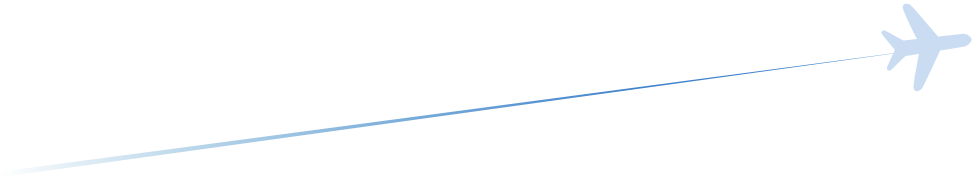

D3.1 Risk framework: scope and SoA

Deliverable ID:	**D3.1**
Dissemination Level:	**PU**
Project Acronym:	**SafeOPS**
Grant:	**892919**
Call:	**H2020-SESAR-2019-2**
Topic:	**SESAR-ER4-06-2019**
Consortium Coordinator:	**TUM**
Edition date:	**24 December 2021**
Edition:	**00.02.00**
Template Edition:	**02.00.02**

Founding Members
EUROPEAN UNION EUROCONTROL
SESAR JOINT UNDERTAKING

Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
Elizabeth Humm	Project Team Member	01/08/21
Carlo Abate	WP3 Leader	22/10/21

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
Lukas Beller	Project coordinator	26/10/21
Josef Hartl	Project Team Member	

Approved for submission to the SJU By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
Lukas Beller	Project coordinator	
Carlo Abate	WP3 Leader	

Rejected By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
-	-	-

Document History

Edition	Date	Status	Author	Justification
00.00.01	02/08/21	Draft	Elizabeth Humm	Contents review
00.00.02	14/10/21	Draft	Elizabeth Humm	Collaborators review
00.00.03	22/10/21	Draft	Carlo Abate	Revision of Sections 2 – 5
00.00.04	26/10/21	Draft	Carlo Abate	Overall revision after internal review
00.01.00	28/10/21	Initial Release	SafeOPS Consortium	
00.01.01	21/12/21	1st Revision	Carlo Abate Elizabeth Humm	Update with changes as requested by SJU
00.02.00	24/12/21	Second Issue Version	SafeOPS Consortium	

SafeOPS

D3.1 – RISK FRAMEWORK: SCOPE AND SOA

This Deliverable is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 892919 under European Union's Horizon 2020 research and innovation programme.

Abstract

The SafeOPS project aims at investigating the impact of possible artificial-intelligence-based decision-support systems on routine air-traffic operations. The context selected for this investigation is the missed approach initiated by the flight crew of a landing flight and the subsequent go-around. The go-around scenario has a number of uncertainties and therefore makes it an ideal candidate for the integration of a predictive technology to support air traffic controllers (ATCOs) in managing aircraft in this situation.

One aspect of the incorporation of a predictive technology in the air traffic operating environment is the risk associated with the technology integration, management and use. Therefore Work Package 3 of the SafeOPS project is assigned to the investigation of this risk, structured as a 'Risk Framework'. The Risk Framework presented by SafeOPS will analyse the impact of the technology and the information presented to the ATCOs. A first process step in the Risk Framework will be developing a risk model. The risk model will be used to determine the resulting risks of each specific operational context under assessment.

This document addresses the initial phase of the process in the compilation of the Risk Framework, namely a systematic review of current risk models available for application in an aviation context. The review aims to provide a critical assessment of existing risk models and their suitability for use in the SafeOPS Risk Framework. The review will also make a recommendation on the most appropriate model for use in the SafeOPS context.

The review will select risk models for review which are widely used and validated in an aviation environment, inclusive of models applied to air traffic control or aircraft operations. Each model will be described in terms of the theoretical basis of model, the method used to apply the model, the technique used in analysis of the risk and the output of the model. A set of acceptance criteria will be defined according to the requirements of a model for use in SafeOPS. Each model will thus be evaluated against this set of criteria and adherence with the criteria will then be used for cross comparison between models. The model that meets the most criteria will be recommended for use in the Risk Framework. If any criteria are not met by any model, then the consequences of this will be assessed and reported. For the recommended model, strength, weaknesses and limitations will also be presented.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	
A/C	*Aircraft*
AI	*Artificial Intelligence*
AIM	*Accident Incident Model*
ALA	*Approach and Landing Accidents*
ANSP	*Air Navigation Service Provider*
AP	*Airport*
ASCOS	*Aviation Safety and Certification of New Operations and Systems*
ATC	*Air Traffic Control*
ATCO	*Air Traffic Controller*
ATM	*Air Traffic Management*
B	
BBN	*Bayesian Belief Nets*
C	
CFIT	*Controlled Flight Into Terrain*
CW	*Calendar Week*
D	
DBL	*Deep Blue SrL*
DFS	*Deutsche Flugsicherung*
E	
ESD	*Event Sequence Diagrams*
F	
FC	*Flight Crew*
FRAM	*Functional Risk Analysis Method*
FT	*Fault Tree*
G	
GA	*Go-Around*
I	
IMC	*Instrument Meteorological Condition*
INX	*Innaxis*
L	
LDG	*Landing*
LOC-I	*Loss of Control in Flight*
M	
MA	*Missed Approach*
MAC	*Mid-Air Collision*
MAP	*Missed Approach Procedure*
MARIA	*Model of ATM Reality in Action*
N	
NOTAM	*Notice(s) to Airmen*
R	
RWY	*runway*
S	
SID	*Standard Instrument Departure*
STAMP	*Systems-Theoretic Accident Modelling and Processes*
T	
T/O	*Take-Off*
TOPAZ	*Traffic Organization and Perturbation AnalyZer*
TUM	*Technical University of Munich*
U	
UA	*Unstabilized Approach*
W	
WP	*Work Package*

1 Introduction

1.1 Project SafeOPS

SafeOPS investigates the impact of possible artificial-intelligence (AI) based decision-support systems on routine air-traffic operations. The scenario selected in SafeOPS for this investigation is the *missed approach* of a landing aircraft and the subsequent *go-around*[1]. The go-around scenario has a number of uncertainties and safety critical factors associated with it. It is therefore an ideal candidate for studying the integration of a predictive technology, with the aim of providing greater support to Air Traffic Controllers (ATCOs) in managing aircraft.

1.2 Go-Arounds

According to a recent study by the Flight Safety Foundation [1], go-arounds occur with an average rate of 1-3 per 1000 approaches. This and similar studies raise awareness on the risks associated with unstable approaches and about the possible benefits for safety of encouraging crews to go-around more often [1]–[4]. However, evidence shows that only 5-10% of all unstable approaches ended in a go-around. The remaining 95% of unstable approaches that continued to land were subject to an increased risk for runway excursion [1]. Even though performing a go-around is encouraged in the unstable case, one in ten go-around reports still shows potentially unsafe outcomes, including problems with fuel endurance and exceeding performance limits.

1.3 Potential Consequences of Missed Approach

From the safety perspective, deciding to go-around is the ideal solution when approaches are unstable. ATCOs are informed by the pilots after they have initiated a go-around procedure or in some cases earlier, when they communicate the intention to perform this manoeuvre and before this decision becomes apparent from the change in flight trajectory on the radar screen. Although airlines publish specific procedures on when to perform a go-around under unstable approach conditions, the final decision lies with the cockpit crew and the captain responsible for the flight. The decision to continue a slightly unstable approach to achieve an overall higher level of safety from the cockpit perspective is justifiable in some situations for example due to the fuel shortage on board or when adverse weather conditions are predicted to deteriorate.

From the Air Traffic Management (ATM) perspective, go-arounds are well established but can be complex to manage in some cases, e.g. when a go-around occurs close to the preceding aircraft which has just taking-off, or when two aircraft go-around consecutively. Additionally, go-around procedures are not always followed strictly by the pilot in certain situations and even the ATCO could, in situations of high traffic congestions, provide instructions which differ from the authorised go-around procedure.

[1] When, for any reason, it is judged that an approach cannot be continued to a successful landing, a missed approach or go-around is flown. A go-around occurs when an aircrew makes the decision not to continue an approach, or not to continue a landing, and follows procedures to conduct another approach or to divert to another airport [27].

As a result, more complex situations could arise which could potentially lead to separation infringements or causing the following aircraft to also perform a go-around. The unpredictability of go-around scenarios and the potential consequences therefore represent an overall uncertainty for ATCOs.

1.4 Theoretical basis of this study

One of the objectives of SafeOPS is to implement a data-driven tool capable of predicting the occurrence of a go around based on real-time flight and weather parameters. The idea is that a timely forecast of a highly probable go-around, combined with a confidence interval for that prediction, could be used by ATCOs in their decision-making process [5]. For example they could choose against giving clearance authorisation to another aircraft to take-off in front of the potentially unstable aircraft, and thus avoid challenging coordination efforts in the case of a go-around.

The initial activities of SafeOPS consisted in a series of workshops with ATCOs dedicated to identifying the circumstances in which an unexpected go-around increases the complexity of the traffic situation and consequently the workload for the Tower Control. SafeOPS deliverable 2.1 [5] reports on the outcomes of these workshops. In particular, it defines a set of Scenarios and Use Cases in which probabilistic predictions of upcoming go-arounds can give the ATCOs more time to plan their subsequent actions. Timely predictions would thus reduce the time pressure and workload of the ATCOs, thus bringing increased safety and resilience through higher predictability of the overall system. In the upcoming months, SafeOPS will investigate the performance of such predictive tool and how it can be designed to maximise the accuracy and reliability of its forecasts. An additional aspect that is crucial to understand is how this tool modifies the level of safety of the system. In particular, it is fundamental to investigate how the ATCOs interact with the tool, how they reply on and exploit the information it generates, and what are the consequences for the overall systems of an unexpected outcome, e.g. an aircraft not performing a go-around when it was predicted to do so. All these aspects are to be analysed to have a complete picture of the impact of such a tool on the safety of the ATM system.

1.5 Purpose of using a risk model approach

One aspect, of the incorporation of a predictive technology in the air traffic operating environment, is the risk associated with the technology insertion, management and use. Therefore it is critical to assess and manage this risk. Work Package 3 (WP3) of the SafeOPS project is assigned to the investigation of this risk, structured as a 'Risk Framework'. The Risk Framework to be developed by SafeOPS aims at analysing the impact of the technology on the safety of the current system.

In general, a risk assessment is made up of many, potentially complex, parts with connections and dependencies in between. There are ways of assessing these individual components, but in reality they are not strictly separate parts and so should be represented more like the system that they originate from. For this purpose, *models* have been developed to better represent an approximation of reality, which encompass the interactions and relationships between components in the system. Moreover, the increasing complexity of systems, and the high criticality of safety in industries like Nuclear, Aviation and the medical arena, has driven the need of representing these systems in simpler terms for the purpose of safety analysis.

Risk assessment processes and risk frameworks usually take on a standardised approach or have considerable components in from a standardised approach. According to the international standard for Risk Management, namely ISO 31010, the main stages of the risk assessment process are [6]:

- RISK IDENTIFICATION
- RISK ANALYSIS
- RISK EVALUATION

The following paragraphs detail the tasks that can be conducted in the process of a risk assessment. The stages of risk identification, analysis and evaluation should make up the core structure of any risk assessment. However the tasks (listed below) are required according to the scope, context and criteria of the risk assessment and so not all tasks will necessarily be included in the risk assessment. Alongside each task definition there are examples of structured techniques that can be used to assist in undertaking each component task:

Risk Identification

Risk identification enables the articulation of the uncertainties in the system. This can be carried out through the collection of information which can take the form of accident/incident analysis, past data, observations or expert opinion. Indeed, this stage makes significant use of the knowledge and experience of a variety of stakeholders. It can involve the following activities:

a) Eliciting views from stakeholders and experts – This ensures that the risk assessment is valid and applicable. Furthermore, it brings together different areas of knowledge and expertise, which can contribute to identifying and understanding risk. Structured techniques for the purpose of this task include brainstorming, the Delphi technique, the Nominal group technique, structured or semi-structured interviews and surveys;

b) Identifying Risk – A methodical and iterative technique of recording all uncertainty in the system. Examples are checklists/classifications/taxonomies, HAZOPS, FMEA/FMECA, Scenario Analysis, and Structured ‘What-If’ Technique (SWIFT);

c) Determining causes, sources and drivers of risk – This allows the estimation of the likelihood of an event or consequence and helps to identify treatments that will modify risk. Examples are the Cindynic approach and the Ishikawa analysis (fishbone) method.

Risk Analysis

Analysing the identified risks gives them meaning and context and helps articulate them in quality or quantity. It can also highlight how each risk is connected to the larger system of risk and safety by examining interactions and dependencies among the different components of the system. It should be noted that some of the techniques used for risk identification also encompass the risk analysis component; therefore, there is some overlap in term of the structured techniques that can be used. The stage of risk analysis can involve the following activities:

a) Investigating The Effectiveness Of Existing Controls - Risk is affected by the overall effectiveness of any controls that are in place. Controls can change the likelihood, consequences or both, of the risk. Examples are Bowtie Analysis, Hazard Analysis and Critical Control Points (HACCP) and Layers of Protection Analysis (LOPA);

b) Understanding consequences and likelihood – Consequence is the impact of a risk source and event. It can be positive or negative. Likelihood is the chance of something happening. Examples are Bayesian analysis, Bayesian networks and influence diagrams, Business impact analysis, Cause-consequence analysis, Event tree analysis, Fault tree analysis, Human reliability analysis (HRA), Markov Analysis and Monte Carlo simulation;

c) Analysing interactions and dependencies - There can be many interactions and dependencies between risks. For example, multiple consequences can arise from a single cause or a particular consequence might have multiple causes. Examples are Causal mapping and Cross impact analysis;

d) Providing a measure of risk – It can be useful to provide a measure of risk which usually takes the form of a combination of the exposure to a certain threat, the potential consequences and the likelihood of those consequences. This can be done through qualitative, semi-quantitative or quantitative measures. Structured techniques that can be used for this tend to be specific to certain scenarios e.g. The Toxicological risk assessment, Value at risk (VaR) and Conditional value at risk (CVaR) or expected shortfall (ES), the latter two being related to financial loss.

Risk Evaluation

Risk evaluation involves deciding whether and how to treat the risks that have been identified in the previous steps. This could be in terms of its tolerability or acceptability, the relative significance of a risk or the ranking of risks against each other in priority. It can involve the following activities:

a) Significance of risk - The outputs from risk identification and analysis can be used to draw conclusions about whether the risk should be accepted and the comparative significance of the risk relative to the objectives and performance thresholds of the organization. Structured techniques that can be used for this include 'As low as reasonably practicable' (ALARP) and 'so far as is reasonably practicable' (SFAIRP), Frequency-number (F-N) diagrams, Pareto charts, Reliability centred maintenance (RCM) and Risk indices.

b) Selecting between options – This involves weighing the potential advantages and disadvantages of each option in terms of risks, to inform decision making. Examples of structured techniques are Cost/benefit analysis (CBA), Decision tree analysis, Game theory and Multi-criteria analysis (MCA).

1.6 Context within SafeOPS

Within WP3 of SafeOPS, the Risk Framework task covers the majority of the process described in ISO 31010. Selecting the most appropriate Risk Model to use will also take into consideration the advice provided in the ISO to ensure that total coverage is guaranteed in any risk assessment conducted in SafeOPS. WP3 of SafeOPS comprises the following tasks:

Task 3.1: Benchmarking of existing risk models

Existing risk models are investigated, considering (but not exclusively) models currently used in ATM safety management system and previous research carried out in SESAR and other European research projects. The user requirements from previous SafeOPS work packages will serve as valuable inputs to this task. The outcome will be a portfolio of models encompassing best practices, which are summarised in the present document.

Task 3.2: Development of an integrated risk model which is focussed on the technological solution and its operational deployment

This task will pick up the recommended risk model resulting from task 3.1, develop it and integrate into it the existence of the predictive technological solution and all aspects related to its operational deployment within Air Traffic Control (ATC). Choosing the model to employ will require the definition of criteria in terms of suitability and efficiency: in fact, the goal is to develop a model for the evaluation of risk, which achieves an optimal balance between thoroughness and effort required for its usage. A third aspect to consider will be the "integrability" of the risk model in the existing Safety Management System within the end-users organisation, e.g. Air Navigation Service Provider (ANSP). Both, selection of the model and its development will require the involvement of safety experts in order to evaluate the fitness to both the technology solution and the environment in which it will operate. Furthermore, a close connection to the development of the predictive AI algorithms will be important, as these algorithms will provide one major input to the risk framework.

Task 3.3: Analysis of the Human Factors impact of real time risk information provision

This task will assess the risks concerning the provision of real time information to end users. It will focus on assessing possible dysfunctional interactions between humans, tools and procedures by looking at aspects like reliability, missed/nuisance alert ratio and meaningfulness of visual feedback. A starting point for this task will be a narrow-focused review of the HF studies already performed in this area.

This document addresses the initial phase of the process (Task 3.1), namely a systematic review of currently available risk models. This review aims to provide a critical assessment of risk models and their suitability for use in the SafeOPS Risk Framework. The review will conclude with a recommendation on the most appropriate model for use in the SafeOPS project.

1.7 Structure of this Review

The review will select risk models for review which are widely used and validated in an aviation environment, inclusive of models applied to air traffic control or aircraft operations. Each model will be described in terms of the theoretical basis of model, the method used to apply the model, the technique used in analysis of the risk and the output of the model.

Section 2 presents the methodology and acceptance criteria adopted to select the risk model for use in SafeOPS. Section 3 provides a basic description of a set of risk models currently in use in the aviation domain. The fundamental strengths, weaknesses and limitations of each model are also summarised. Section 4 provides the results of the analysis of these models against the acceptance criteria and identifies the recommended model that will be at the core of the SafeOPS Risk Framework, while Section 5 concludes.

2 Methods

2.1 Selecting risk models

The idea of assessing risk has a long history which, according to some authors, dates back to the ancient Athenians [7]. While the fundamental ideas and principles on how to assess risk are old, the systematic study of risk emerged in the 20th century with the growth of industrialization to understand why and how certain incidents and accidents occur. Different domains developed a variety of methodologies to assess and manage risk, which resulted in an ever-growing wealth of models and techniques to identify, qualitatively describe and, in some cases, quantitatively evaluate the risk of a system or an operation. In particular, the aviation industry has always been a leader in this field of research and the use of quantified fault trees and event trees is well established and deployed to support new equipment design, safety assessment and incident investigation.

The most extensive catalogue of the existing tools and methods for the analysis of risk, to the best of our knowledge, is the Safety Method Database [8], which includes more than 800 entries from various domains. It is beyond the scope of this deliverable to provide a critical review of all of them. The aim of the present deliverable is to describe in some detail risk models that can be adapted and used within SafeOPS to quantify the impact that an eventual data-driven, predictive tool has on the workflow of ATCOs in the situation of having to manage a potential pilot-induced missed approach.

In order to down-select a manageable number of risk models to review in this deliverable, three guidelines were drawn up. It was judged that these guidelines would focus the review on realistically applicable models, whilst maintaining a scope that encompasses different risk philosophies. The guidelines were as follows:

1. The present review focuses on systemic risk models, not the individual structured techniques used for conducting component parts of the risk assessment and referenced in 1.5.

2. The models have been adopted in aviation and are well established and are presently in use in this environment.

3. The model is sufficiently flexible to be adapted to the scope of the SafeOPS project and the necessary expertise to modify and use the model is available to the Consortium.

The first guideline is justified by the fact that there is a requirement in the SafeOPS project to assess the risk, of integrating a new technology, across all aspects of the ATM system that the technology has an interaction with. This is because a novel technology would potentially encourage a change in the way an ATCO undertakes the ATM role, and thus will have an impact on many aspects of the socio-technical and operating environments. This can be illustrated by use of the ICAO SHELL model (Figure 1) in which the interactions within a human-centric system are many and varied and a change in one component of the system, will have an impact on many or all of the other components in the system, largely by means of the human-component [9]. Thus assessing risk by using individual structured techniques would not necessarily cover all relevant system components and would not capture the relationships, dependencies and interactions between the individual system components.

S – Software: the rules procedures written documents etc. which are part of the standard operating procedures.

H – Hardware: the Air Traffic Control suites their configuration controls and surfaces displays and functional systems.

E – Environment: the situation in which the Liveware-Hardware-Software system must function, the social and economic climate as well as the natural environment.

L – Liveware: the human beings - the controller with other controllers, flight crews, engineers and maintenance personnel management and administration people - within in the system.

Figure 1. Components of the ICAO SHELL model (source: [9], [10]).

The second guideline originates from the fact that the aviation industry has a long and robust history of safety management through risk assessment and analysis. It is therefore diligent and beneficial to exploit this knowledge and expertise, partly because this would make any model used in SafeOPS more applicable and familiar to the ultimate 'users' of the model e.g. the operatives, but also because it would be more aligned and support integration with the existing safety models used in ATM. Moreover, an aviation specific risk model will consider aspects of the system that are wholly unique to aviation, which might be absence in models specific to, for example the nuclear or rail industries.

The third guideline is necessary due to the nature of risk models and their complexity. Some risks models represent intellectual property in some organisation and consequently are not available. Some models have very intricate and complicated calculations and algorithms involved, the understanding of which might not be easily accessed by a secondary user (e.g. not the originating author/data scientist). There can be associated software for some risk models that is not readily available or in some cases there is an extensive need to understand the idiosyncrasies of applying rules and assumptions within models, or conducting the risk analysis. Lastly, some models are simply not mature enough to be applied in a functional manner and remain in a theoretical stage in terms of their design. All these factors restrict the use of some models in the SafeOPS project.

Using the Safety Method Database as a starting point, we scanned the literature on risk modelling and assessment and applied this set of rules to the results of the search. This process enabled us to trim the list down to seven risk models, which are presented in Section 3. It should be noted, however, that this selection comes to the price of excluding from the present review a variety of widely-used methodologies and tools with many application in aviation. For example, techniques focusing on the human performance within a system, such as Hierarchical Task Analysis [11] and Human Reliability Analysis [12], were not considered. Similarly, individual descriptive tools which are only one of the many components of a risk model were not reviewed, for example the bow-tie, fault-tree, or event-tree diagrams, which enable to visualise the threats, hazards, consequences and mitigation actions related to an event. While these techniques and tools are not presented in Section 3 of the present review, some of them might nevertheless be used throughout the project as building blocks of the SafeOPS risk framework.

2.2 Acceptance criteria for assessing the selected risk models

Acceptance criteria for selecting a single, appropriate model to use in SafeOPS were developed from the overarching project objectives for the integration and use of the technological solution. These cover aspects of the solution itself, but also those required from its integration into the operational system and within high level Human Factors requirements.

The criteria have been developed within the context of the Scenarios and Use Cases defined in Task 2.1 of SafeOPS. The workshops with the ATCOs contributed to set out the requirements that need to be met for a complete description of the risk picture. Specifically, D2.1 [5] clarifies that the model needs to capture not only the events immediately related with the go-around itself (e.g. the dynamics of the interactions between different actors, the verbal and non-verbal means of communications, the instruments used, etc.) but also a wider space of related events that include the possible chain of consequences potentially following from an action, e.g. a certain instruction or decision might lead the onset of a separation infringement, a runway incursion, or in principle more dramatic situations such as mid-air collision or a controlled flight into terrain. To fully capture the risk involved in the use of the predictive tool proposed by, and developed within, SafeOPS, the risk framework has therefore to include all these aspects. Also, the workshops showed that different ways of using the predictive tool can be expected depending on how the probabilistic information about the go around is presented to the ATCOs. Consequently, an additional requirement to be met by the model is the ability to assess the risk associated with the characteristic of the tool's design, and how this design impacts on the interpretability of the conveyed information and thus the likelihood of a human error or of a misuse of the tool. In addition to the requirements emerged from the workshops with the ATCOs, other criteria were set out to meet the scope of the project, for example the ability to quantitatively measure the changes in probability of adverse event to occur when the predictive tool is in use, the suitability of the model for use in an ATM context, and the possibility to adopt the tool both during real time operations and for desktop safety analyses.

The considerations summarised in the previous paragraph led to the definition of twelve acceptance criteria. The model that meets the most criteria will be recommended for use in the Risk Framework. If any criteria are not met by any model, then the consequences of this will be assessed and reported. The selection criteria are as follows:

1) The model shall be able to capture all aspects of the operating environment and allow for the examination of the predictive technology within that.

2) The model shall be able to assess how the technology changes the level of extant risk when it is being used as a real-time prediction tool.

3) The model shall be able to assess how the technology changes the level of extant risk when it is being used for offline provision of analytics.

4) The model shall be able to integrate into the current safety management system of the organisation of use.

5) The model shall be able to capture any risk associated with misleading, incorrect or untimely presentation of predictive data from the technology.

6) The model shall be able to capture any risk associated with the display design and usability.

7) The model shall be able to capture any risk associated with human error related to information processing, decision making and acting on the predictive data.

8) The model shall be able to capture any risk associated with violations in the use of the technology e.g. inappropriate use (misuse) or rejection of use (disuse).

9) The model shall consider risk in all scenarios related to a standard and problematic approach.

10) The model shall consider all risk in the context of the potential source of harm (the resultant hazard) e.g. Controlled Flight into terrain, airborne conflict, loss of control.

11) The model should have the capacity to quantitatively articulate the likelihood and consequences of risk associated with the integration of the technology.

12) The model should be validated in an ATM environment.

3 Review

3.1 Risk Models Considered in this Review

As described in section 2.1, seven models where preselected to be evaluated by the criteria defined in section 2.2. The following list presents the seven models, including a brief summary:

1. Accident Incident Model (AIM) is a risk model which provides a set of individual models (one for each accident type) that all SESAR Solutions shall use to identify where and how an operational change brought by a specific solution will impact on the safety of ATM/ANS provision. AIM is a risk model that shows the risks of aviation accidents and provides a structured breakdown of the causes, with particular emphasis on ATM/ANS contributions (both positive and negative).

2. Functional Risk Analysis Method (FRAM) is a model which is aimed at describing the dynamic and non-linear nature of the interactions in a system. This model is quite unique in the sense that it doesn't consider accidents in a sequential manner, or as connecting latent conditions as more tradition models do.

3. Causal Model for Air Transportation Safety (CATS) is a causal model for air transport safety. Its purpose is to establish in quantitative terms the risks of air transport. The design calls for the combining of three modelling techniques into a single model, these techniques are Event Sequence Diagrams (ESD), Fault Trees (FT) and Bayesian Belief Nets (BBN).

4. Aviation Safety and Certification of New Operations and Systems (ASCOS) is based on the previous accident model development work carried out to create CATS however the focus in ASCOS is more towards certification than risk assessment.

5. Model of ATM Reality in Action (MARIA) is a model describing the whole ATM system and the interdependencies between its functions, for the purpose of safety analysis.

6. Systems-Theoretic Accident Modelling and Processes (STAMP) is an accident *causality* model in which accidents are examined in terms of why the controls that were in place, did not prevent or detect the hazard(s) and why these controls were not adequate in enforcing the system safety constraints.

7. Traffic Organization and Perturbation AnalyZer (TOPAZ) is a quantitative and qualitative risk model utilizing an agent-based model, Monte Carlo simulation and bias/uncertainty analysis.

3.2 Risk Models Descriptions

In the following, each model under review is explained in further detail, describing its theoretical basis, how it is applied and the analyses used within the model.

3.2.1 Accident-Incident Model (AIM)

3.2.1.1 Theoretical basis of AIM

The Accident-Incident Model (AIM) is a top down, barrier-based quantitative model designed to capture the increase or decrease of risk introduced by a change in an Air Traffic Management (ATM) system, or part of it. This change could include a new technology *solution* or operating procedure. The AIM model provides individual templates for a number of accident types which can be used as a basis for identifying where and how the change to the system will impact on the safety being achieved in the existing system. The use of this model requires identification of the parts in the relevant templates that would be impacted on by any change to the system. From this it is possible to explore how safety could be increased, decreased or remain the same with the addition of a new solution.
The latest release of AIM, dated 2020, includes the following templates [13]:

- Mid Air Collision Risk in En Route
- Mid Air Collision Risk in Terminal Manouvering Area
- Mid Air Collision in Final Approach
- Mid Air Collision in Initial Departure
- Mid Air Collision Risk in Oceanic Environment
- Runway Collision Risk
- Runway Excursion
- Taxiway Collision Risk
- Controlled Flight Into Terrain
- Wake Induced Risk on Final Approach
- Wake Induced Risk in Initial Departures
- Wake Induced Risk in En Route
- Mid Air Collision in En Route with RPAS
- Mid Air Collision in TMA with RPAS
- Mid Air Collision in U-SPACE VLL
- Airspace Excursion for Drones
- Airspace Infringements

3.2.1.2 Method of applying AIM

The AIM model is incorporated into a larger safety assurance framework as defined in the SESAR Safety Reference Material (SRM) [14]. Within this AIM is used as a mean to define the safety criteria for a *solution* and also it can also be used to support the identification of the operational services related to a specific change. AIM has also been used to set up the 'Severity Classification Schemes'[2] and 'Risk Classification Schemes'[3].

With regards to setting safety criteria for a solution, this task is required in a 'Safety and Change Assessment' in order to define what is considered tolerably safe for the change being introduced. As

[2] Severity of the effects from each operational hazard.

[3] Classification of ATM safety occurrences according to the severity of their effect.

such, safety criteria are 'explicit and verifiable criteria that must be satisfied in order to ensure tolerable safety is achieved following the change'.

AIM supports the task of defining safety criteria in the following ways:

1. Identify pre-existing Hazards: relevant pre-existing hazards for the Solution are identified in order to select afterwards the relevant AIM models

2. Safety impact assessment of the operational change: the safety impact of the Solution is qualitatively assessed at the level of the identified AIM model (at the barrier, contributors and precursors level). This impact is then quantified as a % of improvement / reduction in barriers performances or contributors and thus on precursors occurrences.

3. Setting the Safety Assurance Criteria: once the safety impact of the change is understood, the Safety Assurance Criteria, defining the tolerable safety following the change, are to be set at the level of the corresponding precursors in the models, indicating the expected benefit, neutrality or degradation, and considering the potential traffic increase.

3.2.1.3 Analysis used in AIM

The AIM risk model templates are barrier-based models based on fault trees, event trees and influence diagrams. As such the terminology used is recognisable from these technical tools, namely; accident risk, precursors, barriers, base events, and induced events. The main components are combined to have and their main elements are shown in Figure 2.

The barriers are represented in the figure with green background, the events (or precursors) are depicted as yellow bubbles and the grey cells are the contextual factors. Each barrier can fail in multiple ways, represented by the adjacent Fault Trees. Each purple cell of the tree is the so-called 'base event', while black cells can be further developed into base events. Each precursor has a probability of occurrence, and each barrier has a failure rate. The failure rate of each barrier is determined by two aspects: the event contribution percentage and the event failure rate. Since the AIM follows the Boolean logic with AND and OR gates, the failure rates of the base events are easy to multiply or add, depending on the gate type.

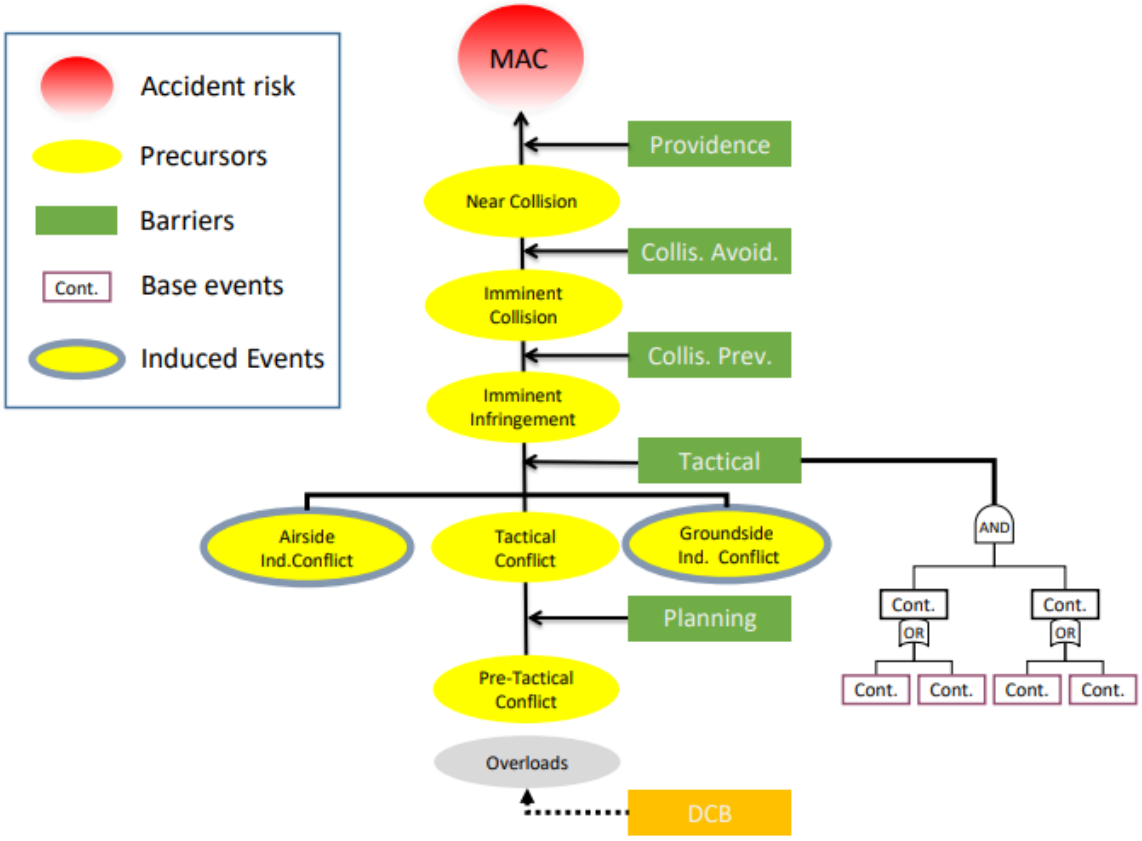

Figure 2. Main elements of the AIM barrier models illustrated with Mid-Air Collision

3.2.1.4 Conclusions drawn by AIM

As AIM is incorporated in the SRM, it is utilized in safety assurance exercises related to the integration of most SESAR solutions. The more recent iterations of the AIM include *Safety II*[4] aspect of what happens when particular barriers work as intended. Due to its aviation-oriented design and application, AIM has not been used in other industries. This represents an advantage, as its approach has been extensively validated in ATM and ATC-related solutions. However, this is also a limitation in that it has not benefited from the different viewpoints and the approaches that might be co-developed by different industries.

Further advantages of AIM include the fact that as it is ATC/ATM specific, it has extensive coverage of all elements and interactions specific to the ATC and ATM environment, including Human Factors aspects. Moreover, its quantitative nature and logical structure make it easy to construct and follow, and the clear temporal axis (bottom -> top) provides additional support to comprehension. A bi-product of the temporal axis is the hazard increase in severity, which is implied in the format of the AIM. Finally, the AIM is capable of showing the change in safety risk with the addition/change of a

[4] Safety-II: Safety management that is aimed at ensuring that 'as many things as possible go right' (in contrast Safety-I is that which is aimed at ensuring that 'as few things as possible go wrong')

technological tool or procedure, simply by *plug-and-play*, i.e. adding a barrier failure or base event is intuitive and the probabilities and failure rates can easily be adapted.

Its major downside is that in some instances, the probabilities and failures rates are just an estimation or do not exist at all. However with the evolution of safety intelligence[5] and data collection, such disadvantages will most likely disappear.

3.2.2 The Functional Resonance Analysis Method (FRAM)

3.2.2.1 Theoretical basis of FRAM

Developed by Erik Hollnagel [15], FRAM is a model that focuses on system interdependencies, their dynamics and complexity. This model originates from the principles found in Resilience Engineering, in which the characteristic performance of a system is considered as a whole, rather that the cause-effect mechanisms of more simple liner models. In resilience engineering, accidents and failures are not seen as a breakdown or malfunctioning of normal systems functions, but instead are represented as the converse of the adaptations necessary to cope with real world complexity. The FRAM model is therefore focussed on describing the dynamic and non-linear nature of the interactions in a system. This is a unique model and differs from more traditional models that consider accidents in a sequential manner, or as connecting latent conditions.

FRAM describes system failures or 'adverse events' as the outcome of a 'functional resonance' which arises from the 'variability' of normal performance. The model represents individual and or organisational 'functions'[6], where each function makes its own, potential 'performance variability' contribution to the whole system. The concept of 'resonance' is used to explain how large effects can arise from small variations. In FRAM the emphasis is more on dependencies than failure probabilities. 'Common Conditions' (CC) are factored into the model which are analogous with performance shaping factors in other models, that is people must adjust their activity to the working conditions or context of the system in order to accomplish a task. CC's are partly responsible for the performance variability observed in each function.

Functions (either individual humans or an aggregation of people) can potentially 'couple', however there are no cause-effect relationships in the model. These couplings are described in terms of six types of dependencies, namely input, output, time, control, pre-conditions and resources (cf. Figure 3). Exploring the dependencies between functions, according to these categories, provides a method of assessing the possibility of a coupling (intended or not) between two functions. The 'couplings' are then analysed to show where 'coincidences' (rather than causal relationships) may arise from performance variability and hence where risk exists in the system.

[5] Safety Intelligence: usable and actionable safety information arising from a systematic process for processing safety data and information.

[6] Functions: A description of real work (work as is) as functional elements of a sociotechnical system (what must be carried out to achieve a given goal).

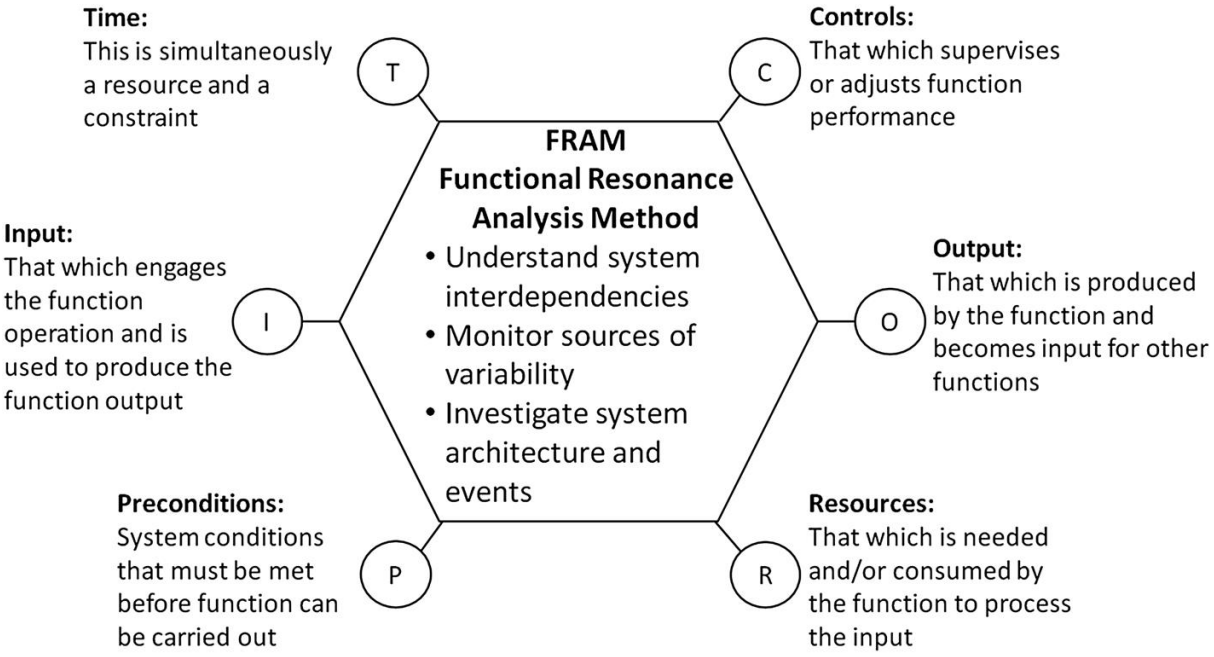

Figure 3 Functional Units of FRAM

3.2.2.2 Method of applying FRAM

The method of applying FRAM is described in 5 steps, as follows:

A. Define the purpose of the analysis since FRAM has been developed for use in accident investigation (examining past events) and safety assessment (future events).

B. Identify and describe the system functions. A system function is something of either a human, technological or organisational nature, which transforms the state of the system towards fulfilling the operational purpose of this system.

C. Assess and evaluate the potential variability for each singular function. Variability is largely ascertained according to two dimensions and based on the output of the functions, e.g. variability in time: the output is on time or within an acceptable timeframe, too early, too late and variability in quality: the output is up to expected standards, out of expected standards but adequate, unsuitable, etc.

D. Identify functional resonance. This step aims to show the possible ways in which variability from one function could spread in the system and how it might combine (couple) with the variability of other functions to cause a risk.

E. Identify effective countermeasures to be introduced into the system to dampen the negative performance variability and also measures that can sustain or amplify functional resonance that results in a desired/improved outcome.

3.2.2.3 Analysis used in FRAM

After defining the scope of the assessment and classifying components of the system in steps A. and B., the majority of the risk analysis is carried out in steps C. and D. The FRAM model does not result in a numerical value of risk; there is no 'counting' of errors or calculation of failure probabilities. It instead evokes the theory that disproportionately large effects can arise from small or insignificant variations in the system [15]. Moreover in FRAM, the emphasis is on the dynamic dependencies within the system rather than failure probability.

Step C is the analysis of the variability associated with the functions and acts to highlight area of the system that are showing different levels of performance, good and bad. This step also allows for the assessment of the systems capability to cope with the differences in performance. Furthermore, the assessment of the variability involves the identification of which part of the system is experiencing this variability, namely human, technological and organisational aspects of the system. Once the variability in performance output of the functions is understood, step D. aims to analyse how this variability propagates throughout the system, not in a linear fashion but in a dynamic and changeable manner. This step evaluates how the variability associated with the functions has an impact on the system through two mechanism, firstly through the direct coupling with the variability of another function e.g. if the output of a function comes too late, it will result in a reduction of the time that is available for the other functions to produce their output. Secondly the variability from a function can have a second order effect by causing a change in one or more the 'common conditions' e.g. increased variability may lead to an increased use of resources or less time availability in the system. It is suggested that this last step is supported with a software tool. The final step E aims to determine if and how the risk identified in steps C. and D. can be mitigated or eliminated.

3.2.2.4 Conclusions drawn concerning FRAM

The FRAM model presents a contrasting approach to more traditional risk assessment models in that all outcomes of the system functioning (negative or positive) are due to performance variability. In this same system safety is seen as an ability to succeed under varying conditions and thus applying the model provides an tool to improve resilience in the system, hence the use of the term 'resilience engineering'. In contrast, a large number of risk assessment models suggest that negative outcomes of a functioning system are caused by failures and malfunctions and that providing 'safety' is the act of reducing the number of adverse events in the system. The aspiration in these models is to eliminate failures and malfunctions as far as possible.

Due to its versatility, FRAM has been used in the healthcare sector, where human performance variability highly influences the outcomes at a level of fine granularity, for example medicines must be administered in precise amounts and at the right time. The amount of equipment that hospital staff are using, but also the reliance on automation, is significant. In this setting FRAM has been applied with good success as it has been able to represent very many 'couplings' in such a hierarchical type of organisation. FRAM has also been applied in aviation, to model the highly complex and intensive work of Air Traffic Controllers [16] and Air Traffic Management (ATM), in the context of a Mid-Air collision risk. In this study, the system is broad with many interactions and dependencies, e.g. individuals in this system must communicate with a large number of other agents in the system, whilst paying attention to multiple screen, making challenging calculations and monitoring a multitude of aircraft. This level of interconnecting components makes FRAM a good candidate for conducting a risk assessment in this situation.

As such, the advantages of FRAM are that it is holistic, has a hierarchically-free approach and is scale-invariant, has the capability to capture all and interactions between all elements in a complex socio-technical system, it points to the resilient elements of the system and has the capability to underline what in an operation works as intended. Its disadvantages are that the resonance and its variability cannot always be defined due to the non-linear couplings between elements in the model. The model can become hard to understand once multiple interactions are added, meaning that some parts of the operations have to be considered separately. There is no timeline of the event, however steps can be added by number the elements, which in turn has the disadvantage of adding some rigidity to the interpretation of the model.

3.2.3 Causal Model for Air Transportation Safety (CATS)

3.2.3.1 Theoretical basis of CATS

The CATS model provides insight into the cause-effect relationships in the *event sequences* leading up to potential incidents and accidents. Its purpose is to establish, in quantitative terms the risks specific to air transport. These 'event sequences' cover all potential failure modes of the operations during the different gate to gate flight phases. It was developed specifically for aviation by a consortium including Delft University of Technology (TUD), National Aerospace Laboratory (NLR), White Queen Safety Strategies (WQ), the National Institute for Physical Safety (NIFV) in The Netherlands, Det Norske Veritas (DNV) and JPSC UK.

3.2.3.2 Method of applying CATS

Its structure and approach is similar to that of AIM, in that failures of various agents and/or components are shown by Fault trees (FT), which have top events that progress towards incidents or accidents through the ESD (Event Sequence Diagram). The two major differences compared to AIM are that CATS does not use barriers and also it employs Bayesian Belief Networks (BBN) to quantify the risk from an available database. **Figure 4** shows the basic components of CATS.

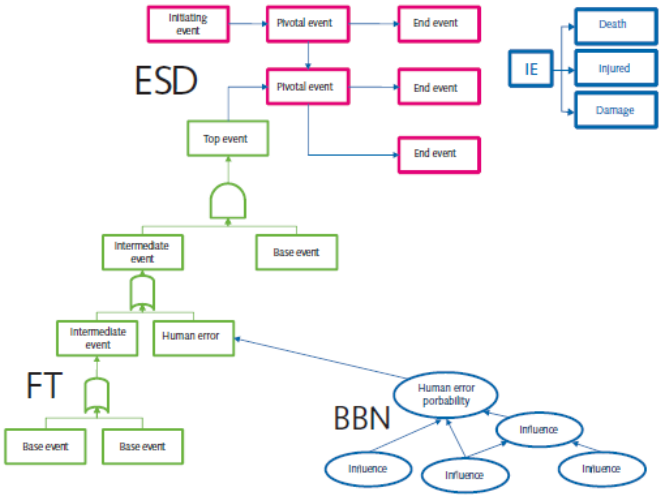

Figure 4. Basic components of the CATS model.

3.2.3.3 Analysis used in CATS

In CATS, each Fault Tree is quantified by a BBN, and all BBNs are interlinked to form one giant BBN with over 5000 nodes and 1440 links [17]. Thus 'all' dependencies are taken into account and the probability of different accident/incident types are generated, which is the main purpose of the CATS. For such a high magnitude of inputs, a considerable number of rules and large amount of data must be input into proprietary software, making this tool greatly labour intensive.

3.2.3.4 Conclusions drawn concerning CATS

By utilising BBN's, CATS has the advantage of allowing the use of distributions rather than point values, which makes the model more realistic and reliable. Other advantages of using CATS are that it has a clear logic, it is easy to understand and has been developed for aviation.

Disadvantages in the model include the need for a large amount of input data and thus a lack of data renders it useless as it is only focused on probability, it is labour intensive to input all the data and it is not easily understandable due to its large size.

3.2.4 Aviation Safety and Certification of New Operations and Systems (ASCOS)

3.2.4.1 Theoretical basis of ASCOS

ASCOS is based on the previous accident model development work carried out to create CATS but has a specific focus in defining an harmonised approach towards an improved certification process in aviation [18]. The ASCOS model is comprised of Event Sequence Diagram (ESD) and fault trees (FT) that represent the total aviation system under examination. Indeed the ESDs and FTs from CATS are the starting point of ASCOS accident model. The ASCOS accident model has a fault tree for each initiating event, and for most pivotal events. It is noted in ASCOS literature that FT's are generally developed to a level of detail according to the best failure probability data available but since detailed failure information on non-critical events is often lacking in aviation, the fault trees on these are not overly detailed.

3.2.4.2 Method of applying ASCOS

Similar to other models, ASCOS is quantified by assigning probabilities of occurrence to each of the different pathways in the scenarios, based on historic or expert opinion-derived data. It can be used to analyse the risk of individual events, but it can also be used to assess the impact on safety of changes to the system. By considering a change within the system, the model can be used to determine the quantitative influence of the change on accident risk. New events can also be added and quantified that are specific to the change. ASCOS also aims to provide a continuous monitoring process in which safety occurrences will be recorded and integrated into the model and used as safety performance indicators. The basic components of ASCOS are represented in Figure 5.

3.2.4.3 Analysis used in ASCOS

One characteristic of ASCOS that AIM does not provide is 'representation of emerging and future risks'. The model does this by identifying precursors, linked to a specific base event of the fault tree, which are associated with emerging and future hazards. These precursors are defined as "identifiable event that may be used as an indicator for known or potential hazards". A common method for identifying emerging and future risks is by creating a series of possible uses and scenarios capturing how the

system of interest might develop. The second step in ASCOS to integrate emerging risks is then to link the precursors with elements of the model and if this not immediately possible, then the applicable part of the model should be modified or extended to allow a connection between the precursor and the model for emerging risk.

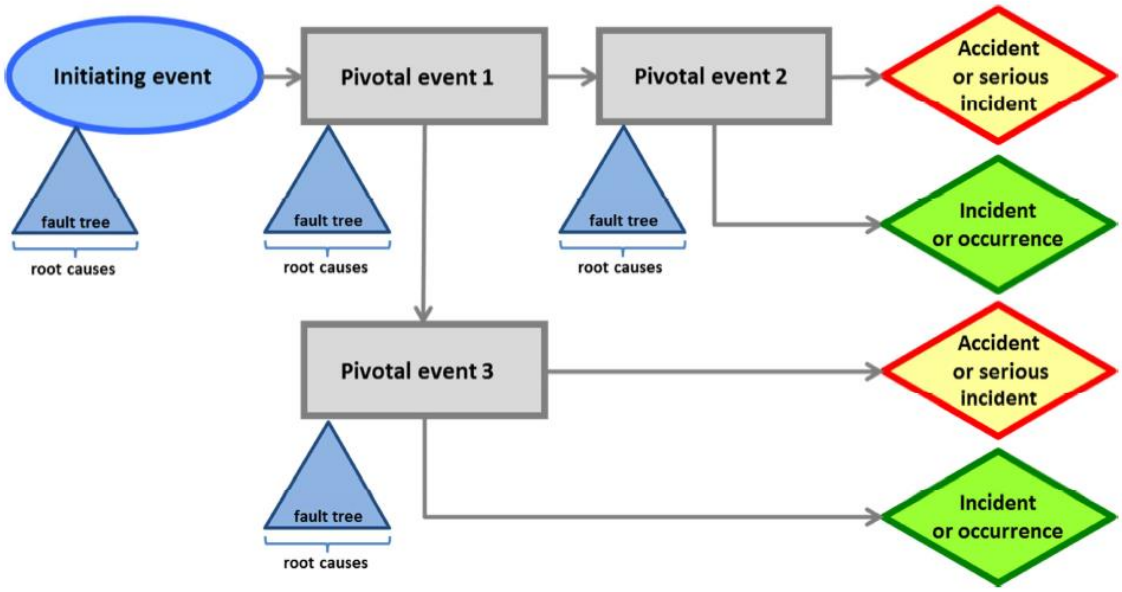

Figure 5. Basic components of the ASCOS model

The risk model supports safety management in several ways, especially by improving oversight over different operations and understanding the safety significance of a service or supporting service.

3.2.4.4 Conclusions drawn concerning ASCOS

ASCOS is a quantitative model for risk assessment and its structure similar to AIM is an advantage as it makes it familiar to the SESAR context. Also, the additional component which enables to evaluate emergent risk is beneficial as it supports the Safety- II approach. In addition, ASCOS can provide insight on the overall aviation risk picture because hazardous situations for each type of operations have already been modelled. However, disadvantages include the fact that ASCOS is based on extensive Event Sequence Diagrams and Fault Trees that are inherently complex to follow because of their size. Also, because of its ambition to consider aviation as a whole, the description of the causal relationships often captures only part of the interactions between actors or events, and it is missing the component of barriers and their associated failure rates.

3.2.5 Model of ATM Reality in Action (MARIA)

3.2.5.1 Theoretical basis of MARIA

The Model of Air Traffic Management (ATM) Reality In Action (MARIA) is a knowledge database and an automation framework which has been developed specifically for ATM by NAV Portugal [19]. It is neither a business nor safety model, but aims to examine how safe and resilient current ATM systems

are, and what the current knowledge and understanding of these systems are. It has also the ambition to provide a sound base for safety analysis by describing the whole system and the interdependencies between its functions. It may be used in simulations to identify strong and weak points in procedures, to evaluate risk mitigation strategies and for other purposes.

The aims of the model are to:

- Give a global view of the system (people, procedures and equipment)
- Show the dependencies between functions / processes
- Build a description of the system architecture
- Use it as a reference for future safety assessments
- Come up with a systematic and reproducible hazard identification method
- Help the definition of risk mitigation strategies

3.2.5.2 Method of applying MARIA

NAV Portugal suggests that although several relevant models are already available for risk assessment in ATM, none provided an adequate description of the whole system, claiming that they either lack a description of the big picture or a description of how everything in the system interacts.

A first version of the generic ATM architecture was released at the start of 2014. MARIA is based on the day-to-day activities modelled top-down, starting with very high level functions (Figure 6) which are systematically decomposed down to the level where the resources performing them can be identified.

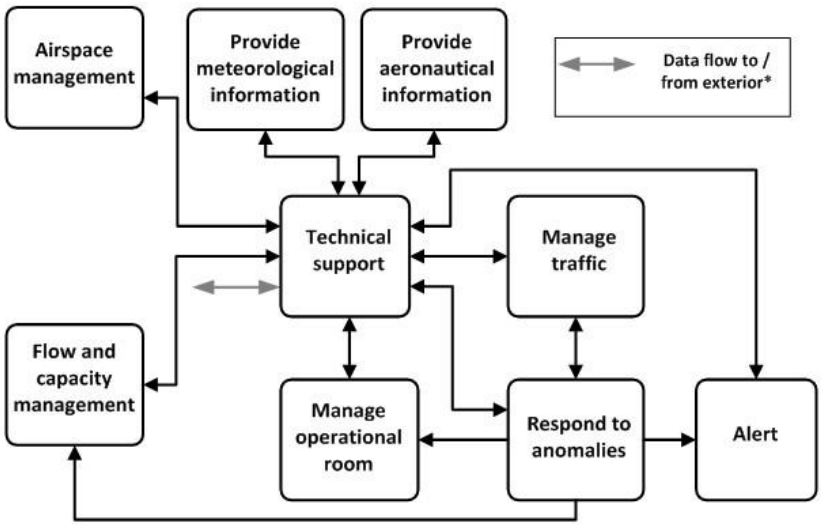

Figure 6. Top Level Functions in MARIA

3.2.5.3 Analysis used in MARIA

The Structured Analysis and Design Technique (SADT) was used to represent the top level functions and the associated inputs and outputs. Details of some of the functions were captured using Business Process Modelling Notation (BPMN), a notation which permits recording the way work is done. As an example, Figure 7 illustrates one of the sub-function ('Conflict resolution') in which the high-level functions can be decomposed [19]. Each box indicates the information that will be processed as input (top and left) and the results that will be produced as output (right). The lower edge receives the resources that perform the function, be it people or equipment, and usually called mechanisms or enablers.

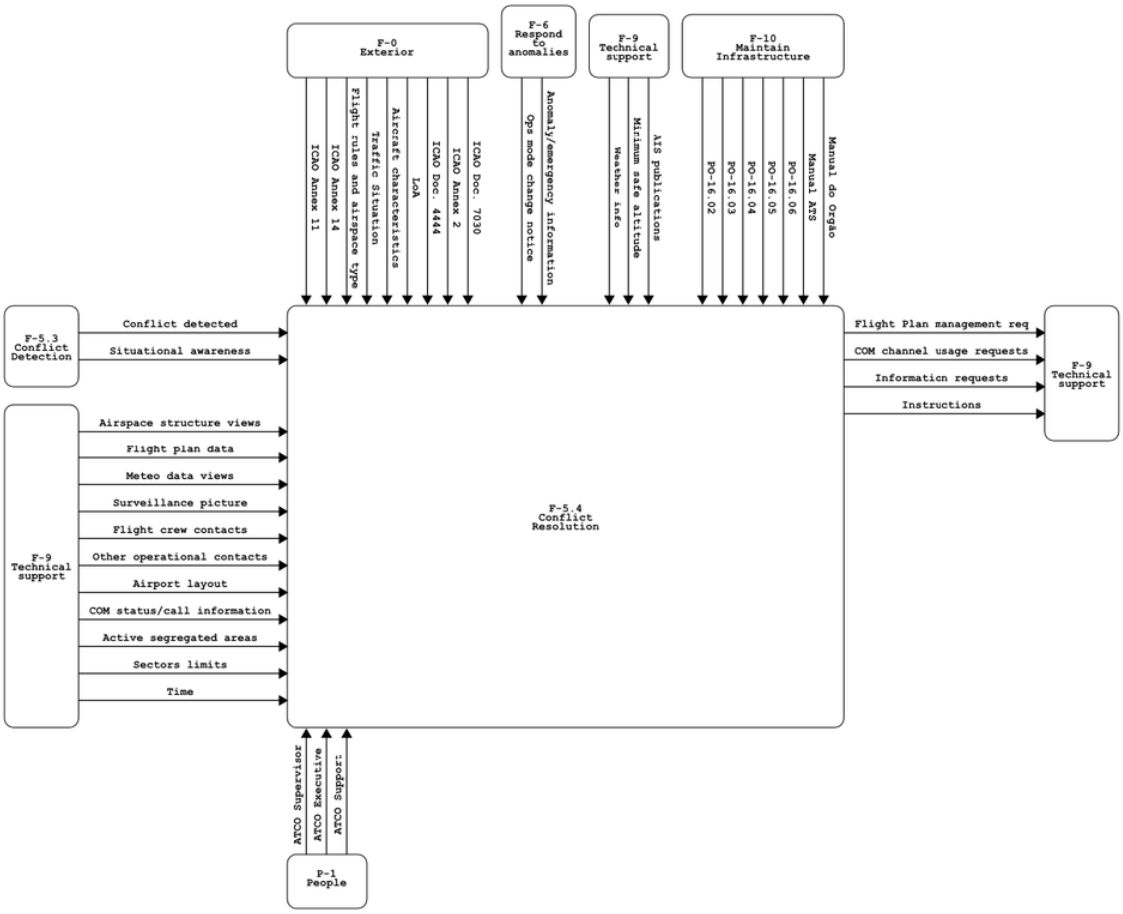

Figure 7. Diagram of a 'Conflict resolution' in the MARIA framework.

3.2.5.4 Conclusions drawn concerning MARIA

The systemic use of MARIA in time yields a repository covering all that is done to provide an ATM service, in a systematic and uniform way, structured in a top-down manner, with the human as an integral part of the system. In addition, allowing automation of analysis and representation are significant advantages to improve understanding, safety and resilience of the ATM services.

Advantages of the model are that it can be used for safety assessments and as its primary focus is on functional modelling, it is relatively easy to build. Moreover it is easy to understand at the high levels and the flow of information is show very well. Finally the model framework is open allowing the easy addition of new properties to the existing descriptions, and new areas or systems

In terms of disadvantages, the model is not visible in its entire form due to its hierarchical structure and so it is difficult to appreciate the whole structure/composition. There are no quantitative aspects of the model. Moreover, it requires extensive effort to build it as it requires an extensive and exact understand of how the process works, talk with ATCOs etc. MARIA deals with providing a framework, a static picture, than modelling the dynamics and the unknown areas of planned study.

3.2.6 System-Theoretic Accident Model and Process (STAMP)

3.2.6.1 Theoretical basis of STAMP

The System-Theoretic Accident Model and Processes (STAMP) is a qualitative accident causation model created by Dr. Nancy Leveson (Massachusetts Institute of Technology) to analyse accidents in systems [20]. Although it is not strictly a safety model, it is a noteworthy model to present in the review as it demonstrates a different approach to looking at safety in a complex system, albeit after an accident in the case of this model. In this model, accidents are examined in terms of why the controls that were in place did not prevent or detect the hazard(s) and why these controls were not adequate in enforcing the system safety constraints.

STAMP is unlike traditional accident causation models where the root cause consist of an event or chain of events, STAMP focuses on investigating the cause of an accident by identifying the safety control that were inadequately enforced. More traditional accident causation models generally try to identify the first adverse event in the chain and prevent it from happening without considering environmental, organizational, or human contributions. FMEA, FTA, ETA, and Cause-Consequence Analysis are based on this approach. Leveson claims that these older models don't work well in complex systems involving human behaviour, as they are based on a linear chain of events and assume an accident is a result of a component failure, whilst not accounting for components are compromised without failure. In contrast, STAMP aspires to consider accidents as a result of interactions among system components and a lack of control of safety related constraints. In this model no blame is attributed to a single component or individual human.

3.2.6.2 Method of applying STAMP

The three main principles of STAMP are safety constraints, hierarchical control structure, and process models.

First, *safety constraints* are enforced through safety controls, which if adequately implemented, will prevent adverse events from happening. Second, *hierarchical control structure* represent an essential step in applying STAMP where each level of the system contributes to the safety or to accidents in a system. Each level of the hierarchy enforces safety constraints to the level below it, and each level below have to give feedback on how these constraints are successfully implemented or have failed. Consequently, higher levels of hierarchy are responsible of the performance of the lower levels. Missing constraints, inadequate safety control command, commands not executed properly at lower level, or inadequate feedback communications about constraints are the main reasons of inadequate controls. Third, four conditions must exist for a process to be controlled under the STAMP model. These are:

- Goal (enforcing safety constraints in each level of the hierarchy structure by controllers),
- Action Condition (implement actions downward the hierarchy structure),
- Observatory condition (Upwards in the hierarchy),
- Model condition (the controller's model of the process being controlled), which in our case is the *process model*. Without this a process would not adequately be controlled.

3.2.6.3 Analysis used in STAMP

As mentioned, STAMP is not based on a chain of events, but instead based on system-theory, where each level of the organization contributes to an accident or to attaining successful system safety controls. Thus STAMP accounts for organizational factors, human error, and adaptation to change over time. In STAMP, system safety is not achieved by preventing component failure measures; instead it is achieved by enforcing safety constraints continuously. As a result, accidents do not occur because of failure of components, they occur because of ineffective safety constraint. Therefore, the main focus is not on how to prevent failure, but on how to design better safety controls.

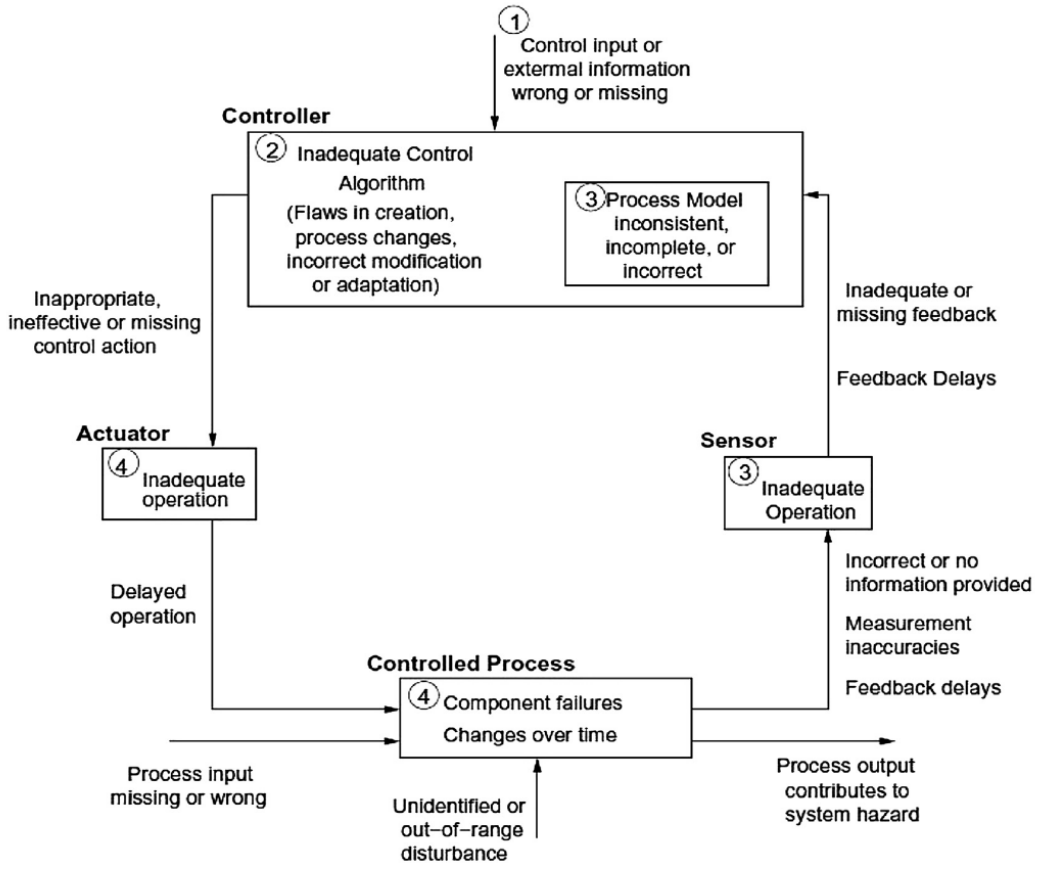

Figure 8. Diagram a control loop's components which can become contributory factors to an accident, if improperly operated ([20])

Figure 8 shows the STAMP taxonomy of contributory factors to accidents. Examining the components in the control loop helps examine how each one can add to the inadequacy of safety control, if improperly operated. Causal factors have been divided into three categories: controller operation,

behaviour of actuators and controlled processes, and communication and coordination among controllers and decision makers [21].

Figure 9 shows the general classification of the flaws in the components of the system development and system operations control loops during design, development, manufacturing, and operations according to STAMP. This classification can be applied to all levels of the organization under investigation during accident analysis or as an accident prevention to prevent future or potential adverse events.

1. Inadequate enforcements of constraints (control actions)
1.1. Unidentified hazards
1.2. Inappropriate, ineffective or missing control actions for identified hazards
1.2.1. Design of control algorithm (process) does not enforce constraints
—Flaws in creation process
—Process changes without appropriate change in control algorithm (asynchronous evolution)
—Incorrect modification or adaptation.
1.2.2. Process models inconsistent, incomplete or incorrect (lack of linkup)
—Flaws in creation process
—Flaws in updating process (asynchronous evolution)
—Time lags and measurement inaccuracies not accounted for
1.2.3. Inadequate coordination among controllers and decision makers
2. Inadequate execution of control action
2.1. Communication flaw
2.2. Inadequate actuator operation
2.3. Time lag
3. Inadequate or missing feedback
3.1. Not provided in system design
3.2. Communication flaw
3.3. Time lag
3.4. Inadequate sensor operation (incorrect or no information provided)

Figure 9. Hierarchical classification of control flaws leading to hazards

For each level of the hierarchy, the three main categories should be investigated to determine their contribution to the accident:

- Control actions: inadequate handling of control actions by controllers
- Execution of control action: inadequate execution of action
- Feedback: missing or inadequate feedback and communication

Another category can be added if humans are involved in the organization being investigated.

3.2.6.4 Conclusions drawn concerning STAMP

Studies have shown that utilizing STAMP to analyse accidents has revealed more hazards and potential failures in systems than other traditional hazard analysis or accident causation models [21]–[23].

This model's strength lies in the fact that it uses systems thinking and, like FRAM, tries to integrate the social and technical parts into one, as opposed to a linear change of events. STAMP certainly accounts for organizational factors, human error, and adaptation to change over time

However, this model does not use quantitative analysis and so has limitations as a safety model, according to this it is then better suited to accident investigation.

3.2.7 Traffic Organization and Perturbation AnalyZer (TOPAZ)

3.2.7.1 Theoretical basis of TOPAZ

Motivated by a need to model the dynamics, the stochastics and the interactions of safety-critical multi-agent systems in advanced ATM concepts of operations, the National Aerospace Laboratory (NLR) of the Netherlands developed the TOPAZ (Traffic Organization and Perturbation AnalyZer) safety risk assessment methodology [24], [25].

TOPAZ has a qualitative component which enables the user to define the ATM Concept of Operations (ConOps) to be considered and identifies potential hazards of this ConOps. The identified hazards can be subsequently taken into account into the quantitative part of the TOPAZ methodology.

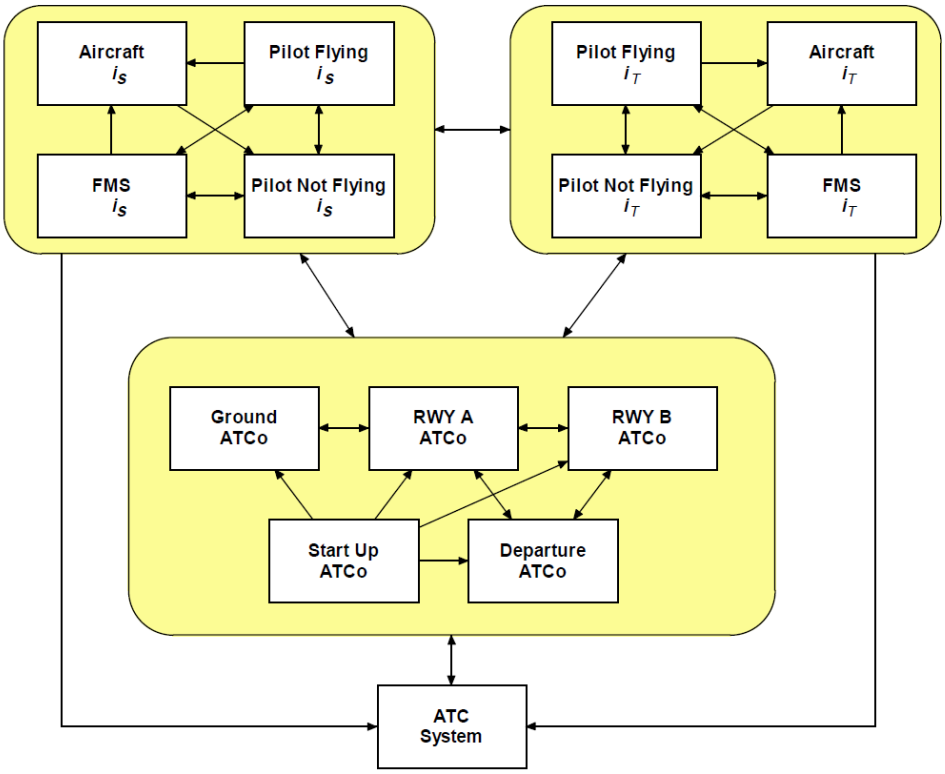

Figure 10. Relations between 'agents' in the active runway crossing operation in the TOPAZ model [26]**.**

This develops an agent-based model of the operation considered, that is, a computational tool for simulating the actions and interactions of autonomous agents (both individual or collective entities such as organizations or groups) in order to understand the behaviour of a system and what governs

its outcomes. Subsequently, TOPAZ uses this model in a rare-event Monte Carlo[7] simulation and bias and uncertainty analysis.

3.2.7.2 Method of applying TOPAZ

The TOPAZ methodology supports two quantitative approaches in taking a hazard into account. The first approach is to assure that the hazard is captured in the agent-based model that is used for the MC simulations. The second approach is to assess the safety risk impact of the hazard during the bias and uncertainty assessment.

For the majority of the hazards identified during the qualitative phase within TOPAZ it is possible to model them in an agent-based setting. However there are hazards for which it is not yet known how to capture them in an agent-based model. For this latter category of hazards the only choice available is to assess them through bias and uncertainty assessment.

3.2.7.3 Analysis used in TOPAZ

TOPAZ uses a 'Multi Agent Dynamic Risk Modelling (MA-DRM)'. This approach a quantitative modelling and analysis one, which integrates the following five computational modelling techniques:

- Agent-based Modelling (ABM): used to conceptualise processes in the world, and in particular open socio-technical systems such as the ATM system;
- Human performance modelling: For the human agent, an agent-based model is used which integrates various human performance sub-models. In addition to this, shared and distributed Situation Awareness (SA) across various agents is modelled using an extension of the situation awareness (SA) model of Endsley to a multi-agent SA (MASA) propagation model;
- Powerful Petri Net modelling syntax: In order to manage the systematic development of a mathematical model of such operation, use is made of a powerful Petri Net (PN)[8] syntactical framework;
- Rare event Monte Carlo (MC) simulation: ATM safety analysis requires covering many magnitude orders in time scales. This is accomplished through making use of rare event MC simulation techniques;
- Sensitivity and Bias and uncertainty analysis: this allows the assessment of the impact of potential differences between the true operation and the agent-based model on the risk level assessed.

[7] In the context of risk assessments, Monte Carlo simulation is a computational technique that allows people to account for risk in quantitative analysis and decision making. This technique is used to calculate, for a range of possible outcomes, the probabilities they will occur for any choice of action.

[8] A Petri net, also known as a place/transition net, is one of several mathematical modelling languages for the description of distributed systems.

3.2.7.4 Conclusions drawn concerning TOPAZ

The advantages in using TOPAZ include the fact that it shows the situation in great details, with high fidelity through quantitative means, it has been used in aviation before and continues to be used, and that the interrelations and interactions are represented in a basic visual manner (however must be described more extensively in text).

Disadvantages of the model include the fact that there is no timeline or events, the approach requires a lot of quantitative data and requires the MA-DRM tool which isn't immediately available, emerging events are hard to extract e.g. the model calculates the probability of the known events in the system, and introducing a system change and tracking its impact is difficult since it uses Monte Carlo simulations.

4 Analysis

4.1 Evaluation of the risk models against the acceptance criteria

The evaluation to establish the most appropriate risk model to use for SafeOPS purposes was done by analysing the models reviewed in this document through the lenses of the acceptance criteria developed in Section 2.2. In this evaluation, a grading was given on a three point scale, namely 'the model meets the criterion', 'the model partially meets the criterion', and 'the model does not meet the criterion' (cf. Table 1). A model is assigned to the category II when it fails to meet the criterion in its totality in terms of quality or quantity, but has some aspect that supports the criterion and could potential be more suitable with further development or application of the model.

Table 1. The three possible suitability categories assigned to the models against each individual criterion.

	Suitability Category	Symbol
I	The model meets the criterion	√
II	The model partially meets the criterion	~
III	The model does not meet the criterion	X

The results of the evaluation of the risk models against the criteria are presented in Table 2. In the following Section 4.2 we discuss the main outcomes of the analysis.

4.2 Discussion

All the considered models present at least a few characteristics that make them suitable for SafeOPS. However, the majority of models did not meet, or only partially met, criteria 2 and 3 of being able to evaluate how new technology *changes* the level of extant risk in the system, e.g. change assessment. This shortcoming is mainly related to the fact that 'baseline' risk-model scenarios have not been developed in these models. Thus, to apply them in the SafeOPS environment would potentially require all applicable scenarios from SafeOPS to be modelled before then looking at the impact of the technological solution.

Another criterion that created a clear division on the suitability of a model was that of qualitative versus quantitative articulation of the risk. Assessing a risk quantitatively is not straight forward and requires a lot of subjective and objective data, including that taken from past incident/accident databases. Models that do not articulate the risk in quantitative terms often have other strengths in being able to model more dynamic relationships between components in the system. The SafeOPS project, however, aims at quantifying the impact of introducing the new, predictive tool in the routine of the ATCOs, and it is therefore necessary to have a scale of reference to measure the consequences of the changes that are introduced. In addition, there is an inherent need to select a model which is validated specifically in the ATM environment, as required by Criterion 12.

Table 2. Evaluation of the risk models against the acceptance criteria

	Acceptance Criteria	AIM	FRAM	CATS	ASCOS	MARIA	STAMP	TOPAZ
1	The model shall be able to capture all aspects of the operating environment and allow for the examination of the predictive technology within that.	√	~	√	~	~	~	~
2	The model shall be able to assess how the technology changes the level of extant risk when it is being used as a real-time prediction tool.	√	~	√	√	X	X	~
3	The model shall be able to assess how the technology changes the level of extant risk when it is being used for offline provision of analytics.	√	~	√	√	X	X	~
4	The model shall be able to integrate into the current safety management system of the organisation of use.	√	~	√	√	~	~	√
5	The model shall be able to capture any risk associated with misleading, incorrect or untimely presentation of predictive data from the technology.	√	√	√	√	X	X	√
6	The model shall be able to capture any risk associated with the display design and usability.	√	√	√	√	√	√	√
7	The model shall be able to capture any risk associated with human error related to information processing, decision making and acting on the predictive data.	√	√	√	√	X	X	√
8	The model shall be able to capture any risk associated with violations in the use of the technology e.g. inappropriate use (misuse) or rejection of use (disuse).	√	√	√	√	X	X	√
9	The model shall consider risk in all scenarios related to a standard and problematic approach.	√	~	√	√	√	~	√
10	The model shall consider all risk in the context of the potential source of harm (the resultant hazard) e.g. CFIT, airborne conflict, loss of control.	√	√	√	~	√	√	√
11	The model should have the capacity to quantitatively articulate the likelihood and consequences of risk associated with the integration of the technology.	√	X	√	√	X	X	√

12	The model should be validated in an ATM environment.	√	√	x	√	√	√	√

Models that have not been validated specifically in the ATM environment are less favourable as they would require an extra development step, and namely to quantify the 'baseline risk' of the operations, before being able to assess the change. By contrast, a model validated in the ATM context ensures that the SafeOPS Risk Framework builds on current knowledge of quantified risk and also more ease in integrating any assessment into the models currently being used by ANSPs.

The AIM model meets all the acceptance criteria presented in Section 2.2. More specifically, it provides thorough coverage of all elements and interactions specific to the ATC and ATM environment, in which it has been extensively validated, and it is capable of showing the change in risk with the addition/change of a technological tool. Safety impact is qualitatively assessed at the level of the individual barriers, contributors and precursors, and is quantified in terms of the variation in barriers performances or precursors' occurrences. Also, AIM is currently the framework of choice at Eurocontrol and is already incorporated into a larger safety assurance framework of SESAR. In addition, AIM allows an extensive coverage of Human Factors aspects (cf. e.g. the work done within the SAFEMODE project [13]), although at present there is no information available specifically on the modelling of the provision of predictive data in real time.

The FRAM model has been validated in the ATM environment and can be adapted to account for all types of hazard that are relevant to the missed approach scenario, also incorporating in great detail the human component, as it was validated in the healthcare sector, where human performance variability highly influences risk modelling. One of the main limitations of FRAM is that not all aspects of aircraft approach have yet been modelled. Consequently, the current level of risk for some possible outcomes of the go around scenario (e.g. controlled flight into terrain, runway incursion and separation infringements or collisions in the terminal manoeuvring area) would need to be modelled in the form of FRAM before examining the change in risk. Also, unlike the majority of existing models that represent risk as a linear sequence of events and failures, FRAM focuses on the dynamic and non-linear interdependences of elements in a system. Consequently, its integration with existing models is possible but could be challenging.

CATS and ASCOS describe the evolution of a system in terms of a sequence of events, and the failures of various agents and components are schematised with fault trees. Similarly to AIM, these two models are well-suited to cover all potential failure modes of the operations during the different gate to gate flight phases. However, the main limitation of CATS is that it has not been validated in an ATM context, as it was developed in the air transport environment. ASCOS is largely based on CATS and has been used also in the ATM environment; however, it was developed with the ambition to contribute to removal of certification obstacles and support implementation of technologies by providing an integrated approach to risk modelling and safety for the current total aviation system in accident scenarios. With such a broad focus, ASCOS necessarily captures only part of the causal relationships, type of interactions between actors or events, and component of barriers and their associated failure rates. For these reasons it is less suitable for the scope of the SafeOPS project.

The main limitations of MARIA and STAMP are that they are better suited to performing accident investigation and analysing possible events, safety, and resilience in the current ATM systems rather than the integration of new technology and the changes introduced in the system in terms of risk. Also, while the human is considered as an integral part of the system, human aspects related to the introduction of new tools (particularly novel predictive technology) is not covered. Finally, despite providing a powerful and quantitative methodology to assess risk and risk changes for a variety of ATM-relevant scenarios, TOPAZ is a rather complex model, in which system changes are not straightforward to implement and their impact is not easy to monitor.

In conclusion, based on the above-described arguments, the most suitable model to be used within the scope of the SafeOPS project, according to the criteria, is the AIM model. It was judged as meeting all criteria and had particular strengths in its ability to assess change in the system, its use in ATM and the possibility to be relatively-easily integrated into the current Safety Management System adopted by the ANSPs.

5 Summary and Conclusions

This deliverable documents the first phase of the Work Package 3 (WP3) of the SafeOPS project. WP3 is devoted to the development of an integrated Risk Framework for evaluating the risks concerning the use by Air Traffic Controllers (ATCOs) of a novel, data-driven tool to predict pilot-induced go arounds. The work conducted in this initial phase consisted in a systematic review of the state of the art of risk models available for application in the context of Air Traffic Management. The end goal of this review was to identify the risk model best suited for adaptation and use in the context of the SafeOPS project.

In the literature, there are more than 800 documented methods, techniques and tools for the analysis of risk in various industries, from aviation and transport in general to nuclear, healthcare, information technology and manufacturing. It was therefore necessary to make a selection of the models to analyse in detail. The starting point was to only consider *systemic models* that enable to analyse *all* components of the go-around scenario including all relevant aspects of current operations, the design and performance of the predictive tool, and the Human Factors aspects involved in managing traffic during missed-approach procedure. As a consequence, this choice led to the exclusion of a variety of individual structured techniques and tools used for describing and analysing specific elements of the overall scenario. The other requirements for the selection were for the models to be presently in use in aviation and adaptable to the scope of the SafeOPS project. These initial criteria led us to a list of seven models, namely AIM, FRAM, CATS, ASCOS, MARIA, STAM and TOPAZ, which were reviewed to identify their main advantages and limitations. In this review, twelve acceptance criteria were laid down to determine which of these seven models is best suited for describing the impact on safety of the new predictive tool in the go-around scenario.

In conclusion, this analysis identified the Accident-Incident Model (AIM) as the most appropriate model for further use in the SafeOPS project. AIM meets all twelve acceptance criteria. In particular, it is well established and widely used for modelling ATM operations, it already covers all relevant aspects and hazards involved in go arounds, and it enables to consider Human Factors aspects and to capture the impact of variations to the current standard of operations in a relatively straightforward manner. AIM is therefore recommended for adoption in the next phases of WP3 to describe qualitatively and quantitatively the impact of the novel predictive tool in the ATCOs' routine and in particular to analyse the Human Factors aspects of real-time provision of data-driven, probabilistic information.

6 References

[1] T. Blajev, “Final Report to Flight Safety Foundation Go-Around Decision-Making and Execution Project,” 2017.

[2] SKYbrary Aviation Safety, “Go-around Decision Making | SKYbrary Aviation Safety.” https://skybrary.aero/articles/go-around-decision-making (accessed Dec. 21, 2021).

[3] Airbus, “Descent Management Being Prepared for Go-Around.”

[4] Iata, “Unstable Approaches: Risk Mitigation Policies, Procedures and Best Practices, 3rd Edition,” 2017.

[5] SafeOPS Consortium, “D2.1 - Development of Use Cases , User Stories and Requirements,” 2021.

[6] “ISO - IEC 31010:2019 - Risk management — Risk assessment techniques.” https://www.iso.org/standard/72140.html (accessed Oct. 26, 2021).

[7] P. L. Bernstein, “Aganst the Gods - The Remarkable True Story of Risk,” pp. 1–383, 1996.

[8] M. H. C. Everdij and H. A. P. Blom, “Safety Methods Database. Version 1.1,” *Netherlands Aerosp. Cent. NLR*, no. August, pp. 1–261, 2016, [Online]. Available: http://www.nlr.nl/documents/flyers/SATdb.pdf.

[9] International Civil Aviation Organization (ICAO), “International Civil Aviation Organization Safety Management Manual (SMM).” 2012, Accessed: Dec. 22, 2021. [Online]. Available: www.icao.int.

[10] SKYbrary Aviation Safety, “ICAO SHELL Model | SKYbrary Aviation Safety.” https://skybrary.aero/articles/icao-shell-model (accessed Dec. 22, 2021).

[11] N. A. Stanton, “Hierarchical task analysis: Developments, applications, and extensions,” *Appl. Ergon.*, vol. 37, no. 1, pp. 55–79, Jan. 2006, doi: 10.1016/J.APERGO.2005.06.003.

[12] M. Philippart, “Human reliability analysis methods and tools,” *Sp. Saf. Hum. Perform.*, pp. 501–568, Jan. 2018, doi: 10.1016/B978-0-08-101869-9.00012-1.

[13] SAFEMODE Collaboration, “D4.2-Risk models of major accident types in both domains Document Details.”

[14] Eurocontrol, “INDUSTRIAL RESEARCH SESAR Safety Reference Material,” 2018.

[15] E. Hollnagel, *FRAM - The Functional Resonance Analysis Method Centre for Quality*, no. june. 2014.

[16] P. N. P. Ferreira and J. J. Cañas, “Assessing operational impacts of automation using functional resonance analysis method,” *Cogn. Technol. Work 2019 213*, vol. 21, no. 3, pp. 535–552, Feb. 2019, doi: 10.1007/S10111-019-00540-Z.

[17] B. Ale *et al.*, "Causal Model for Air Transport Safety: Final report." Ministerie van Verkeer en Waterstaat, Directoraat-Generaal Luchtvaart en Maritieme Zaken, 2008, Accessed: Oct. 26, 2021. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3Ae4c5e6b0-9e20-4e61-993e-1d9bd0e23d07.

[18] A. L. C. Roelen *et al.*, *Risk models and accident scenarios in the total aviation system*. National Aerospace Laboratory NLR, 2016.

[19] P. L. Santos and P. Monteiro, "Modeling ATM day to day operations A functional model of the ATM system."

[20] N. Leveson, "Engineering a safer world : systems thinking applied to safety," 2012.

[21] H. Altabbakh, M. A. AlKazimi, S. Murray, and K. Grantham, "STAMP - Holistic system safety approach or just another risk model?," *J. Loss Prev. Process Ind.*, vol. 32, pp. 109–119, 2014, doi: 10.1016/j.jlp.2014.07.010.

[22] N. A. Stanton, C. Harvey, and C. K. Allison, "Systems Theoretic Accident Model and Process (STAMP) applied to a Royal Navy Hawk jet missile simulation exercise," *Saf. Sci.*, vol. 113, pp. 461–471, Mar. 2019, doi: 10.1016/J.SSCI.2018.12.020.

[23] C. K. Allison, K. M. Revell, R. Sears, and N. A. Stanton, "Systems Theoretic Accident Model and Process (STAMP) safety modelling applied to an aircraft rapid decompression event," *Saf. Sci.*, vol. 98, pp. 159–166, Oct. 2017, doi: 10.1016/J.SSCI.2017.06.011.

[24] H. A. P. Blom, S. H. Stroeve, and T. Bosse, "Modelling of potential hazards in agent-based safety risk analysis," *Proc. 10th USA/Europe Air Traffic Manag. Res. Dev. Semin. ATM 2013*, 2013.

[25] H. A. P. Blom and A. Sharpanskykh, "Modelling situation awareness relations in a multiagent system," *Appl. Intell.*, vol. 43, no. 2, pp. 412–423, 2015, doi: 10.1007/s10489-015-0651-4.

[26] S. Stroeve, H. Blom, and M. van der Park, "Multi-agent situation awareness error evolution in accident risk modelling," *5th USA/Europe Air Traffic Manag. R&D Semin.*, no. June, pp. 23–27, 2003.

[27] "SKYbrary Aviation Safety." https://www.skybrary.aero/index.php/Main_Page#operational-issues (accessed Oct. 26, 2021).