A Study of Human-Factors based Anomalies of Flight Data Recorder's Data of Commercial Aircraft during descent.

Qaiser Mehmood (PhD Student)
Department of Electrical.Electronics and Cyber Physcial Systems
Istanbul Medipol University
Istanbul, Turkey
qaiser.mehmood@std.medipol.edu.tr

***Abstract*—This document is a study of anomalies based on human factors in commercial aviation in terms of pilot's (pilot in command) inputs. We studied data from Flight Data Recorder (FRD) of a real commercial jet, retrieved from NASA website. We experimented using MATLAB tool to study anomalies in human (person in command) inputs in perspective of pilot's suicide scenario. The data was experimented to create a sudden altitude drop scenario like a pilot's suicide and algorithm was tested to identify the anomaly. Our Model successfully identified the anomalies for a potential suicidal human input from a potential suicidal pilot.**

Keywords—Human Factors, Anomaly Detection, Data Science, Avionics Systems, Cyber-Physical Systems

I. Introduction

Since the first commercial flight in 1930s [1], the aviation industry has grown rapidly. According to the Federal Aviation Administration (FAA), in 2017, average daily flights per day in USA alone are more than 24,000 and there are about 2,250,000 people in the air on a given day in USA alone [2]. As the air traffic grows every day, the safety measures and enhancements of the aviation systems also improves. Although these safety measures make air travel as one of the reliable sources of transportation yet, despite all the safety measures during the last few decades, still there have been many accidents in aviation and many people lost their lives [3]. Aviation accidents have been categorized to different categories [4]. One of the categories is related to human factors [3][4].

Human factors-based accidents are the situation when a human being intentionally or unintentionally makes a mistake that triggers the cause of the accident [1]. This person can be a pilot, an Air Traffic Controller (ATC), a maintenance engineer or a fuel truck handler and so on. Most of the human factors-based accidents are caused by the people inside the cockpit during the flight. Some of them were unintentional human mistakes, while others were intentional acts of crime [2]. The accident does not occur until in both cases, unwanted inputs are initialized by the pilots. Unwanted inputs may include a wrong push of a button at wrong time and situation, a wrong maneuver and so on. The human factors that may involve in such inputs may include carelessness, fear, stress, suicidal tendencies, mental health, a hijack situation and so on.

Out of all the aviation accidents, 70 % of the accidents are caused by the human factors [5]. There are different levels of human factors that are described in the literature and are summarized in these studies [5] and [6]. These levels are defined as 'Human Factors Analysis and Classification Systems' (HFACS) [5] [6] [7]. There are four levels of human factors as describes by [5] [6] [7] as Acts, Preconditions, Supervision and Organizational Influence. The first level, named as 'acts', describes different acts done by the inflight humans, including pilots, that has caused different commercial aviation accidents. The deliberate acts done by the authorized persons, such as pilots, are considered as 'Pilot's suicide' [5] [7]. Thus, Pilot's suicide is a human factor that has caused different aviation accidents and it come under the category of 'acts' as classified by HFACS.

There have been different incidents in aviation history that have been categorized as pilots suicided. For example, LAM Mozambique Airline Flight 470, EgyptAir Flight 990, SilkAir Flight 185, Japan Airlines Flight 350 and Germanwings Flight 9525. We shall discuss them in literature review in detail. For this study, we consider Germanwings Flight 9525 as a case study to demonstrate the pilot's input to sudden change in altitude to make the aircraft nose-dive.

Germanwings Flight 9515 accident is one of the most recent cases of pilot's suicide in which more than 150 people died. The cause of the accident was co-pilots suicide when he locked the cockpit door from inside when the captain went out of the cockpit [12]. The recommendations, because of these accidents were regarding improving the procedures, like most of the airlines adopted the procedure of at least 2 crew members in the cockpit all time during the flight, as discussed in the official report [12]. Some 'behavioral recommendations' like this study [8] and other relevant clinical health solutions in perspective of this accidents like [9,10,11,12] have been suggested. Yet there is no such recommendation that suggest an improvement in the system so that such human factors can be automatically handled by the system and the safety and security of flight is insured.

As it is clear from the literature [5] that there cannot be any training, medication or policy solution that can make sure that human factors will not occur again in aviation, therefore there is a need to propose such solution that can intelligently identify a potential wrong input from the pilot and once it is identified, then it can be dealt with further. These wrong inputs can be identified by studying the dataset in perspective of a certain data feature, in this case, a change is altitude throughout the flight, and looking for 'anomalies' in the altitude. In case of a suicidal attempt, the altitude of the aircraft is intentionally decreased abnormally from the suicidal pilot to crash the aircraft. In this study we propose a mechanism to identify such anomalies.

II. Literature Review

So far, to best of our knowledge, FDR's data is analyzed only after an accident or an incident to investigate the cause of the accident. Although, limited access flight data (Flight Operations Quality Assurance FOQA), is being used for different research investigations, for example [13-20], however, publicly available FDR's data [21] has not been studied previously in the perspective of experimentation and investigation of potential causes and avoidance of human factors-based accidents.

However, a most recent study, published in 2022, [51] uses this dataset to investigate anomaly detection in perspective of stable or unstable landing of a commercial jet. They studied this dataset using time series mechanism to identify stability or instability of landing using anomaly detection. They studied anomalies in 'airspeed' and 'altitude' to establish stability or instability of the landing. The data they studied, or 'airspeed' and 'altitude' is based on instrumental inputs. We, on the other hand, need to study data for identification of abnormal inputs by humans (pilots in command) that may not directly include the instruments or devices' inputs.

Other than [51], literature shows that most of the data studied using FOQA dataset is studied for anomaly detection and system's health detection [13-20]. Since FOQA datasets are not publicly available for research and analysis, therefore for the purpose of this study, we shall use publicly available flight data on the NASA website [21].

Flight data is normally recorded in FDR which saves the data of each flight during the flight. This data is retrieved by competent authorities of most of the countries to maintain the FOQA. When the FOQA is maintained, this data can be privately used for experimentation purposes [22]. The most relevant use of FOQA is the use for the anomaly detection and prediction for the flight safety in literature. Anomaly analysis closely matches to our domain because we intend to use the same mechanism for human-factor based anomalies.

Recent literature, that covers anomaly detection, can be studied in two perspectives, one for aviation domain, and the other for non-aviation domains [13]. First attempt to study anomalies from FRD data was done by the Morning Report software package in 2005 [23]. The method was mathematical model using the quadratic equation. It was a simple and innovative method that use simple mathematics technique. Later, data mining techniques were used to detect anomalies in aerospace systems [24-28]. Most of these methods has limitations in perspective of use of labelled data [27], usage of only discrete parameters [24-25], which was improved by [26] to use both discrete and continues parameters. But they made assumption that normal operations always have one type of data, which practically is not the case.

The second perspective of anomaly detection is for the non-aviation domains. In such cases some of the earlier ones are only domain specific methods like [29-30]. Others may include, but not limited to intrusion detection in computers [31-33], fault detection in mechanical systems [34,35], credit card fraud detection [36,37] and so on [38-50]. Most of the above-mentioned literature is relevant to our topic of study but it does not mention anomaly deduction in perspective of human-factors in aviation, especially using the publicly available FDR dataset [21]. As we discussed earlier, the latest published work in June 2022. [51], Ezequiel J. G. et al. worked in anomaly detection of a flight in perspective of stability or instability upon a landing approach of a commercial aircraft.

They selected 2 parameters, altitude, and airspeed. They have set metric for instability as deviation of the measurements of altitude and airspeed while the aircraft is approaching (attempting landing). Any unpredictable changes in altitude and/or deviation is airspeed is defined as the complexity in the dynamics of the aircraft, during landing approach.

This complexity, they calculated, over 15 seconds over a 60 second overlap window, with 45 second of overlap with the previous window. Then standard deviation and mean of each rolling window is also calculated. Therefore, they calculated a total of 3 features, complexity, mean and standard deviation of each reading of a 15 second window sample for a mean estimate time of 75 minutes from take-off to landing. This feature generation is done using NASA dataset of 3000 flights with total of 890000 measurements were takes.

After feature generation, they used k-means clustering to identify states from the distribution of the features. They eventually experimented with the test data to analyze the stable or instable landing approaches of the aircraft.

Our proposed work, however, deviates a bit in a way that we may not be relying on only 2 parameters as they justifiably did for their case (altitude and airspeed). In our case, since we are dealing with human factors, these can occur during any time of flight. Secondly, our final target is to identify and predict a safe or unsafe input from the pilot, during the whole flight, because human factors all rely on the pilot's (person in command) input. For example, and incorrect flap setting during take-off is a wrong input by the pilot, which is an anomaly, and an unsafe input. We want system to be capable of distinguishing all such inputs during all three phases of flight, climb, cruise and descent for landing. Just as they [51] selected 2 parameters from NASA flight recorder dataset, there are about more than 150 other features as well. Some of them are machine inputs and rest are human inputs, machine inputs come from devices as described by [51] about altitude and airspeed. These two parameters are machine/device generated. While flaps setting for example during take-off are set by pilots. Which is a human input. We intend to look for all human inputs among all more than 150 inputs, studying them under dependencies with other related parameters, during a particular phase of flight, climb, cruise, on descent.

However, during this study, due to time limitation, we will study 'pilot's suicide' human factor during the descent of the aircraft. We intend to experiment our approach by using a case study accident scenario and inserting abnormal data that match the accident to analyze the performance and the result of the proposed technique.

The case study we intend to use for this purpose is Germanwings Flight 9515 accident. This accident is a pilot's suicide case in which he locked himself up and rapidly descent the aircraft to the ground. Flight data recorder shows a rapid descent in the altitude of the aircraft. As the accident suggests, this rapid descent was caused by the human input by pushing the nose-down.

We intend to investigate a mechanism to detect anomaly in the decrease of altitude up to a specific thresh-hold, that matches a suicidal scenario only. After anomaly is accurately identified, and no action is taken from any of the pilots to avoid it, we intend to suggest an automatic control of the auto pilot to declare a MAYDAY (extreme emergency call) to the ATC tower and abort the dive and maintain a safe altitude. We focus mainly on detection of anomaly is case of an abnormal drop of altitude and produce a warning. We intend to mimic the exact rapid drop of altitude that is caused by the suicidal pilots as of Germanwings Flight 9515 accident. It is important to consider that in some extreme cases, due to maintenance of air pressure, the pilots must rapidly decrease the altitude legally, so we are aware of such situations while preparing our data and model.

Making autopilot taking control of the aircraft in case of extreme human factors can potentially be done if the autopilot is "Intelligent" enough to "understand" the criticality of the situation. Being able to distinguish between a routine flight input by pilots, and the abnormal risky input by the pilots is the key. This distinguishing mechanism we intend to achieve using machine learning techniques that accurately classify the normal and abnormal inputs from the pilots that can be considered as human factor based.

Machine learning techniques for classification normally need a descent amount of data to be trained for the desired classes to be learned by the system and tested for the accuracy by inserting new data to the classifier. In case of a commercial jet, the data recorded and potentially be capable of usage is the data stored in Flight Data Recorder (FDR), which is the second black-box, other than Cockpit Voice Recorder (CVR).

FDR records all the inputs from pilots to the system, alongside all input to the system provided by all the devices for decision making. For example, airspeed indicator feeds the air speed data continuously throughout the flight into the FDR. So, all the data, with respect to time, starting from the engine start to take-off, cruise, and landing in recorded in FDR. This data is mostly used for anomaly detection as it has been discussed in the previous section, but we intend to configure anomalies related to human factors, so that a system-oriented solution can be proposed to deal with human factors inside the cockpit.

III. Method

The main objective of our methodology is to identify anomaly in the altitude of the aircraft with respect to time, to detect human interference in drop of altitude. However, before doing this, the flight data needs to be prepared to extract altitude parameter in time series relevant to this study.

A. Dataset

We used the same dataset used by [51] that is publicly available on the NASA website [21] and can be downloaded and used. The original dataset contains the data of about 180,000 regional commercial jet flights. Each file of the dataset is a MATLAB file and represents a single flight data. The data is a time series data that shows data of more than 100 parameters, i.e. altitude, airspeed etc.. in contrast to the minutes of flight. Recently on their website they have added cleaned data ready to use for experimentation, some weeks ago at the time of writing of this document at [52].

For this study, we have used the prepared dataset available at [52]. The dataset contains 19 parameters with a time series of 160 seconds. For our experiment, we considered two parameters, altitude (written technically as BARO CORRECT ALTITUDE in figure 1), and time, the first most column in figure 1. The average flight time of a single flight is also about 150-160 minutes. So, while performing our experiments, we assumed the time in minutes rather than seconds. We set out data in a way so that we can compare time with altitude during the descent phase of the aircraft. This is because the altitude is decreasing during the descent, this will help us distinguishing a normal descent and the abnormal descent. In case of a pilot suicide the aircraft rapidly descent, this descent is different than gradual descent. Therefore, studying the anomaly during already in descent phase will give our algorithm a chance to distinguish between a 'gradual descent', which is normal and the 'rapid descent' which is caused by pushing the nose-down a pilot's suicide scenario.

	0	1	2	3
	AILERON POSITION LH	AILERON POSITION LH	CORRECTED ANGLE OFF ATTACK	BARO CORRECT ALTITUDE
0	81.26119	82.652336	-8.111792	1969.6174
1	79.604095	81.0157	-7.6446114	1955.6995
2	81.30211	80.7702	-7.552573	1940.0267
3	82.34547	83.900276	-8.395265	1924.5493
4	81.87493	82.75462	-7.854284	1905.367
5	81.44531	83.28653	-7.0097327	1887.5901
6	80.892944	81.956764	-4.647246	1868.6716
7	81.465775	82.61142	-4.820765	1848.9906
8	80.46333	82.75462	-5.805797	1827.9095
9	83.18425	83.73661	-3.691424	1817.7819
10	76.65814	79.03127	-4.5056286	1800.8936
11	78.928986	76.59677	-7.823943	1787.0592
12	82.01814	79.93143	-6.8519983	1775.7482
13	81.895386	83.12287	-3.1272287	1760.368
14	80.524704	81.11799	-3.8048985	1747.8729
15	81.50669	85.53691	-7.2863693	1736.3134
16	82.652336	84.4731	-7.3579016	1726.7883
17	81.649895	84.04348	-8.170669	1718.6221
18	83.59341	82.447754	-10.417438	1713.8584
19	82.63188	83.85936	-13.02724	1707.0176
20	82.63188	84.12531	-11.419528	1694.7053
21	82.63188	84.18669	-7.959095	1678.0804
22	82.07951	83.450195	-9.489269	1668.9381
23	80.17692	81.465775	-10.560795	1661.23
24	81.30211	81.62944	-8.633039	1650.0541
25	80.46333	82.5705	-10.6125345	1639.2573
26	80.81111	81.93631	-11.74863	1630.3398

Fig .1. Dataset Screenshot

In order to correctly identify the anomaly, our algorithm needs to focus the time at which the rapid descent started. As we prepared dataset for experimentation in terms of time assumed as minutes, therefore it is important to read the descent at a particular time to be an anomaly. For instance, flying at 2000 feet, at 120th minutes of flight, the gradual descent is say 200 meters per minute, but a rapid descent is 1000ft per minutes. So, the algorithm needs to identify the correct 1000 feet at correct time spam as an anomaly because with gradual descent it will eventually reach 1000 ft after some more minutes, and that 1000ft is not an anomaly.

B. Algorithm

The algorithm reads the time and the altitude minute by minute and considers a change in altitude as descent if it goes below a specified altitude threshold. This threshold is measured by taking standard deviation of all 160 values of altitude for 160 minutes. The standard deviation is only measured for altitude, not for time, because the target is to identify the anomaly in altitude but not the time. The threshold of time is selected as anytime less than 10 minutes to land. The algorithm is shown in figure 2.

```
start
read altitude and time in the dataset
    if the altitude < altitude threshold and time < total threshold time
        consider it as an anomaly
        else
        do not consider it an anomaly
```

Since the data shown in figure 1. is the actual data of a normal flight, therefore we inserted some values to mimic the rapid descent like the Germanwings Flight 9515 accident according to the data available in the official accident report of the aircraft [12].

1969.617	1
1955.7	2
1940.027	3
1924.549	4
1905.367	5
1887.59	6
1868.672	7
1848.991	8
1827.91	9
1817.782	10
1800.894	11
1787.059	12
1775.748	13
1760.368	14
1747.873	15
1736.313	16
1726.788	17
1718.622	18
1713.858	19
600	20
500	21
400	22
500	23
600	24
700	25
800	26
900	27
1000	28
1100	29
1200	30
1300	31
1400	32
1545.268	33
1535.448	34
1527.043	35
1516.001	36
1506.975	37

Fig.3. Descent vs Descent start time

The modified values to test the threshold are shown in Figure 3. It is visible in Figure 3 that the descent of the flight started at 1969 ft. Second column represents the time, meaning that, at 1st minute of descent, the aircraft was at 1969.617 ft and so on. We inserted rapid descent values at 20th minute of descent as dropping to 600 ft from 1713.858 ft at 19th minute. This is a sudden drop that can cause by multiple reasons. Since this is well above the threshold value (standard deviation of the entire altitude values of the column) therefore the algorithm will not consider it as an anomaly. Only when it reaches the threshold, then system will consider it as an anomaly. We performed following tests on this data.

C. Test 1. (Altitude drop above threshold)

We set the values more than the altitude to test that whether our algorithm will identify it as anomaly or not. This test was to make sure that the algorithm does not consider intentional legal drop of altitudes by pilots due to certain weather conditions. Also, this includes a drop of altitude because of swear turbulence which is not a human factor, therefore system should not identify a problem.

D. Test 2. (Altitude drop below threshold)

As the suicidal drop of altitude continues to drop, therefore system must understand at some point that this is not a normal descent but it might be a intentional act. The threshold is set to indicate such situation. We tested the system to identify the anomaly if the altitude is dropped below the threshold.

E. Test 3. (Only Rapid descent is considered as anomaly)

The gradual descent of an aircraft, of a normal flight, eventually reaches to zero when the flight lands at zero altitude. As shown in Figure 4. The last 30 minutes of a normal gradual descent.

625.8586	130
617.4502	131
608.9289	132
597.4188	133
587.8344	134
578.4343	135
571.3021	136
559.9043	137
553.566	138
545.1715	139
536.2616	140
528.9753	141
515.9176	142
505.0176	143
492.4743	144
483.6527	145
473.0002	146
462.7638	147
451.4742	148
443.4125	149
432.7402	150
422.8261	151
413.8359	152
403.0644	153
392.5677	154
381.3916	155
369.5183	156
357.7426	157
346.1152	158
335.1151	159
321.7636	160

Fig. 5 Last 30 minutes of descent

At this point, it is also important that when the descent is gradual and the aircraft reaches the threshold value, system should not consider this to be an anomaly because it is not a sudden drop of the altitude. So this is also considered to distinguish gradual drop of altitude below the threshold and rapid drop of altitude below the threshold.

The results of all these tests are discussed in the next section. Since most of the dataset is available in MATLAB, so we used MATLAB tool for the implementation of algorithm. The results are discussed in the next section.

IV. Results

The results show that our algorithm successfully reads the data from dataset and successfully detects the rapid descent as an anomaly as we mimic the similar data as of Germanwings Flight 9515 accident. As discussed in previous section, we performed 3 tests to verify different scenarios and to make sure that the algorithm identifies anomaly only when it is really an anomaly and does not give any false alarms.

A. *Test 1. (Altitude drop above threshold)*

When the test is run for the drop of altitude and the values inserted were kept above the threshold, it was expected from the algorithm that it will not identify and anomaly. Figure 6 is the resulted plot of the graph. And original plot on the left hand side matched the algorithmic plot on the right hand side with no indication of anomaly. Therefore, the algorithm passed this test.

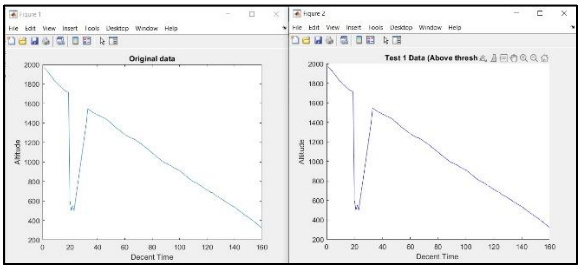

Fig.6. Altitude-drop above threshold

B. *Test 2. (Altitude drop below threshold)*

Secondly, the main purpose of this work is to identify the anomaly when the altitude is dropped more than the threshold. Figure 7 shows the difference between the original data and the algorithmic results. The red coloured dots on the right hand side figure are indicating the anomaly. The green dot represents the threshold. We changed two data points to less than threshold and algorithm successfully identified those.

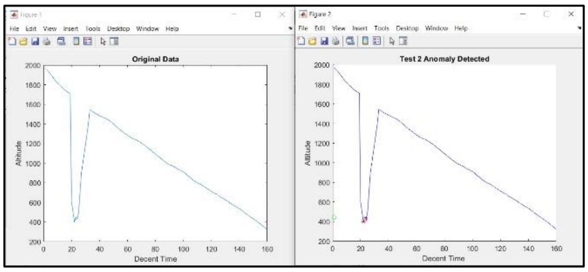

Fig.7. Altitude-drop below threshold

This identification is the rapid drop of altitude, which is not caused by any weather conditions or turbulence. The system identifies it to be a human intentional act and considers it to be an anomaly. When the altitude raises again above the threshold, the system does not show any further anomaly. This also means that if the aircraft is forced to brought back to flight by the other pilot, or, by autopilot which assumingly has taken control after the anomaly is detected by the system, then the system will not give any further warnings post anomaly detection.

C. *Test 3. (Only Rapid descent is considered as anomaly)*

The third point was to handle false indications and avoid algorithm to consider false data as anomaly. The system is required to be intelligent enough to distinguish a normal gradual descent of aircraft, which definitely will cross the threshold before landing, and a rapid abnormal descent below threshold. So we tested the algorithm for false indications.

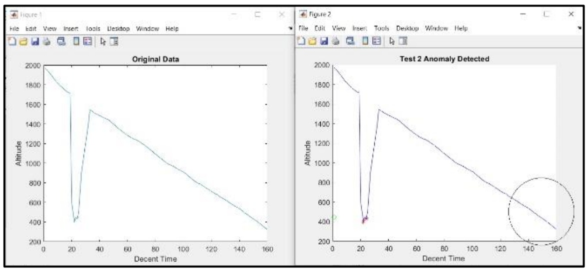

Fig.8. Gradual-drop vs Rapid-drop of altitude

The right-hand side of figure 8 clearly indicates, encircled with black, that the altitude is decreasing below the threshold. But there is no red mark for anomaly. It means that system successfully distinguished between a sudden drop, that indicates a human intervention, and a gradual descent, which is a routine landing.

V. Discussion and Future Work

The time limited nature of this project compelled us to consider only one accident case in specific descending scenario and perform the experiment with the available data. Yet this algorithm can work with any time series data to identify anomaly in case of a sudden altitude drop. The algorithm itself is not much of a data dependent. However, the threshold mechanism is universal for all altitudes of any flight because it is considered in contrast to ‘time to landing’. ‘Time to landing’ can further be identified using total predicted flight time on that particular path on which the aircraft is bound to travel that day. Therefore, this anomaly detection algorithm we proposed and tested in this study can detect abnormal altitude drop anomalies for any flight under any situations.

Due to the time limited nature of this project, we had to limit our experiment to one flight with only landing data and mimicking the Germanwings Flight 9515 accident. However, in future, we intend to investigate potentially all possible human factors that have been playing a role in accidents and incident in recent history of aviation. Also, we studied only one human factor, pilot’s suicide, and the actions a pilot takes to make it possible. Our anomaly detection algorithm puts a check on those potential actions (sudden drop on the altitude) and identifies the anomaly. Once the anomaly is accurately identified, further actions, such as, taking control of the aircraft form human pilot by autopilot and recovering the aircraft from the dive can be considered as future work.

This algorithm can further be tested for 'stall' situation, (especially one similar to the Air France Flight 447 accident, when the aircraft stalled because of human error caused by the co-pilot) by including more parameters such as the reading of the angle of elevation of nose of the aircraft.

Furthermore, it is still needed to be investigated for other potential human factors that have been occurring in past during all phases of flight and prepare a comprehensive auto-turning framework that not only identify such human factors but also take intelligent actions to handle such situations and assist in safely landing the aircrafts.

ACKNOWLEDGMENT

Most of the data used is taken from the NASA website and we really appreciate the efforts of the researchers who provided the data publicly and allowed us to use it. Furthermore, we also want to appreciate National Geographic Channel for the informative documentary on the Germanwings Flight 9515 accident that help us understand the accident in much easier way which eventually helped in this study.

REFERENCES

[1] Anderson, A. J. (2011). Aviation safety: evolution of airplane interiors. *Chicago, US: The Boeing Company*.

[2] Peterson, M. P., Hunt, P., & Weiß, K. (2017). Mapping air population. *Journal of Geovisualization and Spatial Analysis*, *1*(1), 1-16.

[3] Kelly, D., & Efthymiou, M. (2019). An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017. *Journal of safety research*, *69*, 155-165.

[4] Evans, J. K. (2007). An Application of CICCT Accident Categories to Aviation Accidents in 1988-2004.

[5] Kelly, D., & Efthymiou, M. (2019). An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017. *Journal of safety research*, *69*, 155-165.

[6] Lower, M., Magott, J., & Skorupski, J. (2018). A system-theoretic accident model and process with human factors analysis and classification system taxonomy. *Safety science*, *110*, 393-410.

[7] Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2007). Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system. *Human factors*, *49*(2), 227-242.

[8] Molan, G., & Molan, M. (2021). Theoretical model for accident prevention based on root cause analysis with graph theory. *Safety and health at work*, *12*(1), 42-50.

[9] Pasha, T., & Stokes, P. R. (2018). Reflecting on the Germanwings disaster: a systematic review of depression and suicide in commercial airline pilots. *Frontiers in psychiatry*, *9*, 86.

[10] Hardie, J. A., Green, G., Bor, R., & Brennan, P. A. (2021). Cutting edge selection: learning from high reliability organisations for virtual recruitment in surgery during the COVID-19 pandemic. *The Annals of The Royal College of Surgeons of England*, *103*(6), 385-389.

[11] Cahill, J., Cullen, P., Anwer, S., Wilson, S., & Gaynor, K. (2021). Pilot Work Related Stress (WRS), Effects on Wellbeing and Mental Health, and Coping Methods. *The International Journal of Aerospace Psychology*, *31*(2), 87-109.

[12] Germanwings Flight 9515, March 2015, France, https://www.bea.aero/uploads/tx_elydbrapports/BEA2015-0125.en-LR.pdf

[13] 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.

[14] Li, L., Gariel, M., Hansman, R. J., & Palacios, R. (2011, October). Anomaly detection in onboard-recorded flight data using cluster analysis. In *2011 IEEE/AIAA 30th Digital Avionics Systems Conference* (pp. 4A4-1). IEEE.

[15] Stolzer, A. J., & Halford, C. (2007). Data mining methods applied to flight operations quality assurance data: a comparison to standard statistical methods. *Journal of Air Transportation*, *12*(1), 6.

[16] Stolzer, A. J. (2002). Fuel consumption modeling of a transport category aircraft using flight operations quality assurance data: a literature review. *Journal of Air Transportation*, *7*(1), 93.

[17] Sheridan, K., Puranik, T. G., Mangortey, E., Pinon-Fischer, O. J., Kirby, M., & Mavris, D. N. (2020). An application of dbscan clustering for flight anomaly detection during the approach phase. In *AIAA Scitech 2020 Forum* (p. 1851).

[18] Puranik, T. G., & Mavris, D. N. (2018). Anomaly detection in general-aviation operations using energy metrics and flight-data records. *Journal of Aerospace Information Systems*, *15*(1), 22-36.

[19] Das, S., Matthews, B. L., & Lawrence, R. (2011, June). Fleet level anomaly detection of aviation safety data. In *2011 IEEE Conference on Prognostics and Health Management* (pp. 1-10). IEEE.

[20] Sun, Z., Ma, C., Li, W., & Shen, C. (2014). Flight Operations Quality Assurance Based on Clustering Analysis. In *Proceedings of the First Symposium on Aviation Maintenance and Management-Volume II* (pp. 413-422). Springer, Berlin, Heidelberg.

[21] Official NASA website for dataset, Tail 687, https://c3.ndc.nasa.gov/dashlink/projects/85/resources/?type=ds

[22] Beaubien, J. M., & Baker, D. (2003). Toward generalized human factors taxonomy for classifying ASAP incident reports, AQP performance ratings, and FOQA output.

[23] Amidan, B. G., & Ferryman, T. A. (2005, March). Atypical event and typical pattern detection within complex systems. In *2005 IEEE Aerospace Conference* (pp. 3620-3631). IEEE.

[24] Budalakoti, S., Srivastava, A. N., & Otey, M. E. (2008). 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.

[25] Budalakoti, S., Srivastava, A. N., & Akella, R. (2006, March). Discovering atypical flights in sequences of discrete flight parameters. In *2006 IEEE Aerospace Conference* (pp. 1-8). IEEE.

[26] Sriastava, A. (2005). Discovering system health anomalies using data mining techniques.

[27] Schwabacher, M., Oza, N., & Matthews, B. (2009). Unsupervised anomaly detection for liquid-fueled rocket propulsion health monitoring. *Journal of aerospace computing, information, and communication*, *6*(7), 464-482.

[28] Das, S., Matthews, B. L., Srivastava, A. N., & Oza, N. C. (2010, July). 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* (pp. 47-56).

[29] Hodge, V., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial intelligence review*, *22*(2), 85-126.

[30] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys. *vol*, *41*, 15.

[31] Hofmeyr, S. A., Forrest, S., & Somayaji, A. (1998). Intrusion detection using sequences of system calls. *Journal of computer security*, *6*(3), 151-180.

[32] Lee, W., & Stolfo, S. (1998). Data mining approaches for intrusion detection.

[33] Mukherjee, B., Heberlein, L. T., & Levitt, K. N. (1994). Network intrusion detection. *IEEE network*, *8*(3), 26-41.

[34] Bishop, C. M. (1994). Novelty detection and neural network validation. *IEE Proceedings-Vision, Image and Signal processing*, *141*(4), 217-222.

[35] Worden, K., Manson, G., & Fieller, N. R. (2000). Damage detection using outlier analysis. *Journal of Sound and vibration*, *229*(3), 647-667.

[36] Guo, T., & Li, G. Y. (2008, July). Neural data mining for credit card fraud detection. In *2008 International Conference on Machine Learning and Cybernetics* (Vol. 7, pp. 3630-3634). IEEE.

[37] Brause, R., Langsdorf, T., & Hepp, M. (1999, November). Neural data mining for credit card fraud detection. In *Proceedings 11th International Conference on Tools with Artificial Intelligence* (pp. 103-106). IEEE.

[38] Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern recognition*, *38*(11), 1857-1874.

[39] Gaber, M. M., Zaslavsky, A., & Krishnaswamy, S. (2005). Mining data streams: a review. *ACM Sigmod Record*, *34*(2), 18-26.

[40] Gariel, M., Srivastava, A. N., & Feron, E. (2011). Trajectory clustering and an application to airspace monitoring. *IEEE Transactions on Intelligent Transportation Systems*, *12*(4), 1511-1524.

[41] Chan, P. K., & Mahoney, M. V. (2005, November). Modeling multiple time series for anomaly detection. In *Fifth IEEE International Conference on Data Mining (ICDM'05)* (pp. 8-pp). IEEE.

[42] Keogh, E., Lonardi, S., & Ratanamahatana, C. A. (2004, August). Towards parameter-free data mining. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 206-215).

[43] Kumar, M., Patel, N. R., & Woo, J. (2002, July). Clustering seasonality patterns in the presence of errors. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 557-563).

[44] Möller-Levet, C. S., Klawonn, F., Cho, K. H., & Wolkenhauer, O. (2003, August). Fuzzy clustering of short time-series and unevenly distributed sampling points. In *International symposium on intelligent data analysis* (pp. 330-340). Springer, Berlin, Heidelberg.

[45] Au, W. H., Chan, K. C., Wong, A. K., & Wang, Y. (2005). Attribute clustering for grouping, selection, and classification of gene expression data. *IEEE/ACM transactions on computational biology and bioinformatics*, *2*(2), 83-101.

[46] Rousseeuw, P. J., & Leroy, A. M. (2005). *Robust regression and outlier detection*. John wiley & sons.

[47] Fox, A. J. (1972). Outliers in time series. Journal of the Royal Statistical Society: Series B (Methodological), 34(3), 350-363.

[48] Bianco, A. M., Garcia Ben, M., Martinez, E. J., & Yohai, V. J. (2001). Outlier detection in regression models with arima errors using robust estimates. *Journal of Forecasting*, *20*(8), 565-579.

[49] Smyth, P. (1994). Markov monitoring with unknown states. *IEEE Journal on Selected Areas in Communications*, *12*(9), 1600-1612.

[50] Smyth, P. (1996). Clustering sequences with hidden Markov models. *Advances in neural information processing systems*, *9*.

[51] NASA, "DASHlink – Sample Flight Data," [Online], 2022 https://c3.ndc.nasa.gov/dashlink/resources/1018/