Apriori Algorithm for Association Rules Mining in Aircraft Runway Excursions

Natalia Distefano[*], Salvatore Leonardi

Department of Civil Engineering and Architecture, University of Catania, Italy

Abstract Safety is one of the International Civil Aviation Organization's (ICAO) strategic objectives to foster a global civil aviation system. Statistically, almost 40% of aviation accidents occur at airports, the most frequent type is runway excursion. The assessment of accident severity is an essential part of safety assessment methods. In this study, a set of influencing factors which would affect the severity of different types of runway excursions were investigated in order to determine those factors that typically occur together in runway excursion accidents. In order to achieve this aim, a large database was created, which contains information on all the conditions surrounding the runway excursion event, in a period between 2006 and 2016, for a total of 434 runway excursions. Associate Rules method with Apriori algorithm was used. The Apriori algorithm was used separately for each type of runway excursion. The results of this study show that different variables are associated with different types of runway excursions and different categories of severity. The most significant variable for all types of runway excursion is the class of the aircraft. Events with Major and Hazardous severity are associated with small aircraft, while events with catastrophic severity are associated with aircraft of medium-large dimensions. The least significant variable for runway excursion accidents is "Potential causes". The knowledge of the runway excursions severity based on analysis of their causes is essential to priorities safety budgets and safety risk mitigation measures, as required by ICAO regulations.

Keywords Safety Airport, Risk Assessment, Severity, Overrun, Veer-off

1. Introduction

Safety is one of the International Civil Aviation Organization's (ICAO) strategic objectives to foster a global civil aviation system that consistently and uniformly operates at peak efficiency and provides optimum safety, security and sustainability [1].

Because of the complexity of airport system and related operations, the airport runway has proven to be vulnerable and at risk of failure with the consequence that accidents and serious incidents may occur. To ensure that safety risks (including e.g. accidents/incidents) are identified, assessed and appropriately mitigated, aviation stakeholders are required to implement a Safety Management System (SMS). A SMS is a systematic approach to managing safety that is based on the four cornerstones of safety policy and objectives, risk management, assurance, and safety promotion. A SMS is a framework that provides an organization with the adequate tools to ensure that any drift by the organization towards a lower safety performance is prevented. Establishing a risk management mechanism for airports to monitor and improve these risks is the only solution to lower latent risks efficiently and to achieve the goal of airport safety. Safety risk management in practical terms is concerned with hazard and occurrence identification through reporting and data collection, investigation, and subsequent data analysis [1]. In particular, airport surface risk management is concerned with the collection, investigation, and analysis of four main accident/incident types: excursions, incursions/collisions, wildlife strikes, and Foreign Object Damage (FOD).

Statistically, almost 40% of civil aviation accidents occur at airports, the most frequent type is runway excursion with 55% of all runway safety accidents [2]. There are at least two runway excursions each week worldwide. Runway excursions are persistent problems and their numbers have not decreased in more than 20 years. Runway excursions can result in loss of life and/or damage to aircraft, buildings or other items struck by the aircraft. ICAO identifies the Runway excursion among typical examples of safety indicators in the aviation system [1]. As the transformation of the civil aviation operation environment of aircraft and the insight research on runway

safety, runway excursion prevention had gradually become the priority of the worldwide runway safety fields [3].

A runway excursion event occurs when an aircraft on the runway surface departs the end or side of the runway surface during take-off or landing.

They consist of two types of events:

- Veer Off: A runway excursion in which an aircraft departs the side of a runway
- Overrun: A runway excursion in which an aircraft departs.

The definition of risk may be different in research but it always emphasizes the expected value of combining probability and severity. The assessment of accident severity is an essential part of safety assessment methods. Classifying the severity is a prerequisite to establish the safety objectives and the safety requirements that an organization decide to put in place. The safety requirements are then substantiated into mitigation actions – like technical adjustments, innovative features, procedural changes and training programmers – in order to achieve an acceptable level of risk.

In this study, a set of influencing factors which would affect the severity of different types of runway excursions were investigated, in order to determine those combinations of factors that characterize the different severity of runway excursion accidents. In order to achieve this aim, runway excursions data from Aviation Safety Network (ASN) Database was collected and the Association Rules method with Apriori algorithm (AR) has been applied to them.

2. Literature Review

An airport is a complex system, and each facility in the airport is an important component of the system. Any component influences the airport operation to some extent and may lead to aviation accidents if it fails.

Most of past research in aviation safety focuses on the safety of aircraft operation, traffic control system, crew management, aviation safety system of airlines organization and culture, and logistics issues such as apron operation and security check; less attention has been paid to runway risk management.

Although airport surface safety has been addressed in the past, previous research shows major limitations. To date no research exists that analyses surface safety in an integrated manner. Safety mitigation strategies are developed for single occurrence types from the perspective of single aviation stakeholders reflecting only pieces of surface operations and associated failures/errors. Studies of runway incursions, likely due to their more frequent nature, are more numerous in the literature [4-9] than runway excursions. As such, incursions are often considered 'near misses,' or potentially precursors to aviation accidents, but they are not necessarily accidents.

Excursions, while less frequent, more readily represent an unsafe situation. Compared with other kinds of runway unsafety occurrences, there are three significant characteristics of runway excursion occurrences, which include complex influencing factors, high severity of consequences and latent failures hard to detect [3].

Therefore, identifying the risk factors leading to these accidents and creating strategies and undertaking actions to mitigate runway excursions are of great urgency.

Since "risk" is a function of the probability of an event and the severity of the consequences of the event, a valid and reliable measure of the severity of the outcome of a runway excursion is essential for measuring the risk of runway operations. Assessing risk required both specific tools, which need to assign probability values to specific accidents, and models, which are able to estimate consequences of such events. Several accident probability models have been developed in the last decades. Moreover, most of past researches in aviation safety risk management focus on estimate of the probability of occurrence of aircraft accidents; less attention has been paid to estimate of the severity of the consequences of aircraft accidents. Hale [10] examined airport risk evaluation using models based on historical, causal data. In historical models, risks are calculated separately for each type of aircraft using the airport, and accident probabilities are classified into six scenarios, i.e., during landing: veer-off, overrun and undershoot; and during take-off: veer-off, overrun and overshoot.

Kirkland et al. [11] discussed the need for models for evaluating risk at any airport, using available data on past accidents for that purpose. They developed models showing the annual probability of aircraft overruns occurring as a result of aborted landings and take-offs, as well as the distance from the runway end to where the wreckage is located.

Valdés et al. [12] proposed a risk model for runway overrun and landing undershoot using probability analysis as their technical support. They determined whether or not risk levels at a given airport were acceptable. For that purpose, they used historical data on accidents in the vicinity of the runway. Also, the Airport Cooperative Research Program (ACRP) has produced two studies on runway excursions [13,14], primarily applying traditional logistic regression to predict the likelihood of a runway excursion occurring.

Other studies identify the causes of Runway Excursion in order to define mitigation interventions, for example, [15] make a multivariate analysis of historical data on accidents of runways excursion in order to quantify the effect that various factors have upon runway excursions. Okafor et al. [16] show that environment casual factors are most significant contribution to runway excursion accidents when compared to system and human induced casual factors. For which, environmentally induced accident mitigation should focus on efficient monitoring

of the weather condition in order to make decision such as delay or cancellation of flight. Distefano and Leonardi [17] found that the most common cause of runway excursion on take-off is aircraft system faults, while on landing is weather conditions.

From the literature review above, it can be seen that there are still relatively few studies on the assessment of the severity of aircraft accidents.

Severity of the Aircraft accidents is defined in terms of their effect on both the aircraft and on persons. Safety risk severity is defined by [1] as the possible consequences of an unsafe event or condition. Besides the financial and economic impact, an incident or accident may cause several other subjective consequences such as human well-being impacts, environmental impacts, political implications, reputation of the players involved and media interest.

Many studies have addressed this issue of severity assessment, but they all tend to make it in a general and simple way. Wong et al. [18] evaluate the probability of a catastrophic consequence to happen during a landing event. In their work, a non-catastrophic event is defined as the one with small chances of causing hull loss and injuries to its occupants, and, under this assumption, four categories of obstacles were characterized on the maximum speed that an aircraft may collide on it still causing a non-catastrophic event.

In the risk analysis study that supported the Norwegian Civil Aviation Authority to define its requirements for physical design of aerodrome [19], the severity of an unsafe event consequence was generally determined by the type of event in analysis. For example, the overrun is assumed to be a catastrophic event or the deviation from the runway onto the graded area of the strip results in comparatively minor consequences.

Even the International Civil Aviation Organization (ICAO) on its Safety Management Manual [1], has a not much clear guideline on categorizing severity. On this manual, for assessing severity, there is a table with the five levels of severity of an unsafe condition and their respective possible consequences (Table 1), but there is no recommended methodology on how to connect these consequences with the unsafe event itself, i.e. an event that caused to multiple deaths must be considered a Catastrophic event, but it is not described how to determine if an event will cause multiple deaths.

As to the consequences of runway excursion occurrences, they have the feature of high severity of consequences [20]. Compared with other kinds of runway unsafe occurrences, such as runway incursion occurrences, runway excursion occurrences always busted out in a sudden with high serious consequences. For this reason, only the runway excursions with Major, Hazardous and Catastrophic consequences were considered in this study.

Table 1. ICAO's severity classification [1]

Severity of occurrence	Meaning
Catastrophic	• Equipment destroyed • Multiple deaths
Hazardous	• A large reduction in safety margins, physical distress or a workload such that the operators cannot be relied upon to perform their tasks accurately or completely • Serious injury • Major equipment damage
Major	• A significant reduction in safety margins, a reduction in the ability of the operators to cope with adverse operating conditions as a result of increase in workload, or as a result of conditions impairing their efficiency • Serious incident • Injury to persons
Minor	• Nuisance • Operating limitations • Use of emergency procedures • Minor incidents
Negligible	• Little consequences

3. Data

With runway safety management practices, runway safety management agencies/organizations in field had gathered numerous runway safety data, and gained management experiences as well. But almost all of these data and experiences focused on runway incursion, only a small part of them were related to runway excursion prevention.

For the purpose of this study, runway excursions data from Aviation Safety Network (ASN) Database was collected. The ASN Safety Database contains detailed descriptions of some 20,300 incidents, hijackings and accidents to airliner, military transport category aircraft and corporate jet aircraft safety occurrences since 1921. Most of the information are from official sources (civil aviation authorities and safety boards), including aircraft production lists, ICAO ADREPs, and country's accident investigation boards.

Previous works have already been conducted for safety analysis based on historical aircraft accidents data. Das et al. [21] published results on anomaly detection based on NASA records, known as the Distributed National FOQA Archive (DNFA). This archive contains many continuous and discrete data from various on-board systems (propulsion systems, landing gears, cockpit switch positions, etc.), yet they do not offer a comprehensive view of the context in which aircraft evolve. Sherry et al. [22] also presented risk assessment analyses based on surveillance track data provided by the FAA National

Offload Program.

In this study the data used contains solely runway excursion accidents (overrun and veer-off), in a period between 2006 and 2016, for all categories of aircraft, and in all world regions. This period was considered to be sufficient to obtain statistically relevant results.

The database created for this analysis includes 434 runway excursions that occurred in the eleven-year period. Military flights were not considered in this study. Runway excursions occurred most often during the landing phase (354 events) with a slightly lower division for landing overruns (154 events) respect to veer-offs (200 events). Take-off runway excursions (80 events) present a slightly higher number for overruns (43) respect veer-offs (37). All events were classified according to severity.

Figure 1 shows the percentage distribution of these events by severity and flight phase. Most runway excursions occur during landing, but the events with higher severity are those that occurred during take-off.

The original database was arranged at 8 categorical variables: Year, Accident type, Airport's country, Accident severity, Potential cause, Aircraft class, Runway code, Aircraft age.

Runway code refers to the runway where the accident occurred and it corresponds to runway code ICAO defined in Annex 14 [23]. It has two 'elements', the first is a numeric code based on the Reference Field Length for which there are four categories and the second is letter code based on a combination of aircraft wingspan.

Table 2 shows the distribution of the 8 categorical variables of the resulting database.
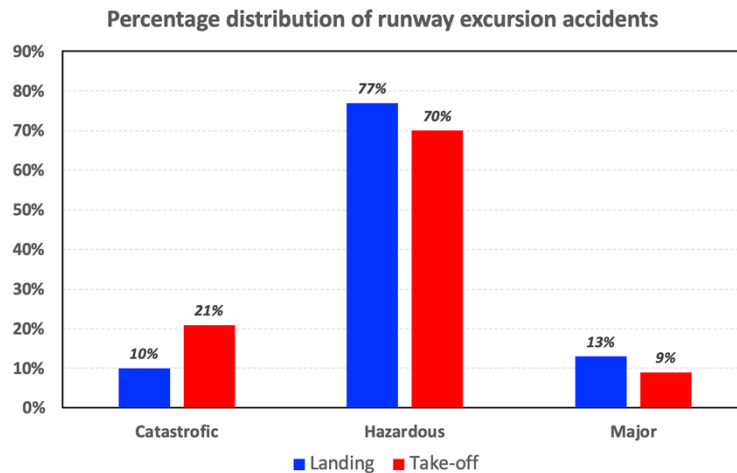
Figure 1. Distribution of runway excursion accidents by severity and flight phase

Table 2. Descriptive statistics of accident data

	N	%
Accident severity		
Catastrofic	53	11,42
Hazardous	328	70,69
Major	53	11,42
Accident year		
2006	42	9,05
2007	51	10,99
2008	35	7,54
2009	33	7,11
2010	41	8,84
2011	31	6,68
2012	39	8,41
2013	41	8,84
2014	39	8,41
2015	37	7,97
2016	45	9,70

Table 2 Continued

2017	30	6,47
Aircraft class		
GA: General Aviation	80	17,24
CA : Corporate Aircraft	112	24,14
Com A: Commuter Aircraft	104	22,41
S-R: Short-range	96	20,69
M-R: Medium-range	28	6,03
L-R: Long-range	14	3,02
Potential cause		
Aircraft System Faults: Engines, brake (wheel brakes, spoilers or reversers), hydraulic, electric, main gear, tire, other.	104	22,41
Human error: Incorrect flight planning, communication/coordination, pilot error, visual illusion, excessive speed, loss of control, other.	123	26,51
Weather Conditions: Low visibility, rain, wind shear, tailwind, ice, crosswind, low ceiling, strong wind, turbulence, freezing rain, other.	67	14,44
Runway Conditions: Wet, contaminated (standing water, rubber, oil, ice, slush, snow), FOD, wildlife hazards, down slope.	32	6,90
Unknown	108	24,88
Accident type		
Landing veer-off (LDVO)	200	46,08
Landing overrun (LDOR)	154	35,48
Take-off veer-off (TOVO)	37	8,53
Take-off overrun (TOOR)	43	9,91
Airport's country		
NAM: North American	117	26,96
EUR: Europe	52	11,98
LATAM: Latin America & the Caribbean	76	17,51
AFI: Africa	45	10,37
ASPAC: Asia Pacific	79	18,20
CIS: Commonwealth of Independent States	24	5,53
MENA: Middle East & North Africa	41	9,45
Runway code		
1A-B	30	6,91
2B-C	51	11,75
3B	6	1,38
3C	80	18,43
3D	7	1,61
4C	46	10,60
4D	172	39,63
4E	42	9,68
Aircraft age		
<=10	104	23,96
11-20	90	20,74
21-30	134	30,88
31-40	69	15,90
>40	37	8,53

4. Methodology

This study uses the Association Rule with Apriori algorithm to attempt to find associations that exist between the Accident severity variable and the other database variables for each runway excursion type. The goal of Association Rule analysis is to investigate a group of items that typically occur together in a given event.

The variables used to achieve the aim of this study are shown in table 3. Since all the countries where the accidents occurred are part of the ICAO and therefore they have similar regulations and that the relative distribution of database events is strongly linked to the number of annual movements that occur in them (data that we do not know), it was decided not to take the Airport's country variable into account for the current analysis.

Table 3.   Items of the variables for Association Rule

AT-Accident Types	RC-Runway Code
AT1=Catastrophic	RC1=1B-1A
AT2= Hazardous	RC2=2B-2C
AT3=Major	RC3=3B
AC-Aircraft Class	RC4=3C
AC1=GA	RC5=3D
AC2=CA	RC6=4C
AC3=Com A	RC7=4D
AC4=S-R	RC8=4E
AC5=M-R	
AC6=L-R	
PC-Potential Cause	**AA-Aircraft Age (years)**
PC1=Aircraft system faults	AA1=<10
PC2=Human error	AA2= 11-20
PC3=Weather conditions	AA3= 21-30
PC4=Runway conditions	AA4=31-40
PC5=Unknown	AA5=>40

The Association Rule is a suitable technique to discover interesting relations between variables in large databases.

Apriori algorithm AR is one of the most popular data mining techniques, having been first introduced in 1993 for discovering buying patterns [24]. In recent years, the AR method in data mining has been successfully applied to uncover potential patterns or rules in a variety of fields, such as road traffic safety [25-27]. AR analysis is the method of effectively identifying sets of items that occur together in a given event. It is based on the relative frequency of the number of times the sets of items occur alone and jointly in a database. AR is a standard approach that starts with a dataset containing transactions and aims to construct frequent item sets by setting up user specified thresholds, namely support, confidence, and lift.

The Support for a particular association rule A ⇒ B is the proportion of transactions in the database containing both A and B and is formulated as equation (1):

$$Support(A \rightarrow B) = \frac{P(A \cap B)}{N} \qquad (1)$$

Where P(A∩B) is the number of transactions containing both A and B, and N is the total number of transactions.

The confidence of the association rule A ⇒ B is a measure of the accuracy of the rule, which is determined by the percentage of transactions in the database containing A that also contains B and is defined as equation (2):

$$Confidence\ (A \rightarrow B) = \frac{P(A \cap B)}{P(A)} \qquad (2)$$

Where P(A∩B) is the number of transactions containing both A and B, and P(A) is the number of transactions containing A.

Lift is defined as a simple correlation that measures if A and B are independent or dependent and correlated events and is expressed by the equation (3):

$$\boldsymbol{Lift}\ (\boldsymbol{A} \rightarrow \boldsymbol{B}) = \frac{P(A \cup B)}{P(A)P(B)} \qquad (3)$$

Where P(A∪B) is the number of transactions containing A or B, P(A) is the number of transactions containing A, and P(B) is the number of transactions containing B.

If a particular rule has a lift of one, it indicates that the probabilities of A and B are independent. When two events are independent, there is no rule drawn involving these two events. In contrast, if a particular rule has a lift greater than one, it indicates that A and B are dependent and positively correlated. The higher the lift, the greater is the strength of the association rule.

It is desirable for the rules to have a high level of support, a large confidence, and a lift value considerably greater than one. Since we have interest also in rare accident characteristics (such as catastrophic accidents), the support for some rules of interest could be much lower than the support typically used in other applications, such as the market basket analysis. Furthermore, to ensure that the patterns identified by the rules are observed with reasonable frequency and that the rules are sufficiently accurate, minimum thresholds for support and confidence are also needed.

To identify strong associations, threshold values for support (S), confidence (C), and lift (L) were set as follows: S ⩾ 5%, C ⩾ 20%, and L ⩾ 1.

Analyses were performed using the software SPSS Modeler.

5. Results

Association rule analysis was further applied to investigate the combinations of factors that typically occur together in the runway excursion accidents. The Apriori

algorithm was used separately for each type of runway excursion. The association algorithm identified 822 total rules with support greater than 5%, confidence greater than 20%, and lift greater than 1 (110 rules for LDVO, 94 rules for LDOR, 152 rules for TOVO, and 466 rules for TOOR). Among these rules, only the top ten with greater confidence value in each accident type were selected. The result of the selected AR was the accident severity (consequent), thus providing statistical evidence that different accident severities of various runway excursion types are dependent on different contributory factors.

In Tables 4–7, the 2-item, 3-item, and 4-item rules are reported, along with their support, confidence, and lift values. In each table, the rules are ranked according to the confidence value. For each table the 10 most significant rules have been reported but in such a way that all three consequent types (accident severity) appear.

5.1. Landing veer-off

As regards LDOR, 79% of them have a hazardous severity (AT2), 14% of them have major severity (AT3) and 7% of them have catastrophic severity (AT1). Among the 10 most significant rules, those that have AT2 as consequence are 5, which have high Confidence and Support values but low Lift values (Table 4). A high Confidence value (C = 95%) and a high Support value (10%) characterize the association between Hazardous event and aircraft to Short-range and aircraft aged between 21 and 30 years (3-item rule). Minor values of Confidence (93.75%) and of Support (8%) characterize the association between Hazardous event and event caused by weather conditions and aircraft between 21 and 30 years (3-item rule). The only rule that associates the AT2 consequence with the Runway code (1A or 1B) is the No. 3.

The most significant association rule for events with consequence Major (AT3) is that with Corporate Aircraft and 4D Runway code (3-items rule), this rule has low values of Confidence and Support, but a high value of Lift (2.55%). Another association rules for Major severity include the association with the 4D Runway code and events caused by unknown factors.

The only association rule concerning catastrophic severity (AT1) is characterized by low Confidence and Support values, but a high value of Lift (3.06%). This is a 2-Items rule and associates the AT1 consequence with the Aircraft to Medium-range.

Table 4. Association Rules for LDVO

ID Rule	Consequent	Antecedent	S(%)	C(%)	L
1	A T 2	AC4 and AA3	10.0	95.0	1.203
2	A T 2	PC3 and AA3	8.0	93.75	1.187
3	A T 2	CR1	7.0	92.857	1.175
4	A T 2	AC2 and AA3	6.0	91.667	1.160
5	A T 2	PC3	20.5	90.244	1.142
6	A T 3	AC2 and RC7	7.0	35.714	2.551
7	A T 3	PC5 and RC7	9.0	27.77	1.984
8	A T 3	AA1 and RC7	10.5	23.809	1.701
9	A T 3	PC1 and AA3	10.5	23.809	1.701
10	A T 1	AC5	7.0	21.429	3.061

Table 5. Association Rules for LDOR

ID Rule	Consequent	Antecedent	S(%)	C(%)	L
1	A T 2	AA4 and AC2	6.494	100	1.351
2	A T 2	AA4 and PC5	7.143	90.909	1.228
3	A T 2	AC2 and RC7	6.494	90	1.216
4	A T 2	AC1 and CR4	5.844	88.889	1.201
5	A T 2	AC1 and PC2 or PC1	5.844	88.889	1.201
6	A T 2	PC3	10.39	87.5	1.182
7	A T 1	AC3 and AA3	9.091	35.714	2.5
8	A T 3	CR2 and PC5	5.844	33.333	2.852
9	A T 3	AC5	7.143	27.273	2.33
10	A T 3	AA1 and PC5	7.143	27.273	2.33

5.2. Landing overrun

74% of LDOR accidents have a hazardous severity (AT2), 12% of them have major severity (AT3) and 14% of them have catastrophic severity (AT1). When landing, the percentage of catastrophic events is greater for overruns than for veer-offs. Among the 10 most significant rules, those that have AT2 as consequence are 6, which have high Confidence and Support values but low Lift values (Table 5). A high Confidence value (C = 100%) characterizes the association between Hazardous event and Corporate aircraft and aircraft aged between 31 and 40 years (3-item rule). Minor values of Confidence (90.90%) but major value of Support (7.14%) characterize the association between Hazardous event and event caused by unknown factors and aircraft between 31 and 40 years (3-item rule). The association between Hazardous consequence and 4D Runway code and Corporate aircraft presents Confidence equal to 90% and Support equal to 6.5%.

The only association rule concerning catastrophic events is characterized by low Confidence and Support values, but a high value of Lift (2.5%). This is a 3-Items rule and associates the Catastrophic event with the Commuter Aircraft and aircraft aged between 21 and 30 years.

The most significant association rule for events with consequence Major is that with 2B or 2C Runway code and unknown cause, this rule has low values of Confidence and Support, but a high value of Lift (2.85%). The other association rules for Major events include the association with the Aircraft to Medium-range, and the association with unknown cause and aircraft with less than 10 years.

5.3. Take-off veer-off

On the TOVO accidents, 76% of them have a hazardous severity (AT2), 2% of them have major severity (AT3) and 22% of them have catastrophic severity (AT1). Among the 10 most significant rules, those that have AT2 as consequence are 6, which have high Confidence (100%) and Support values but low Lift values (Table 6). A high Support value (S = 13.5%) characterizes the association between Hazardous event and Corporate aircraft and aircraft aged between 31 and 40 years (3-item rule). Another rule with the same parameter values is the one that associates Hazardous event with Corporate aircraft and 4D runway code. Minor value of Support (8.10%) characterizes the association between Hazardous event and event caused by unknown factors and General Aviation or Corporate aircraft or aircraft between 31 and 40 years (3-item rules).

The two association rules concerning catastrophic events are characterized by High Confidence, Support and Lift values. The first is a 4-Items rule and it associates the Catastrophic event with the Aircraft to short-range, Human error and 4D runway code.

The most significant association rules for events with consequence Major have low values of Confidence and Support, but a high value of Lift (18.5%). These associate Major events with 2B or 2C Runway code and aircraft older than 40 years or event caused by weather conditions (3-items rules).

Table 6. Association Rules for TOVO

ID Rule	Consequent	Antecedent	S(%)	C(%)	L
1	A T 2	AA4 and AC2	13.514	100	1.321
2	A T 2	AC2 and RC7	13.514	100	1.321
3	A T 2	AC1 and PC5	8.108	100	1.321
4	A T 2	RC2 and PC2	8.108	100	1.321
5	A T 2	PC5 and AA4	8.108	100	1.321
6	A T 2	PC5 and AC2	8.108	100	1.321
7	A T 1	AC4 and PC2 and RC7	10.811	75	3.469
8	A T 1	PC2 and RC7	13.514	60	2.775
9	A T 3	AA5 and RC2	5.405	50	18.5
10	A T 3	PC3 and RC2	5.405	50	18.5

Table 7. Association Rules for TOOR

ID Rule	Consequent	Antecedent	S(%)	C(%)	L
1	A T 2	AC4 and RC7	9.302	100	1.536
2	A T 2	AC4 and RC3	9.302	100	1.536
3	A T 2	AA4 and AC2	9.302	100	1.536
4	A T 2	AA4 and CR7	9.302	100	1.536
5	A T 2	RC4 and AC3	6.977	100	1.536
6	A T 3	AA2 and AC2 and RC3	6.977	66,667	4.778
7	A T 3	RC2 and AC2 and AC5	6.977	66,667	4.778
8	A T 3	RC2 and AC2 and RC3	6.977	66,667	4.778
9	A T 1	AA1 and AC2 and RC3	6.977	66,667	3.185
10	A T 1	AA1 and AC2 and AC5	6.977	66,667	3.185

5.4. Take-off overrun

As regards TOOR accidents, 65% of them have a hazardous severity (AT2), 14% have major severity (AT3) and 21% have catastrophic severity (AT1). Indeed, the percentage of catastrophic events is greater for take-off than for landing. All rules with AT2 as consequence have high Confidence (100%) and Support values but low Lift values (Table 7). Hazardous event is associated with the Aircraft to short-range and 3B Runway code or 4D Runway code (3-items rules). The other rules associate AT2 consequence with aircraft aged between 31 and 40 years and Corporate aircraft or 4D Runway code (3-items rules). The last rule for AT2 (with a minor Support value) associates this at 3C Runway code and Commuter Aircraft.

The most significant associative rules for events with consequence Major have high values of Confidence, Support and Lift. The main rule associates Major severity with aircraft aged between 11 and 20 years and Corporate aircraft and 3B Runway code (4-items rules).

Also, the association rules concerning catastrophic events are characterized by High Confidence, Support and Lift values. Both are 4-Items rules and associate the Catastrophic event with aircraft with less than 10 years, Corporate aircraft and 3B runway code or Aircraft to Medium-range.

6. Discussion and Conclusions

The safety of an airport and in particular the runways and taxiways (i.e. manoeuvring area) is a cause of great concern.

It is impossible to predict when and where a runway excursion will occur because the factors that contribute to an event will be in the hundreds and extremely varied. However, it is possible to identify the factors that have the largest influence for these events.

The present study makes it possible to determine these factors, which will be achieved through the construction of a large database, which contains information on all the conditions surrounding the runway excursion event. Association rules method then allowed to enable the factors that appear most often and have the largest effects to become apparent. The contributing factors were defined by type of incident and by type of consequence severity.

Comparing the association rules related to the 4 types of runway excursion considered in this study, it is possible to draw interesting results. Figure 2 shows the associative rules for each severity of consequences represented by appropriate symbols. From the observation of this chart, it is possible to associate specific aspects for three categories of severity considered.

Figure 2. Runway excursion Association Rules Chart

With regard to the runway excursion with Major severity of consequence, it is possible to note that the events that occur at take-off are always associated with small runways (2B-C or 3B runway code), while runway veer-offs during landing are associated with larger runway (4D runway code). Take-off overruns are strongly associated with small-sized aircraft (Corporate aircrafts). The age of the aircraft is not decisive for any type of runway excursion in any of the two phases, moreover no potential cause is decisive for any type of accident.

Association rules for runway excursion with Hazardous severity of consequence show that the majority of the accidents is associated with small aircraft, and the landing veer-off is associated with aircraft older than 20 years, while other types of accidents are associated with aircraft older than 30 years. The dimensions of the runway are mainly relevant for take-off accidents, and different causes are associated with the various types of accidents with the exception of the take-off overruns, for which in no rule an item belonging to this category appears.

Finally, the association rules for Catastrophic severity of consequence allow us to affirm that the dimensions of

the runway are associated with the runway excursion that occurred during take-off and not with those in the landing phase, moreover classes of smaller aircraft are associated with take-off overruns, while larger aircraft are associated with the other types of runway excursions. Only take-off veer-offs are associated with a potential cause (human error), and the age of the aircraft is not decisive for any type of runway excursion.

The results of this study show that different variables are associated with different types of runway excursions and different categories of severity. According to the discovered association rules, the most significant variable for all types of runway excursion is the class of the aircraft, events with Major and Hazardous severity are associated with small aircraft, while events with catastrophic severity are associated with aircraft of medium-large dimensions. The least significant variable for runway excursion accidents is "Potential causes". This can mean that the causes of the accident do not play an important role in defining the severity of the consequences.

These findings suggest that the occurrence of runway

excursions is a complex phenomenon that involves complex interactions between various factors.

Therefore, the development and implementation of effective safety risk mitigation strategies, in particular to prevent the most severe of these occurrences at controlled airports, are essential. The knowledge of the runway excursions severity based on analysis of their causes is essential to prioritize safety budgets and safety risk mitigation measures, as required by ICAO regulations. Association Rules was helpful in identifying the most important combinations of accidents-contributory factors and can address the design of the safety countermeasures.

REFERENCES

[1] ICAO. Safety Management Manual - 3rd Edition - Doc 9859 AN /460. Montreal, Quebec, Canada, 2013.

[2] IATA - Runway Safety Accident Analysis Report 2010-2014. 1st Edition - International Air Transport Association. Montreal—Geneva, 2015.

[3] Z. Xie , D. Li. Research on Integrated System for Civil Aviation Runway Excursion Prevention. Proceedings of a meeting 3rd International Conference on Management Science and Management Innovation (MSMI 2016). 13-14 August 2016, Guilin, China, 2016.

[4] D. McLean, I. Monro. Runway Incursions and Prevention Systems. Aircraft Engineering and Aerospace Technology, Vol. 76, No. 5: 472–478, 2004. DOI: https://doi.org/10.1108/00022660410555149.

[5] S. Young, J. Vlek. An Analysis of the Causes of Airfield Incursions Attributed to Ground Vehicles. Journal of Airport Management, Vol. 3, No. 3: 299–308, 2009.

[6] Z. Yan, Y. Y. Haimes. Cross-Classified Hierarchical Bayesian Models for Risk- Based Analysis of Complex Systems under Sparse Data. Reliability Engineering and System Safety, Vol. 95, No. 7: 764–776, 2010. http://dx.doi.org/10.1016/j.ress.2010.02.014.

[7] D. Kim, H. Yang. Evaluation of the Risk Frequency for Hazards of Runway Incursion in Korea. Journal of Air Transport Management, Vol. 23, No. 1: 31–35, 2012. DOI: http://dx.doi.org/10.1016/j.jairtraman.2012.01.011.

[8] E. C. Rogerson, J. H. Lambert, A. F. Johns. Runway Safety Program Evaluation with Uncertainties of Benefits and Costs. Journal of Risk Research, Vol. 16, No. 5: 523–539, 2012. http://dx.doi.org/10.1080/13669877.2012.725674.

[9] J. Schonefeld, D. P. F. Moller. Runway Incursion Prevention Systems: A Review of Runway Incursion Avoidance and Alerting System Approaches. Progress in Aerospace Sciences, Vol. 51, No. 1: 31–49. 2012. doi:10.1016/j.paerosci.2012.02.002.

[10] A. Hale. Risk contours and risk management criteria for safety at major airports, with particular reference to the case of Schiphol. Safety Science, Vol. 40, No. 1, 299-323, 2002. DOI: 10.1016/S0925-7535(01)00051-0.

[11] I.D.L. Kirkland, R.E. Caves, I.M. Humphreys, D.E. Pitfield. An improved methodology for assessing risk in aircraft operations at airports, applied to runway overruns. Safety Science, Vol. 42, No. 10: 891-905, 2004. http://dx.doi.org/10.1016/j.ssci.2004.04.002

[12] R.M.A. Valdés, F.G. Comendador, L.M. Gordún, F.J. Sáez Nieto. The development of probabilistic models to estimate accident risk (due to runway overrun and landing undershoot) applicable to the design and construction of runway safety areas. Safety Science, Vol. 49, No. 5: 633e650, 2011. DOI: 10.1016/j.ssci.2010.09.020.

[13] ACRP (Airport Cooperative Research Program). Analysis of Aircraft Overruns and Undershoots for Runway Safety Areas. ACRP Report 3. Washington, DC: ACRP, Transportation Research Board, 2008.

[14] ACRP (Airport Cooperative Research Program). Improved Models for Risk Assessment of Runway Safety Areas. ACRP Report 50. Washington, DC: ACRP, Transportation Research Board, 2011.

[15] N. Distefano, S. Leonardi. Investigation of the causes of runway excursions. Proceedings of the AIIT International Congress on Transport Infrastructure and Systems, TIS 2017, 127-134, 2017. DOI: 10.1201/9781315281896-19.

[16] E.G. Okafor, P. O. Jemitola, M. A. Soladoye. Assessment of runway excursion causal factors and mitigation strategies. Nigerian Journal of Technology (NIJOTECH), Vol. 37, No. 3: 619 – 625, 2018.

[17] N. Distefano, S. Leonardi. Aircraft runway excursion features: a multiple correspondence analysis. Aircraft Engineering and Aerospace Technology, Vol. 91, No. 1: 197-203, 2018. DOI: 10.1108/AEAT-11-2017-0244.

[18] D. Wong, M. Jr. Ayres, H. Shirazi, R. Carvalho, J. Hall, R. Speir, E. Arambula, R. David, J. Gadzinski, R. Caves, D. Pirfield. Modelling the location and consequences of aircraft accidents. Safety Science, Vol. 51, No. 1: 178–186, 2013. DOI: 10.1016/j.ssci.2012.05.012.

[19] M. Eddowes, J. Hancox, A. MacInnes. Final Report on the Risk Analysis in Support of Aerodrome Design Rules - A Report produced for Norwegian Civil Aviation Authority. United Kingdom: AEA Technology, 2001.

[20] Z. Hu. Statistical Analysis on Runway Excursion of Transport Aircraft in China. Journal of Safety Science and Technology, pp.129, 2012.

[21] S. Das, B.L. Matthews, A.N. Srivastava, N.C. Oza. Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study. 16th SIGKDD International Conference on Knowledge discovery and data mining. ACM, pp. 47–56, 2010.

[22] L. Sherry, Z. Wang, H.K. Kourdali, J. Shortle. Big data analysis of irregular operations: aborted approaches and their underlying factors. Integrated Communications, Navigation and Surveillance Conference (ICNS), IEEE, H5–1–10, 2013.

[23] ICAO. Annex 14 - Aerodrome Design and Operations - Eighth Edition, July 2018

[24] R. Agrawal, T. Imielinski, A. N. Swami. Mining Association Rules between Sets of Items in Large Databases. Proc.,

ACM SIGMOD International Conference on Management of Data, 1993. https://10.1145/170036.170072.

[25] A. Montella. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. Accident Analysis and Prevention, Vol. 43, No. 4: 1451–1463, 2011.https://doi.org/10.1016/j.aap.2011.02.023.

[26] P. Wu, X. Meng, L. Song, W. Zuo. Crash Risk Evaluation and Crash Severity Pattern Analysis for Different Types of Urban Junctions: Fault Tree Analysis and Association Rules Approaches. Transportation Research Record 1–14, 2019. DOI: 10.1177/0361198118822817.

[27] G. Prati, M. De Angelis, V. Marìn Puchades, F. Fraboni, L. Pietrantoni. Characteristics of cyclist crashes in Italy using latent class analysis and association rule mining. PLoS ONE, Vol. 12, No. 2: e0171484, 2017. https://doi.org/10.1371/journal.pone.0171484.