Results of trials of the Risk Observatory

S. Metge (AI-F), B. Hickling (EUROCONTROL), P. Bieber (ONERA),
W.F.J.A. Rouwhorst, B. van Doorn (NLR), and R. Peixe (CEiiA)

Short abstract: Future Sky Safety (FSS) is a Joint Research Programme (JRP) on Safety, initiated by EREA, the association of European Research Establishments in Aeronautics. The Programme contains two streams of activities: 1) coordination of the safety research programmes of the EREA institutes and 2) collaborative research projects on European safety priorities.

This deliverable (D4.11) is produced by the Project P4 “Total System Risk Assessment”. The main objective is to provide the results of user assessments and feedback sessions with the Risk Observatory prototype tool that were performed by AIRBUS and KLM safety experts.

Programme Manager	Michel Piers, NLR
Operations Manager	Lennaert Speijker, NLR
Project Manager (P4)	Wilfred Rouwhorst, NLR
Grant Agreement No.	640597
Document Identification	D4.11
Status	Approved
Version	2.0
Classification	Public

This page is intentionally left blank

Contributing partners

Company	Name
Airbus (AI-F)	Sylvain METGE
CEiiA	Rui PEIXE
CAA-UK	Matthew WEEKS
CIRA	Domenico PASCARELLA
EUROCONTROL	Brian HICKLING
NAVBLUE	Mathieu TERRISSON
ONERA	Pierre BIEBER
NLR	Bas van DOORN
NLR	Wilfred ROUWHORST

Document Change Log

Version	Issue Date	Remarks
1.0	28-06-2019	First formal release
2.0	28-06-2019	Second formal release

Approval status

Prepared by: *(name)*	**Company**	**Role**	**Date**
Sylvain METGE	Airbus (AI-F)	Author	28-06-2019
Rui PEIXE	CEIIA	Author	28-06-2019
Brian HICKLING	EUROCONTROL	Author	28-06-2019
Bas van DOORN	NLR	Author	28-06-2019
Wilfred ROUWHORST	NLR	Author	28-06-2019
Pierre BIEBER	ONERA	Author	28-06-2019
Checked by: *(name)*	**Company**	**Role**	**Date**
Alex Rutten	NLR	Quality Assurance	28-06-2019
Approved by: *(name)*	**Company**	**Role**	**Date**
Wilfred Rouwhorst	NLR	Project Manager (P4)	28-06-2019
Lennaert Speijker	NLR	Operations Manager	28-06-2019

Acronyms

Acronym	Definition
ANSP	Air Navigation Service Provider
ATCO	Air Traffic Controller
BB	Back Bone
CF	Contributing Factors
CFIT	Controlled Flight Into Terrain
EASA	European Aviation Safety Agency
EC	European Commission
ECAC	European Civil Aviation Conference
FCOM	Flight Crew Operations Manual
FCTM	Flight Crew Training Manual
FDM	Flight Data Monitoring
FSS	Future Sky Safety
GUI	Graphic User Interface
HMI	Human Machine Interface
IF	Influencing Factors
IRP	Integrated Risk Picture
JRP	Joint Research Program
LOC-I	Loss Of Control-Inflight
MAC	Mid Air Collision
MRO	Maintenance, Repair and Overhaul
OAT	Outside Air Temperature
OEM	Original Equipment Manufacturer
RE	Runway Excursion
RO	Risk Observatory
ROO	Risk Observatory Organisation
SMS	Safety Management System
SPI	Safety Performance Indicators
SW	Software
TMA	Terminal Manoeuvring Area
UI	User Interface

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Acronym	Definition
VR	Virtual Reality
WP	Work Package

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

This page is intentionally left blank

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

EXECUTIVE SUMMARY

Problem Area

Safety is a dynamic characteristic of the aviation system, whereby safety risks must be continuously mitigated. In order to be able to do so there is a need for a holistic, total system approach to aviation safety integrated across all components and stakeholders, as set forth in Flight Path 2050. A key pillar to this need is a pro-active and predictive approach to safety management. This implies that being able to monitor risks is a prerequisite for the definition and implementation of appropriate measures for safety management. This has led to the P4 (Total System Risk Assessment) project in which the main aim is to create a Risk Observatory (RO concept and a RO software prototype) employing risk assessment models that are fit for purpose, not overly complex and that can eventually cover the total aviation system (i.e. cover all accident categories, all relevant safety events) and can be easily scoped to stakeholder's need, like per aircraft type, per airport and so on. P4 has created an early RO prototype based on stakeholder requirements. This was followed by a development of the selected risk models, i.e. for Mid Air Collision (MAC) and longitudinal Runway Excursion (RE). The RO prototype was iterated in various drafts and brought to a first and second software release. Part of this process was a final RO assessment by stakeholders.

Description of Work

Within FSS the project has defined so-called "Exploitation Activities" for the final year in the project. For P4 the main objective of these exploitations was the preparation of a RO assessment workshop with the following specific objectives:

- Evaluate how a stakeholder would use the RO and what is the added value in comparison to current practice
- Determine how a lead user can use its own data within the RO
- Identify what is needed to deploy the RO and integrate it in an operational environment for a specific organisation

This has led to organizing these exploitation activities which were held with two different stakeholders, an airline (KLM) and an aircraft manufacturer (Airbus). Kick–off meetings were prepared in which the Risk Observatory concept was presented to (safety) experts in the organisations. Furthermore a prototype version of the RO software tool was demonstrated at these meetings. Both the RO concept and the RO prototype were discussed in detail. Jointly by partners and KLM several use cases were created for the final assessment of the RO software prototype. Subsequently further improvements to the RO software itself were made. This led in May-June to the actual RO assessment where user feedback sessions were held at KLM and Airbus offices separately. The results of those sessions are documented in this report.

Results & conclusions

The main results are described below:

The RO tool itself gave a very good insight in the foreseen RO-concept. It was very well understood that the risk modelling combined (in the future) with occurrence data in one tool is a strong combination to show the risk elements. However the top–event risk for Mid Air Collisions or Runway Excursion is less important than the occurrence risk of the precursor or barriers. The latter show more directly where the problems are for Safety management, but the collection of the proper data on it is much more difficult to organize. Also the RO-models need to be trusted. This implies that there is a need for an indication on the validity of the data used inside the models. Especially in case high-fidelity in-house data are coupled to the integrated risk framework, the other risk modelling aspects should also be stepping up, both in accuracy as well as in reliability.

Enhanced GUI aspects, improved model presentation within the tool, as well as flexibility on the GUI development of the RO software tool suite selection is crucial. Different stakeholders will prefer or even demand different GUIs. Better back-office elements (exporting data functions, storage functions, etc.) are also necessary for an industrial use of the concept and tool. Therefore RO to become an industrial operational tool will still require a lot of research and development work. But the RO concept and its prototype development has a high potential for future safety risk management, see also [14].

Airbus recommends making the RO more non-linear and more dynamic. This is yet a big challenge that will require a big effort that can only be made in the frame of future European R&T projects on safety modelling.

Airbus and also KLM would appreciate to expand the RO-modelling to other Risk Aspects. P4 has shown no show-stopper on that point thanks to discussions made with all concerned P4 partners on the other type of major risks (CFIT, lateral runway excursions, LOC-I, etc). Both Airbus and KLM recommend expanding the RO to other domains: MRO, National Safety Authorities, etc., which is in line with the future outlook provided in [14].

It has been strongly recommended to disclose more accurate / confidential databases. This will require an actual RO organisation to deliver and maintain RO services. In view of this, the key elements of a business model – value proposition, customers, customer relationships, communication channels, key activities, resources, partners and cost and revenue streams – are described in [14].

KLM recommends improving the RO-coupling with the Bowtie's and with other representation of the risks (CATS models…). This would be a good way for improvement.

Lastly, all persons involved in the trial sessions would appreciate to continue the RO-developments towards a future tool for Safety Management.

Applicability

This report is applicable to Safety Management and Safety Analysts in Air Transport stakeholder organisations with a focus on using safety intelligence to allow the implementation of appropriate measures to positively influence safety - i.e. reducing the serious incident and accident probability.

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1. The Programme

The European Commission's (EC) Flight Path 2050 vision aims to achieve the highest level of safety to ensure that passengers and freight as well as the air transport system and its infrastructure are protected. However, trends in safety performance over the last decade indicate that the ACARE Vision 2020 safety goal of an 80 % reduction of the accident rate is not being achieved. A stronger focus on safety is required. Therefore, a Joint Research Programme (JRP) on Aviation Safety - Future Sky Safety (FSS) - is started at the beginning of 2015, aiming for Coordinated Safety Research as well as Safety Research Coordination. Future Sky Safety has goals to coordinate safety research of the involved EREA research establishments and perform safety research and innovation actions targeting the highest levels of safety for European aviation [1],[2].

1.2. Project context

In the FSS project P4 "Total System Risk Assessment", a prototype Risk Observatory (RO) is developed as an enabling tool for safety management, see [1]. The main objectives of the RO concept are translated into software features built into the RO prototype.

The risk observatory will acquire, fuse and structure safety data and translate it into actionable safety information: output that helps the user to distil safety intelligence to allow the implementation of appropriate measures to positively influence safety - i.e. reducing the serious incident and accident probability. The core of the risk observatory is formed by a risk assessment framework that integrates risk assessment models specifically developed to represent a certain domain. The framework is fed by different safety data inputs: e.g. normal operation data from the aircraft operator domain (e.g. originating from Flight Data Monitoring (FDM)) and ANSP domain, but also occurrence and incident data. The risk observatory will offer important insights in safety performance to safety analysts, which can be used in the risk assessment of new aircraft and systems and in safety assurance by identifying safety trends, key risk areas, and efficient mitigation measures. The risk observatory's scope is currently limited to the EASA Member States and the operations performed by service providers within the EASA Member States. Project P4 has as main objective to develop a working and practical Risk Observatory prototype to assess and monitor safety risks throughout the Total Aviation System and allow frequent update of the assessment of risks [2].

1.3. Research objectives

The main purpose of this document is to report on the user feedback on the developed Risk Observatory (RO) prototype as part of the Exploitation Activities within P4. Two different stakeholders were involved in this activity: KLM airlines and Airbus SAS.

1.4. Approach

This document is part of the WP4.4 of the FSS project P4, see [1] - [2] for more detailed information. This work package has developed a framework able to integrate various domain specific models into the RO. This is based on the backbone model approach developed within WP4.2. The RO prototype software has been created in multiple steps using a scrum type of software development process. The RO software and its functionalities has been independently assessed with two different stakeholders, i.e. KLM airlines respectively AIRBUS SAS, a mayor aircraft manufacturer. This document reports on that user assessment. Some proposed modifications have already been introduced inside the latest version 2 of the RO. The new proposed aspects that would not be ready in time before end of the project are regarded recommendations for the future.

1.5. Document structure

The document is structured as follows:

- Section 1 introduces the background and main purposes of this document;
- Section 2 describes the exploitation actions performed.
- Section 3 describes the usecases developed and executed.
- Section 4 describes the results from the KLM RO assessment sessions and the feedback received from Airbus internal sessions.
- Section 5 contains the concluding remarks.
- Appendix A provides the questionnaires used.
- Appendix B provides the individual results on the usecase questionnaires and discussions.

2 EXPLOITATION ACTIONS

2.1. Introduction

As part of the extended P4-activities, named Exploitation Actions, the main Risk Observatory development work performed by P4 (Total System Risk Assessment), consisting of development of both the RO concept and a RO prototype software tool, have been brought to two different stakeholders for exploitation:

- KLM
- AIRBUS SAS

The general idea behind this exploitation activity was to give to various experts on-site exposure and in-depth explanations on the RO-concept, meaning to the conceptual idea behind it, and on the RO prototype developments. Furthermore its purpose was to show the current RO software prototype status, the modelling used and to gather feedback on the usefulness of the RO, and the future RO needs of the end user/stakeholder having each a different interest: an airliner and an OEM.

The objectives defined for the Exploitation Actions were preparation of workshops/seminars:

- To evaluate how a lead user would use the RO and what is the added value in comparison to current practice
- To determine how a lead user can use its own data within the RO
- To identify what is needed to deploy the RO and integrate it in an operational environment for a specific organisation

As a Deliverable/Output of this Exploitation Activity it was defined: To have an evaluated RO software application (alpha version), including recommendations for further development and exploitation.

As will become clear in this report, the first two objectives have been fulfilled. The two stakeholders have evaluated the use of the RO and the impact on their way of working. Given the fact, as confirmed by KLM and Airbus, that the RO tool should first be further expanded and enhanced and provided with some more features to work with, the identification of the deployment aspects could only be limitedly executed, hence the third objective has only be partly met. Based on the stakeholder's feedback the RO was updated and has been successfully delivered in a second version [10]. As such the promised output has been delivered.

2.2. First meeting at KLM offices

At 19th February 2019 a kick-off meeting annex Workshop was held with KLM at their offices near Amsterdam Airport Schiphol, to introduce some of their operational safety experts to the RO-concept and the RO prototype software tool. Extensive presentations and discussions were held on the RO concept, its evolution over time and on the details of the Back Bone (BB) models used on the chosen risk aspects worked out inside the RO: longitudinal Runway Excursion (RE) and Mid Air Collision (MAC). See [3],

A first set of RO assessment usecases were launched, presented and discussed. KLM suggested some potential content for a few new usecases. Based on this meeting the initial set of usecases were further refined and it was jointly explored what could and what could not be done with the RO software in relation to these usecases, and whether still some RO software modifications could be introduced to facilitate these usecases.

2.3. Second meeting at KLM offices

At the 21th of May the actual RO assessment day was held at KLM offices were the usecases set up were presented to KLM's involved safety persons, see [5]. Most of the usecases were executed with the running version of the RO tool existing at that time.

2.4. Meetings held at Airbus SAS

At 10th of April a similar Exploitation event took place at AIRBUS offices in Toulouse. The RO concept, the models and the RO-tool were presented and a demonstration of the tool was provided. See [4].

It was the opportunity to get direct feedback from Airbus safety specialists both involved in SMS and in aircraft safety assessment. A specialist in Human Factors involved in accident investigations took partly part in the meeting.

Feedback from this session is provided in sections 2.4 and 4.2.

Mid of June, further internal discussion sessions were held within Airbus. Results of these discussions are summarized below.

During these meetings the following three topics were addressed:

- Feedback from Airbus on Human errors
- Feedback from Airbus on the modelling approach
- Feedback from Airbus on P4 scope

Details on these three topics will be provided next.

Feedback from Airbus on Human errors

Human errors (actions) as defined in ICAO Annex 13 def. (i.e. actions, omissions…which if eliminated, avoided or absent would have reduced the probability of the accident/ incident) fits the notion of 'Influencing Factors" used in P4 to address Human behavior. However Human Factors in P4 project was partly addressed through some 'general' and limited number of influencing factors both for flying crew and ATCo. However this cannot be considered as a strong limitation since new IF's can easily be defined.

Human errors are very difficult to predict; when considered in the risk models it may introduce biases in the final results.

The correlation between human errors and system failures are not indeed inherently linked together. This can lead again to erroneous results if considered in combination. This is why we should use human errors in a very careful way in the P4 safety models.

Instead of considering the Human Errors in the models, Airbus specialist in Human Factors recommends to work on the way to characterize the exposure of the flight crew to unsafe situations, by identifying the combination of events (failures, adverse weather conditions) that lead to a degraded (unsafe) situation. Such result (precursors) could be used to define training scenarios to be used in sessions dedicated to Flight crew from airlines to be familiar and even qualified to face all such situations.

Feedback from Airbus on the modelling approach

Airbus has considered that P4 method and tool (RO) are more dedicated to SMS and thus more appropriate to Airlines and ANSP than to aircraft manufacturers and system providers. However Airbus may use the P4 approach in the near future for quick safety assessment of new concepts of flying an aircraft in the future where the flying crew would rely on a strong support from the ground or should interact with other aircraft.

P4 modelling approach has been considered too much static. However the today safety models used in the design offices for the validation of aircraft architectures as defined in ARP4754 are static models. More dynamic models are more dedicated to performances and handling quality assessment.

Feedback from Airbus on P4 scope

The two following main comments from Airbus are related to the P4 scope and on possible extension of the initial scope as defined in the project Plan.

- MEL aspect could have been taken into consideration. P4 partners pointed out that MEL could be yet partly managed thanks to the use of the Local Reference Data, which is an improvement discussed with KLM and that have been implemented in the last version of the RO. Setting some parameters with a rate of occurrence equal to 1 is a way to consider NOGO equipment as per MEL.
- Maintenance errors could have been taken into consideration. P4 partners pointed out that P4 has considered maintenance out of scope of the project in the project plan. Airport Operators and Maintenance Operators stakeholders of the Aviation Transport System were not in in the FSS P4 consortium. This is why P4 partners had no possibility to address maintenance errors in absence of available data managed by MRO's and no expertise on modelling aircraft maintenance safety impacts.

3 KLM USECASE DESCRIPTIONS AND MAIN RESULTS

Nine usecases were set up for the RO assessment at KLM, see Table 1.

Table 1 Usecase Overview

#No	Title
UC-1	General User Feedback on RO software (as a whole)
UC-2	RE-Model (2 IFs): influences the Tail Wind and Crew Fatigue Assessment
UC-3	RE-Model (3 IFs; Combined 2 previous IFs with IF500.1 “RWY surface quality”
UC-4	RE-Model Flap Lock Out combined with Adverse (Stormy / Gusty) Weather
UC-5	MAC-Model Loss of Separations Safety analysis for High Density TMA
UC-6	MAC-Model : Perform Sensitivity Analysis
UC-7	Baseline Versus KLM Reference Data Comparison
UC-8	RO versus KLM Bow-tie model on RE
UC-9	KLM’s Airprox Bow-Tie versus MAC-modelling

Prior to running the usecases by the KLM users (assessors) themselves firstly a full introduction was given to the RO software tool, essentially demonstrating all main features to them first. These feature consisted of:

- login
- home page
- Risk pictures and occurrence dashboard
- RE and MAC Risk models
- What-if? Analysis
- Options Menu
- The way to select and set Contributing Factors (CF) Influencing Factors (IF)
- The way to Manage Risk Scenarios
- The Set Local Reference data (instead of using default / baseline data inside the risk models

Descriptions on all these functions can be found in D4.9, see [6].

The usecases presented in Table 1 will be further detailed in the next sections.

3.1. Usecase-1 General User Feedback

3.1.1. Description of UC-1

The aim of Use Case 1 was to assess the basic functions and Graphic User Interface (GUI) of the RO software prototype as a whole.

KLM users (i.e. assessors) involved in the RO trial sessions played with the RO, its risk models and set parameters (modification of some default values of the Contributing Factors and Influencing Factors). Then they ran the backbone models and saw how it all worked.

UC-1 has been defined to manage the risk scenarios features, the ability to create user scenarios for a given context as for example a specific airport or for a specific aircraft type.

The Risk Observatory HMI, also named Graphic User Interface (GUI) was assessed regarding ease of use and Intuitiveness.

Using the RO prototype enables KLM users to identify 'Missing' (function) aspects to be derived for the prototype and/or determine how to improve the future RO concept. UC-1 was thus intended to collect KLM users' feedback on required corrections and expected improvements (not only the ones to be implemented in short-term but also nice to have in next versions).

3.1.2. Main Assessment Results

The explanation on how the RO precisely worked and the GUI aspects involved took considerable more than planned time, but this was regarded a necessity to allow the KLM assessors to understand the basic principles of the RO-tool at hand.

KLM Users expressed that they could easily change values for each contributing factor and values for each influencing factor. It was reminded to KLM that only failure rates of the IF can be modified by using the GUI, not the default values of the weights; see [8] and [9]. This is not a limitation of the RO but a requirement linked to the method that has been developed. Fixed IF weight values must indeed be shared by all users whatever the Organization they belong to. The reason is that the weight represents a penalty associated to each IF. It must be the same for all users as for example the weight defined for a contaminated runway that does not depend on a specific airport. Contaminated runways are strictly defined in standard documents like for example: 'Medium quality' corresponds to:

- Dry snow: More than 3 mm up to 100 mm
- Wet snow: More than 3 mm up to 30 mm
- Compacted snow: OAT above -15°C
- Dry snow over compacted snow
- Wet snow over compacted snow
- Slippery when wet

Consequently, braking deceleration is noticeably reduced for the wheel braking effort applied. Directional control may be reduced. Only the rate of occurrence of such IF's can be customized since the frequency of encountering such a runway quality may vary from one airport to another ones especially when they are located in different countries.

KLM has commented and reported (via the questionnaires, see Appendices) that the RO Graphic User Interface (GUI) needs improvements. Some of these comments have been implemented directly after the trial session like for example adding information fields for each parameter that can be modified by the user. This gives a more clear understanding of the parameters. The following KLM requests for improvement were also taken into account:

- What is actually precisely presented on the horizontal axes? This was clarified
- Data/Results out of Baseline, default and local references were clarified
- Colors used in risk pictures / What-if picture was harmonized.

Lastly the following suggestions for improvement have also been recorded during the trial session:

- Enhance the way of managing the scenarios: Delete, copy/ paste (duplicate) a scenario, renaming it.
- The function of the horizontal axis of the risk plots had to be revised and this was incorporated into the RO directly after the session.
- To add an explanatory Legend for each precursor plot
- Harmonization of the Menu options & features
- The headings of the graphs and the Scenarios need to be improved
- Pop up text on Occurrence Dash board needs to be improved
- Occurrence Dashboard now has risk pictures, see[11],[12] and [13] but Safety Management may need other data
- FDM data incorporation for Occurrence Dash board: more SPI's needed and more interfaces to couple own data
- Improvement of the mental picture of the risk models for the RO users: how does the model (fault three) structure look like, where am I precisely working inside the model. How many risk branches and risk parameters are there involved, etc. It was possible to take this request into account directly after the session; Hence a backbone models structure presentation field has been added in Information field. But it was also suggested for the future to work with something like the SANKEY diagram representation, or even more wildly: use the incorporation of VR-tools.

3.2. Usecase-2 RE-Model (2 IFs); Tail Wind and Crew Fatigue Assessment with RO

3.2.1. Description of UC-2

The goal of UC-2 is to assess the potential of using Influencing Factors (IFs) via a tailwind & crew fatigue example. By changing relevant Influencing Factors (IFs), the following matters are:

- Individual effect of tailwind
- Individual effect of flight crew fatigue (long-haul flight has been considered)
- The combined effect of tailwind and flight crew fatigue on the increase risk of doing a longitudinal runway excursion.

The following six different scenarios have been considered:

Scenario 1 (UC-2.1): use of the baseline rates of occurrences for both IF's

Scenario 2 (UC-2.2): Baseline rate for flight crew fatigue combined with an intermediate rate of occurrence for the tail wind

Scenario 3 (UC-2.3): Baseline rate for flight crew fatigue combined with a pessimistic rate of occurrence for the tail wind

Scenario 4 (UC-2.4): Use of pessimistic rate of occurrence for the flight crew fatigue combined with baseline value for the tail wind

Scenario 5 (UC-2.5): Use of pessimistic rate of occurrence for the flight crew fatigue combined with an intermediate value for the tail wind

Scenario 6 (UC-2.6): Use of pessimistic rate of occurrence for both IF's.

Note: The values of the Contributing Factors are unchanged (baseline)

The following Table 2 gives the values of the rates of occurrences used in the 6 scenarios (i.e. sub-use cases).

Table 2 Scenarios of Usecase-2

Scenario	**IF503.4** **Crew fatigue long haul operation** L: Low M: Moderate H: High	**IF501.1 Tail wind** MH: Moderate head wind SH: Strong head wind MT: Moderate tail wind ST: Strong tail wind
UC2.1	Baseline rates: L (80%) – M (15%) – H (5%)	Baseline rates: MH (80%) – SH (14%) – MT (5%) – ST (1%)

Scenario	**IF503.4** **Crew fatigue long haul operation** L: Low M: Moderate H: High	**IF501.1 Tail wind** MH: Moderate head wind SH: Strong head wind MT: Moderate tail wind ST: Strong tail wind
UC2.2	Baseline rates: L (80%) – M (15%) – H (5%)	Intermediate case: MH (30%) – SH (10%) – MT (10%) – ST (50%)
UC2.3	Baseline rates: L (80%) – M (15%) – H (5%)	Worst case: MH (0%) – SH (0%) – MT (0%) – ST (100%)
UC2.4	Pessimistic rates: L(5%) – M(15%) – H (80%)	Baseline rates: MH (80%) – SH (14%) – MT (5%) – ST (1%)
UC2.5	Pessimistic rates: L(5%) – M(15%) – H (80%)	Intermediate case: MH (30%) – SH (10%) – MT (10%) – ST (50%)
UC2.6	Pessimistic rates: L(5%) – M(15%) – H (80%)	Worst case: MH (0%) – SH (0%) – MT (0%) – ST (100%)

3.2.2. Main Assessment Results

The numerical (risk) results of the above scenarios are given in Table 3. The top row shows the derived RE main risk values in blue (the top event risk). The five rows below show the precursors and the associated risk values.

Clearly the risk outcome of the three last rows (with numerical values in red) show, for UC2.1 till UC2.6, which precursor has the highest contribution to the RE risk.

Table 3 RO RE-Risk Assessment Results (long haul operations)

Risk \ Sub-Use cases	UC2.1	UC2.2	UC2.3	UC2.4	UC2.5	UC2.6
BB000a - Longitudinal Runway Excursion	6.53E-06	7.10E-06	7.79E-06	8.04E-06	8.94E-06	9.75E-06
BB000b - Non-decelerated incorrect touchdown	6.52E-06	7.09E-06	7.78E-06	8.03E-06	8.93E-06	9.74E-06
BB000c - Non-rejected incorrect touchdown	2.17E-04	2.33E-04	2.56E-04	2.45E-04	2.73E-04	2.98E-04
BB000d - Incorrect touchdown	2.17E-01	2.31E-01	2.53E-01	2.25E-01	2.50E-01	2.73E-01
BB000e - Non-corrected unstable approach	3.13E-02	3.25E-02	3.40E-02	3.51E-02	3.69E-02	3.86E-02
BB000f - Unstable approach	3.13E-01	3.22E-01	3.37E-01	3.22E-01	3.39E-01	3.54E-01

The next two figures are related to graphical results.

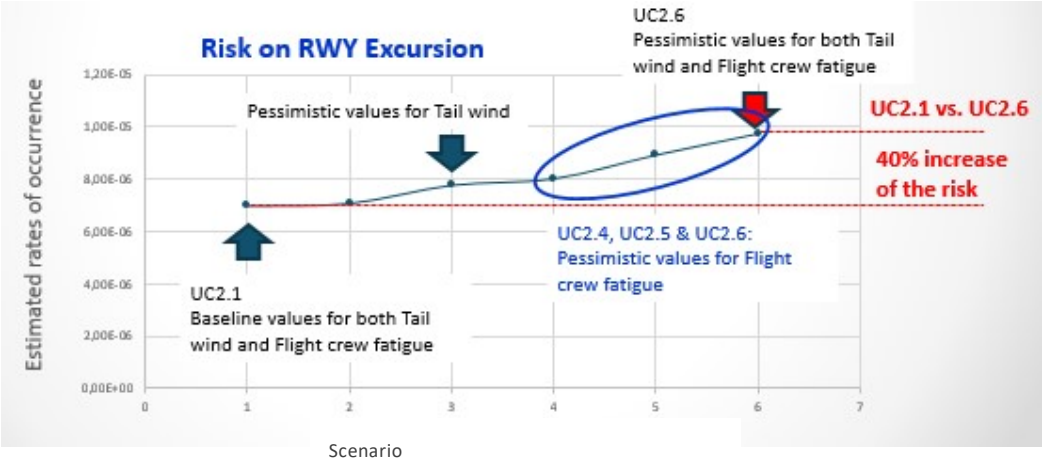

Figure 1 – RE– main risk values for Scenario 1 till 6 (i.e. for UC-2.1 till UC-2.6)

This Figure 1 graph highlights the calculated probability of a runway excursion. An interesting result is the comparison between scenario UC-2.1 and UC-2.6. Considering pessimistic values for both IF's show the estimated increase of the risk. The graphs inside Figure 2 focus on the precursors. The graphs on the left part are related to the incorrect touchdown while the graphs of the right part deal with the risk of unstable approach.

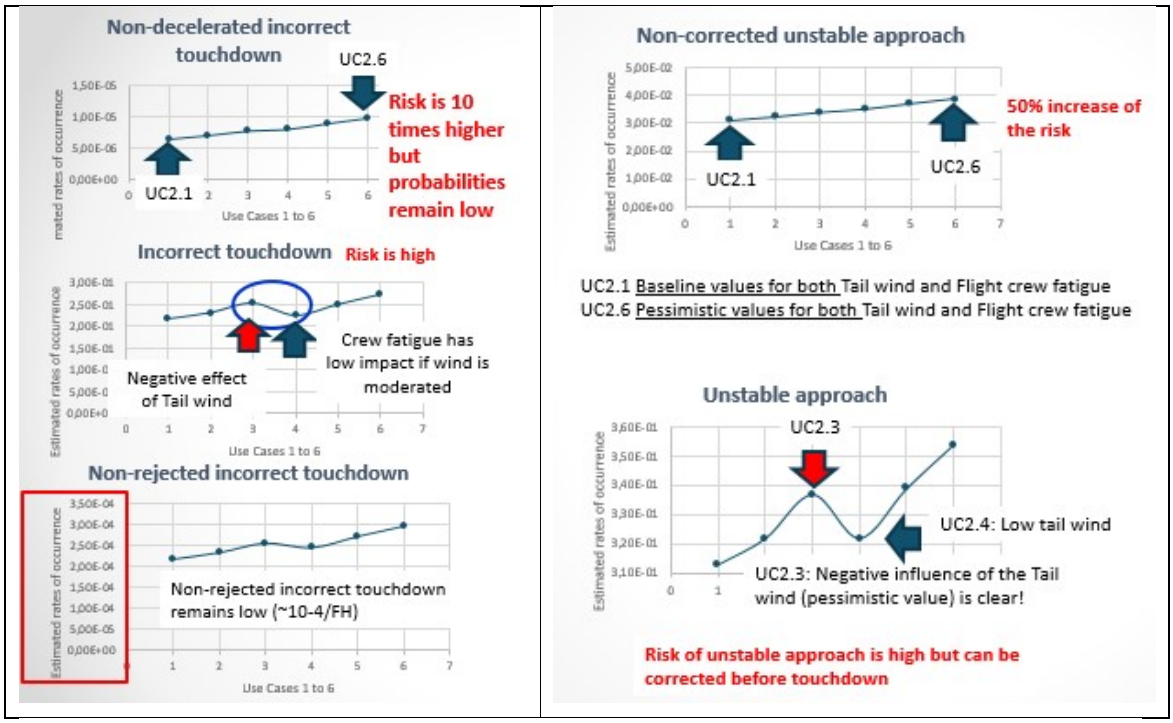

Figure 2 – RE–risk values of precursors for Scenario 1 till 6 (UC-2.1 till UC-2.6)

Afterwards, also the scenario results related to the short haul operations were produced for sake of completeness, see Table 4.

Table 4 RO RE-Risk Assessment Results (short haul operations)

Risk \ Sub-Use cases	UC2.7	UC2.8	UC2.9	UC2.10	UC2.11	UC2.12
BB000a - Longitudinal Runway Excursion	**6.53E-06**	**7.27E-06**	**7.94E-06**	**7.84E-06**	**8.71E-06**	**9.54E-06**
BB000b - Non-decelerated incorrect touchdown	6.52E-06	7.26E-06	7.93E-06	7.83E-06	8.70E-06	9.53E-06
BB000c - Non-rejected incorrect touchdown	2.17E-04	2.37E-04	2.5E-04	2.42E-04	2.68E-04	2.94E-04
BB000d - Incorrect touchdown	2.17E-01	2.32E-01	2.54E-01	2.24E-01	2.48E-01	2.72E-01
BB000e - Non-corrected unstable approach	3.11E-02	3.30E-01	3.45E-02	3.46E-02	3.64E-02	3.81E-02
BB000f - Unstable approach	3.11E-01	3.23E-01	3,38E-01	3.21E-01	3.37E-02	3.53E-02

The KLM assessors provided positive feedback during and after executing this usecase:

- Many factors are included. Very positive.
- The potential with good and accurate data is great.

However they were more circumspect regarding the following points:

- How to validate the numerical (risk) results?
- The RO tool has to gain confidence by users; it would be useful to show a kind of "result confidence indicator" and to improve data source information.

In term of way for improvement they recommended to:

- To provide a mental picture of the Risk Model used in the RO. This has been implemented in the latest version of the RO prototype via an information graphics.
- To consider new technologies for the GUI like e.g. 3D graphical results.

3.3. Usecase-3 RE-<Model (3xIFs ;Tail Wind and Crew Fatigue combined with IF500.1 "RWY surface quality")

3.3.1. Description of UC-3

The goal of UC-3 is to extend the UC-2 scope in order to show the potential of using multiple (max. three) Influencing Factors (IFs) via a tailwind & crew fatigue example combined with a third IF, namely IF 500.1 'Runway surface quality'.

Expected result was to assess the effect of the 3 IFs on the runway excursion (RE) risk and on the probability of precursors for runway excursions. Twelve scenarios were defined, The first six scenarios, see Table 5, were related to the long haul crew operation, and are considering a baseline rate of occurrence for IF 500.1 'RWY quality' combined with the two other IF's. For Crew Fatigue, the rates are varied for Low (L), Medium (M) and High (H) occurrences. For Tailwind the rate selections are Moderate Headwind (MH), strong Headwind (SH), Moderate Tailwind (MT) and Strong Tailwind (ST), while for RWY quality the Good (G) and Poor (P) selections should be made.

Table 5 Scenarios for long haul operations crew fatigue under different tail wind and runway conditions

Scenario	IF503.4 Crew fatigue long haul operation	IF501.1 Tail wind	IF500.1 RWY quality
UC3.1	Baseline rates: L (80%) – M (15%) – H (5%)	**Baseline rates**: MH (80%) – SH (14%) – MT (5%) – ST (1%)	Baseline IF500.1 rates: G (90%) – P (10%)
UC3.2	Baseline rates: L (80%) – M (15%) – H (5%)	Intermediate case: MH (30%) – SH (10%) – MT (10%) – ST (50%)	Baseline IF500.1 rates: G (90%) – P (10%)
UC3.3	Baseline rates: L (80%) – M (15%) – H (5%)	Worst case: MH (0%) – SH (0%) – MT (0%) – ST (100%)	Baseline IF500.1 rates: G (90%) – P (10%)
UC3.4	Pessimistic rates: L(5%) –M(15%) – H (80%)	**Baseline rates**: MH (80%) – SH (14%) – MT (5%) – ST (1%)	Baseline IF500.1 rates: G (90%) – P (10%)
UC3.5	Pessimistic rates: L(5%) –M(15%) – H (80%)	Intermediate case: MH (30%) – SH (10%) – MT (10%) – ST (50%)	Baseline IF500.1 rates: G (90%) – P (10%)
UC3.6	Pessimistic rates: L(5%) –M(15%) – H (80%)	Worst case: MH (0%) – SH (0%) – MT (0%) – ST (100%)	Baseline IF500.1 rates: G (90%) – P (10%)

RWY surface quality: G: Good P: Poor

3.3.2. Main Assessment Results

The KLM assessors did not run all scenarios with the RO tool. They divided the work in performing three each to get the hands-on experience with three IFs and showing the basic idea. Instead, given the time available, it was decided to focus on the RE model-related use cases UC-5 and UC-8.

For KLM the initial interest was in risk of overrun due to crew fatigue, bad weather on approach after long haul operations. But it is known that short haul operations also could generate a lot of crew fatigue due to the high amount of take-off and landings performed during one day. While the other two IFs (tailwind and a poor runway condition) would then contribute to an increased RE risk. Therefore it is remarked that a similar RO exercise could be performed for the crews performing short haul operations. This formed part of the other 6 scenarios (not listed in this report) belonging to the Usecase-3. Executing all 12 scenarios would therefor also allow a RO-comparison between the results of long haul and short haul operations with the RO-tool.

Figure 3 and Figure 4 provide the numerical outcome of the foreseen exercise.

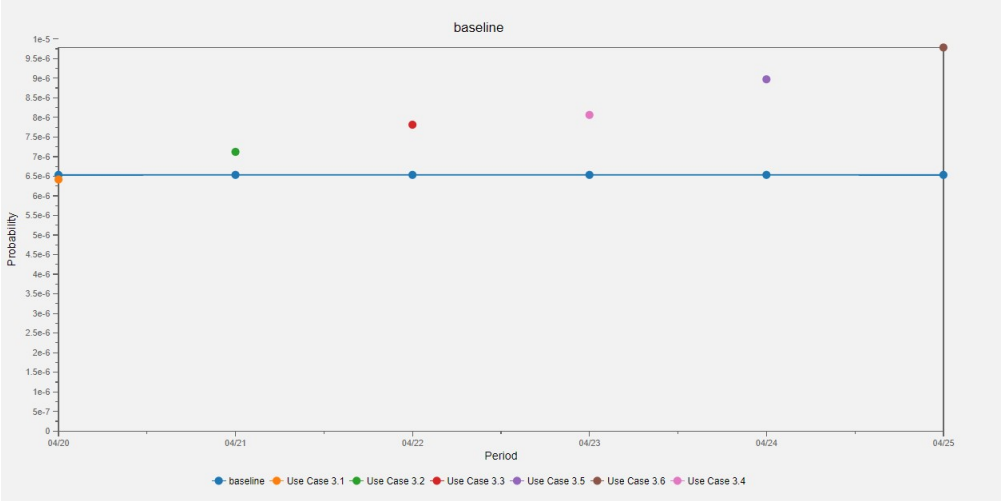

Figure 3 – RE–risk values of precursors for UC-3.1 till UC-3.6 (long haul operations)

Table 6 RE-Risk and precursor values of UC-3.1 till UC3.6 (long haul operations)

Risk \ Sub-Use cases	UC3.1	UC3.2	UC3.3	UC3.4	UC3.5	UC3.6
BB000a - Longitudinal Runway Excursion	**6.53E-06**	**7.12E-06**	**7.81E-06**	**8.06E-06**	**8.97E-06**	**9.78E-06**
BB000b - Non-decelerated incorrect touchdown	6.52E-06	7.11E-06	7.80E-06	8.05E-06	8.96e-06	9.77E-06
BB000c - Non-rejected incorrect touchdown	2.17E-04	2.33E-04	2.56E-04	2.45E-04	2.73E-04	2.98E-04
BB000d - Incorrect touchdown	2.17E-01	2.31E-01	2.53E-01	2.25E-01	2.50E-01	2.73E-01
BB000e - Non-corrected unstable approach	3.13E-02	3.25E-02	3.40E-02	3.51E-02	3.69E-02	3.90E-02
BB000f - Unstable approach	3.13E-01	3.22E-01	3.37E-01	3.22E-01	3.34E-01	3.54E-01

As can be observed from the above table, the last three rows with precursors contribute the most to the overall risk.

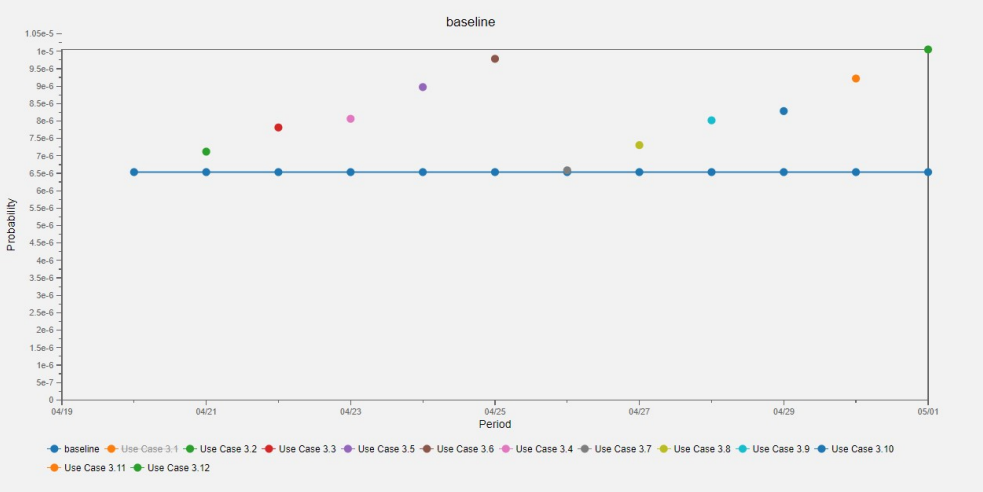

Figure 4 – RE–risk values of precursors for all usecases including UC-3.7 till UC-3.12 (short haul)

Table 7 RE-Risk and precursor values of UC-3.7 till UC3.12 (short haul operations)

Risk \ Sub-Use cases	UC3.7	UC3.8	UC3.9	UC3.10	UC3.11	UC3.12
BB000a - Longitudinal Runway Excursion	**6.58E-06**	**7.31E-06**	**8.01E-06**	**8.28E-06**	**9.21E-06**	**1.0E-05**
BB000b - Non-decelerated incorrect touchdown	6.57E-06	7.3E-06	8.01E-06	8.27E-06	9.20E-06	1.0E-05
BB000c - Non-rejected incorrect touchdown	2.1E-04	2.33E-04	2.56E-04	2.45E-04	2.73E-04	2.98E-04
BB000d - Incorrect touchdown	2.08E-01	2.31E-01	2.53E-01	2.25E-01	2.50E-01	2.73E-01
BB000e - Non-corrected unstable approach	3.1E-02	3.25E-02	3.40E-02	3.51E-02	3.69E-02	3.86E-02
BB000f - Unstable approach	3.1E-01	3.2E-01	3.36E-01	3.22E-02	3.39E-01	3.54E-01

As can again be observed in the above Table 7, the last three rows have the highest precursor contributions. Also UC3.12 has a slightly higher RE risk compared to the other UC3.7 till UC3.11.

3.4. Usecase-4 RE-Model Flap Lock Out combined with Adverse (Stormy/Gusty) Weather

3.4.1. Description of UC-4

The goal of UC-4 was to show the potential of using the RO for assessing the effect on longitudinal runway excursion risk of a flap lock out during approach in combination with adverse weather conditions.

Expected result was to assess on runway excursion (RE) risk the probability of precursors of runway excursions with a flap lock out during approach.

"Windshear/ turbulence" IF501.2 has been selected by considering default rates with 2 different flap angles: 15 and 25 degrees.

This use case has used estimated failure rates of flap lock out coming from KLM statistics over a period of 7 years (2012-2018) of reported occurrences. See Table 8. Apart from the number of flights (in the second column), the number of occurrence per flap setting, the derived failure rates are shown in the fourth respectively the sixth column of the table. Also shown are the average failure rate values for Flaps 15 and Flaps 25. See below the last table row.

Table 8 Flap failure rate occurrence statistics

	# of flights (in AGS)	Flaps 15	Failure Rate Flap 15	Flap 25	Failure Rate Flap 25
2012	73221	3	4.10E-05	1	1.37E-05
2013	75429	7	9.28E-05	0	0.00E+00
2014	82968	5	6.03E-05	5	6.03E-05
2015	78363	5	6.38E-05	4	5.10E-05
2016	81324	11	1.35E-04	1	1.23E-05
2017	78009	4	5.13E-05	2	2.56E-05
2018	73157	7	9.57E-05	3	4.10E-05
		Average value	7.72E-05		2.91E-05

Using these average failure rate values, the following four scenarios have been defined for this use case-4, see Table 9 below.

Table 9 Scenarios of Use case-4

Scenario	GCF4 (flap failure)	IF501.2 Wind shear/ turbulence **Wind shear/ turbulence**: N: None/ light M: Moderate S: Severe
UC4.1	2.9E-5/FH	Baseline IF500.1 rates: N (80%) – M (19.5%) – S (0.5%)
UC4.2	2.9E-5/FH	Pessimistic rates: N (0.5%) – M (19.5%) – S (80%)
UC4.3	7.7E-5/FH	Baseline IF500.1 rates: N (80%) – M (19.5%) – S (0.5%)
UC4.4	7.7E-5/FH	Pessimistic rates: N (0.5%) – M (19.5%) – S (80%)

3.4.2. Main Assessment Results

The version of the backbone model implemented in the RO has to be slightly updated considering the flap lock out failure needed a new Contributing Factor. Different ways to modify the model were discussed with KLM.

First option discussed with Airbus was to consider the flap lock issue as an additional potential cause leading to unstable approach as depicted in the following Figure 5:

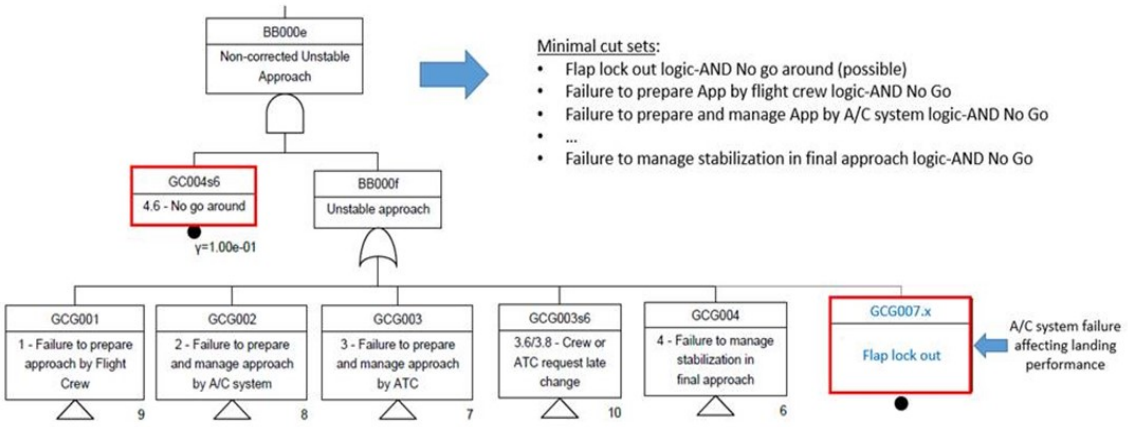

Figure 5 – Contributing Factors in RE-model initially considered relevant for flap lock put

According to KLM, the Flap Lockout issue is typically a performance problem and not a controllability problem. It would therefore be incorrect to chain this issue to the "unstable approach" precursor (see BB000f in Figure 5). In a linear model (which is the case with the RE backbone model), "unstable approach" would not be the proper precursor.

KLM considered that we should instead focus on "inadequate landing performance". See in Figure 6 below the existing and high-lighted Contributing Factor GC001s3. However this CF was defined to focus on flight crew errors and not on technical failures.

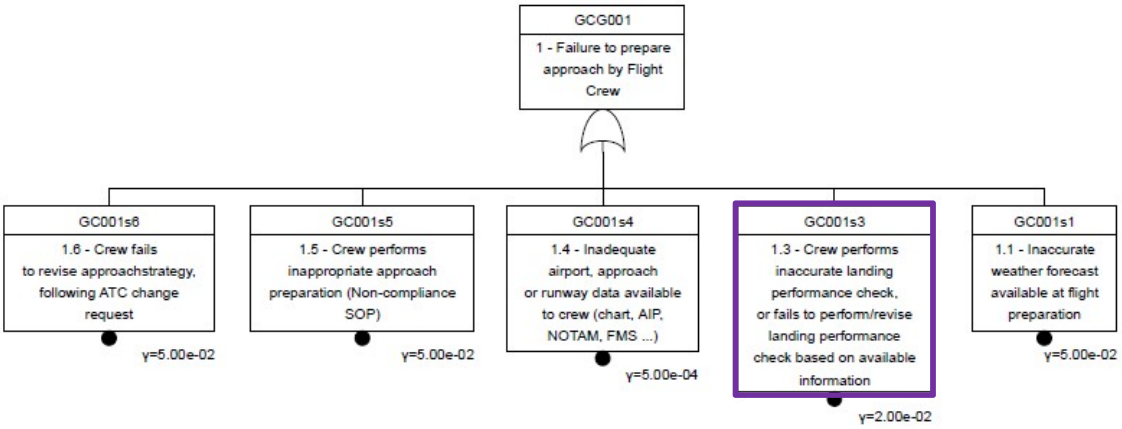

Figure 6 – Other Contributing Factors in RE-model initially considered relevant for flap lock put

Considering the KLM comment, Airbus suggested to combine the new Contributing Factor (Flap lock out) with 'Failure to prepare the approach by Flight crew' as depicted below in Figure 7. This should be considered a suggestion for a future update of the RO software version (beyond the P4 project).

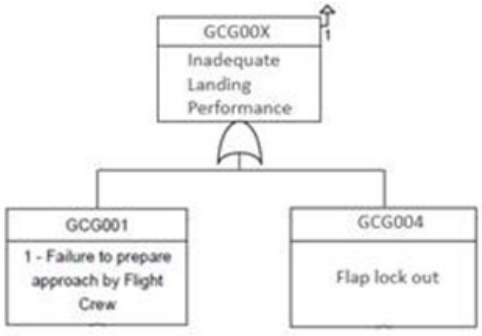

Figure 7 – New proposed Contributing Factors in RE-model considered relevant for flap lock out

Eventually none of the options was considered fully satisfactory. This is why it was agreed to disregard UC-4 in the frame of the trials given that the best change in the model structure would require considerable changes to the modeling, which KLM did not consider feasible. At this point KLM proposed to refrain from quantifying the UC-4 scenarios, which was accepted.

3.5. Usecase-5 MAC-Model Loss of Separations Safety analysis for High Density TMA

Within the RO-tool the user can also select a Mid Air Collision (MAC) risk model. This model has both Ground, (hence ATC) and Air, so cockpit crew, aspects inside.

3.5.1. Description of UC-5

This usecases intended to demonstrate the MAC–model and its risk calculation aspects both for arlines and ANSPs. Special in this usecase was to show the potential of using the RO for:

- Assessing the effect of TMA operations in a busy European TMA (using operational, meaning reported occurrence, data from 3 years) in relation to a mid-air collision and its pre-cursors.
- Determine the impacts of key airborne risks in comparison with the ECAC baseline.

The expected result was to assess the specific pilot/crew related risks in TMA operations on Mid Air Collision (MAC) risk and on the probability of the precursors. Note that this is a special usage of the MAC model, since this model was dedicatedly set up for Enroute airprox situations. See [9].

Three scenarios were set up for a loss of separations safety analysis for High density TMA, see Table 10 using dedicatedly for this purpose received ANSP data to feed the GCF's listed inside the table. Hence an example of coupling user-owned data in combination with the RO- MAC model. showing the RO potential for ANSPs as another stakeholder.

Table 10 MAC-Scenario for High Density TMA

Scenario	MAC-model Risk Outcome for High Density TMA data
UC5.1: (GCF 31.2) Inappropriate crew response to RA	0.25 (baseline is 0.1001)
UC5.2: (GCF 33.4) Communications Issues Misunderstanding)	7.0E-03 (baseline is 5.0E-04
UC5.3: 33.8 Inappropriate Crew	1.9E-03 (baseline is 4.0E-04)

Scenario	MAC-model Risk Outcome for High Density TMA data
response to ATC instructions	
UC5.4: All the negative impacts (UC5.1+ UC5.2 + UC5.3) plus the observed major positive safety impacts in the TMA operation on GCF	GCF (35.1): Pilot Error causes Airspace Infringement; Risk value is 2.6E-05 (baseline is 1.59E-04) GCF (36.1): Pilot Action Induces Deviation; Risk value is 1.9E-04 (baseline is 1.02-04) GCF(36.2): ATC Action/Information Induces Deviation; Risk value is 4.0E-05 (baseline is 2.26E-04) 37.2 Communication Issues – misunderstanding 2.8E-05 (baseline is 6.08E-05)

3.5.2. Main Assessment Results

The RO's MAC-model risk outcome on the first three scenarios is presented in Figure 8. The figure shows (see colored dots on the right side) that the UC5.1 (Inappropriate crew response increases the accident risk with 97%; UC5.1 (Communications Misunderstanding) increases the risk with 300% !!, while UC5.3 (Inappropriate Crew response to ATC) increases the MAC-risk with 44%, all compared to the baseline (blue dot). Hence, this implies that the misunderstanding of ground-air communications (UC5.2) is seen to be the key airborne increase (3 fold) in accident risk in busy TMA operations. While an increased risk due to ignoring or late application of ACAS RA maneuvers (due to a higher prevalence of false alarms in TMA) and the risk of an overreaction leading to knock-on conflicts (higher levels of vertically maneuvering traffic) is also observed.

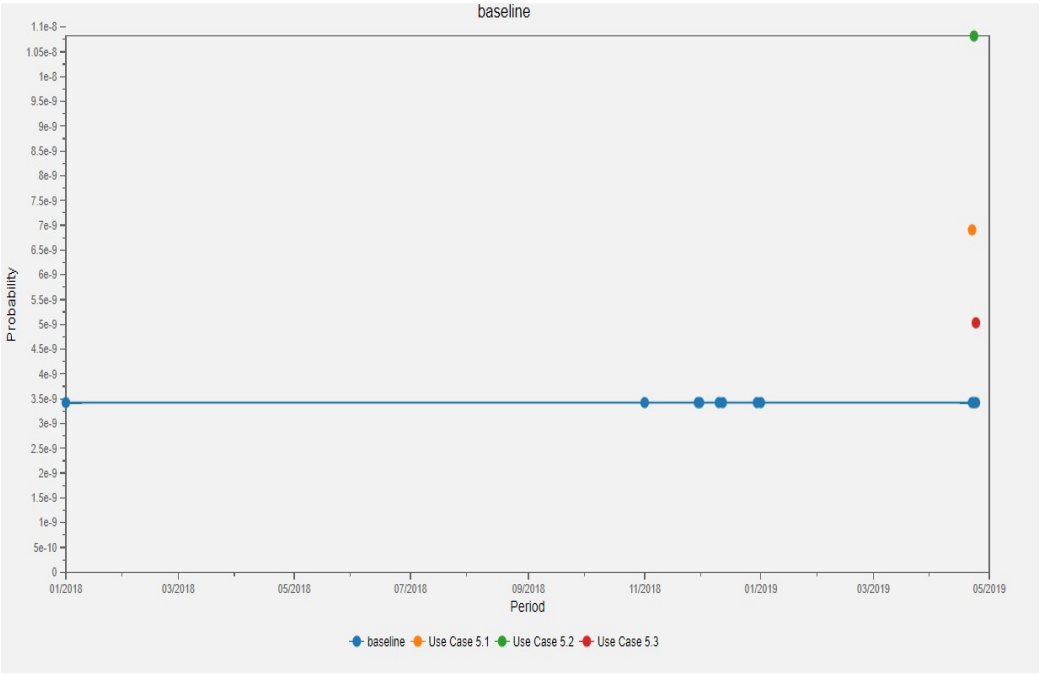

Figure 8 – MAC-model Risk calculations outcome, first three scenarios

Performing the risk calculations for the 4th scenario (thus for UC5.4), where we have all the negative impacts of (UC5.1+ UC5.2 + UC5.3) plus the observed major positive safety impacts in the TMA operation on GCF. Then we get the risk results as presented in Figure 9 and as given in the right side column on the last row of Table 10.

This implies that while the ECAC RO Accident risk from Integrated Risk Picture (IRP) data over 2005-2015 is derived as being 3.5E-09, that the busy TMA Risk from 3 years TMA data (see UC5.4) is 2.8E-08, about a factor 10 higher.

Much of the increased risk observed is because we are comparing a busy TMA operation with overall ECAC MAC risk where the average traffic density and number of maneuvers per flight hour are a small fraction of this complex environment.

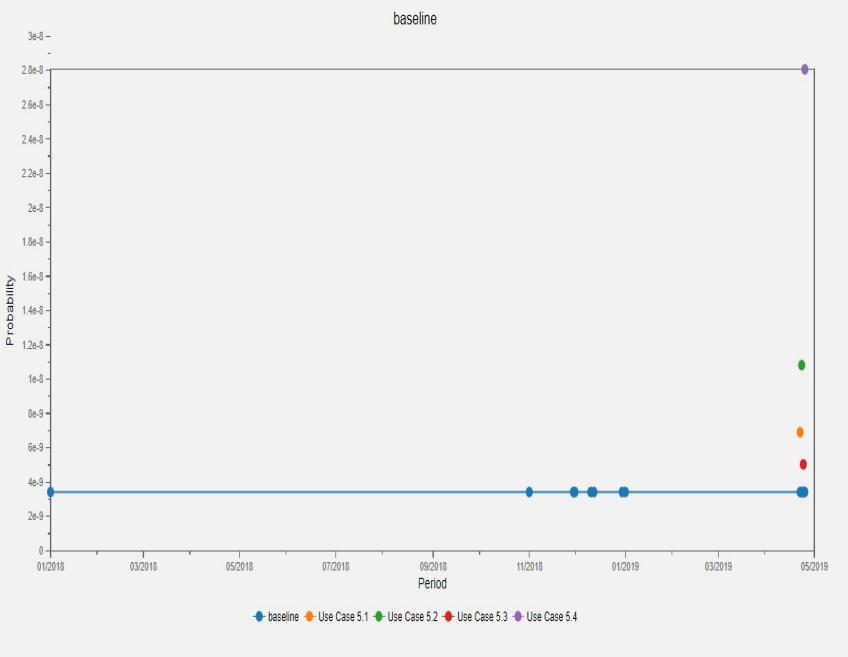

Figure 9 – MAC-model Risk calculations outcome, all four scenarios

The data used is only a representative sample of TMA safety data. It was used here for the purposes of demonstrating the capabilities of the RO. Different TMA will have different risk profiles

It is however representative of the increased MAC accident risk in TMA when compared with En-route operations.

This section provided numerical results on the MAC-model to demonstrate the way of using the RO to the KLM assessors. They tried the MAC-model itself and the main feedback provided by them was:

- the likely hood of occurrence of a MAC is statistically very low. So maybe it would have been more appropriate, or more useful, to look into other aviation risk aspects (/categories), like CFIT, or lateral runway excursions, or LOC-I, etc, since there are many more risks that have a higher occurrence probability than MAC.
- that they appreciated the integration of the airborne and the ground (i.e. ATC) part in the overall MAC risk assessment. However the comment was made how to translate the risk results, like for the precursors, into actionable measure for risk prevention / mitigation for air crews. That was found rather unclear in case the relevant (MAC-model related and needed) data is not available within an airline.

3.6. Usecase-6 MAC-Model: Perform Sensitivity Analysis

3.6.1. Description of UC-6

The intention of this usecase was to perform a sensitivity analysis on the outcome of the MAC-model. Firstly to run the MAC-model with the baseline data. Secondly 16 scenarios were proposed for the sensitivity analysis in which the generic Contributing Factors that were chosen should be provided with a new value. Foreseen where to modify the following GCF's: 31- (31.2 till 31.5) and 33- (33.2 till 33.9), see [9]. Subsequently it was foreseen to run some scenario cases with modifications on the IF's at own selection. However there was too little time to execute all the foreseen cases, In addition it was found out afterwards that the new proposed GCF-values to be inserted where not correct (from a unit point of view) and thus would not have provided correct risk results.

3.6.2. Main Assessment Results

The basic principle of the sensitivity aspects were explained and shown to the KLM assessors. But unfortunately no real feedback results were collected for this usecase due to lack of time.

3.7. Usecase-7 Baseline Versus Reference Data Comparison

3.7.1. Description of UC-7

In this usecase it was the intent to replace the RO's default available model data by own reference data.

The baseline scenario represents a European average scenario which is not specific for one user or type of ser. Hence a baseline set of data is available within the RO's risk models. This has to overcome blocking of risk modelling via confidentiality of data. But in the preparatory sessions held with KLM it was clearly expressed that they would like to have introduced and use their own company data inside the risk models. This was named "Local Reference Data". A user would then be able to integrate own, potentially higher fidelity data. With KLM two sets of data were created for the RE-model, named "Optimistic" data and "Pessimistic" data to assess this concept.

3.7.2. Main Assessment Results

The KLM "Local Reference Data" sets were set up and presented. The RO was adapted to be able to use this local reference data. But at the time of the KLM assessment some software problem occurred, especially with the "Manage Risk Scenario" options in which these "local Reference Data" sets played an important role. Since the KLM had proposed the idea, its intention was well understood and supported positively, but this usecase was not further pursued. And after the KLM assessment had taken place, the RO software was made to work properly with this "Local Reference Data" functionality, that is now available inside the most recent RO software version 2, see [10]. Figure 10 shows the RE-risk result comparison produced with this RO version.

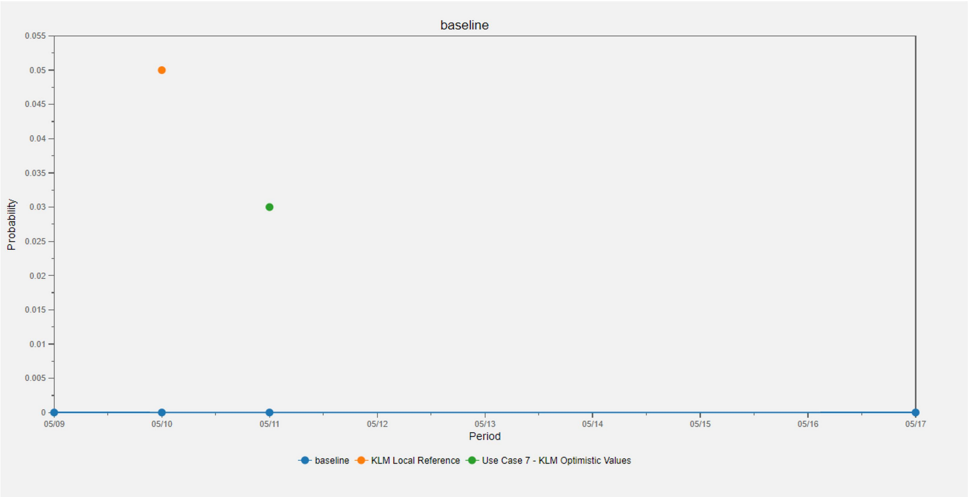

Figure 10 – RE-risk; Baseline versus KLM Local Reference and KLM Optimistic values

Also shown are the RE-risk precursor results, see Table 11, but for illustration purposes only.

Table 11 RE-risk precursor results for KLM Local reference-1 (Pessimistic) and Local reference-2 (Optimistic) values

BackBone		Pessimistic model input data set	Optimistic model input data set
EVENT	**DESCRIPTION**	**PROBABILITY**	**PROBABILITY**
BB000a	Longitudinal Runway Excursion	5.0E-02	3.0E-02
BB000b	Non-decelerated incorrect touchdown	1.65525E-06	3.52962E-07
BB000c	Non-rejected incorrect touchdown	2.36396e-05	8.8196E-06
BB000d	Incorrect Touchdown	1.18198E-01	4.4098E-02
BB000e	Non-corrected Unstable Approach	2.8178E-02	1.4078E-02
BB000f	Unstable approach	2.8178E-01	1.4078E-01

3.8. Usecase-8 RO versus KLM Bow-tie model on RE

3.8.1. Description of UC-8

The goal of UC-8 is to assess the potential benefit of the RO to highlight the bow-tie barriers that may necessitate to be strengthened/added in order to reduce the level of risk of runway longitudinal excursion (RE).

The two following scenarios have been defined in order to quantify the effect of safety barriers:

1st scenario:

- Maximum efficiency of the barriers i.e. rate of GCF related to the Barriers (failure of the barriers) set to 0 and IF's are not used (i.e. any even that may increase the inefficiency of the barriers)
- 2nd scenario:
- Minimum efficiency of the barriers i.e. worst rate allocated to GCF related to the Barriers and worst case IF's (maximum negative influence)

Note:

- Most Bow-Tie barriers are included in the Backbone Model.
- Barriers related with Maintenance & Aerodrome Operators are not included
- Post-runway excursion mitigations are not included in Backbone model.

For this usecase KLM made available their in-house developed Bow-Tie. First step of the work done on this Use case consisted to build a mapping between the contributing Factors and Influencing Factors available in the RO with 'threats" defined in the KLM bow-tie.

The use case was restricted to one threat "Aircraft system Failures" among the 6 threats involved in the risk of experiencing a runway overrun as depicted in the following Figure 11`(KLM bow-tie source).

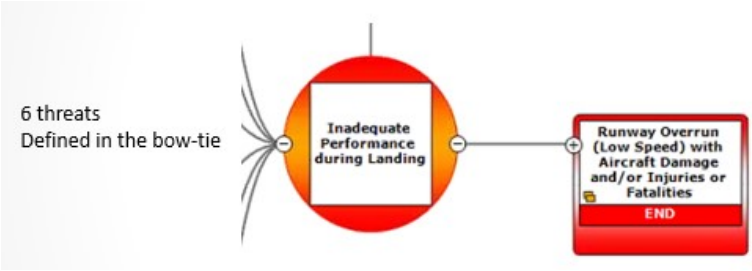

Figure 11 – KLM Bow-Tie modelling – RE Risk

It has been considered that for the purpose of the Proof of Concept it was not necessary to extend UC-8 to other threats.

3.8.2. Main Assessment Results

UC-8 enabled to compare Maximum and Minimum efficiency of the existing barriers with respect to acceptable probability. Running the two scenarios of UC-9 gave indication on whether the efficiency of barriers should be increased or if additional barriers would be required

The 'max efficiency' has provided an estimation on the best efficiency of the existing barriers. However this is a "theoretical" point of view since barriers are not failures-free and cannot assumed to be actually 100% efficient

The 'min efficiency' has provided an estimation of the actual efficiency considering the failures of the existing barriers (provided that inputs used are reliable, i.e. trustable).

- If the probability of the 'max efficiency' is greater than the acceptable probability of the risk then additional barriers should be added in the bow-tie since improving their efficiency in not sufficient
- If the probability of the 'min efficiency' is greater than the acceptable probability of the risk then the stakeholder should try to improve the efficiency of the existing barriers in the Bow-Tie.

KLM has recognized potential of the approach based on UC-8 results that could be increased with good data. They have raised the following remarks

- Bowties need good explanation before to do the mapping with the Backbone
- Mapping requires Expert advice
- RO mapping with Bowtie-"threats" is potentially more promising than mapping with Bowtie-"barriers"

As way for improvement they recommend to avoid mapping with aspects to deep inside the Bowtie model. Grouping of "barriers" and mapping those groups to the RO-model is perhaps the way to go.

3.9. Usecase-9 KLM's Airprox Bow-Tie versus MAC-modelling

3.9.1. Description of UC-9

Similar to the previous one it was aimed at with this particular usecase to try to couple parts of the Bow-tie of an Airprox to the Risk Observatory usage. Figure 12 shows the attempt as presented to KLM. At the bottom half of this figure the clustered GCF's can be found, which are grouped and coupled to one "barrier".

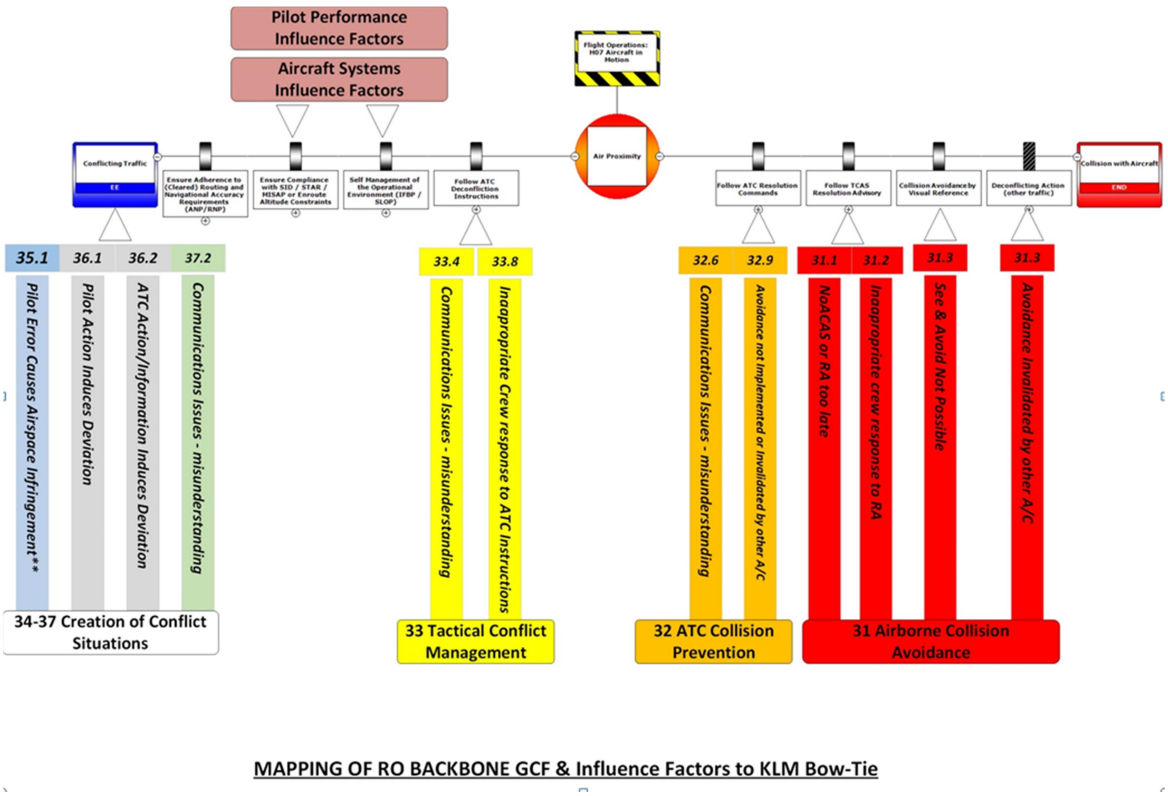

Figure 12 – KLM Bow-Tie modelling – MAC Risk

3.9.2. Main Assessment Results

The usecase 9 was discussed with the KLM assessors in conjunction with the usecase 8. They mainly commented that, according to them the:

- Usecase 9 Bowtie (related to MAC-model) seems easier to perform than mapping the RO with the RE-Bowtie since Bowtie model & MAC-model are more similar
- Grouping of Contributing Factors, as performed in this usecase, maybe can be seen as the way forward in this type of research. Based on this insight, this was therefor also recommended in the previous usecase.

3.10. Questionnaire results

Appendix A shows the questionnaire that were prepared and actually used in the KLM assessment. Appendix B records the collected questionnaire inputs (results).

Given the fact that only two KLM assessors were present, while a third one needed to urgently work on other matters, and the invited back-up person (from outside KLM) fell ill one day before the assessment,

the minimum number of three persons needed for statistical analysis was not met. Hence it was decided to record all the ratings provided and written comments but not to process the ratings any further.

4 GENERAL FEEDBACK FROM AIRBUS AND KLM

The main comments derived from the usecases, the filled questionnaires (see appendix B), and the discussions held, are summarized, in the following sub-sections.

4.1. Overview on KLM feedback comments

- The ROs GUI was not always found easy to understand. It is noted that based on the detailed KLM feedback, various GUI aspects, like use of color coding and styling of the risk pictures, was directly improved inside the final release version 2 of the RO.
- Importing and exporting functionalities, seem not present, or are too much hidden so not easy to use.
- Expansion of RO proposed with other Risk models, potentially of other domains: Lateral RWY excursion model, CFIT-model, but also MRO-data, etc.
- FDM coupling with Occurrence Dash board to be expanded and with more SPI's. The FDM data coupling process is currently rather unclear to a user.
- Coupling the risk models with the Occurrence Dashboard within the RO tool itself is necessary => it will show directly the overall Risk Picture, but insight into precursor risk values is more essential for safety management
- To be able to reproduce results from older RO-versions all input data needs to be stored as well (for version control)
- The RO is a good concept. It is helpful for safety focus and comparison with global initiatives. Structured quantification and sensitivity analysis are a valuable supplement to Qualitative Safety Systems. The use of the RO requires specialists due to the complexity of the models. The RO tool requires detailed risk model expertise and therefore it is of use for safety experts. It is not yet a tool for Safety managers.
- KLM values the many IFs that are included in the model as very positive. The current RO prototype limits the number of IFs to be changed in one scenario by 3. KLM doesn't want to be restricted by that number, and would like to be able to changes the IFs they need for a scenario.
- KLM pointed out that the potential of the RO is great when good data is used. Difficulty is the validity of the calculation given the fact that the quality of the underlying data may vary. It should be clarified in the RO on wat kind of data the results are based on hard data (and what scope does this data cover), calculations, expert judgement, etc. KLM stressed that the users must have confidence in the results and may have to build up confidence. The judgement of the validity of the model and its results seems to be easier using relative risk such as risk ratio instead of the low accident risk numbers.
- The RO tool provides the option to include User Reference data and a comparison with Baseline data (e.g. European average data). KLM experts see this as a handy tool for risk comparison and possible prioritisation.
- The RO concept is found useful.

- It is positive that it seems feasible to perform a mapping of the RO models on KLM Bowties. The mapping between the RO's contributing factors and the Bowties' barriers looks reasonable. It is a kind of two-way model check: one for the RO and one for the Bowtie. Based on this mapping you can see which contributing factors are more important than others. With the mapping it can be identified which Bowtie branches could be important, or have weaknesses.
- Accident analyses rely on occurrence data so it would be valuable for the future to invest in it, {thus to invest in data access and data use, WR}. Since data needs to be obtained outside of the organisation, investment is justified.

The following suggestions for improvement of the GUI have been recorded during the trial session:

- Enhance the way of managing the scenarios: Delete, copy/ paste (duplicate) a scenario, renaming it.
- To add explanatory Legend for each precursor plot
- Harmonization of the Menu options & features
- The headings of the graphs and the Scenarios need to be improved
- Pop up text on Occurrence Dashboard needs to be improved
- Improvement of the mental picture of the models for the user. It was possible to take this request into account; backbone models structure has been added in Information field.
- A confidence indication for the results
- Data source information should be included
- To consider new technologies for the GUI like e.g. 3D graphical results, or VR-technology.
- The software interface needs upgrade. Hard to trace inputs that lead to outputs.

4.2. Overview on Airbus feedback comments

- Airbus suggested improving visibility and trust on the models that were built. Airbus also asked to better clarify the liability and industrialization aspects. P4 partners pointed out that Visibility was considered in the last version of the RO. Structure of the Backbone model has been included in the RO prototype for information. The Backbone model dealing with Mid Air Collision En-route which is available in the RO prototype has already been validated in SESAR project. Regarding the model on longitudinal runway excursion Airbus pilots have also validated the functional description of the backbone model based on Airbus Operating procedures (incl. FCOM / FCTM).

- A way for improvement would consist of extending the RO models that are currently focusing on ECAC zone (European data used). The IRP models from Eurocontrol and partly re-used in the frame of P4 project are in use by the FAA (North America) using their own data. Therefore extending the scope of P4 to a worldwide perimeter is not at all a limitation of the approach.

- Airbus intends to continue the dissemination activity and will identify and work on synergies with other national organization like ASD (European Aeronautics, Space, Defense and Security industries) and International Working Groups especially on possible links between P4 and Data4Safety.

5 CONCLUDING REMARKS AND RECOMMENDATIONS

5.1. Concluding Remarks

KLM and Airbus recognize and appreciate the efforts that the P4 team has put in to get the most out of this project.

The RO tool itself gave a very good insight in the foreseen RO-concept. It was very well understood that the risk modelling combined with occurrence data in one tool is a strong combination to show the risk elements. However the top–event risk for Mid Air Collisions or Runway Excursion is less important than the occurrence risk of the precursors or barriers. The latter show more directly where the problems are for Safety management, but the collection of the proper data on it is much more difficult to organize. Also the RO-models need to be trusted. This implies that there is a need for an indication on the validity of the data used inside the models. Especially in case high-fidelity in-house data are coupled to the integrated risk framework, the other risk modelling aspects should also be stepping up, both in accuracy as well as in reliability. So there is a need for more insight in these data validity matters. As such the RO tool outcome could be made more reliable by coupling more accurate stakeholder databases / models. But this, most likely can only be organized via an (independent) Risk Observatory Organisation (ROO).

Enhanced GUI aspects, improved model presentation within the tool, as well as flexibility on the GUI development of the RO software tool suite selection is crucial. The Pentaho tool suit (used to develop the RO tool), may have to be reconsidered given the high GUI demands on flexibility: different stakeholders will prefer or even demand different GUIs. Better back-office elements (exporting data functions, storage functions, etc.) are also necessary for an industrial use of the concept and tool. Therefore, the RO to become an industrial operational tool will still require a lot of research and development work. Industrialization and commercialization are not yet to be expected, more functionalities need to be developed first and more prototyping, with other stakeholders will be required. But the RO concept and its prototype development has a high potential for future safety risk management.

5.2. Recommendations

Airbus recommends making the RO more non-linear and more dynamic. This is yet a big challenge that will require a big effort that can only be made in the frame of future European R&T projects on safety modelling.

Airbus and also KLM would appreciate to expand the RO-modelling to other Risk Aspects. P4 has shown no show-stopper on that point thank to a discussions made with all concerned P4 partners on the other type of major risks (CFIT, lateral runway excursions, LOC-I, etc).

Both Airbus and KLM also recommend expanding the RO coverage to other domains: MRO, National Safety Authorities, etc.

It has been strongly recommended to disclose more accurate / confidential databases. But this will require a ROO as mentioned before to cover liability, ownership, access rights and other legal issues

More coupling of data and data-sources, like FDM, is needed within the RO prototype, but this requires more precise definition of inputs- and outputs. Also the occurrence / risk dash boards need to be further specified, more towards the individual stakeholder, and tailored to management needs, and developed.

KLM recommends improving the RO-coupling with the Bowtie's and with other representation of the risks (CATS models, etc.) This would be a good way forward for improvement.

Lastly, all persons involved in the trial sessions would appreciate to continue the RO-developments towards a future tool for Safety Management.

6 REFERENCES

[1] Future Sky Safety - Annex 1 - Description of Action (part A) – 2018 version.

[2] FSS_P4_NLR_D8.1 - Project Plan P4 Total System Risk Assessment, D8.1-Appendix P4, Version V3.0, issued 08-12-2015.

[3] Minutes of FSS-P4: RO Usecase definition Workshop at KLM offices, Amsterdam-Schiphol, The Netherlands, held at 19 Feb 2019, v1 issued 07th of March 2019.

[4] Minutes of KO Meeting held at 10th of April'19 at Airbus Offices, Toulouse, France. Version v1 issued 2nd of May-2019.

[5] Minutes of the KLM assessment Meeting held at 21st of May, at KLM offices Amsterdam-Schiphol, The Netherlands, to be published in July 2019.

[6] FSS_P4_CIRA_D4.1 - Risk Observatory Requirements, Version v2.0, issued 26-10-2015.

[7] FSS_P4_INTA_D4.4.1 - Risk model gap analysis, Version v1.0, issued 24-10-2015.

[8] M. Llobet Lopez, L. Carbo, S. Metge, P. Bieber, C. Wang, C. Schmidt-Moll, F. Kaakai, J. Ple, M. Morel, T. van Birgelen, J. Verstraeten; Risk assessment models Future Sky Safety D4.4, 2017.

[9] S. Metge (AI-F), B. Hickling (EUROCONTROL), P. Bieber, J. Morio, X. Olive (ONERA), F. Kaakai (TR6), J. Plé (NAVBLUE), C. Sannino (TAV), W.F.J.A. Rouwhorst, Bas van Doorn (NLR), F. Oliveira and R. Peixe (CEiiA) ; Integrated Risk Assessment Framework, D4.7, Version 4.0, June 2019.

[10] R. Peixe, F. Oliviera, J. Arantes; Risk Observatory User Manual release 2, Future Sky Safety D4.5, June 2019.

[11] A. Balk, R. Wever, G. Greene, Total Aviation System Risk Picture 2016, FSS-P4 D4.3, version 2.0, December 2016.

[12] A. Balk, Total Aviation System Risk Picture 2017, FSS-P4 D4.6, version 2.0, 2018.

[13] B.A. van Doorn, A. Balk; Total Aviation System Risk Picture 2018, Future Sky Safety D4.8, 2019.

[14] L. Speijker, J. Choi, V. de Vries, Y.S. Cheung, d. Pascarella, R. Peixe; Risk Observatory for data-driven safety analysis, D4.10, June 2019.

APPENDIX A RO ASSESSMENT QUESTIONNAIRE

Date: 21-May-2019

Organisation: KLM

Subject Number:

A1. Questions about yourself

Age. What is your age? :

Gender. You are? : male / female / ...

Occupation. What is your current occupation/function inside the organisation? :

Experience-1. How many years are you active in the field of Safety? List your experience:

Experience-2. What experience do you have in using Safety-related software tools? Please indicate below:

Very Much				Very Little
1	2	3	4	5

Please explain:

Involvement and Preparation. What was your involvement in the development of the current RO-software?

Very Much		None		Very Little

Please explain if there was involvement:

A2. USECASE Questionnaires

Provide a cross at your choice and explain.

USECASE-1

Q.1: What is your first impression of the RO-software tool as a whole?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 How would you rate the easiness of use of the RO-software tool?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the RO-software tool?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Q.4 Are you pleased with the RO-software functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Q.5 Open comments/Remarks (any first observations, suggestions, etc) related to the RO-tool:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

USECASE-2

Q.1: What is your impression of the RE-risk model?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 Given the tailwind/crew fatigue example exercised with 2 Influencing Factors (IFs), how would you rate the usefulness of the RO and specifically the RE-risk model for your Safety domain purposes?

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the RE-risk model?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Q.4 Are you pleased with the RE-risk model functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.5 Open comments/Remarks (any first observations, suggestions, etc) related to the RO-tool and/or RE-model:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

USECASE-3

Q.1: What is your impression of the RO-software tool related to the RE-risk model?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 Given the tailwind/crew fatigue examples exercised with 3 Influencing Factors (IFs), how would you rate the usefulness of the RO and specifically the RE-risk model for your Safety domain purposes?

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the RE-risk model?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Q.4 Are you pleased with the RE-risk model functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.5 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool and/or RE-model:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

USECASE-4

Q.1: What is your impression of the RO-software tool related to the RE-risk model?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 Given the flap lock-out examples exercised with adverse weather, how would you rate the usefulness of the RO and specifically the RE-risk model for your Safety domain purposes?

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the RE-risk model?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Q.4 Are you pleased with the RE-risk model functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.5 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool and/or RE-model:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

USECASE-5

Q.1: What is your impression of the RO-software tool related to the MAC-risk model?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 Given the examples exercised, how would you rate the usefulness of the RO and specifically the MAC-risk model for your Safety domain purposes?

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the MAC-risk model?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Q.4 Are you pleased with the MAC-risk model functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.5 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool and/or MAC-model:

USECASE-6

Q.1: What is your impression of the RO-software tool related to the MAC-risk model **and specifically its Influencing factors?**

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 What is your impression of the RO-software tool related to the MAC-risk model **and specifically it's Generic Contributing Factors?**

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of the MAC-risk model?

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Q.4 Are you pleased with the MAC-risk model functionality offered?

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.5 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool and/or MAC-model:

USECASE-7

Q.1: What is your impression of the RO-software tool related to the (KLM) **reference data set** up and comparison with the baseline values?

Very Good		Neutral		Very Poor
1	2	3	4	5

Please explain:

Q.2 What is your impression of the RO-software tool related to the RE-model **and specifically it's Generic Contributing Factors?**

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.3 How would you rate the complexity of **inserting the new (optimistic) values for the GCF's?**

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Q.4 What is your impression of the RO-software tool related to the RE-model **and specifically it's Influencing Factors?**

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

Q.5 How would you rate the complexity of **inserting the new (optimistic) values for the Influencing Factors?**

Very Complex		Neutral		Very Simple
1	2	3	4	5

Please explain:

Q.6 Are you pleased with the possibility **to self-create user scenarios?**

Yes indeed, very pleased		Neutral		No, not at all pleased
1	2	3	4	5

Please explain:

Q.8 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool and/or its RE-model options:

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

USECASE-8

Q.1: What is your impression of the RO-software tool related to the RE-risk model **and its ease of use to be applied for bow-tie coupling/linking ?**

Very easy		Neutral		Very difficult
1	2	3	4	5

Please explain:

Q.2 What is your impression of the RO-software tool related to the RE-risk model and its **usefulness to be applied for bow-tie coupling/linking ?**

Very useful		Neutral		Not very useful
1	2	3	4	5

Please explain:

USECASE- 9

Q.3: What is your impression of the RO-software tool related to the MAC-risk model **and its ease of use to be applied for bow-tie coupling/linking?**

Very easy	Neutral	Very difficult

1	2	3	4	5

Please explain:

Q.4 What is your impression of the RO-software tool related to the MAC-risk model and its **usefulness to be applied for bow-tie coupling/linking ?**

Very useful **Neutral** **Not very useful**

1	2	3	4	5

Please explain:

Q.5 Would you see other ways of using the RO-software tool (i.e. models) in relation to the bow-tie models?

Choose : Yes / No

Please explain:

Q.6 Would you see other ways of using the RO-software tool (i.e. models) in relation to flight safety risk analysis

Choose : Yes / No

Please explain:
Q.7 Open comments/Remarks (any new observations, suggestions, etc) related to the RO-tool for risk investigations

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

A3. Final Questionnaires

To be answered as a final judgement after all usecases have been performed. The answers may differ from the answers given per usecase. See question 3.1 and 3.2 on next two pages.

A3.1 System Usability Scale (SUS)

Please read carefully through the list of statements below on the current system. Indicate to which extent you agree with this statement by putting a cross on the scale from 1 (strongly disagree) to 5 (strongly agree). Use the space below the question to explain your final choice.

	Strongly disagree				Strongly agree
1. I think that I would like to use this RO-software frequently.	1	2	3	4	5
2. I found the RO-software unnecessarily complex.	1	2	3	4	5
3. I found the RO-software easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this RO-software.	1	2	3	4	5
5. I found that the various functions in this RO-software were well integrated.	1	2	3	4	5
6. I found too much inconsistency in this RO-software.	1	2	3	4	5
7. I would imagine that most people would learn to use this RO-software very quickly.	1	2	3	4	5

8. I found the RO-software very cumbersome (difficult) to use.					
	1	**2**	**3**	**4**	**5**
9. I felt very confident using the RO-software.					
	1	**2**	**3**	**4**	**5**
10. I needed to learn a lot of things before I could start working with the RO-software	1	2	3	4	5

A3.2 SATI Questionnaire (SHAPE Automation Trust Index)

Put your cross in a rating cell

		Never						Always
1	The RO-software was useful	0	1	2	3	4	5	6
Please explain:								
2	The RO-software was reliable	0	1	2	3	4	5	6
Please explain:								
3	The RO-software worked accurately	0	1	2	3	4	5	6
Please explain:								
4	The RO-software as understandable	0	1	2	3	4	5	6
Please explain:								
5	The RO-software worked robustly	0	1	2	3	4	5	6
Please explain:								

6	I was confident when working with the RO-software	0	1	2	3	4	5	6
Please explain:								

A4. Open questions

The RO-Concept aims to bring an overarching of the risk of different domains, hence aims to enrich a specific risk domain.

Q.1- What do you think of the RO-CONCEPT, so independent of the RO-software implementation?

Q.2- What do you regard as the benefits of the RO-concept?

Q.3- What do you regard as the drawbacks of the RO-concept?

Q.4- Would you consider to "invest" (in terms of time, budget, etc.) to bring the RO-concept further in the future?

APPENDIX B RESULTS ON QUESTIONNAIRES AND USER FEEDBACK

The usecases were executed in the sequential order. Still due to time limitations, the usecases 6, 7 and 9 were not fully performed as intended and given a lower prioritisation during the day. Listed values x between {x} provide the ratings given per participant, Subject S1 respectively S2.

B.1 Feedback on Usecases

B.1.1 Feedback on Usecase-1

The usecase-1 was first shown to the Subjects without them acting only listening to the explanations given and watching the various actions performed. After it has been shown to them they could run the usecase themselves.

This usecase led to the most comments and discussions since this was the first familiarisation of the KLM persons involved in the assessment with the actual use of the RO-tool. Many questions were asked about the design reasons of particular HMI choices and this led to explanatory discussions.

Discussion during/after the Usecase:

When login into the RO-tool the opening screen shows a Occurrence (Risk) dashboard. Once moving with your mouse over these figures there appears additional data information. It was observed that there were a few errors present: the labels presented were cut-off and not always visible and some data cells values were presented as being “null” while clearly there were occurrences in the cell. This software aspect has been fully revised after the session.

Related to the *Manage Scenarios* functionality it was remarked that it was found rather strange that the existing scenarios could not be copied from the existing to a new one in its entirely. One had to select a minimum of one and a maximum of three IFs. [This aspect has been fully revised after the session].

A discussion took place on the “orange dot” and the connecting lines presentation inside the main risk calculation and what-if risk picture. The KLM subjects found it fully unclear what the meaning was of these dots and lines. These “orange” dots should be totally decoupled from the scenarios” [This aspect has been fully revised after the session].

What-if: colours to be changed. Also change title of risk picture and its legend.

Colour use was not consistent between “What-if” elements and “Precursor” elements. [This aspect has been fully revised after the session].

When deselecting “baseline” it disappeared correctly from the main risk picture, but not from the precursor pictures.

The list of Precursors indicated and selectable inside the Option menu and the ordering of the presented precursor figures are not the same. This could be further improved.

Easier access to and modification of the weights was proposed.

Vertical axis with probability scales were found okay.

Keeping track of the original data values after insertion of a new value, for instance for a GCF or IF, was considered very important. Also the derived risk result would require a historical tracking of all the values used to derive it.

A long discussion took place about the meaning and difference between the “baseline data”, the local reference data and the EU-average data, as well on what was actually presented on the horizontal axes of the risk picture. Since this horizontal axes served multiple purposes (data varying over time as well as creation dates of models and dates of using the risk-models it was concluded that (in conjunction with the various risk-picture scenarios created) that it did not make sense and that it needed to be revised. [This presentation aspect has been fully revised after the session].

There was the proposal to be able to recalculate existing scenarios (or usecases) in case the baseline data would change over time. However this was regarded too complicated and the idea was discarded.

The RO-software developer explained to the attendees that Pentaho was seen as too limited to address quickly all new GUI wishes, especially on new desired graphical items like buttons, etc.

Questionnaire rating results and/or comments related to this usecases are provided next per question.

Question	Subject	Score	Comment
Q1: What is your first impression of the RO-software tool as a whole?	S1	2	Many Options, somewhat limited by inconsistencies in GUI
	S2	3	<No comments provided>.
Q2: How would you rate the easiness of use of the RO-software tool?	S1	3	See comments [?] ; Few tooltips not [?]; Back/Forward navigation menu missing
	S2	3	Options hidden
Q3: How would you rate the complexity of the RO-software tool?	S1	3	Apart from a few UI [?] the modelling and use is quite understandable.

			The [?] is the sensitivity to and reliability of underlying data
	S2	2	Not obvious what options are available. Some options are "hidden"
Q4: Are you pleased with the RO-software functionality offered?	S1	2	Although improvements are positive it provides a way to present relative [information ?].
	S2	3	Scenarios are interesting and useful but now laborious work. You have to take notes between scenarios for changed entered data
Q5: Open Comments / Remarks, etc	S1	n.a.	- Why is the local reference limited to max 3 IF values? What if we want to compute everything according to baseline except for a few modifications? In that case we would like to keep all IF as they are currently [?]. - - Provide UI to indicate relative changes per scenario / [local?] reference. Make line connection between usecases [optional?] in the diagram. - Data quality is essential. Could we get a Quality figure of the underlying data? 1= Hard data, .2 ....., 3 ..... .4....., 5= Estimated data
	S2	n.a.	<No comments provided>

Note: [?] means a non-readable written word

B.1.2 Feedback on Usecase-2

Discussion during/after the Usecase:

There was a question asked if the risk scenario and risk outcome data could be exported to other tools somehow. The answer provided was that this is not yet the case, but that it is on the wish list. Results have to be manually exported into excel or similar tools.

Related to the Risk Dashboard it was found that the usecase did not provide (=present) correct risk results for scenarios with modified IF-values. The top risk calculation was correct though. Hence since we could

not use the scenario option so we continued this and following usecases with the "What-if" option only. [This aspect has been corrected after the session].

Still the subjects found it difficult to be able to trace-back on what had already been inserted and modified (or inserted) during the "What-if's" performed, especially in case this had to be done multiple times after each other. The proposal was made to put somewhere on the screen the new inserted values and the original (default) values with an indication (for instance via coloring, or via appearance) to make visible and clear that differences were created inside the IF's and GCF's.

Concern about traceability of input values and risk results was raised. How to make sure that results can be reproduced?

Subjects both indicate that it is an interesting usecase in itself but that it is found difficult to check if the outcome values are correct. The wish was expressed to have some kind of sensitivity tool to be able to "validate" or check the model outcome.

Not executed were the sub-usecases related to the "short-haul" type of operation. However it was explained to the subjects that the RO therefor has the potential to compare long- haul and short haul results.

Questionnaire rating results and/or comments are provided next per question.

Question	Subject	Score	Comment
Q1: What is your impression on the RE-risk model?	S1	2	What-if provides easy way to analyze effects of different IF
	S2	2	Many factors are included. It seems possible to evaluate contributions of different factors
Q2: Given the tailwind/crew fatigue example exercised with 2 IFs, how would you rate the usefulness of the RO and specifically the RE-risk model for your Safety domain purposes?	S1	3	At this moment the risk ration for 100% TW (tailwind) seems to be less than 2. This seems not in line with the specific [research?] on the subject
	S2	2	Useful, but how reliable is the result?
Q3: How would you rate the complexity of the RE-risk model?	S1	3	Sufficient detail seems to be present. Not overly complicated. Data should be [tested?] whether specific scenarios are in the proper range by comparing with [?].
	S2	2	At the moment the mental picture of the risk model is not supported by the interface. That makes the model more complex.

Question	Subject	Score	Comment
Q4: Are you pleased with the RE-risk model functionality offered?	S1	3	A specific sensitivity interface would be nice.
	S2	2	Rather high potential, has to gain confidence from users
Q5: Open Comments/ Remarks, etc.	S1	n.a.	The potential with good data is great. Difficulty is the validity of the calculation given the fact that the quality of the underlying data may vary
	S2	n.a.	[still busy with?] confidence building by: - confidence indication for results - data source information

Note: n.a. = not available; [?] means a non-readable written word

B.1.3 Feedback on Usecase-3

This usecase follows on the previous usecase by adding a third IF. It was needed to use the "What-if" option again instead of the scenario option.

Only the sub-usecases 3.1, 3.6 and 3.13 were executed to grasp the principle.

Questionnaire rating results and/or comments are provided next per question.

Question	Subject	Score	Comment
Q1: What is your impression on the RE-risk model?	S1	3	It would be nice to access the [default, wr] input values again [after having modified them, wr]
	S2	4	The software interface needs upgrade. Hard to trace inputs that lead to outputs.
Q2: Given the tailwind/crew fatigue example exercised with 3 IFs, how would you rate the usefulness of the RO and specifically the RE-risk model for your Safety domain purposes?	S1	-	Flexibility and capability of model very good. Results seem suspicious. Extreme values [result in?] relative small risk ratio
	S2	2	Analysis of these factors is interesting
Q3: How would you rate the complexity of the RE-risk model?	S1	-	No additional comments
	S2	3	<No comments provided>

Question	Subject	Score	Comment
Q4: Are you pleased with the RE-risk model functionality offered?	S1	n.a.	-
	S2		In principle has high potential. Currently bit cumbersome user experience
Q5: Open Comments/ Remarks, etc.	S1	n.a.	From an expert perspective judgement of validity seems to be easier using relative risk such as risk ratio
	S2	n.a.	<no comments provided>

Note: n.a. = not available; [?] means a non-readable written word

B.1.4 Feedback on Usecase-4

Questionnaire rating results and/or comments are provided next per question.

Question	Subject	Score	Comment
Q1: What is your impression of the RO-software tool related to the RE-risk model?	S1	n.a.	No comments provided
	S2	3	Would be interesting to see effect of flap lock on RE, now not (yet) possible
Q2: Given the flap lock-out examples exercised with adverse weather, how would you rat the usefulness of the RO and specifically the RE-risk model for your Safety Domain purposes?	S1	4	The current model does not account for the effect of flap setting on approach speed and landing roll.
	S2	1	Flap lock at short runway destination is risky
Q3: How would you rate the complexity of the RE-risk model?	S1	n.a.	No comments provided
	S2	4	Assume the flap lock influence would be easy to evaluate when implemented
Q4: Are you pleased with the RE-risk model functionality offered?	S1	n.a.	No comments provided
	S2	2	No comments provided
Q5: Open Comments/ Remarks, etc.	S1	n.a.	No comments provided

Question	Subject	Score	Comment
	S2	n.a.	No comments provided

Note: n.a. = not available

B.1.5 Feedback on Usecase-5

Questionnaire rating results and/or comments are provided next per question.

Question	Subject	Score	Comment
Q1: What is your impression of the RO-software tool related to the MAC-risk model?	S1	2	Software UI adequate for model complexity
	S2	3	Values can easily be entered. Several local settings will be useful.
Q2: Given the examples exercised, how would you rate the usefulness of the RO and specifically the MAC-risk model for your Safety domain purposes?	S1	n.a.	<no comments provided>
	S2	4	See answers provided under question 5
Q3: How would you rate the complexity of the MAC-risk model?	S1	4	Also in Bowtie, the escalation chain is relatively simple
	S2	3	<no comments provided>
Q4: Are you pleased with the MAC-risk model functionality offered?	S1	2	Sufficient detail, seems to match our [model?]
	S2	4	Not clear yet
Q5: Open Comments/ Remarks, etc.	S1	n.a.	- MAC model seems to be the easiest to map to the Bowtie model. Mapping the RO-model to the high-level Bowtie elements should be feasible as a [minimum?]. For Barrier mapping direct linking to the RO-model may be too difficult /inaccurate. - Grouping of Barriers and mapping those groups to the RO-model perhaps is the way to go.

Question	Subject	Score	Comment
	S2	n.a.	Big challenge is how to derive lessons for pilots from those model results

Note: n.a. = not available; [?] means a non-readable written word

B.1.6 Feedback on Usecase-6

Due to lack of time this Usecase was given a lower priority and only demonstrated quickly near the end of the day. The questionnaire on it was not filled.

B.1.7 Feedback on Usecase-7

Questionnaire rating results and/or comments are provided next per question.

Question	Subject	Score	Comment
Q1: What is your impression of the RO-software tool related to the (KLM) reference data set up and comparison with the baseline values?	S1	2	Handy tool for risk comparison and possible prioritisation
	S2	2	The concept/ idea is useful, the software is not yet useful to learn from this
Q2: What is your impression of the RO-software tool related to the RE-model and specifically its Generic Contributing Factors?	S1	n.a.	No comments provided
	S2	3	No comments provided
Q3: How would you rate the complexity of inserting new (optimistic values) for the GCF's?	S1	4	After one day [?] good feel to handle the software and different options to arrive at the result
	S2	4	Simple
Q4: What is your impression of the RO-software tool related to the RE-model and specifically its Influencing Factors?	S1	n.a.	No comments provided
	S2	3	It is not easy to see how the IF influences the model outcome

Question	Subject	Score	Comment
Q5: How would you rate the complexity of inserting new (optimistic) values for the influencing Factors?	S1	n.a.	No comments provided
	S2	4	No comments provided
Q6: Are you pleased with the possibility to self-create user scenarios?	S1	1	Self-created user scenarios show great potential
	S2	1	No comments provided

Note: n.a. = not available

B.1.8 Feedback on Usecases-8 and -9

The aim of this usecase was explained first. This was making relationship between the RO, the Back bone model and your own Bowtie model, and how this combination could produce a good result. Together with Pierre these aspects were set up. Initially it was tried to work on the CAA-UK model, but for the purpose of the exercise the KLM-model was found more interesting as it would show more relevance to the usefulness of the RO tool for an airline like KLM. We compared the two models and compared the barriers. We looked at the so-called “threats” inside the KLM bowtie and how to map those with our precursors and/or GCFs of the BackBone model and IFs. So the GCFs present the possible failures of the barriers. We could not look at all the branches inside the KLM-bowtie, so we limited ourselves to the branch 4 (aircraft / system failures). A few sub-scenarios prior to the assessment and two (preliminary) results were already present in the slideware used during the assessment day.

S2: positive is that it seems feasible to perform such a mapping at all and that the mapping between the RO’s contributing factors and the Bowties’ barriers looks reasonable .So that is a kind of two-way model check: one for the RO and one for the Bowtie. Based on this mapping you can see with contributing factors that some are more important than others, so how the end-result is sensitive to certain GCFs.

Indeed and this is a way to identify which Bowtie branches could be important, or have weaknesses, by playing a bit with input values, even in case the accuracy of the actual input values is questioned. This could give steering where to put more emphasis to safety means inside the organization.

S2: If you are missing a Contributing Factor, or suppose that a Bowtie-barrier exists that cannot be mapped yet, is it then difficult to add such a factor to the RO?

Answer provided: It is not so easy indeed in the framework of this project. Since we are close to the end of the project. However in another context, say in bi-lateral activities, this could definitely be achieved. Then a backbone model could be set up dedicatedly for a given risk, say CFIT.

S2: Quantification of the barrier effectiveness is something for the future.

Inside the current set up, the result leads to the top-even but not yet to the actual risk, since there are more branches leading up to the risk. So there is actually more in the model preventing the RE risk occurring than only this top-event in this set up.

S1: The intent is excellent. But maybe we should not link everything to the barriers. Maybe you went to deep down. The first thing to do would be in my opinion to quantify the escalation chain from threats to top events to the final accident scenarios. Intuitively I would expect that the correlation between the backbone model, the GCFs and IFs and the mapping to the Bowtie-barriers are less accurate than the mapping to the threats. I would guess these to be more accurate.

It was concluded from the discussions that took place that the RO-coupling should take place at a higher Bowtie level.

Use frequencies of the Bowtie-treats and calculate the

The questions on these two usecases were combined in one set of questions Q.1 till Q.5.

The questions on these two usecases were combined in one set of questions Q.1 till Q.5.

Question	Subject	Score	Comment
Q1: What is your impression of the RO-software tool related to the **RE-risk model** and its ease of use to be applied for bow-tie coupling/linking?	S1	3	Bowties need explanations before mapping to the barrier. This requires expert advice
	S2	4	Bowtie linking not in software now, but possible
Q2: What is your impression of the RO-software tool related to the **RE-risk model** and its usefulness to be applied for bow-tie coupling/linking?	S1	1	When high-level mapping is desired and data sufficiently accurate. In Bowtie [no?] constant threat / top event/ consequence This is very valuable. Avoid too deep mapping if not guaranteed accurate
	S2	2	Will be useful; we thinking the quality of the mapping of RO model with Bowtie elements is critical.
Q3: What is your impression of the RO-software tool related to the **MAC-risk model** and its ease of use to be applied	S1	n.a.	No comments provided
	S2	4	The mapping between backbone GCF and

Question	Subject	Score	Comment
to for bow-tie coupling/linking?			bowtie elements seems in principle possible
Q4: What is your impression of the RO-software tool related to **the MAC-risk model** and its usefulness to be applied to for bow-tie coupling/linking?	S1	n.a.	No comments provided
	S2	2	But not possible [with the RO-tool yet,wr]
Q5: Would you see other ways of using the RO-software tool (i.e. models) in relation to the bowtie models?	S1	n.a.	<no comments provided>
	S2	Yes	Only linking between threats, top-event, consequences and the RO-model
Q6: Would you see other ways of using the RO-software tool (i.e. models) in relation to flight safety risk analysis?	S1	n.a.	<no comments provided>
	S2	Not Yet	At this stage of understanding of the model
Q7: Open Comments/ Remarks	S1		<no comments provided>
	S2		<no comments provided>

Note: n.a. = not available

B.2 System Usability Scale (SUS)

Question	Subject	Score	Comment
Q1: I think that I would like to use this RO-software frequently	S1	4	
	S2	3	Risk analyst maybe 1 day [per] month
Q2: I found the RO-software unnecessarily complex	S1	1	
	S2	2	
Q3: I found the RO-software easy to use	S1	4	
	S2	2	
Q4: I think that I would need the support of a technical person to be	S1	1	
	S2	4	

Future Sky Safety has received funding from the EU's Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 640597.

Question	Subject	Score	Comment
able to use this RO-software			
Q5: I found that the various functions in this RO-software were well integrated.	S1	3	
	S2	2	
Q6: I found too much inconsistency in this RO-software	S1	2	
	S2	4	
Q7: I would imagine that most people would learn to use this RO-software very quickly	S1	4	
	S2	1	Only for risk expert
Q8: I found the RO-software very cumbersome (difficult) to use	S1	1	
	S2	4	
Q9: I felt very confident using the RO-software	S1	4	
	S2	2	
Q10: I needed to learn a lot of things before I could start working with the RO-software	S1	2	
	S2	4	Model structure

B.3 SATI Questionnaire Results (SHAPE Automation Trust Index)

Question	Subject	Score	Comment
Q1: The RO-software was useful	S1	4	High level risk estimators (without IF) are always useful
	S2	2	Now
Q2: The RO-software was reliable	S1	5	one unmapped GCF was present
	S2	5	No comments provided
Q3: The RO-software worked accurately	S1	3	See answer above. Accurate demand on data
	S2	3	Not able to check
Q4: The RO-software as understandable	S1	5	When changing reference values the limitations on the number of IFs was not intuitively clear

	S2	2	Some more than others. E.g. graph has unnecessary items. E.g. x-as is strange and not intuitive
Q5: The RO-software worked robustly	S1	3	Too limited experience in one day to conclude anything
	S2	5	No crash , no loss of data
Q6: I was confident when working with the RO-software	S1	5	Main menu structure: comprehendible overview of the model could be handy
	S2	1	When the output was shown I could not easily relate it to the input. So you are unsure of correct input.

B.4 Final Open Questions

The RO-concept aims to bring an overarching of the risk of different domains, hence aims to enrich a specific risk domain.

Question	Subject	Comment
Q1: What do you think of the RO-concept, so independently of the RO-software implementation	S1	Helpful for safety focus and comparison with global initiatives
	S2	Good concept. Hard to achieve, requires specialists
Q2: What do you regards as the benefits of the RO-concept?	S1	Structured quantification and sensitivity analysis as a valuable supplement to Qualitative Safety Systems
	S2	Risk Change ratio's for factors. Sensitivity of factors
Q3: What do you regard as the drawbacks if the RO-concept?	S1	Inevitably data is average data. How reliable are the estimates when getting into the details? The current implementation does not provide clues where and how large in-accuracies are

	S2	Complexity of model. Requirement of experts. Hard to get trust and confidence of user
Q4: Would you consider to "invest" (in terms of time, budget, etc.) to bring the RO-concept further in the future?	S1	Accident analyses rely on occurrence data so it would be valuable for the future to invest. Since data needs to be obtained outside of the organization, investment is justified
	S2	In time as SMS safety consultant