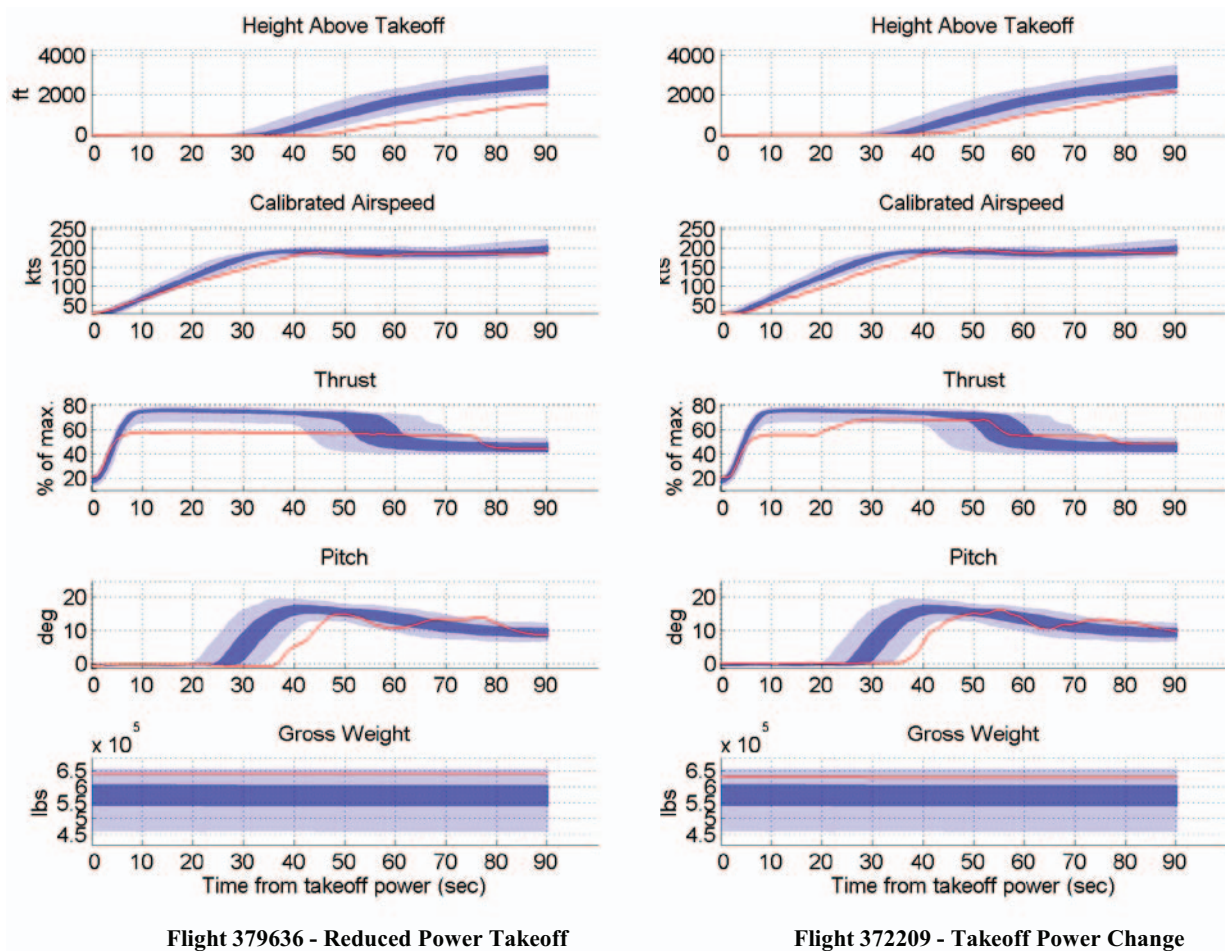


**Figure 7. Example Anomalous Flights in Approach Phase:
Approach Low and Slow (Left) and Unusual Flap Setting (Right)**

Example takeoff anomalies are shown in Figure 8. Flight 379636 show in the left column is the most significant anomaly. It used a much lower takeoff power than most other flights although it was at a relatively heavy weight. The degraded takeoff performance is apparent as the aircraft accelerates slowly and the rotation is not made until the airspeed reaches the required level. In addition, at 80 seconds after applying takeoff, the pitch reaches 15 degree that is similar to the angle during initial rotation. As

the aircraft is relatively underpowered, the climb rate is much lower than other flights as well.

Flight 372209 displays behavior similar to Flight 379636 for the first 20 seconds after applying takeoff power, as shown in Figure 8 right column. However, the power setting is changed back to normal level before rotation happens. As a result, the takeoff performance is better than Flight 379636. However, the climb rate and the acceleration are still lower than most other flights.



Legend: 90% 50% All Flights Outlier

Figure 8. Example Anomalous Flights in Takeoff Phase: Reduced Power Takeoff (Left) and Changed Power Takeoff (Right)

Nominal Data Patterns

The cluster analysis method can also be used to recognize different nominal patterns in the data which can be identified by different clusters. Each cluster represents a type of nominal data pattern. Typical operational behaviors can be characterized by retrieving flights from these clusters.

In this dataset, a single dominant cluster was found in the approach phase; while in the takeoff phase, a large cluster and two small clusters were identified. The result shows that most takeoffs shared a common data pattern and two small groups of takeoffs involved other patterns in this dataset. Table 4 summarizes the cluster structure identified using different density criterion. Cluster 2 was labeled as a separate cluster by all three outlier criteria, which indicates that flights in Cluster 2 were distinctive from most flights. Flights in Cluster 3 were identified to belong to a separate cluster only when the 3% outlier criterion is used. The flights in Cluster 3 were merged into Cluster 1 using the 1% outlier criterion and were classified as outliers using the 5% outlier criterion. Cluster 3 can be viewed as a sub cluster at the border of Cluster 1.

Table 4. Number of Flights in Clusters by Outlier Criterion in Takeoff Phase

	Outlier Criterion		
	1%	3%	5%
Cluster 1	353	341	335
Cluster 2	8	8	8
Cluster 3	--	7	--
Outliers	4	9	22

The differences between the clusters can be seen in Figure 9. Flights belong to Cluster 2 were takeoffs at OR Tambo International Airport (ICAO: FAJS), near the city of Johannesburg, South Africa. Due to the high altitude (5558 ft MSL), the takeoff performance is degraded compared to most other flights, as shown in green in Figure 9. Flights belonging to Cluster 3 were reduced power or de-rated takeoffs. They are shown in orange in Figure 9. They show reduced power settings with subsequently late rotations and lower climb rates.

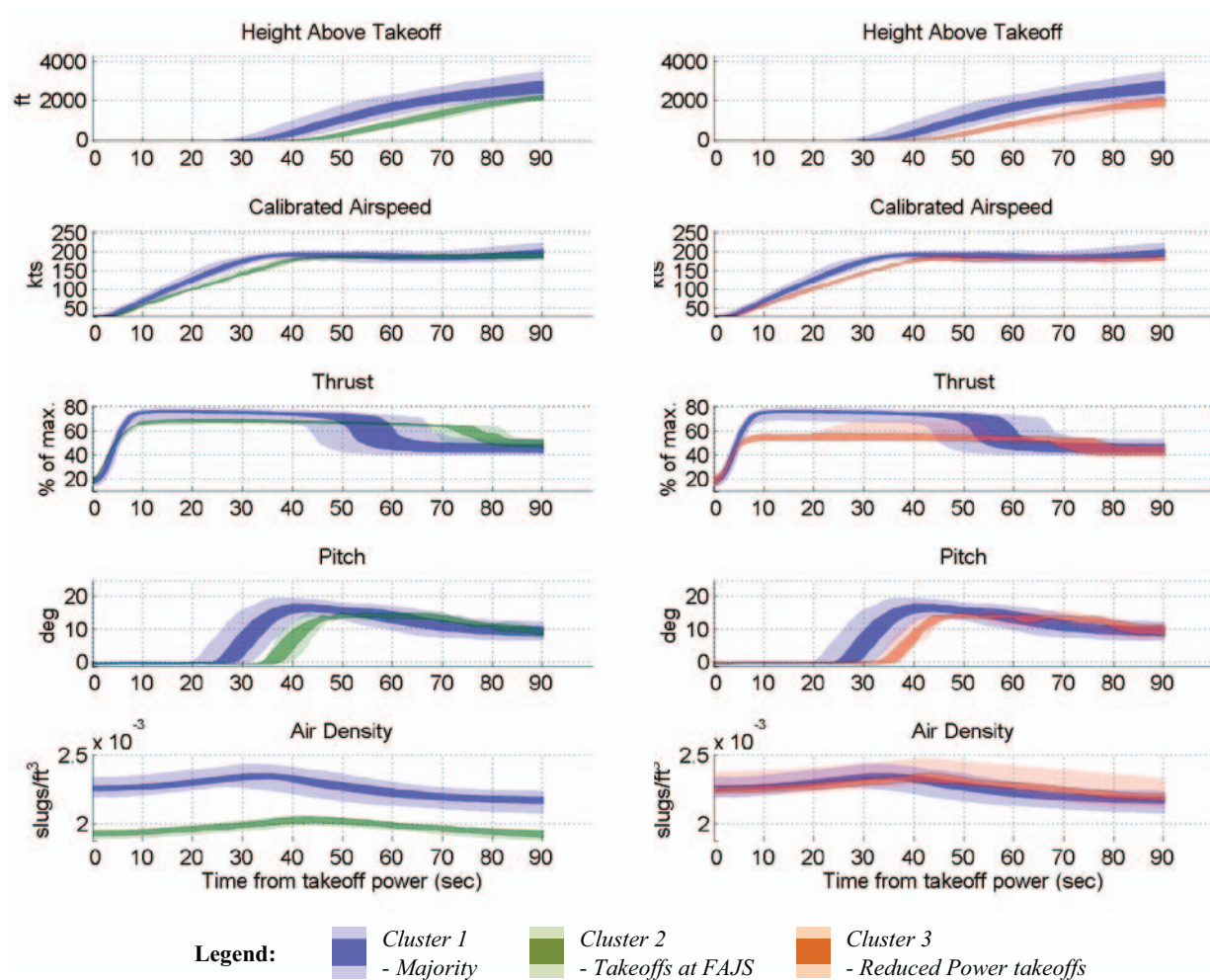


Figure 9. Patterns of Flights in Cluster 2 (Left) and Cluster 3 (Right) in Takeoff Phase

Conclusions

A method has been developed to support Flight Operations Quality Assurance (FOQA) by identifying anomalous flights based on onboard-recorded flight data using cluster analysis techniques. The method converts time series data from multiple flight parameters into a high dimensional data vector. Each vector captures the available information for a single flight. Cluster analysis of the vectors is performed to identify nominal flights which are associated with large clusters and anomalous flights that do not belong to a specific cluster.

The method was applied to a representative Digital Flight Data Recorder (DFDR) dataset from an international airline. Detailed analysis was performed on takeoff and approach for 365 B777

flights. Abnormal flights were detected using the cluster technique which was able to identify anomalous behaviors including: high and low energy states, unusual pitch excursions, approaches/takeoffs, abnormal flap settings, high wind conditions. In addition, data clusters representing nominal conditions were also detected. Three distinct takeoff clusters were identified in the B777 data: one represented a majority of the takeoff cases, one correlated with a specific high altitude airport, one correlated with reduced power takeoffs. This initial evaluation indicates that cluster analysis is a promising approach for the identification of anomalous flights from onboard-recorded flight data

References

- [1] Federal Aviation Administration, 14 CFR 121.344 Digital Flight Data Recorders for Transport Category Airplanes, 2011.
- [2] Federal Aviation Administration, Advisory Circular. 120-82 Flight Operational Quality Assurance, Washington, DC: 2004.
- [3] Budalakoti, S., A.N. Srivastava, and M.E. Otey, "Anomaly Detection and Diagnosis Algorithms for Discrete Symbol Sequences with Applications to Airline Safety," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 39, Jan. 2008, pp. 101–113.
- [4] Budalakoti, S., A.N. Srivastava, and R. Akella, "Discovering Atypical Flights in Sequences of Discrete Flight Parameters," Aerospace Conference, 2006 IEEE, Big Sky, Montana: IEEE, 2006, pp. 1–8.
- [5] Amidan, B.G., T.A. Ferryman, "APMS SVD Methodology and Implementation," Battelle Report, PNWD-3026, 2000.
- [6] Amidan, B.G., T. a Ferryman, B.P.N. Div, and W.A. Richland, "Atypical Event and Typical Pattern Detection Within Complex Systems," 2005 IEEE Aerospace Conference, Big Sky, MT: IEEE, 2005, pp. 3620–3631.
- [7] Chidester, T.R., "Understanding Normal and Atypical Operations Through Analysis of Flight Data," Proceedings of the 12th International Symposium on Aviation Psychology, Dayton, OH: 2003, pp. 239–242.
- [8] Iverson, D.L., "Inductive System Health Monitoring," Proceedings of The 2004 International Conference on Artificial Intelligence (IC-AI04), Las Vegas, Nevada: 2004.
- [9] M. Ester, H.P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Proceedings of the 2nd international Conference on Knowledge Discovery and Data Mining, Portland, Oregon: AAAI Press, 1996, pp. 226–231.
- [10] Gariel, M., A. Srivastava, E. Feron, "Trajectory Clustering and an Application to Airspace Monitoring", IEEE Transaction on Intelligent Transportation Systems, vol. pp, 2011, pp.1-14.
- [11] Pearson, K., "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine Series 6, vol. 2, Nov. 1901, pp. 559-572.
- [12] Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components.," Journal of Educational Psychology, vol. 24, 1933, pp. 417-441.
- [13] Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components.," Journal of Educational Psychology, vol. 24, 1933, pp. 498-520.

Acknowledgements

This work was supported by the FAA Joint University Program under grant FAA 95-G-017.

*30th Digital Avionics Systems Conference
October 16-20, 2011*