A Survey of Flight Anomaly Detection Methods: Challenges and Opportunities

Vivian Igenewari[1], Zakwan Skaf [2], and Ian K Jennions[3]

[1,2,3]*IVHM Centre, Cranfield University, Bedfordshire, MK43 0AL, United Kingdom*

v.igenewari@cranfield.ac.uk
z.skaf@cranfield.ac.uk
i.jennions@cranfield.ac.uk

ABSTRACT

Safety enhancement is a major goal of the aviation industry owing to the predicted increase in air travel. There is also the need to prevent fatalities, increase reliability and reduce monetary costs suffered as a result of delays and accidents that still occur. Accidents today are complex as a result of many causal factors acting alone but more often as a combination with other contributing factors. In tackling this trend, proactive measures have been put in place to find hazardous combinations that occur during flights in order to mitigate them before accidents occur. Flight Anomaly Detection (AD) methods are aimed at highlighting abnormal occurrences of a flight, that are different from the norm. As an improvement on the current state-of-the-art method, previous works have proposed different AD techniques for detection of previously unknown flight risks such as component faults, aircraft operational inefficiencies and some abnormal crew behaviour. However, current AD methods individually have limitations that prevent them from detecting certain significant anomalies in flight data. This paper surveys current flight AD approaches, their strengths and limitations as well as brings to light the benefits of a hybrid AD method to extend previous work and find safety-critical events, particularly those related to abnormal crew activity: a class of events known to amount for a substantial number of accidents/incidents today. It also highlights another emerging AD application opportunity, its challenges and how AD is beneficial in addressing them.

1. INTRODUCTION

The sharp decline in aviation accident rates from 5 to 0.22 per million departures (Boeing, 2017) over the last 50 years is the reason air travel is deemed the safest means of transportation today as shown in Figure 1. However, accidents and severe incidents still occur which affect the reputation of the industry. An **Accident** refers to an occurrence associated with the operation of an aircraft in which any person suffers death or serious injury, or in which the aircraft receives substantial damage. An **Incident** is "an occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation." (FAA, 2015). Examples of incidents are: high-speed approach, improper flap configuration, controlled flight into terrain only marginally avoided, aborted take-offs/landings on a closed or engaged runway. Improper actions often lead to a hazardous system state which in turn lead to an incident effect and if not addressed in time, will eventually lead to accidents.

In addition to the loss of lives involved in air accidents, airlines suffer huge monetary costs as a result of aircraft damage, delays, and other accident-related costs (Čokorilo, Gvozdenović, Vasov, & Mirosavljević, 2010). For example, Boeing publicises it costs approximately $300 million to buy one of the newer commercial aircraft models (Boeing) and no less than $10,000 per hour may be expended for any delay in getting the aircraft airborne (Pohl, 2015). With the growing demand for air travel (IATA, 2016), there is a possibility that accident numbers could increase if the safety levels are not enhanced simultaneously. According to the International Civil Aviation Organisation (ICAO) which is responsible for providing the standards for aviation safety, performance monitoring of the aircraft systems, maintenance operations, as well as the monitoring of the flight crew are some of the plans in place to achieve the desired safety level (ICAO, 2013). This monitoring aims to detect risks well in advance of serious accidents thereby preventing loss of lives and lowering the financial losses in the aviation sector (Cavka & Cokorilo, 2012).
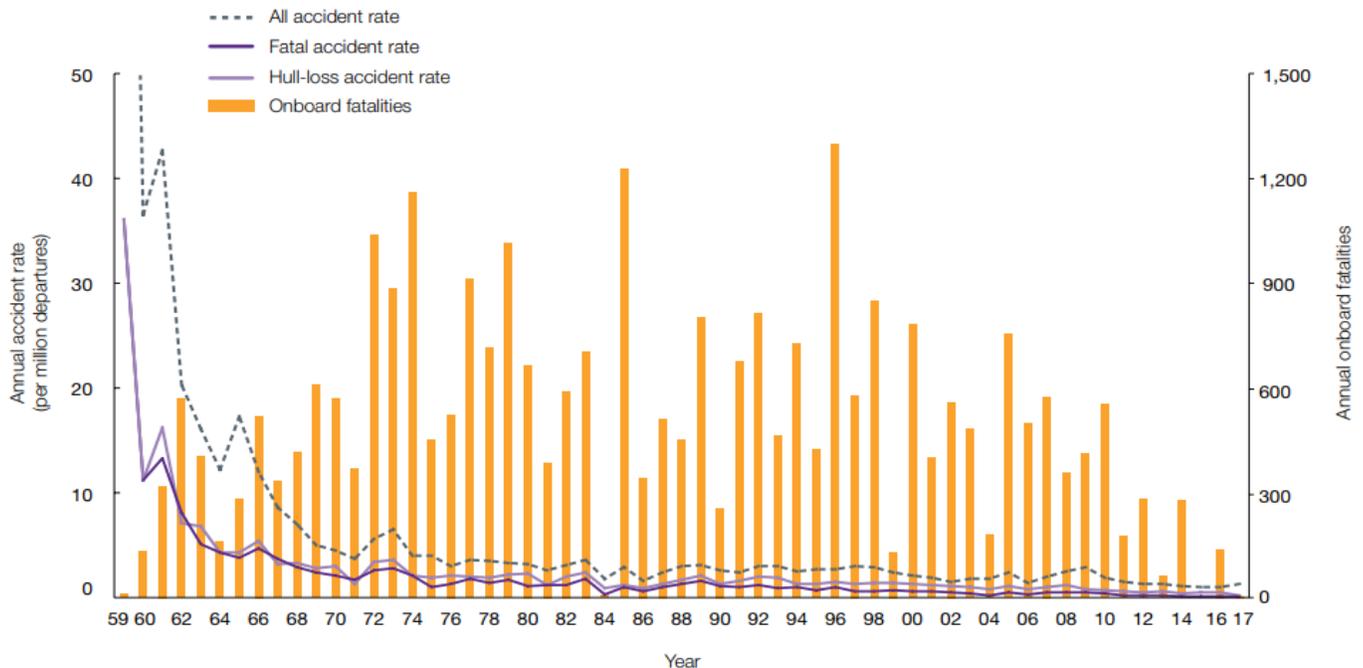
Figure 1. Accident rates and fatalities by year (Boeing, 2017)

Within the aerospace industry, an aircraft is able to gather up to 300,000 parameters from its sensors, depending on aircraft type and flight hours (Emmanuell Badea, Alin, & Boncea, 2018). A practical example in Pohl (2015) shows that a cross-country flight in the United States is able to produce 240 terabytes of data from both engines of a Boeing 737 aircraft after a 6-hour flight. This size of data increases rapidly depending on the number of aircraft in the sky and for say, a year. Such data is referred to as "Big Data" and the above example shows that the aerospace industry understands more than other industries, the challenges involved in making use of technologies to extract value from these huge volumes of data. Big data is often in varied formats that change at a high velocity and vary in context. Data Analytics (DA) is the science that acts on Big Data by using automated algorithms to uncover hidden patterns and insights for better decision making. In the aerospace industry, data analytics can be generally useful for cost reduction and revenue maximization which is achieved through means like better maintenance scheduling, performance monitoring and optimization of resources. Performance monitoring of systems, processes and people has become essential to safety in the aerospace industry (ICAO, 2013).

Improvements made to enhance safety have mostly been reactive as a response to accidents that have occurred. Flight Data Monitoring (FDM) programmes were introduced as a proactive approach for gathering and analysing routine flight data to improve flight crew performance, operating procedures as well as aid pilot training. These programmes

were aimed at identifying hazards before they led to accidents thereby aiding the establishment of appropriate mitigation strategies to prevent such occurrences. Data gathering for FDM programmes are available from onboard aircraft recorders, voice systems, maintenance reports, safety reports, cockpit display messages, or weather systems. The onboard data recorder sources, however produce the "Big Data" mostly useful for the analysis phase of FDM.

1.1. The Flight Data Recorder – Data Variability

The Flight Data Recorder (FDR) provides access to aircraft system data with which data analytics techniques extract the information required to carry out their processes. Figure 2 shows the flow of data to and from the FDR. All monitored aircraft systems pass their measurements through the information bus for the Flight Data Interface Unit (FDIU) and Data Management Unit (DMU) to select the parameters they have been programmed to process. Different recording units such as the current Solid-State Flight Data Recorder (SSFDR) and the Digital FDRs (DFDRs) store the appropriate parameters and produce different data frames which in turn require decoding programs to output useful data formats for analysis.

The SSFDR has expanded recording capacity and improved recorder reliability as an improvement to the DFDR which in turn replaced the older magnetic tape versions. The FDIU provides the required interface between the various sensors and the DFDR thereby solving the problem of the DFDRs' inability to process the larger amounts of incoming sensor

data. It converts analogue signals from the sensors to digital signals suitable for recording by the DFDR. However, the FDIU is not programmable by the operator as it sends only basic regulatory information to the DFDR/SSFDR. The DMU, on the other hand, can be programmed by an operator to record any convenient parameters ultimately producing the DAR (Digital ACMS - Aircraft Condition Monitoring System Recorder) data as also shown in Figure 2.

For analysis purposes, both the QAR (Quick Access Recorder) and DAR data are used based on ease of information retrieval and/or variability of parameters they are able to provide. The QAR is a copy of the FDR on a non-crash resistant recorder which has the specificity to be easy to download and interchangeable (ATR, 2016). DARs are like QARs (ICAO, 2016) however, while the QAR data originates from the FDIU with minimum regulatory parameters as the FDR, the DAR data, on the other hand, is the product of the DMU able to include many more convenient parameters. Both systems were developed to ease data recovery for the airlines' needs and contain high-frequency sensor data.
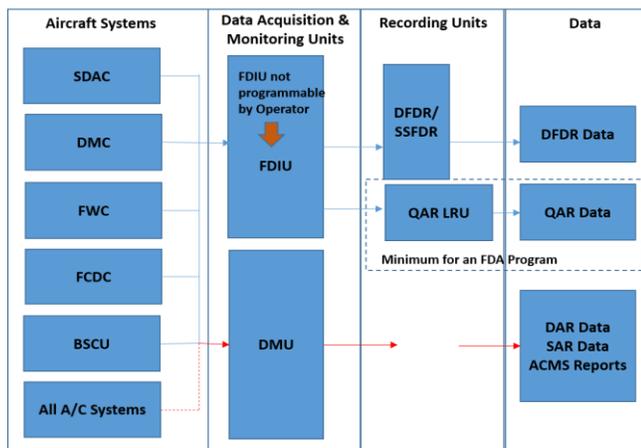

Figure 2. FDR formats and flow (ICAO, 2016).

These different aircraft devices record parameters of different volumes, different data types and at varied speeds thereby giving rise to multiple data dimensions and data inconsistencies among other challenges of flight data analytics. Examples of varied FDR parameter types are **discrete variables** such as landing gear state, landing flap position, autopilot modes, flight director status and **continuous variables** like altitude, wind speed, airspeed, roll, pitch, angle of attack and so on. Though the varied nature of the parameters that recorders are now able to record can provide more insights into occurrences during a flight, data-driven AD techniques have to deal with these data variability challenges in various ways in order to achieve desired results.

The following sections reveal the effectiveness of current and emerging flight data analytics techniques, their strengths and weaknesses with regards to detecting anomalies while dealing with data variability from FDR.

1.1. Flight Data Analysis

The success of FDM depends on insights that can be gained from existing flight data through Flight Data Analysis (FDA) also known as Flight Operations Quality Assurance (FOQA) in the improvement of aviation safety. The effective analysis of flight data has been enhanced by the increased number of parameters that can be measured and the emerging techniques in data analytics.

Exceedance Detection (ED) is one of such data analytics techniques. It is the state-of-the-art method used across the industry for flight data analysis. ED is implemented as a knowledge-based technique which relies on input from domain experts to make its inference within the data. It is, therefore, only able to highlight known risks within the data. More recent studies have built on the limitation of the ED method giving rise to the Anomaly Detection (AD) approach aimed at identifying unknown abnormalities during the flight which cannot be detected by the current method. Highlighting unknown combinations within the aircraft environment is useful in detecting operationally significant events that could potentially lead to accidents. Anomaly Detection approaches could either be **Model-based** relying on domain knowledge to build models or **Data-Driven** which focuses on automatically learning the general rules to build models directly from the data. However, for a complex system like the aircraft, it is not possible to model every combination of factors using model-based methods. Therefore, the advancement in technologies like data mining, data analytics and machine learning are moving the focus from knowledge-based approaches to data-driven methods of improving FDA.

Data-driven approaches are used to describe system behaviour using only the data. Data-driven methods for data analysis exploits large amounts of data available to better understand behaviours and operational faults. These approaches learn by observing and analysing system behaviour of complex systems. The goal here is to build systems that can observe, diagnose, recognize unusual events and inform operators for more effective decision making (Biswas et al., 2016).

This paper thus explores existing data-driven flight AD techniques, their applications, strengths, limitations and proposes a more robust hybrid AD technique to detect a wider range of significant anomalies aimed at enhancing flight safety. The paper is organised as follows. An introduction of FDM, the FDR data variability and FDA are provided in this section. Section 2 reviews the flight AD evolution, its architecture, aviation applications and its methods as well as the strengths and limitations of these methods. Section 3 summarises limitations in existing flight AD techniques and proposes a hybrid AD opportunity to mitigate some of them. Also, the application of AD to the area of Standard Operating Procedure (SOP) deviation detection was introduced as another opportunity. Lastly, section 4 concludes the paper.

2. REVIEW OF FLIGHT ANOMALY DETECTION METHODS

Anomaly Detection refers to the task of identifying new or unknown patterns which, in many cases are abnormal or inconsistent from the norm. AD finds patterns in data that do not conform to expected behaviour (Chandola, Banerjee, & Kumar, 2009). These patterns are called anomalies or outliers and they are often sufficiently different from the majority of data points in the dataset. Since the aircraft operations are standardized, it is assumed that the majority of flights represent normal operations at that given time. Flight data analysis has given rise to a number of AD techniques for discovering anomalies.

This section details the general architecture of an AD method. It also provides a review of current flight AD application areas, the different AD techniques developed for use in each application as well as the strengths and limitations of these techniques.

2.1. AD Architecture

Each AD technique has a slightly different way of approaching the AD problem depending on available data and anomalous behaviour they wish to detect. The available flight data could be FDR, voice, maintenance data, safety reports, cockpit display messages, or weather data. Analysts may wish to detect anomalies in flight operations, components, crew actions or even predict system failures.

Figure 3 gives an overview of the generic AD architecture. Flight AD is generally based on the assumption that most of the flights in the dataset represent normal operations. The DA process thus involves preparing the data for analysis and the AD process begins with the output from the DA into a detection phase and then followed by the characterisation of the anomalies found as the output of the AD process. The detection phase generally entails performing similarity measures between data points thereby determining the nominal group and using this to ascertain the outlier groups;

giving each output an anomaly score say, a zero score referring to a nominal event. The characterisation phase involves identifying the dominant features associated with each anomaly then a domain expert validates its operational significance.

2.2. Anomaly Detection Applications and Techniques

AD has been proposed for use in different industries for diverse applications such as intrusion detection and fraud detection on networks/systems. In this work, we focus only on the aviation sector and how various AD techniques have been applied to FDR data to detect unknown flight profiles and abnormal events. Applications of AD for flight data are condition-based monitoring, condition-based maintenance and safety performance. Some of the current flight AD application areas and techniques are presented here.

Aircraft Condition-based Monitoring (ACM). This refers to the measuring of system parameters over time and while in operation. It is majorly used for predicting failure and calculating the remaining life trends of components. With ACM, replacement, for example, is only performed when the analysis of the data reveals degradation in the monitored components. Some AD methods have been proposed to help anticipate aircraft system failure risks as summarised in the next paragraphs.

The Aviation Performance Measuring System (APMS) was developed by NASA aimed at going one step further than ED by examining trends in multiple flights rather than single flight analysis (Statler, 2007). Further to this was the Morning Report (Chidester, 2003) developed in the late 90s as a software tool developed by NASA. The Morning Report relies on statistical means to identify atypical flights from multivariate comparison with baseline typical flights and tells which parameters contributed to an anomaly in a flight phase. This tool was designed for individual airlines attempting to address the problem of balancing between analysis for each flight and across multiple flights. It also clustered the resulting
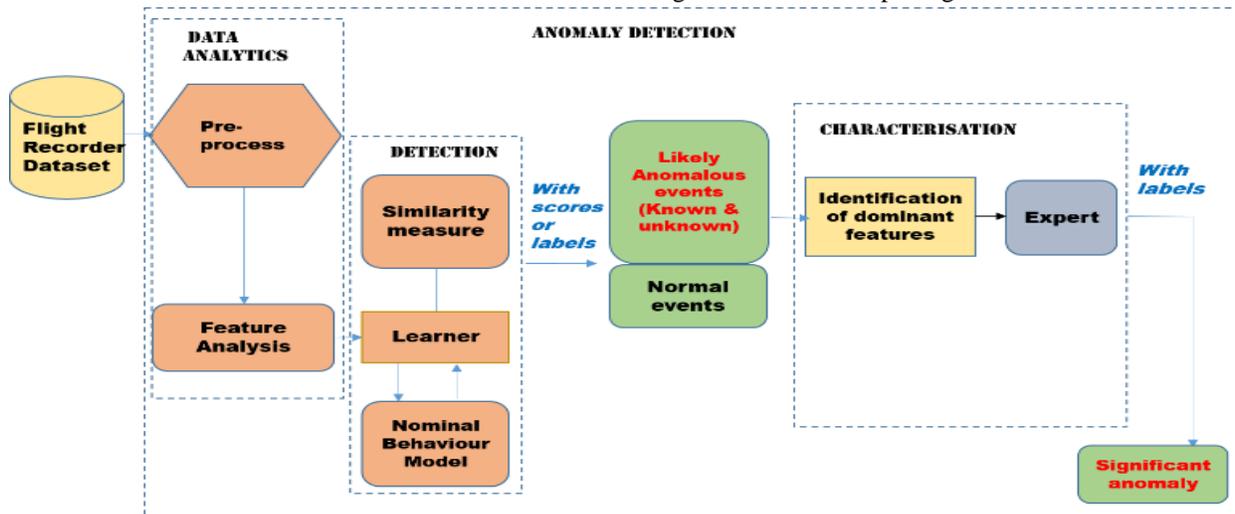

Figure 3. General AD Architecture

flights to help interpret the anomalous event. Morning Report, however, only focuses on continuous data.

Other clustering techniques and data analysis approaches were introduced such as Inductive Monitoring System software (IMS) (Iverson, 2004). It is a distance-based AD technique applying an unsupervised algorithm and using incremental clustering to build models of the expected operation of the system on a set of normal data. The model is then used to test new data and the distance score determines if it is classed as a nominal mode or anomalous mode. A limitation of this method is its vector space-based method which becomes computationally expensive as point by point analysis is performed. IMS does not consider temporal patterns in data hence sequential dependencies in the data are not incorporated. Also, the Euclidean distance used for measuring the similarity between discrete values is limited.

Developed by (Bay & Schwabacher, 2003), Orca is a k-nearest neighbour based unsupervised AD algorithm. It has a nested loop structure to calculate pairwise distances between data points but uses a simple pruning rule to keep the time complexity significantly less than the number of data points achieving near-linear performance with high dimensional data thereby making it appropriate for analysing large datasets. It can process both continuous and binary data format. A limitation of this method is that since each data point is scored independently, anomalies in the time domain cannot be detected (Das, Sarkar, Ray, Srivastava, & Simon, 2013). iOrca (Bhaduri, Matthews, & Giannella, 2011) is a scalable version of Orca since the developers of iOrca introduced an indexing strategy and an early termination criterion to make it scalable to extremely large datasets during processing (Das et al., 2013) but still suffers from the time domain limitation of Orca.

Aircraft Condition-Based Maintenance (CBM). AD for Maintenance Operations was motivated by the possibility that large flight datasets could be useful in understanding undesirable events like aircraft failures and need for scheduled maintenance. Maintenance activities ensure that the system does not fail, is functional and available for operation.

A new data analysis algorithm (C.-H. Lee, Shin, Tsourdos, & Skaf, 2017) for FDR data was proposed focusing on improving airline maintenance operations. This new approach categorizes the FDR time series data into continuous, discrete and warning signals according to their parameter characteristics. A type of k-nearest neighbour approach is used to detect abnormal patterns in FDR coupled with the basic data analytics process involving appropriate techniques for Data Filtering, Data Transformation, Feature generation, Correlation Relaxation and Dimensionality Reduction. The performance of this algorithm was tested with NASA data. Results showed that it can detect unusual patterns within FDR regardless of parameter types thereby helping domain experts interpret data in less time for high-level fault diagnosis.

Safety Performance. This refers to the reduction to an acceptable level, risks associated with the operation of the aircraft and/or other aviation activities. Its continuous enhancement is deemed a fundamental objective of the ICAO

and other aviation organisations. Many AD techniques have been developed to tackle the requirements for safety enhancements.

Sequence Miner technique (Budalakoti, Srivastava, & Otey, 2009) was developed as an unsupervised method to analyse discrete parameters and reveal atypical flight sequences. Only discrete parameters were considered, and it solves the problem of detecting and describing anomalies in large datasets of high-dimensional symbol sequences. It does this by using the nLCS (normalized Longest Common Subsequence) based distance measure. It is able to detect anomalies across a fleet of aircraft. It is, however, limited to analysing similar flights and aircraft since discrete sequences can vary greatly depending on the route, airport and equipment. It is unable to integrate its findings of discrete actions with other continuous flight parameters since its performance degrades with interaction with continuous data.

One of the more recent techniques is the Multiple Kernel Anomaly Detection commonly known as MKAD (Das, Matthews, Srivastava, & Oza, 2010). It has been extended and investigated for small and medium-sized medical and finance datasets (Gautam, Balaji, Sudharsan, Tiwari, & Ahuja, 2019). MKAD is a semi-supervised algorithm proposed for AD. Heterogenous parameters (continuous and discrete) were handled together. Discrete binary switch sequences are used in the discrete kernel and discretized continuous parameter features are used to form the continuous kernel. MKAD combines multiple kernels into a single optimization function using the One-Class Support Vector Machine (OCSVM) framework. The OCSVM uses a training set of nominal examples (in this case, flights) and evaluates test examples to determine whether they are anomalous or not. It combines the strength of both vector space and sequential AD techniques to allow a wider range of anomaly detection. Its ability to combine information from both data types makes it able to identify more complex behaviours of the crew in the anomaly detection process, unlike previous methods. It is, however, more able to find anomalies in discrete data than it is able to detect those from continuous data. For example, while is it was able to detect anomalies in an auto-landing state due to an autopilot mode anomaly, it was unable to detect a high-speed approach based on relevant continuous parameters (Das, Li, Srivastava, & Hansman, 2012). MKAD also needs nominal or near nominal data to learn proper hyperplanes hence, is not appropriate for including multiple nominal data patterns since the trained model could identify other nominal patterns as anomalous.

ClusterAD (L. Li, Das, John Hansman, Palacios, & Srivastava, 2015) is a type of unsupervised learning method to detect an abnormal operation of an aircraft using FDR data. Euclidean distance was used as the similarity measure for both continuous and discrete parameters thereby giving rise to the possibility of incorrectly computed discrete parameters. A density-based clustering algorithm called Density-Based Spatial Clustering of Applications with Noise (DBSCAN) was used also to automatically determine the number of clusters, progressively find clusters based on density criterion

and find outliers in the feature space, but an output of this algorithm is sensitive to its design parameters (Ester, Kriegel, Sander, & Xu, 1996). ClusterAD has also been used for flight AD through its combination with the LoOP – Local Outlier Probability machine learning technique (Oehling & Barry, 2019). ClusterAD is tuned to find anomalies in FOQA type data but does not clearly define the similarity measure of discrete data thereby producing a low performance in finding issues in discrete data (S. Lee, Hwang, & Leiden, 2014). It also loses temporal information during the reduction step thereby making interpretation of anomalies harder and even missing some vital short-length anomalies.

Mack, Biswas, Khorasgani, Mylaraswamy, and Bharadwaj (2018) contributed to improving aviation safety by proposing a general methodology of combining completely unsupervised learning techniques with supervised expert anomaly identification for large volumes of data. The method involves using diverse techniques for data pre-processing, data reduction, applying clustering algorithm, determining nominal and outlier cluster groups, identifying significant feature parameters associated with the anomalous group, domain expert characterization of special nodes and anomalous behaviours within the system. It further details how experts work with the anomalies produced and how they may extract information for decision-making. This work focused on the take-off phase of flight and was able to discover interesting anomalies such as a dead engine, near stall and a pilot choice to delay switching to autopilot mode soon after take-off. This work seeks to use a similar method developed by the author but proposes a hybrid AD technique to address the limitations of existing methods while combining their individual strengths. A summary of these application areas and respective techniques used is summarised in Figure 4.

3. FLIGHT AD LIMITATIONS AND OPPORTUNITIES

In this section, the limitations of techniques highlighted in the previous section are summarized and a more robust data-driven hybrid AD opportunity for detecting a larger range of anomalies is proposed. Also, the challenges SOP deviation detection and the opportunity of using AD to address them are reviewed.

3.1. Hybrid AD Opportunity

The previous section detailed the strengths and limitations of existing flight AD methods. Many AD techniques have been able to find component faults and operational faults but only a few of these (W.-C. Li, 2011; Shappell et al., 2017) are able to detect some type of anomalies of crew activity. Here, we propose a hybrid technique to combine the strengths of multiple AD techniques with the aim of detecting a more varied range of anomalies than individual techniques.

For example, while some techniques are able to detect anomalies influenced by deviations in discrete parameters, some are better at effectively detecting anomalies of continuous variables. Moreover, some techniques are unable to identify nominal instances and some techniques that can, are only able to do so for certain portions of flight. Furthermore, while all the flight AD techniques mentioned were able to detect anomalies of either discrete, continuous, sequences, long-length anomalies, there is no evidence that any would detect short-length anomalies. This could be due to the need for dimensionality reduction by these techniques which often filters out vital information (Nanduri & Sherry, 2016). Detecting short-length delays are important because (Donald L. Wood, 2018) detailed how a 4.5s delay in putting down the lift dump at touchdown, caused the aircraft to overrun the runway leading to other incurred costs. A hybrid AD technique will be useful in combining the strengths of these multiple methods into one thereby finding anomalies in
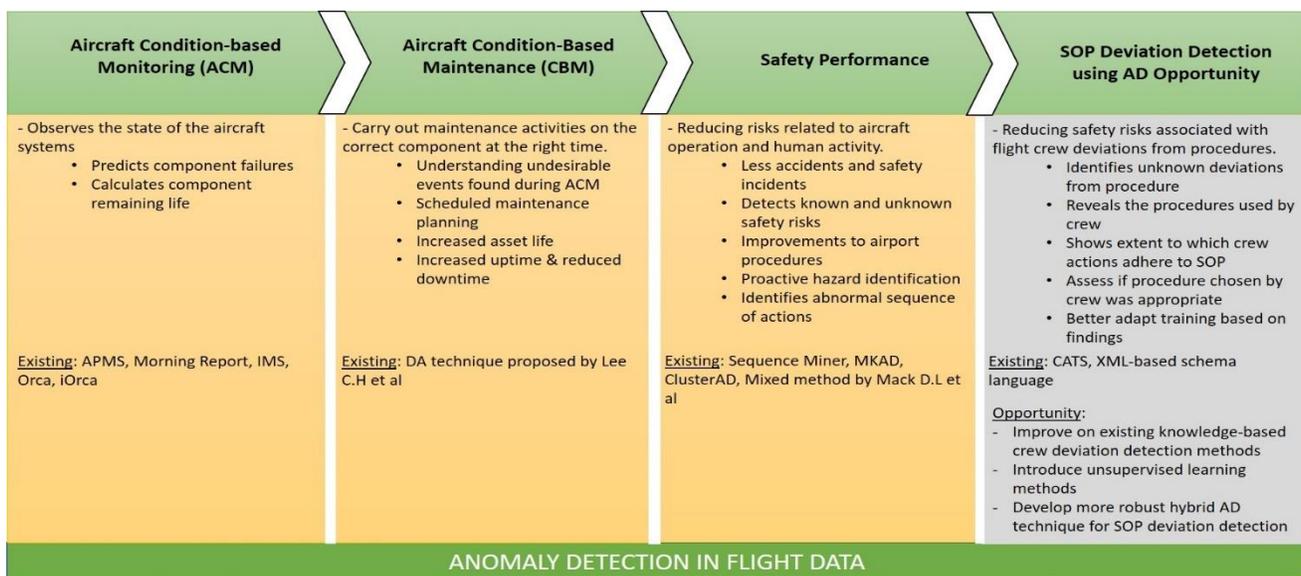

Figure 4. Summary of AD techniques and opportunity

discrete, continuous, sequence parameters as well as inculcating the capability for detecting long and short-length abnormalities in the data.

Figure 5 gives an overview of this mixed approach. The FDR data is pre-processed using appropriate methods including the selection of relevant features to ensure that vital information in the data is retained as much as possible. It is

Boeing reports that 80% of accidents today are caused by human errors (Boeing, 2007). In aviation, flight crew error could be in the form of pilot mistakes, oversights/omissions, commissions and failure to adhere to Standard Operating Procedures (Key Dismukes, 2004). As commercial flights are highly standardized, aircraft operations are described through SOPs. Depending on the operational situation, the appropriate procedure is expected to be applied by the crew.
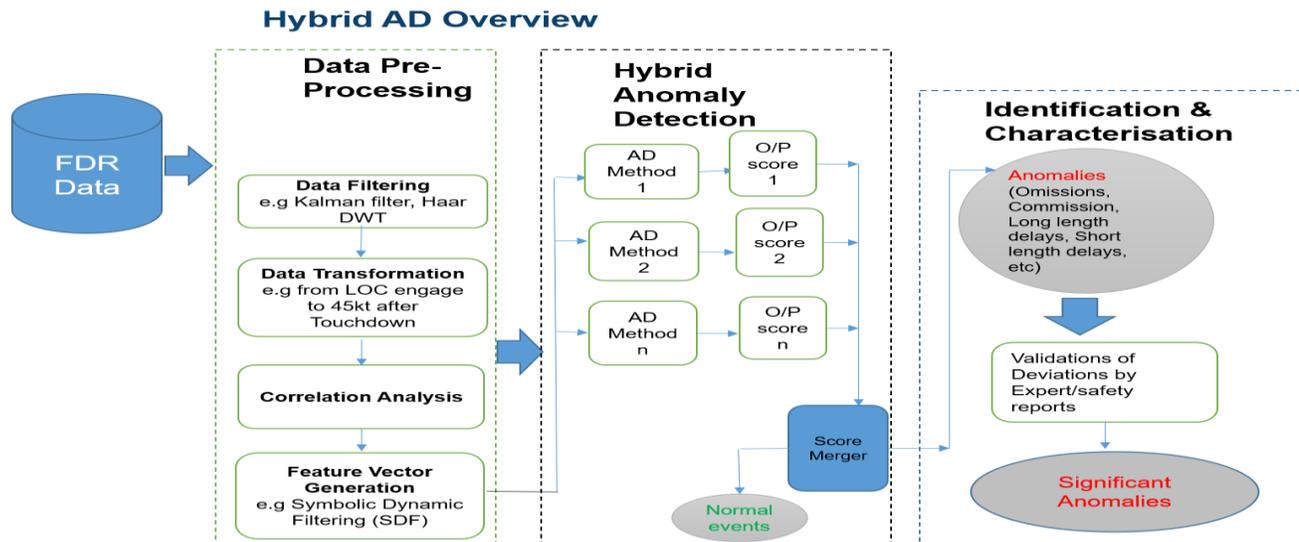

Figure 5. Hybrid AD Overview.

filtered for small portions of certain phases of flight to ensure all flights are comparable. The detection phase combines multiple data-driven techniques which output individually-detected anomalies with their respective scores. The anomaly labels/scores can be numerical representing the extent at which a technique perceives the anomaly as significant. The hybrid AD technique includes a score merger function which takes the output scores of the multiple algorithms and outputs a final score that effectively represents how normal or abnormal an event is from the fleet-wide analysis. The scoring algorithm will use a detection threshold set at different sensitivity levels to ascertain the number of flights classed as anomalous thereby detecting, for example, the top y% anomalies in a set of flights. The score merger should also be able to make accurate judgements on anomaly classification in cases of conflicts in output between individual techniques. The final output of the hybrid algorithm is measured against baseline techniques and validated using domain experts and/or available safety reports which in turn validate the actual significance of the anomalies found.

3.2. AD-SOP Opportunity

A Standard Operating Procedure or Checklist is a series of steps laid out to guide flight crew in appropriately and safely controlling the functioning of the aircraft.

However, it is found that deviations from SOP form a significant part of aviation accidents today involving loss of lives and other accident-related costs (CIAIAC, 2008).

Possible scenarios that occur in the operation of a flight in relation to SOP adherence are cases where: 1) Procedures are followed but no accidents occur. 2) Procedures are not followed, and an accident occurs. 3) Procedures are followed but accidents still occur. 4) Procedures are not followed but no accident occurs. In other to proactively tackle the occurrence of SOP non-adherence, it is critical to focus on the latter scenario with a view to also detect incidents that occur due to SOP deviations during such flights so that mitigation strategies can be put in place to prevent any impending accidents.

There are however, challenges in detecting SOP deviation. One of these is the difficult task of building a representation for every procedure (e.g operator, airline and/or airport procedures) which requires a good knowledge of the domain. It is also challenging to analyse and isolate the source of every crew anomaly algorithmically. Furthermore, a case of normalization of deviance where an SOP deviation error among the majority of the crew over a long period has become perceived as normal, could occur.

The state-of-the-art outlier detection method is able to detect known procedure deviations but does not allow the discovery of emerging/unknown deviations from procedure. Some research (Callantine, 2001; Maille, 2013, 2015, 2017) have

addressed this gap by introducing expert-based methods for modelling SOP normal behaviour and then detecting unknown procedure deviation anomalies that arise. However, an improvement on modelling SOP with more unsupervised learning methods is desired.

There is, therefore, the need to explore flight data using Anomaly Detection framework to detect subtle underlying misunderstandings of the flight crew in relation to deviations from laid down procedures, which do not lead to accidents or incidents under most conditions or even hard to detect by the existing methods. (Jingru Yan, 2013). It is beneficial in not only implicitly considering the crew behaviour (in terms of a sequence of switches the crew may have triggered) as highlighted in Figure 4, but also the relation between pilot-related parameters and existing standard operating procedures. A data-driven AD approach can address the highlighted SOP deviation detection challenges by using patterns within the flight data to better identify and understand significant crew deviations that occur during portions of flights, simplify the validation work of experts by giving scores to anomalies, give more insight into understanding pilot aircraft operations and highlight ways to improve aircraft design or procedures.

The use of data-driven AD techniques in the detection of crew procedure deviations is a step in the direction of identifying leading indicators in safety management. This could then serve as advance warnings of latent problems in training, procedure design or understanding automation to enhance safety. It also complements the state-of-the-art method by highlighting new exceedances related to crew activity by moving from a "was the lift dump deployed 1s after touchdown?" to "how many seconds after touchdown was the lift dump deployed?".

4. CONCLUSION

As also summarised in Figure 4, this paper presented a review of flight anomaly detection techniques highlighting the application areas where they have been useful. The paper provides the strengths and limitations of these techniques as well as an emerging challenge of detecting deviations from standard operating procedures. It proposes the opportunities of using a data-driven hybrid AD method to address the limitations in existing techniques and the benefits of applying an AD approach to tackling SOP deviation challenges. The hybrid AD framework combines the strengths of multiple flight AD techniques to detect a wider range of anomalies than individual methods. A general overview of this approach is given from data pre-processing to the identification and characterisation of the detected anomalies by experts.

The primary contribution of this work is the discussion on the opportunities of AD approaches in improving existing limitations and emerging challenges for flight safety by carrying out a review of flight AD techniques. Specifically, it proposes a more robust data-driven AD method for the detection of a varied range of anomalies in high dimensional flight data by using a score merging function. The use of

anomaly detection for aviation applications described in this paper provides avenues to complement existing detection methods and ultimately contribute to global safety enhancement initiatives.

REFERENCES

ATR. (2016). Flight Data Monitoring on ATR Aircraft. In (pp. 16). Blagnac Cedex - France: ATR Product Support & Services.

Bay, S. D., & Schwabacher, M. (2003, August 24-27). *Mining distance-based outliers in near linear time with randomization and a simple pruning rule.* Paper presented at the Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining.

Bhaduri, K., Matthews, B. L., & Giannella, C. R. (2011, August 21 - 24). *Algorithms for speeding up distance-based outlier detection.* Paper presented at the Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining.

Biswas, G., Khorasgani, H., Stanje, G., Dubey, A., Deb, S., & Ghoshal, S. (2016, October 3-6). *An application of data driven anomaly identification to spacecraft telemetry data.* Paper presented at the Annual Conference of the Prognostics and Health Management Society 2016, Denver, Colorado.

Boeing. (3-05-2019). Major Production Facilities. *Boeing Commercial Aeroplanes.* Retrieved from https://www.boeing.com/company/about-bca/

Boeing. (2007). Aero Magazine. *Boeing Aero Magazine, qtr_0207.*

Boeing. (2017). *Worldwide Operations| 1959 - 2017.* Retrieved from Seattle, USA: https://www.boeing.com/resources/boeingdotcom/company/about_bca/pdf/statsum.pdf

Budalakoti, S., Srivastava, A. N., & Otey, M. E. (2009). Anomaly detection and diagnosis algorithms for discrete symbol sequences with applications to airline safety. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 39*(1), 101-113.

Callantine, T. J. (2001). *The crew activity tracking system: Leveraging flight data for aiding, training and analysis.* Paper presented at the 20th DASC. 20th Digital Avionics Systems Conference (Cat. No. 01CH37219).

Cavka, I., & Cokorilo, O. (2012). Cost-benefit assessment of aircraft safety. *International Journal for Traffic & Transport Engineering, 2*(4).

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR), 41*(3), 15.

Chidester, T. R. (2003, April 2003). *Understanding normal and atypical operations through analysis of flight data.* Paper presented at the Proceedings of the 12th international symposium on aviation psychology, Dayton, Ohio.

CIAIAC. (2008). *Accident involving a McDonnell Douglas DC-9-82 (MD-82) aircraft, registration EC-HFP, operated by Spanair, at Madrid-Baraja Airport* (Report A-032/2008). Retrieved from Madrid, Spain: https://www.fomento.gob.es/recursos_mfom/2008_032_a_eng_3.pdf

Čokorilo, O., Gvozdenović, S., Vasov, L., & Mirosavljević, P. (2010). Costs of unsafety in aviation. *Technological and economic development of economy, 16*(2), 188-201.

Das, S., Li, L., Srivastava, A., & Hansman, R. J. (2012). *Comparison of algorithms for anomaly detection in flight recorder data of airline operations.* Paper presented at the 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

Das, S., Matthews, B. L., Srivastava, A. N., & Oza, N. C. (2010, July 25 - 28). *Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study.* Paper presented at the Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, Washington DC, USA.

Das, S., Sarkar, S., Ray, A., Srivastava, A., & Simon, D. L. (2013, Jun 17 - 19). *Anomaly detection in flight recorder data: A dynamic data-driven approach.* Paper presented at the American Control Conference (ACC), Washington, United States.

Donald L. Wood, T. T. T. (2018). The Effect of piloting Practices Upon Actual as Opposed to Scheduled Landing Field Performance. *AIAA*(January). doi:10.2514/6.2018-1756

Emmanuell Badea, V., Alin, Z., & Boncea, R. (2018). Big Data in the Aerospace Industry. *Informatica Economica, 22*, 17-24. doi:10.12948/issn14531305/22.1.2018.02

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996, 1996-08-02). *A density-based algorithm for discovering clusters in large spatial databases with noise.* Paper presented at the KDD'96 Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, Oregon.

FAA. (2015). *Safety Management System Manual*. FAA Retrieved from https://www.faa.gov/about/initiatives/sms/reference_library/policy_and_requirements/media/AST_SMS_Manual_V2.0.pdf.

Gautam, C., Balaji, R., Sudharsan, K., Tiwari, A., & Ahuja, K. (2019). Localized Multiple Kernel learning for Anomaly Detection: One-class Classification. *Knowledge-Based Systems, 165*, 241-252.

IATA. (2016). IATA Forecasts Passenger Demand to Double Over 20 Years [Press release]. Retrieved from https://www.iata.org/pressroom/pr/Pages/2016-10-18-02.aspx

ICAO. (2013). *State of Global Aviation Safety*. Retrieved from Montreal, Canada: https://www.icao.int/safety/State%20of%20Global%20Aviation%20Safety/ICAO_SGAS_book_EN_SEPT2013_final_web.pdf

ICAO. (2016). *FDA Limit*. Retrieved from https://www.icao.int/NACC/Documents/Meetings/2016/SEMFDAP/FDAP-P02.pdf

Iverson, D. L. (2004). Inductive system health monitoring.

Jingru Yan, J. H. (2013). Flight Data Monitoring and Human Factors Risks Identification: A Review of Best Practices. Retrieved from uwaterloo.ca website: https://uwaterloo.ca/humans-complex-systems-lab/sites/ca.humans-complex-systems-lab/files/uploads/files/523_yan_etal_aero_2013.pdf

Key Dismukes, L. L. (2004). *The Limits of Expertise: The Misunderstood Role of Pilot Error in Airline Accidents*. Retrieved from https://human-factors.arc.nasa.gov/flightcognition/article2.htm

Lee, C.-H., Shin, H.-S., Tsourdos, A., & Skaf, Z. (2017, 16 - 18 Nov, 2017). *Data analytics development of FDR (Flight Data Recorder) data for airline maintenance operations.* Paper presented at the MFI 2017- International Conference on Multisensor Fusion and Integration for Intelligent Systems, Daegu, Korea.

Lee, S., Hwang, I., & Leiden, K. (2014). *Flight deck human-automation issue detection via intent inference.* Paper presented at the Proceedings of the International Conference on Human-Computer Interaction in Aerospace.

Li, L., Das, S., John Hansman, R., Palacios, R., & Srivastava, A. N. (2015). Analysis of flight data using clustering techniques for detecting abnormal operations. *Journal of Aerospace Information Systems, 12*(9), 587-598.

Li, W.-C. (2011). The Causal Factors of Aviation Accidents Related to Decision Errors in the Cockpit by System Approach. *J. Aeronaut., Astronaut. Aviat, 43*(3), 147-154.

Mack, D. L., Biswas, G., Khorasgani, H., Mylaraswamy, D., & Bharadwaj, R. (2018). Combining expert knowledge and unsupervised learning techniques for anomaly detection in aircraft flight data. *at-Automatisierungstechnik, 66*(4), 291-307.

Maille, N. (2013, 2-9 March 2013). *On the use of data-mining algorithms to improve FOQA tools for airlines.* Paper presented at the Aerospace Conference, 2013 IEEE, Big Sky, Montana.

Maille, N. (2015). *On the use of Flight Operating Procedures for the Analysis of FOQA data*. Paper presented at the 6TH EUROPEAN CONFERENCE FOR

9

AERONAUTICS AND SPACE SCIENCES (EUCASS), Krakow, Poland.

Maille, N. (2017, 4-11 March). *Modelling airline crew activity to improve flight safety analysis.* Paper presented at the 2017 IEEE Aerospace Conference, Big Sky, MT, USA.

Nanduri, A., & Sherry, L. (2016). *Anomaly detection in aircraft data using Recurrent Neural Networks (RNN).* Paper presented at the Integrated Communications Navigation and Surveillance (ICNS), 2016.

Oehling, J., & Barry, D. J. (2019). Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data. *Safety science, 114*, 89-104.

Pohl, T. (2015). Why big data can help to keep planes in the air. Retrieved from https://blogs.sap.com/2015/01/20/why-airlines-need-to-keep-planes-in-the-air/

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2017). Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system. In *Human Error in Aviation* (pp. 73-88): Routledge.

Statler, I. C. (2007). The Aviation System Monitoring and Modeling (ASMM) Project: A Documentation of Its History and Accomplishments: 1999-2005.

BIOGRAPHIES

Vivian R. Igenewari is a PhD researcher in the Integrated Vehicle Health Management (IVHM) at Cranfield University, since November 2017. Her Bachelors degree was in Computer Science from Babcock University, Nigeria in 2004 and her Masters degree was in Advanced Computer Science (Cloud Computing) from the University of Leeds in 2014. Vivian has worked as a Systems Analyst/Administrator and Application Support Engineer/Analyst at several companies including Vanfrank Limited, TEAM Energy and Cognisco Limited. As an Application Support Analyst, Vivian was part of business improvement projects working with querying databases for the purpose of finding useful patterns and developing intelligent reporting tools which provided additional sources of income for the organisations. Her current research interests involve big data analytics, machine learning and anomaly detection techniques for aerospace-related datasets.

Dr Zakwan Skaf is currently a Lecturer in Diagnostics and Prognostics in the Integrated Vehicle Health Management (IVHM) at Cranfield University. Before joining Cranfield University, he held several positions with Loughborough University, Sheffield University, Warwick University and the University of Manchester. Zakwan received his B.S. degree from the faculty of Mechanical Engineering in 2001, and his MSc and PhD degrees from the University of Manchester, Control Systems Centre, Manchester. He is an active researcher with over 35 publications in many world classes peer-reviewed journals and conferences. His research interest involves control engineering, data analytics, machine learning, condition monitoring, diagnostics, and prognostics related to aerospace, railway, marine, energy, and automotive applications.

Professor Ian K. Jennions. Ian's career spans some 40 years, working mostly for a variety of gas turbine companies. He has a Mechanical Engineering degree and a PhD in CFD both from Imperial College, London. He has worked for Rolls-Royce (twice), General Electric and Alstom in a number of technical roles, gaining experience in aerodynamics, heat transfer, fluid systems, mechanical design, combustion, services and IVHM. He moved to Cranfield in July 2008 as Professor and Director of the newly formed IVHM Centre. The Centre is funded by a number of industrial companies, including Boeing, BAE Systems, Thales, Meggitt, MOD, DRS, Alstom Transport and Novartis. He has led the development and growth of the Centre, in research and education, since its inception. Ian is on the editorial board for the International Journal of Condition Monitoring, a Director of the PHM Society, chair of the SAE IVHM Steering Group and contributing member of the HM-1 IVHM committee, a Chartered Engineer and a Fellow of IMechE, RAeS and ASME. He is the editor of five recent SAE books on IVHM and a co-author of the book: 'No Fault Found – The Search for the Root Cause'.