Validation exercises execution

Simone Rozzi, Luca Save, Mario Torelli (Deep Blue), Rombout Wever, Bas van Doorn (NLR), Heiko Udluft (TU Delft), Reinhard Menzel, Wietse Post (JRC), Nuno Aghdassi (Avanssa)

This study reports on the results of 3 validation exercises executed in the context of ASCOS WP5.3 Validation Execution. The exercises addressed the initial proposed ASCOS certification approach, the ASCOS Safety Performance Indicator (SPI) framework and the ASCOS Tool for Continuous Safety Monitoring (ATCSM), and the ASCOS Risk Assessment Model and ASCOS Tool for Risk Assessment.

Coordinator	L.J.P. Speijker (NLR)
Work Package Manager	L. Save (Deep Blue)

Grant Agreement No.	314299
Document Identification	D5.3
Status	Approved
Version	1.1
Date of Issue	6/25/2015
Classification	Restricted

This page is intentionally left blank

Document Change Log

Version	Author(s)	Date	Affected Sections	Description of Change
1.0	S. Rozzi	22-04-15	All	Version for approval by PMT
1.1	S. Rozzi	25-06-15		Comments PMT processed

Review and Approval of the Document

Organization Responsible for Review	Name of person reviewing the document	Date
NLR	J. Scholte, A.L.C. Roelen	29-06-2015
CAAi	S. Long	10-04-2015
JRC	W. Post, R. Menzel	10-04-2015
CertiFlyer	G. Temme, M. Heiligers	10-04-2015
TU Delft	R. Curran, H. Udluft	10-04-2015
Thales Air Systems	B. Pauly	10-04-2015
Avanssa	N. Aghdassi	10-04-2015
APSYS	S. Bravo Munoz	10-04-2015
Organization Responsible for Approval	Name of person approving the document	Date
Deep Blue	L. Save	22-04-2015
NLR	L.J.P. Speijker	30-06-2015

Document Distribution

Organization	Names
European Commission	M. Kyriakopoulos
NLR	L. Speijker, A. Rutten, M.A. Piers, P. van der Geest, A. Roelen, J.J Scholte, J. Verstraeten, A.D. Balk, E. van de Sluis, M. Stuip
Thales Air Systems GmbH	G. Schichtel, J.-M. Kraus
Thales Air Systems SA	B. Pauly
Airbus Defence and Space APSYS	S. Bravo Muñoz, J.P. Heckmann, M. Feuvrier
Civil Aviation Authority UK	S. Long, A. Eaton, T. Longhurst
ISDEFE	M. Martin Sanchez, I. Etxebarria, M. Sánchez
CertiFlyer	G. Temme, M. Heiligers
Avanssa	N. Aghdassi
Ebeni	A. Simpson, J. Denness, S. Bull
Deep Blue	L. Save
JRC	W. Post, R. Menzel
JPM	J. P. Magny
TU Delft	R. Curran, H. Udluft, P.C. Roling
Institute of Aviation	K. Piwek, A. Iwaniuk
CAO	P. Michalak, R. Zielinski
EASA	K. Engelstad
FAA	J. Lapointe, T. Tessitore
SESAR JU	P. Mana
EUROCONTROL	E. Perrin
CAA Netherlands	R. van de Boom
JARUS	R. van de Leijgraaf
SRC	J. Wilbrink, J. Nollet
ESASI	K. Conradi
Rockwell Collins	O. Bleeker, B. Bidenne
Dassault Aviation	B. Stoufflet, C. Champagne
ESA	T. Sgobba, M. Trujillo
EUROCAE	A. n'Diaye
TUV NORD Cert GmbH	H. Schorcht
FAST	R. den Hertog

Table of Abbreviations

Acronym	Definition
ANSP	Air Navigation Service Provider
ATM	Air Traffic Management
ATCMS	ASCOS Tool for continuous safety monitoring
AVG	Average
CAA	Civil Aviation Authority
CATS	Causal model for Air Transport Safety
CMA	Continuous Monitoring Approach
CNS	Communication, Navigation, Surveillance
CPS	Certification Process study
CS	Certification Specifications
D	Deliverable
DOW	Description of Work
EASA	European Aviation Safety Agency
ECCAIRS	European Coordination Centre for Accident and Incident Reporting Systems
ECF	ECCAIRS Common Framework
ECR	European Common Repository
E-OCVM	European Operational Concept Validation Methodology
ESD	Event Sequence Diagram
EU	European Union
HMI	Human Machine Interface
FAA	Federal Aviation Administration
FAST	Future Aviation Safety Team
FDM	Flight Data Monitoring
FT	Fault Tree
GPWS	Ground Proximity Warning System
ICAO	International Civil Aviation Organization
KPA	Key Performance Area
NASA	National Aeronautics and Space Administration
OEM	Original Equipment Manufacturer

RIMCAS	Runway Incursion and Collision Avoidance System
SESAR	Single European Sky ATM Research
SPI	Safety Performance Indicator
SMS	Safety Management System
SUS	System Usability Scale
TAS	Total Aviation System
TCAS	Traffic Collision Avoidance System
UG	User Group
UPRT	Upset Prevention and Recovery Training
WP	Work Package

List of Tables

Table 1. Formative vs summative evaluation.....18
Table 2. List of ASCOS results [19]. The results in bold are those that have been included in the scope of the validation.....21
Table 3. Summary table of the three Exercises.....23
Table 4. KPAs for the ASCOS proposed certification approach (Exercise 1).....29
Table 5. KPAs for the proposed ASCOS SPI framework and Tool for Continuous Safety Monitoring (Exercise 2).....29
Table 6. KPAs for the proposed ASCOS Risk Model and the Tool for risk assessment (Exercise 3).....30
Table 8. EXERCISE 1: Profile of the expert group.....42
Table 9. EXERCISE 1: Summary of the validation results for the proposed ASCOS certification approach.....53
Table 10. EXERCISE 2: Profile of the expert group.....55
Table 11. EXERCISE 2: Summary of the validation results for the ASCOS SPI framework.....61
Table 12. EXERCISE 2: Summary of the validation results for the ASCOS Tool for Continuous Safety Monitoring.....61
Table 13. EXERCISE 3: Profile of the expert group.....62
Table 14. EXERCISE 3: Summary of the validation results for the ASCOS Risk Model.....74
Table 15. EXERCISE 3: Summary of the validation results for the ASCOS Tool for risk assessment.....75
Table 15 CATS (2009) ESDs.....110
Table 16 Top ten ESD 19 base events, sorted on their contribution to ESD 19 accident risk.....120
Table 17 Top ten ESD 19 base events, sorted on their frequency.....121

List of Figures

Figure 1. Operational concept lifecycle and maturity levels according to E-OCVM [22]. 19
Figure 2. EXERCISE 1: Experience of the participating experts in certification...43
Figure 3. EXERCISE 1: Perspectives represented by the expert group 1. 43
Figure 4. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA1-Efficiency. 44
Figure 5. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA2-Soundness.........46
Figure 6. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA3: Cross-domain integration. 47
Figure 7. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA4-Harmonization. .49
Figure 8. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA5-Accomodation of Innovation..50
Figure 9. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA6-Acceptability. 51
Figure 10. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA7-Flexibility.........52
Figure 11. EXERCISE 2: Perspectives represented by the expert group. 56
Figure 12. EXERCISE 3: Perspectives represented by the expert group. 63
Figure 13. EXERCISE 3: ASCOS Risk model questionnaire. 64
Figure 14. EXERCISE 3: Tool for risk assessment questionnaire 69
Figure 15 CATS toolset [33] 112
Figure 16 Highest accident frequency per ESD end node 114
Figure 17 Highest fatal accident frequency per ESD end node 115
Figure 18 Highest accident frequency per end event type 116
Figure 19 Highest frequency of accident types..117
Figure 20 Events (nodes) of ESD 19 (AL19)...119
Figure 21 Base events of ESD 19 (selection)..120
Figure 22 Relative changes in accident frequency for accident types wrt the baseline 130
Figure 23 Relative changes in accident frequency for end events wrt the baseline...131
Figure 24 Relative changes in accident frequency for ESD end nodes wrt the baseline..................................131
Figure 25 Detailed relative changes in accident frequency for ESD end nodes wrt the baseline......................132
Figure 26 Accident frequencies as a result of case study 4 138

This page is intentionally left blank

Executive Summary

ASCOS is an EU funded project aiming at developing certification enhancements that should make more effective and flexible the certification of aeronautical products and services [1]. Approaching certification from a Total Aviation System perspective, ASCOS promotes the considerations of the safety impact of a change across different certification domains and across the entire system lifecycle.

The present deliverable reports on the outcome of the three validation exercises carried out in the context of ASCOS Work Package 5.3: Validation Execution. Consistently with the definition of validation as proposed by the European Operational Concept Validation Methodology (E-OCVM [2]), validation has been interpreted as an issue of investigating whether and how a proposed concept or solution delivers its intended benefit. Therefore, each exercise focused on investigating the fit-for-purpose of the main ASCOS results. In particular:

1. Exercise 1 addressed the initial proposed ASCOS certification approach, which was developed in the context of ASCOS Work Package (WP) 1.3;
2. Exercise 2 addressed the ASCOS Safety Performance Indicator (SPI) framework and the ASCOS Tool for Continuous Safety Monitoring (ATCSM), which were developed in the context of ASCOS WP2.1[3] and WP2.4[4] respectively;
3. Exercise 3 addressed the ASCOS Risk Assessment Model and the ASCOS Tool for Risk Assessment, which were developed in the context of ASCOS WP 3.2.3 [5] and WP3.3 [6] respectively.

Each validation exercise consisted of a one-day workshop with a group of selected experts, during which data was collected by means of focus group discussions and individual questionnaires. The profile of each group differed depending on the exercises: exercise 1 expert group included mostly certification experts; exercise 2 expert group, safety experts; exercise 3 expert group, risk assessment and risk management experts.

The collected results provide initial support for the proposed solutions. The intents motivating ASCOS were acknowledged as ambitious, important, and worthwhile by the participating experts. At the same time, however, it was not fully clear to the participating experts how the ASCOS solutions could actually deliver their proposed benefit. This could be related to the lack of prior insight and experience of the participants with the available ASCOS material. Further work on the development of guidance material is recommended. Addressing this point would require, in particular, a better definition of the context of use of the proposed certification approach—i.e. roles and responsibilities. This emerged as a common concern across the three exercises.

In reporting this result it needs to be acknowledged that ASCOS targets a wide and complex domain of human activity (certification), a domain that has developed incrementally over a period of several decades. In other words, the complexity of the target domain is such that it is very unlikely that any proposed solution in this area can be defined in full from the outset. Developing comprehensive enhancements in certification practices is not like developing a new piece of software, a product or a service—something that can be fully specified in advance. The number of involved stakeholders, their different political, cultural contexts and professional backgrounds, and the number of involved certification domains implies that predictability is inevitably low.

That being said, it can be noted that no evidence has been collected that points against the potential value of the proposed approach—i.e. no significant show-stopper emerged. The exercises, instead, allowed identification of a number of relevant areas for consideration and improvement. These are the main contribution of this report, and their summaries are provided in the tables available under sections 4.9, 5.4, and 6.5. Overall, these areas create the basis for advancing the maturity level of the proposed solutions—so to work towards a more comprehensive and more repeatable view of the proposed ASCOS enhancements.

This is what will be done in the next phase of work: to integrate the results presented here with the results collected from WP4 case studies in order to consolidate recommendations for improvement. In particular, these recommendations will target ASCOS WP1.5, which delivers an improved version of the initial proposed ASCOS approach, based on the validation and evaluation activities carried out in the context of the project.

Table of Contents

Document Change Log 1

Review and Approval of the Document 1

Document Distribution 2

Table of Abbreviations 3

List of Tables 5

List of Figures 6

Executive Summary 8

1 Introduction 15

1.1 ASCOS project background 15

1.2 Purpose and structure of this document 16

2 Validation of the proposed ASCOS enhancements 17

2.1 Definition of validation 17

2.2 Validation approach 18

2.3 Maturity assessment of ASCOS proposed solutions 19

2.4 Validation scope: Selection of products to validate 21

2.5 Validation objectives 22

3 Methodology 23

3.1 Methodology Overview 23

3.2 Rationale for the selected validation format 24

3.2.1 Consideration of constraints in accessing experts 24

3.2.2 Adequate to maturity level 24

3.2.3 Rationale for the sequence of the Exercises 25

3.3 Recruiting of experts 25

3.4 Performance frameworks 25

3.5 Familiarization and training 30

3.5.1 Exercise 1 Familiarization and training 31

3.5.2 Exercise 2 Familiarization and training 31

3.5.3 Exercise 3 Familiarization and training 32

3.6 Feedback Gathering 32

3.6.1 Exercise 1 Feedback gathering 33
3.6.2 Exercise 2 Feedback gathering phase 33
3.6.3 Exercise 3 Feedback gathering phase 34
3.7 Validation scenarios 34
3.7.1 Exercise 1 scenarios 34
3.7.2 Exercise 2 scenarios 36
3.7.3 Exercise 3 scenarios 37
3.8 Materials 39
3.8.1 Questionnaires 39
3.8.2 SUS questionnaire (Exercise 3) 39
3.9 Data analysis 39
3.10 Limitations 40
4 Results Exercise 1: The ASCOS certification approach 42
4.1 Expert group profile 42
4.2 Efficiency 44
4.3 Soundness 46
4.4 Cross-domain Integration 47
4.5 Harmonization 48
4.6 Accommodation of innovation 49
4.7 Acceptability 50
4.8 Flexibility 51
4.9 Summary of Exercise 1 validation results 53
5 Results Exercise 2: ASCOS Safety Performance Indicator Framework and ASCOS Tool for Continuous Safety Monitoring 55
5.1 Expert group profile 55
5.2 The ASCOS Safety Performance Indicator Framework 57
5.2.1 Soundness 57
5.2.2 Completeness 58
5.2.3 Standardization 59
5.3 The ASCOS Tool for Continuous Safety Monitoring 59

5.3.1 Usefulness 59
5.4 Summary of Exercise 2 validation results 61
6 Results Exercise 3: ASCOS Risk Model and the Tool for risk assessment 62
6.1 Expert group profile 62
6.2 The ASCOS Risk Model 63
6.2.1 Soundness. 65
6.2.2 Completeness. 66
6.2.3 Standardization. 67
6.2.4 Acceptability 68
6.3 The ASCOS Risk Model and the Tool for risk assessment 68
6.3.1 Manipulability 70
6.3.2 Quantification capability 70
6.3.3 Cross-domain integration 71
6.3.4 Standardization 72
6.3.5 Acceptability from an Applicant and Certifier perspective 72
6.4 ASCOS Tool for risk assessment 72
6.4.1 Usability 73
6.4.2 Suggested areas for improvement 73
6.5 Summary of Exercise 3 validation results 74
7 Conclusions 77
References 79
Appendix A Definitions 81
Appendix B EXERCISE 1: Agenda 84
Appendix C EXERCISE 1: Questionnaire 86
Appendix D EXERCISE 2: Agenda 92
Appendix E EXERCISE 2: QUESTIONNAIRE 1 93
Appendix F EXERCISE 2: Questionnaire 2 97
Appendix G EXERCISE 3: Agenda 99
Appendix H EXERCISE 3: Risk model questionnaire 100
Appendix I EXERCISE 3: System Usability Scale (SUS) questionnaire 102

Appendix J EXERCISE 3: Risk Assessment Tool Questionnaire 103
Appendix K Causal model for Air Transport Safety (CATS) case studies 105
Appendix A.1 Introduction 105
Appendix A.1.1 Background 105
Appendix A.1.2 Objective and approach 105
Appendix A.1.3 Structure of the appendix 106
Appendix A.2 Identification of case studies 106
Appendix A.2.1 DGB 106
Appendix A.2.2 EASA 107
Appendix A.2.3 ASCOS 108
Appendix A.2.4 Case studies 109
Appendix A.3 CATS model and toolset 109
Appendix A.4 Case study results 112
Appendix A.4.1 Case study 1: Most frequent accidents and their probabilities 112
Appendix A.4.2 Case study 2: Identify precursors to accidents 118
Appendix A.4.3 Case study 3: Bird strikes 122
Appendix A.4.4 Case study 4: Remote Virtual Tower 133
Appendix A.5 Overall findings 139

This page is intentionally left blank

1 Introduction

1.1 ASCOS project background

Technological development and changes in regulatory structures in the aviation and ATM domains call for the improvements of current certification practices. A project funded by the European Commission (EC), ASCOS (Aviation Safety and Certification of new Operations and Systems) aims at bringing improvements in the certification practices of aeronautical products, operations and systems in a way that is consistent with the EU ACARE Vision 2020 [7], [8] and Flight Path 2050 [9]. In particular, ASCOS proposes an approach to certification that:

- is more flexible with regard to the introduction of new products and operations;
- is more efficient, in terms of cost, time and safety, than the current certification processes;
- Considers safety impact on all the relevant domains of the Total Aviation System and across the entire system life-cycle in a complete and integrated way.

Collectively, such enhancements are expected to bring a reduction of fatal accidents due to loss-of-control in flight, aircraft system or component failure or malfunction, aircraft ground handling damage, and Air Traffic Management related events. These are the five top commercial air transport accident categories that include the higher number of fatal accidents.

In previous work packages (WP) the ASCOS project teams have delivered the following results:

- WP1 results: An analysis of the existing European certification and rulemaking processes [10], the development and evaluation of a list of novel certification approaches [11], and the definition of an initial newly proposed certification approach [12];
- WP2 results: The development of a process and supporting tools for continuous safety monitoring, using a baseline risk picture for all parts of the Total Aviation System. This included the development of a safety performance indicator framework [3] and the baseline risk picture, i.e. the establishment of the current risk level of the various parts of the Total Aviation System.
- WP3 results: The development of a total aviation system risk assessment method and supporting tools that can be used for safety based design of new systems, products and/or operations. This included the development of a risk model based on accident scenarios and an approach to assess future and emerging risks. ASCOS WP3 also developed recommendations to improve aviation safety standards, including processes for safety assurance in operation and lessons learned requirements.

The established ASCOS User Group is involved in various stages of the project to keep the project focused and to facilitate the uptake of project results. They have an important role to play in the validation effort of the ASCOS products.

1.2 Objectives of WP5 Validation

The objective of ASCOS WP5 is to validate the main ASCOS results developed by the above work packages. To do so, this work package has been structured around the following five sub-work packages:

- WP5.1 Validation Strategy;
- WP5.2 Validation Plan and Scenario;
- WP5.3 Validation exercises and execution;
- WP5.4 Results analysis and reporting.

To date, the WP5.1 and WP5.2 have been completed and their results are available under deliverables D5.1 : *Validation Strategy* [13] and D5.2: *Validation Plan and Scenarios* [14] respectively. The former document applied the seven steps promoted by E-OCVM[2]—the EUROCONTROL standard validation framework—to the definition of ASCOS validation strategy. Deliverable D5.2 further refined this strategy by breaking it down into three validation plans for three different validation exercises. For each exercise the document reported aspects such validation objectives, methodology, roles, etc. Essentially, D5.2 completed the preparatory phase of the ASCOS validation exercises, and in doing so it provided the foundations for executing the ASCOS validation exercises in WP5.3 that are addressed by this deliverable.

1.3 Purpose of this study

This document presents the results of the validation activities carried out in the context of Work Package 5 of the ASCOS project. Such activities consist of three validation exercises within ASCOS WP5:

- Exercise 1 addresses the initial proposed ASCOS certification approach (as described in ASCOS D3.1);
- Exercise 2 addresses the ASCOS Safety Performance Indicator (SPI) framework and the ASCOS Tool for Continuous Safety Monitoring (ATCSM);
- Exercise 3 addresses the ASCOS Risk Assessment Model and the ASCOS Tool for Risk Assessment.

1.4 Structure of this document

The report is organized as follows:

- Section 2 sets the broader validation context for the ASCOS project, and present the validation goals;
- Section 3 describes and justifies the chosen validation methodology;
- Sections 4 to 6 reports the results of each validation exercise;
- Section 7 reports on the conclusions.

2 Validation of the proposed ASCOS enhancements

2.1 Definition of validation

Validation may mean different things to different people. A useful document to clarify its meaning is E-OCVM [2], a standard reference framework defined by EUROCONTROL for supporting ATM R&D projects during the validation of novel operational concepts. E-OCVM defines validation as "the iterative process by which the fitness-for-purpose of a new system or operational concept being developed is established" [2]. According to this definition validation investigates whether and to which extent a new system, product or service satisfies the originally intended purpose for which it was designed. In other words, validation answers the question "have we built the right system?"[2]. Implicit in the scope of this question is the consideration of the usage of the proposed solution—that is, the way in which the intended solution can actually be used by the intended user(s), in the intended field of practice, to deliver the expected benefit.

This focus is important because it allows to distinguish validation from its direct cousin—verification—which focuses, instead, on understanding whether the final product meets the initial specified requirements. Verification aims at understanding the technical quality and performance of the proposed solution. Verification's driving question is in fact "have we build the system right?" [2].

Distinguishing between verification and validation[1] is helpful to highlight the importance of capturing the target users' perspective and performances in the context of validation. Validation places a greater emphasis, in fact, on areas of performance such as end users' acceptability and suitability. In other words, validating a new system, service, or product requires appropriate means to capture effectively the end user's perspective.

This is necessary to bridge the gap existing between the view of the proponents of a novel solution (i.e. the developers and the stakeholders other than the end user) and the view of the end users. It is quite normal for the former group to hold a view about the potential offered by the new solution that differs from the view hold by the latter [16]–[18]. On the one hand, the proponents of the new solution work with an idealized abstracted view—which is, a view about the envisaged (and desired) benefit mechanism(s), the enabling conditions, the required components, etc. This view is essentially the view of a projected, idealized reality, and therefore it is imperfect: It usually fails to capture accurately all the (low level) implications related to the introduction of the proposed concept into the target field of practice. On the other hand, the target users, although lacking a complete understanding of the proposed solution, have a direct and concrete understanding of their work environment. Hence they are, by virtue of their (professional) experience, a valuable source of expert knowledge—expert knowledge that can help system's proponents to better understand how the proposed solution may actually fit into the target field of practice, and what value it can bring there.

For ASCOS, these considerations suggest that validating the ASCOS results requires a comprehensive consideration of the target user—certification and safety experts, possibly from different certification

[1] E-OCVM definition of validation and verification coheres with definitions of the same terms provided by the Project Management Book of Knowledge [15], a foundational project management standard.

domains. Failing to do so would risk to produce seriously flawed conclusions about the potential value that the ASCOS solutions can deliver. Of course, considering target users feedback does not mean to limit oneself to the collection of possible issues and expectations the experts in question may express—users may well have unrealistic expectations. Rather, the point is to interpret their feedback, understand the deeper rationale behind what they say, confront this with the projects developers view points, and understand the implications of this for the development of the project. The specific mode of involvement of the ASCOS target users in the context of ASCOS validation is addressed under the Methodology section (§ 3).

2.2 Validation approach

In devising the validation of ASCOS, it needs to be reminded that validation can be approached by two different evaluative perspectives, which essentially vary depending on purpose, i.e. how the results of the evaluation activities will inform project's stakeholders, and methodology, i.e. the means by which user and performance data is to be collected and processed.

When approached from a *formative evaluation* perspective, validation entails diagnosing the main limitations and strengths of the evaluated system. These are considered for the purpose of diagnosis and improvement. In other words, a formative perspective implies finding out what works and what does not with the proposed solution; and from here devise recommendations for improvement accordingly. Because of this focus, formative validation makes use, essentially, of qualitative research methodologies—such as a workshops with experts, heuristic or expert evaluations, open ended interviews, and questionnaires. The focus is on learning and refinement of the proposed solution. This type of validation activity does not require, necessarily, a formally defined baseline, although the validation activities can be designed to factor in the expert knowledge and experiences of the target users of the solution. And this user knowledge and experience provide an implicit baseline against which to assess the value of the proposed solutions.

When approached from a *summative evaluation* perspective, validation focuses on determining how the proposed solution compares against a baseline. The baseline can consists of an earlier version of the same product, when the focus is on understanding how much has the proposed product gained in maturity compared with its earlier version. Similarly, the baseline can consist of a comparable solution available on the market or in operation, when the emphasis is on understanding how well does the product perform compared to the state of the art. While formative evaluations focus on ways to improve the current concept, summative evaluations focus on proof and demonstration. Table 1 summarizes the main differences between the two evaluation perspectives.

Table 1. Formative vs summative evaluation

Formative evaluation	Summative evaluation
-Diagnosis, learning and improvement -Investigative	-Demonstration, and testing -Comparative

The selection of the focus of the validation depends on the maturity level of the proposed solution(s). While a summative evaluation would be the most appealing approach in evaluating the value of a novel concept, it cannot be neglected that exposing a novel concept too soon to such an evaluation would expose it to overly conservative reactions—i.e. premature rejection—by the relevant stakeholders.

2.3 Maturity assessment of ASCOS proposed solutions

In the case of ASCOS, a maturity assessment of its results was run during the definition of ASCOS Validation Strategy [13]. The purpose of this assessment was to ensure that the proposed validation activities were compatible with the actual maturity level of the proposed ASCOS enhancements.

The maturity level of ASCOS results was assessed by using the EUROCONTROL E-OCVM. Although devised for "classic" operational concepts such as the introduction of new operational procedures, or novel automation or displays, E-OCVM was deemed appropriate also for ASCOS. As note by D5.1, the notion of operational concept could be interpreted as "the new ASCOS proposed certification approach, the proposed continuous safety monitoring enhancements, and the proposed risk assessment enhancements" [13].

E-OCVM proposes a framework for maturity assessment that determines maturity levels of a solution by comparison with a reference maturity scale. This scale consists of six phases, which cover the entire system lifecycle, from the early needs identification phase to decommissioning. As E-OCVM is intended for the validation of R&D project only, the reference phases to be considered for the maturity assessment of ASCOS are the V0 to V3 levels (besides the V3 level the maturity scale enters industrialization and assumes that a consolidated model is in place (see Figure 1)).

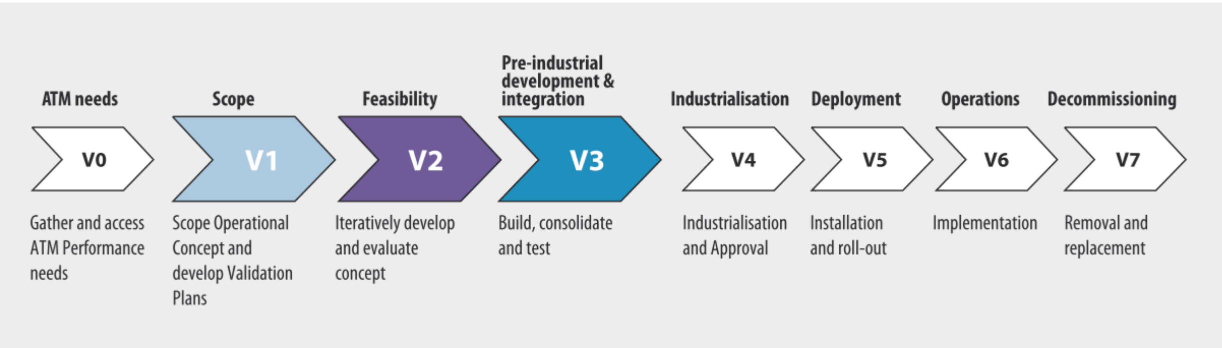


Figure 1. Operational concept lifecycle and maturity levels according to E-OCVM [22].

The ASCOS maturity assessment was done as part of the WP5.1. When mapping the ASCOS results against the E-OCVM maturity levels V0-V3 it was observed that:

- Work conducted in the ASCOS project prior the start of the WP5 addressed the need for change in current certification practices based on a description of the shortcomings and bottlenecks of current certification practices, see D1.1 [3]. This material provided an initial understanding of the needs and opportunities for improving current certification practices. However, it was noted that a performance framework was not defined in the previous phases of the project, and hence such a framework had to be developed in the context of the WP5 (the framework is described under § 3.4).

- Regarding the E-OCVM phase V1 objective, an initial description of the ASCOS products and how they deliver the intended benefit was available in various ASCOS deliverables, for example D1.2 [13], D1.3 [4], D2.1 [6], D2.3 [9], D3.1 [7], and D3.2 [8]. However, it was noted that the link between the expected benefits and the relevant problems and needs identified in V0 was difficult to demonstrate explicitly due the lack of a performance framework.
- Regarding the E-OCVM phase V2, the ASCOS project needed to further specify and refine the description of the proposed approach, process and supporting tools. In particular, the stakeholders and personnel involvement had to be addressed in more detail, so to further specify the roles involved in certification, responsibilities and tasks, together with any changes of procedures, team structure, communication and coordination between certification domains and organisations. This level of description is usually best covered by diagrams such as a use cases.
- By comparing the ASCOS project against maturity level V3 objectives, it was noted that no evaluation exercise was conducted to collect the evidence that the proposed ASCOS concept will be operationally feasible when integrated into ongoing operations.

In conclusion, the level of maturity of the ASCOS products was established to correspond to the E-OCVM levels V1 and V2. This made sense since ASCOS is a R&D project that aims to explore preliminary concepts of a novel certification approach.

The maturity level of the ASCOS solutions has a direct impact on the validation objectives and exercises, which should be appropriate to the level of the project maturity. In particular, the estimated maturity level of the ASCOS solutions (“prototype”, “feasibility”) was such that one cannot expect that they will bring at that stage the full benefits and meet all user expectations. Therefore, it was anticipated that a summative evaluation would have produced overly conservative results, which, in turn, would have not done justice to the actual value of the proposed ASCOS certification enhancements. In other words, running a summative, quantitative evaluation could have led to overly negative reactions before ASCOS has fully expressed its benefit. This kind of validation would be typically done at higher maturity levels.

On the other hand, it was anticipated that a formative evaluation would have been more beneficial for the project. A formative, qualitative evaluation would have allowed the collection of qualitative data, user feedback and recommendations useful for further concept exploration and refinement. This would support further development of the ASCOS solutions toward higher maturity levels. Furthermore, a formative evaluation looked much more feasible to the research team: In ASCOS it is difficult to establish a proper baseline, due to the available effort, lack of access to proprietary information, and other practical considerations (more details on this are available in § 3.3).

In conclusion, it is for these reasons that the chosen validation approach will not yet be about fully demonstrating or proving that all ASCOS solutions work well in practice, fully meet user expectations and bring benefits. The proposed validation approach will aim to provide feedback for the further development and evolution of the proposed products.

2.4 Validation scope: Selection of products to validate

The ASCOS project has developed several results (See Table 2. List of ASCOS results [19]. The results in bold are those that have been included in the scope of the validation.Table 2Figure 5). A selection of the results to validate was made considering the project resources and constraints. In making this selection, the following considerations applied:

- WP1's result consists of one item only: the ASCOS proposed certification approach. And this approach is the main result of the project, therefore it was included in the scope of validation by default;
- WP2 and WP3 have developed several results. The selection of the ASCOS P1 and 3 to validate has been based on their relative maturity level. It was decided to include in the scope of the validation those results that were sufficiently defined to be presented to a group of certification and safety management experts external to the ASCOS project. This selection has been coordinated with the leader of both WP2 and WP3.

These considerations have led to the identification of the results indicated in bold in the Table 2.

Table 2. List of ASCOS results [19]. The results in bold are those that have been included in the scope of the validation.

WP	Product/outcome	Responsible Partner
1	**Proposed certification approach**	EBENI
2	**Framework of Safety Performance Indicators (SPIs)**	NLR
	Process safety performance monitoring *(including list of 500 precursors associated with SPIs)*	ILOT
	ASCOS Tool for Continuous Safety Monitoring (ATCSM)	JRC
3	**Risk model *(improved CATS)***	NLR
	Tool for risk assessment	TU DELFT
	Safety assurance process in operation	APSYS
	Lessons Learned Requirements Process	APSYS
	Overall safety impact assessment method	ISDEFE
	E-learning environment	-

2.5 Validation objectives

The broader aim of ASCOS WP5 is to establish the fitness-for-purpose of the ASCOS project results. Consistently with the definition of validation reported under section 2.1, this aim implies understanding to which extent ASCOS can deliver the expected improvements to current certification practices. Basically this implies to understand whether ASCOS meet the needs and expectation of the certification stakeholders—either applicants or certifier authorities—of the different certification domains of the Total Aviation System (TAS).

As ASCOS comprises several results, this aim cannot be addressed directly but needs to be broken in three objectives. As discussed under section 2.4, the following ASCOS results have been selected:

- The ASCOS certification approach (WP1);
- The ASCOS Safety Performance Indicators framework and the ASCOS Tool for Continuous Safety Monitoring (WP2);
- The ASCOS Risk Model and the Tool for risk assessment (WP3).

The focus on these three sets of results translates into the following three validation objectives:

1. Investigate the value that the proposed certification approach may bring with regard the certification of novel aeronautical products and services;

2. Investigate the value that the proposed ASCOS SPI Framework and the ASCOS Tool for Continuous Safety Monitoring may bring to the area of continuous safety monitoring, with particular reference to safety monitoring in the context of certification;

3. Investigate the value that the proposed ASCOS Risk Model and the Tool for risk assessment may bring to the area of risk assessment, with particular reference to risk assessment in the context of certification.

The specific methodology by which these three objectives have been pursued is described next.

3 Methodology

3.1 Methodology Overview

The three objectives described under section 2.5 have been addressed by three separate validation exercises, each exercise evaluating the selected results of WP1, WP2, and WP3. Table 1 matches each exercise to the corresponding evaluated product, the relevant ASCOS deliverable describing the particular product, the structure of the exercise, and the day the exercise took place.

Table 3. Summary table of the three Exercises

Validation Exercises	Evaluated ASCOS Product	ASCOS Reference	Structure of Exercise	Dates
Exercise 1	Initial proposed ASCOS certification process	D1.2 [12]	-Feedback gathering	10th Oct 2014 (2nd Day of ASCOS UG meeting 3)
Exercise 2	• ASCOS SPI framework • ASCOS software Tool for continuous safety monitoring	D2.1 [3] D2.4 [4]	-Familiarization -Feedback gathering	28th Nov 2014
Exercise 3	• ASCOS Risk Assessment Model • ASCOS Tool for risk assessment	D3.2.3 [5] D3.3 [6]	-Familiarization -Interactive session -Feedback gathering	14th Jan 2015

The three exercises shared the same basic qualitative data gathering format: a one-day validation workshop with selected certification experts. This format included two phases:

1. **A familiarization phase,** which gave an opportunity to the participating experts to familiarize with the products and solutions under evaluation. Familiarization consisted of presentations of the ASCOS solutions and presentations of scenarios demonstrating the application of the ASCOS solutions. Only in Exercise 3 the familiarization phase included also an interactive session in which the experts had a chance to use the Tool for risk assessment. The familiarization phase was instrumental to prepare experts to the discussion groups and questionnaires used in the second phase: feedback gathering.

2. **A feedback gathering phase**. This phase represented the core of the workshop. It collected experts' subjective feedback through (i) guided focus group discussions and (ii) individual questionnaires. Both the discussions and the questionnaires were structured around the Key Performance Areas of the performance frameworks that were developed by WP5.1, and WP5.2 for the purpose of validation. The performance frameworks are presented in section 3.4. They cover important performance areas in which these solutions are expected to bring an improvement.

3.2 Rationale for the selected validation format

The use of the chosen data gathering format was justified by three considerations. These are reported below.

3.2.1 Consideration of constraints in accessing experts

The planning and the selection of the validation activities depended heavily on the availability of the certification and safety experts. These experts had to be recruited from the members of the ASCOS User Group. Such members are active professionals whose availability is constrained by the job demands of the different organizations they work for. So it was envisaged that it would have been more feasible to request UG members to attend three short 1-day events organized over the period October 2014 to January 2015, allowing about 1 month interval between the events, rather than a single event lasting 3 days. For this reason it was chosen an evaluation exercise format that limited UG attendance to a few dates.

3.2.2 Adequate to maturity level

As discussed under section 2.3, a formative, qualitative focus appeared as the most appropriate for ASCOS validation given the maturity level of the products under evaluation (maturity level was estimated to be in the range V1—V2). The format chosen for the validation exercises is consistent with this focus. It emphasizes, in fact, the elicitation of knowledge and insights from a group of selected experts through controlled group discussions, i.e. focus groups. Because of their professional experience, experts have a unique insider view about certification practices and their inner working logics, existing barriers, and multiple (conflicting) demands and constraints. Their insider view placed them in a unique position to provide the ASCOS partners with an in-depth view about the pros and cons of proposed ASCOS products. In short, the chosen approach emphasized learning and gaining insight, rather than demonstration and testing. This was consistent with the formative, qualitative focus that has been established for ASCOS validation.

It can be observed that the chosen format does not include a baseline. Because of this reason the consortium has also considered an alternative validation plan for the evaluation of the Logical Argument Approach: the retrospective application of the ASCOS Logical Argument Framework in the context of an existing real life certification case. This option looked appealing because it would have provided a realistic baseline against which to compare ASCOS performances. On this basis, a comparative assessment of the pros and cons of the proposed approach and the retrospective case study methodologies was made by the WP5. The assessment is available in Appendix B of D5.2. It was concluded that the retrospective application of the ASCOS approach in the context of an existing certification case would have encountered problems of both (i) feasibility and (ii) generalizability. Regarding feasibility, it was acknowledged both (i) the difficulty of accessing proprietary certification material not available in the public domain, and (ii) the lack of adequate managerial arrangements, resources, and time in the WP5 to do this kind of exercise. Regarding generalizability, it was acknowledged that the produced insights would have been difficult to generalize besides the certification context of the specific case. Certification comprises, in fact, several (reasonably independent) domains—type, airworthiness, continued airworthiness, design organization, production organization, etc.—, and to conduct a reliable, representative comparative assessment one should include more than a single domain, or a single case. The chosen format addresses this limitation by collecting expert feedback coming from a wide range of certification experiences and domains, independent of a particular certification case.

3.2.3 Rationale for the sequence of the Exercises

The specific sequence of the validation workshops has been defined considering that the WP1 product, i.e. the TAS logical argument framework, was the main product developed by ASCOS. Consequently it had a high priority for the consortium. Also, focusing on the evaluation of the WP1 first was expected to ensure that the evaluation of the WP2 and WP3 products could have been made with consideration of their role in the context of certification. This is because by the time of Exercise 2 and 3 occur, the UG members would have been already familiar, to some extent, with the proposed certification process, having encountered it already during Exercise 1. On the other hand, focusing on the results of WP2 and WP3 prior the WP1 would have risked to assess their value mostly as standalone items, rather than as contributory processes and tools to the WP1 product.

3.3 Recruiting of experts

During the preparation of the exercises, the selection of participating experts with a background appropriate for the solutions under evaluation was a major concern. In focus group research participants' professional background is one of the most important factors in ensuring that the discussions are fit for the purpose of the research. Thus, the recruiting of experts proceeded in order to:

- Ensure that the participating experts had a background in (i) certification (Exercise 1), (ii) safety monitoring (Exercise 2), and (iii) risk management (Exercise 3).
- Ensure that the participants represented as many different perspectives of the total aviation system as reasonably feasible. While it was anticipated that it was unlikely to have representatives of all the domains of the TAS, an effort was made to ensure that, as a minimum, each exercise was attended by representatives of both the applicant and the certifier perspective.

The experts were recruited from both the ASCOS User Group and personal contacts of the research team members. This ensured an attendance level compatible with the minimum number of participants requested by a focus group discussion. (The attendance level would not have been obtained by recurring solely to the ASCOS UG only.) The final profile of the expert groups attending the exercises is reported under sections 4.1, 5.1, and 6.1 respectively.

3.4 Performance frameworks

The validation of a novel product or concept is rarely a matter of evaluating the performance of a system along a single dimension, i.e. based on a single evaluative standard or point of view. Rather, the evaluation of the fitness-for-purpose needs to consider the multiple values, perspectives and viewpoints of the stakeholders that will be affected by the introduction of the novel system. This consideration is particularly relevant to the evaluation of complex products, systems, and concepts such as those proposed by ASCOS. The variety of actors involved in the certification domain, either in the certifier or applicant role, calls for the consideration of multiple view points during the validation process. These include supranational and national civil aviation authorities, manufacturers, airports, air navigation service providers, standard development bodies, etc.

For this purpose, three ASCOS specific performance evaluation frameworks, one for each Exercise, have been designed that captures the key areas of performance (KPA) in which the proposed solutions are expected to deliver their potential. An initial definition of the evaluation frameworks has been provided by the deliverable D5.1 [13, p. 1]. The definition of those frameworks was based on a review of the following sources:

- ASCOS dissemination material (brochures and website [20]);
- ASCOS deliverable D1.1 [10]. This deliverable consists of an analysis of existing regulations and certification processes aimed at identifying potential shortcoming and bottlenecks in current certification processes;
- Minutes of meetings with User Group members. In the ASCOS project a User Group (UG) was established with the intent to represent the relevant stakeholders involved in certification. These expectations of these stakeholders were extracted based on a review of past ASCOS technical meetings in which they participated;
- FAA Certification Airplane Certification Process Study (CPS) [21]. The FAA CPS is the most authoritative reference retrieved from the literature that documents the essential problems and bottlenecks in commercial aviation certification. The study is based on a comprehensive review of the processes and procedures associated with aircraft certification, operations and maintenance, commencing from the initial type certification activities, and extending to the continued operational safety and airworthiness processes. Cross-checking the areas of the framework with the findings of the CPS enabled to further ensure that the framework addresses areas that are of relevance for certification.

Table 4, Table 5, and Table 6 reports the final version of the performance frameworks that have been used during the exercises. These versions represent a refinement of their initial version (proposed under D5.1). In particular, while the initial versions were defined over the period Feb-May 2014, the refined versions were prepared immediately prior to each exercise, keeping into account the practical, logistics of the exercise (mostly available time for discussion and number of participants) which were not yet defined when the D5.1 was prepared.

KPA	Definition	Metric
1. Efficiency	The extent to which the proposed ASCOS certification approach allows to reduce the effort (cost, time, and training) needed by the applicant to obtain a certificate.	Questionnaire items
2. Soundness	The extent to which the ASCOS certification approach promotes, in certification, the consideration of relevant hazards and safety requirements that today are not or are poorly considered—with specific reference to cross-domain hazards and safety requirements.	Questionnaire items
3. Cross-domain integration	The extent to which the ASCOS approach promotes integration, coordination, and exchange of information across the different stakeholders that may be involved in the certification of a change.	Questionnaire items
4. Harmonization	The extent to which ASCOS looks compatible with the different certification approaches in use in different domains (e.g. ATC vs aviation) and geographical areas.	Questionnaire items
5. Accommodation of Innovation	The extent to which ASCOS makes more likely the certification of innovative products and systems, i.e. products and systems for which no standard are available.	Questionnaire items
6. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire items
7. Flexibility	The extent to which the proposed ASCOS approach can be applied to a broad range of different types of products, systems, and services, varying in size and complexity.	Questionnaire items

Table 5. KPAs for the proposed ASCOS SPI framework and Tool for Continuous Safety Monitoring (Exercise 2).

Evaluated ASCOS results	KPA	Definition	Metric
ASCOS SPI Framework	1. Soundness	The extent to which the ASCOS SPI framework promotes, in certification, the consideration of relevant hazards that today are not or are poorly considered (with specific reference to the TAS related hazards).	Questionnaire items
	2. Completeness	The extent to which the proposed SPI framework covers the different (certification) domains of the TAS.	Questionnaire items
	3. Standardization	The extent to which the proposed SPI framework can become a standard reference framework in use across the different actors of the TAS.	Questionnaire items
ASCOS Tool for Continuous Safety Monitoring	4. Usefulness	The extent to which the proposed ASCOS is perceived as a tool useful for supporting continuous safety monitoring,	Questionnaire items

Table 6. KPAs for the proposed ASCOS Risk Model and the Tool for risk assessment (Exercise 3).

Evaluated ASCOS results	KPA	Definition	Metric
ASCOS Risk Model	1. Soundness	The extent to which the ASCOS Risk Model promotes, in certification, the consideration of relevant hazards that today are not or are poorly considered (with specific reference to the TAS related hazards).	Questionnaire item
	2. Completeness	The extent to which the ASCOS Risk Model covers, the different hazards of the TAS.	Questionnaire item
	3. Standardization	The extent to which the proposed risk model can become a standard model used by the different actors of the TAS.	Questionnaire item
	4. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire item
ASCOS Tool for risk assessment	5. Manipulability	The extent to which the Tool for risk assessment promotes a useful means to manipulate—i.e. access, edit, modify— the ASCOS ESDs and FTs.	Questionnaire item
	6. Quantification capability	The extent to which the Tool for risk assessment can calculate the risk quantification.	Questionnaire item
	7. Cross-domain integration	The extent to which the Tool for risk assessment promotes integration, coordination, and exchange of information across the different stakeholders that may be involved in certification.	Questionnaire item
	8. Standardization	The extent to which the proposed tool can become a standard reference framework in use across the different actors of the TAS.	Questionnaire items
	9. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire items
	10. Usability	The extent to which the proposed risk model provides a usable means for supporting risk assessment.	SUS index

Table 4. KPAs for the ASCOS proposed certification approach (Exercise 1).

KPA	Definition	Metric
8. Efficiency	The extent to which the proposed ASCOS certification approach allows to reduce the effort (cost, time, and training) needed by the applicant to obtain a certificate.	Questionnaire items
9. Soundness	The extent to which the ASCOS certification approach promotes, in certification, the consideration of relevant hazards and safety requirements that today are not or are poorly considered—with specific reference to cross-domain hazards and safety requirements.	Questionnaire items
10. Cross-domain integration	The extent to which the ASCOS approach promotes integration, coordination, and exchange of information across the different stakeholders that may be involved in the certification of a change.	Questionnaire items
11. Harmonization	The extent to which ASCOS looks compatible with the different certification approaches in use in different domains (e.g. ATC vs aviation) and geographical areas.	Questionnaire items
12. Accommodation of Innovation	The extent to which ASCOS makes more likely the certification of innovative products and systems, i.e. products and systems for which no standard are available.	Questionnaire items
13. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire items
14. Flexibility	The extent to which the proposed ASCOS approach can be applied to a broad range of different types of products, systems, and services, varying in size and complexity.	Questionnaire items

Table 5. KPAs for the proposed ASCOS SPI framework and Tool for Continuous Safety Monitoring (Exercise 2).

Evaluated ASCOS results	KPA	Definition	Metric
ASCOS SPI Framework	5. Soundness	The extent to which the ASCOS SPI framework promotes, in certification, the consideration of relevant hazards that today are not or are poorly considered (with specific reference to the TAS related hazards).	Questionnaire items
	6. Completeness	The extent to which the proposed SPI framework covers the different (certification) domains of the TAS.	Questionnaire items
	7. Standardization	The extent to which the proposed SPI framework can become a standard reference framework in use across the different actors of the TAS.	Questionnaire items
ASCOS Tool for Continuous Safety Monitoring	8. Usefulness	The extent to which the proposed ASCOS is perceived as a tool useful for supporting continuous safety monitoring,	Questionnaire items

Table 6. KPAs for the proposed ASCOS Risk Model and the Tool for risk assessment (Exercise 3).

Evaluated ASCOS results	KPA	Definition	Metric
ASCOS Risk Model	11. Soundness	The extent to which the ASCOS Risk Model promotes, in certification, the consideration of relevant hazards that today are not or are poorly considered (with specific reference to the TAS related hazards).	Questionnaire item
	12. Completeness	The extent to which the ASCOS Risk Model covers, the different hazards of the TAS.	Questionnaire item
	13. Standardization	The extent to which the proposed risk model can become a standard model used by the different actors of the TAS.	Questionnaire item
	14. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire item
ASCOS Tool for risk assessment	15. Manipulability	The extent to which the Tool for risk assessment promotes a useful means to manipulate—i.e. access, edit, modify— the ASCOS ESDs and FTs.	Questionnaire item
	16. Quantification capability	The extent to which the Tool for risk assessment can calculate the risk quantification.	Questionnaire item
	17. Cross-domain integration	The extent to which the Tool for risk assessment promotes integration, coordination, and exchange of information across the different stakeholders that may be involved in certification.	Questionnaire item
	18. Standardization	The extent to which the proposed tool can become a standard reference framework in use across the different actors of the TAS.	Questionnaire items
	19. Acceptability	The extent to which the proposed ASCOS approach looks acceptable to the applicant and the certifying authority.	Questionnaire items
	20. Usability	The extent to which the proposed risk model provides a usable means for supporting risk assessment.	SUS index

3.5 Familiarization and training

Training was a major concern across all the exercises. It was anticipated that the quality of the feedback that could be collected from the experts depended, indeed, on their understanding of the products to be evaluated. For this reason, the research team took appropriate actions to ensure the participants achieved a familiarity level as high as practicably feasible. Prior to each exercise, the relevant ASCOS partners were required to prepare an exhaustive presentation about the result(s) to be evaluated. Besides that, familiarization occurred differently for the three exercise:

- For Exercise 1, familiarization and training took place the day before the exercise;
- For Exercise 2 and 3 it took place on the morning of the day the exercise took place.

The next three sections describe the familiarization and training phase of the three exercises.

3.5.1 Exercise 1 Familiarization and training

Familiarization and training for Exercise 1 occurred on the day preceding the exercise. This was due to the fact that ASCOS Exercise 1 took place during Day 2 of ASCOS UG Meeting 3, while Day 1 of this event was intended for evaluating the case studies developed in the context of ASCOS WP4. So, during Day 1 of ASCOS UG Meeting 3 the participants were given exhaustive presentations about the ASCOS project, the certification approach, and its application in the context of four certification situations—i.e. the WP4 case studies. The cases (succinctly described under § 3.7.1) were presented over the course of a day. They had a dedicated 1h slot each, thus each slot left plenty of time for in-depth discussions about each case and the underlying ASCOS method. So, while serving the purposes of WP4, Day 1 of UG Meeting 3 also served the purpose of familiarization and training for Exercise 1.

3.5.2 Exercise 2 Familiarization and training

Overview of the ASCOS project and the proposed certification approach. In Exercise 2, the participants were introduced to the ASCOS project by means of a short 10 min overview presentation, which covered ASCOS project objectives, duration, funding, and project phases. This was followed by a longer and more in-depth presentation about the proposed ASCOS Total Aviation System certification process. This presentation covered the steps of the approach, its features, the top level argument and its different modules.

Presentation of the ASCOS SPI framework. Next, participants were introduced to the ASCOS framework of Safety Performance Indicators (SPIs) for the total aviation system, which was developed by the WP2.1. The presentation covered the principles behind the framework—the Reason's accident model; the concept of barrier and barrier analysis; general requirements for meaningful SPIs—and then the framework itself—which contains SPIs for components, humans, organization and system of organizations.

Presentation and non-interactive demo of the ATCSM. Finally, the participants were introduced to the ATCSM by mean of a short presentation and a non-interactive demo. The presentation covered basic aspects of the ATCSM, such as the tool's interface, the graph of SPIs trends the tool can deliver, and the building up of the query needed to deliver the SPIs graph. Also it was explained that the tool is compatible with ECCAIRS 5 infrastructure and taxonomy.

The demo was helpful to see how the tool works in practice. Examples of SPI graphs were shown to demonstrate the kind of graphs the tool can generate in relation to the scenarios reported under section 3.7.2. Note that the demo was kept non-interactive, i.e. the participants did not interact with the tool. This choice was intently made for achieving the dual purpose of (i) maximizing the time spent for useful group discussions, while (ii) avoiding discussions related to the structure and quality of the reported occurrence data and to the interface management aspects of the tool. Based on previous workshops, it was anticipated that experts not familiar with ECCAIRS, instead of commenting on the broader concept behind of the tool under evaluation, may have overly focused (i) on the effort requested to adapt their way of expressing issues to the ECCAIRS taxonomy and (ii) on graphical interface issues.

3.5.3 Exercise 3 Familiarization and training

Overview of the ASCOS project and the proposed certification approach. At the start of the day, participating experts were introduced to the ASCOS project by means of a short 10 min overview presentation, which covered ASCOS background, project objectives and relation between these objectives. This was followed by a second presentation that provided an overview of the Total Aviation System (TAS) certification approach. This presentation covered the steps of the approach, its features, the top level argument and its different modules. As for Exercise 2, these two initial presentation concluded the introduction part of the familiarization session.

Presentation of the ASCOS risk model. Subsequently, participants were given a presentation about the ASCOS Risk Model, and its components (the Event Sequence Diagrams, Fault Trees, and Fault Trees base elements). Basic information on the methodology and the model was provided such as how the model relates to the CATS model, and how it compares to comparable risk assessment models, e.g. the EUROCONTROL Integrated Risk Picture., and, most importantly, how it could be used in a risk assessment and certification context.

Demonstration of the ASCOS Risk model in a Remote Virtual Tower scenario. The subsequent presentation demonstrated how the ASCOS Risk Model can be used in current, future, and emerging risk assessment. As an example it was chosen the introduction of a Remote Virtual Tower, whose introduction is foreseen for small and medium airports. Further information on this scenario is provided under section 3.7.3. This presentation was the last one dedicated to the risk assessment specifically.

Guided interactive demo of the ASCOS Tool for risk assessment. The third presentation was instrumental to let the experts understand the basic features and functionalities of the ASCOS Tool for risk assessment. The session consisted of a guided demo. Firstly all the experts were provided with log-in details to allow them to participate in the exercise using their own laptop. Then, the experts were requested to follow the step-by-step instructions provided by the session leader. This demo let the participants familiarize with the following basic tools functionalities: logging into the system, accessing and exploring the risk model, adding/removing an ESD and FT elements and modifying their probability values, saving and exporting to MS Excel.

3.6 Feedback Gathering

The feedback gathering phase of each exercise was structured around discussion sessions. In each session participants' feedback was gathered by means of (i) an individual questionnaires with quantitative scores, and (ii) by means of a focus group discussion. At the beginning of each session, the questionnaire was handled to the participants, who had about 10 to 15 minutes to fill it up individually and return it to the research team. Then, the group was lead into the group discussion by the moderator. Experts were invited to illustrate their opinions about the positive and negative aspects of the ASCOS approach. It must be noted that, considering the qualitative nature of the exercise, the purpose of the questionnaire was more that of triggering the kind of individual thinking that was functional to the exercise rather than generating highly reliable quantitative scores. In other words the quantitative questionnaire items serve the purpose of stimulating participants to think about specific areas (the KPAs), which were then addressed during the group discussions. The next three sections describe the way in which data collection took place during the three exercises.

3.6.1 **Exercise 1** Feedback gathering

As explained earlier (see § 3.5.1), for Exercise 1 the familiarization session took place on the day before the exercise. Thus, a part from an initial one hour time slot intended as a summary of the ASCOS approach, the exercise was structured around a total of three discussion sessions aimed at collecting the experts' views on the proposed ASCOS certification approach. Each discussion session was moderated by a facilitator in order to seek participants' feedback on these areas:

- Discussion 1: Efficiency and Soundness of the proposed approach;
- Discussion 2: Cross-domain integration, Harmonization, and Accommodation of innovation;
- Discussion 3: Acceptability and Flexibility.

These areas reflected the KPAs of the performance framework specifically defined for Exercise 1 (see § 3.4).

The discussions were run in parallel, meaning that prior to the start of the first discussion the experts were split in two separates groups, each working in a dedicated room. This choice was made to keep the group size to a manageable level. Also, in order to maximize the potential for useful discussions, participants were assigned to each group depending on their professional background. The goal was to ensure that each group included both experts representing the applicant perspective (e.g. manufacturers, aircraft operators, and air navigation service providers), and experts representing the certifying authority perspective (e.g. national aviation authorities). At the end of the day, the participants were presented with a list of the main findings of the day for the purpose of (initial) corroboration.

3.6.2 **Exercise 2** Feedback gathering phase

Exercise 2 feedback gathering comprised two discussions sessions. The first discussion session aimed at collecting experts' feedback on the SPIs framework and the ATCSM in a continuous safety monitoring context. The second discussion session aimed at collecting experts' feedback on the SPIs framework and the ATCSM in the context of certification. Altogether, the discussion sessions aimed at collecting feedback on the performance areas of soundness, efficiency, standardization, acceptability, and accommodation of innovation, consistently with the performance framework defined for this exercise (see § 3.4). Eventually, this plan suffered from the lack of continuous participation: About half of the group of participants could only participate in the morning session; hence the planned areas of performance were not covered as initially planned.

3.6.3 **Exercise 3** Feedback gathering phase

Exercise 3 feedback gathering comprised two discussions sessions. The first discussion session aimed at collecting feedback on the ASCOS Risk Model. The second one collected feedback on the ASCOS Tool for risk assessment. The two sessions collected feedback on the areas of soundness, completeness, standardization, acceptability for the ASCOS Risk Model; and manipulability, quantification capability, cross-domain integration, and acceptability for the ASCOS Tool for risk assessment. This was consistent with the performance framework devised for this exercise (see § 3.4). This exercise allowed the collection of useful observations and feedback also during the training and the interactive sessions relative to the Tool for risk assessment.

3.7 Validation scenarios

This section describes the scenarios or use cases that were shown to the experts during the three exercises. The purpose of these scenarios was to demonstrate possible uses of the evaluated concepts, the selection of the ASCOS results, in real life certification contexts.

3.7.1 **Exercise 1 scenarios**

Exercise 1's scenarios exemplified the application of the ASCOS Logical argument framework in the context of certification. These scenarios corresponded to the initial version of the cases that were developed in the context of the ASCOS WP4. The cases demonstrated the execution of the first three stages of the ASCOS proposed logical argument framework—namely (1) define the change, (2) define the certification argument, (3) develop and agree the certification plan—, as executed in the four different scenarios.

Identification of the Scenarios

The selected scenarios were deemed appropriate to ASCOS because they represented one class of certification situation in which improvement in current practices is desirable: the introduction of innovative technological systems and services. This is a sensitive area of improvement due to factors such as high costs, length of the certification process, and domain and lifecycle compartmentalization in certification.

Scenario 1.1: Certification of RPAS Failure Management System

This scenario demonstrated the application of ASCOS proposed approach to the certification of a change involving the introduction of a novel system for a Failure Management System for a Remotely Piloted Aircraft System (RPAS). The RPAS is conceived as an adaptation of a (generic) existing civil fixed wing cargo aircraft with flight crew on-board. The RPAS definition includes the aircraft, the ground station used to pilot the aircraft and the communications link between aircraft and ground station. It shall cover the specific aspects of training and maintenance of the RPAS product.

Scenario 1.2: Certification of an Automated Aircraft Recovery System

This case demonstrated the application of ASCOS proposed approach to the certification of a change involving the introduction of a novel on-board Automatic Aircraft Recovery system (AARS). This system is intended as remedy against loss of control accidents, which are often triggered by insufficient situational awareness on the crew's side in combination with rare failure modes. The purpose of the AARS is to restore automatically the aircraft safe flight condition after the occurrence of any potential upset or any deviation from a normal control regime. In other words, whenever any system failure or any other flight disturbing event leading to loss of crew's situational awareness occurs, this system will bring the a/c to a stable flight condition, within the normal flight envelope of the aircraft. The system will be in control of the restored situation for a sufficient period of time to permit the crew to diagnose the potential problem and to restore situational awareness.

Scenario 1.3: Certification for Ground De-Icing operations

This case demonstrated the application of ASCOS proposed approach to the approval of a change involving a transfer in responsibility for Anti-Icing/De-Icing ground operations from airlines to an individual Anti-Icing/De-Icing service provider external to the airline. Today, airlines are responsible for establishing procedures for ground de-icing and anti-icing and related aircraft inspections to allow the safe operation of the aircraft. The air operator may provide a ground de-icing and anti-icing service itself, or contract out the service to an external provider. In both cases, responsibility for safety and for compliance with applicable regulations remains with the airline operator.

The change considered in this scenario entails that an individual de-icing/ anti-icing service provider becomes certified itself, and becomes responsible and accountable for providing a safe service. The air operator is then no longer responsible for this. A change in regulations would be required for this. The overall goal of the change is to enhance safety. It is principally a change in responsibilities, which may be accompanied by an adaptation of the human roles, responsibilities and procedures. For example, in stage 5 it may be decided to expand the responsibilities of the de-icing/anti-icing service provider to include the decision to apply de-icing or anti-icing. The actual de-icing/ anti-icing activities and involved equipment can in principle remain unchanged.

Scenario 1.4: Certification of a an integrated ground surveillance system

This case demonstrated the application of ASCOS proposed approach to the approval of a change involving an Integrated Surveillance System (ISS) over Frankfurt PAM area (slightly greater than the TMA). This system is intended to be in charge of all the Air Traffic Control (ATC) Surveillance functions as the primary means of surveillance. Its aim is also to progressively replace all current PSR and SSR of the area. In the case of SSR, a possible conclusion of the system deployment maybe that after the attrition period coverage with SSR shall still be kept as a secondary mean of surveillance.

3.7.2 Exercise 2 scenarios

Identification of Scenarios

The validation scenarios for Exercise 2 have been designed consistently with the following criteria so to make them relevant to ASCOS:

- Coverage of different operational issues of the EASA safety plan (EASp) [22], i.e. Runway excursion, Mid-air collision, Controlled Flight Into Terrain, Loss-of-control in flight, Ground collisions;
- Coverage of different category of ASCOS defined Safety Performance Indicators, i.e. Technology, Human, Organization, System of Organization;
- Consideration of different Total Aviation System domains, i.e. Aircraft, Air Traffic Management, Airport Operations, etc.

Based on these criteria several possible scenarios have been considered in preparation for this evaluation. However, the final choice has been constrained by the difficulty to retrieve the data set needed for each scenario. In particular, the retrieval of occurrence data was constrained by the availability of suitable publicly available occurrence data in ECCAIRS format. In addition, exposure data had to be retrieved by public sources such as EUROCONTROL annual performance reports. Hence, the availability and suitability (quality) of the data depended on these sources.

Scenario 2.1: Runway Incursions

This scenario exemplifies the use of the tool on the side of a safety analyst working for a European Air Navigation Service Provider, which is planning to implement the Runway Incursion and Collision Avoidance System (RIMCAS) at one of its major international airports. The safety expert is tasked to investigate the effect of introducing RIMCAS in some of the major European airports in the last years. As part of the exercise, the analyst decides to take as reference the period between January 2011 and December 2013 and to compare the number of Severe Runway Incursions occurrences in 6 European airports: Geneva, Frankfurt, Milan Malpensa, Brussels National Airport, Paris Charles De Gaulle, and Amsterdam Schiphol. The first three have been equipped with RIMCAS starting from 2011, while the last three are not yet equipped with it. Although the number of flight movements is different in the various airports, for the purpose of the demonstration it is decided that the analyst will make this first comparison to determine if any remarkable difference exist.

Scenario 2.2: Monitoring occurrences related to auto-flight systems

This scenario exemplified the use of the system by a safety officer working for the Civil Aviation Authority of a European country. S/he wishes to assess the reliability of auto-flight systems in the context of a study concerning Loss-of-Control in flight incidents/accidents. S/he decides to investigate the number of auto-flight system failure occurrences involving commercial aviation aircraft flying over Europe in the last 5 years and to compare it with the previous 5 years.

Scenario 2.3: Monitoring crew reaction to E-GPWS

This case exemplified the use of the tool by a safety analyst who wishes to monitor the effect of introducing E-GPWS in the nineties and the effectiveness of the existing E-GPWS training practices. For this purpose s/he analyses all the Controlled Flight into Terrain accidents and incidents involving a response to an E-GPWS alert occurred in the last five years in Europe.

3.7.3 Exercise 3 scenarios

Identification of scenarios

The following criteria were defined for identifying and selecting a set of scenarios that could be used in exercise 3 to meet the validation objectives. To avoid confusion it is useful remembering that the scenarios presented here are not the accident scenarios—i.e. the Event Sequence Diagrams, built into the ASCOS Risk Model—but validation scenarios like those defined for the other validation exercises, i.e. scenarios showing how the risk model and the tool can be used.

As for the other exercises, these are demonstrative scenarios intended to show how the risk model and the tool can be used.

The criteria used for selecting the scenarios were the following:

- Be relevant to the context of certification;
- Coverage of different stakeholders: certifying authority and applicant (manufacturer, airport, aircraft operator);
- Coverage of different certification “products”: an aeronautical product (system, equipment), operation.
- Coverage of different level of size and complexity;
- Coverage of new/innovative novel and existing/derivative products and services;
- Consideration of different Total Aviation System domains: Aircraft, Air Traffic Management, Airport Operations;
- Consideration for future (FAST Areas of Change).

One precondition considered in the preparation of the exercise and the scenarios was that the ASCOS risk model would not have to be expanded. In other words no new ESDs or Fault Tree branches could be added to the risk model, but only small modifications to the model could be made. This was due the available time and constraints.

Scenario 3.1: Demonstration of the risk model: Introduction of Remote Virtual Tower

This scenario illustrated how the ASCOS Risk Model can be used in a risk assessment context. The chosen example was the introduction of a Remote Virtual Tower, whose introduction is foreseen for small and medium airports. Facing such a change, an airport operator and the relevant authority may wonder (i) whether the effect will be visible in the TAS; (ii) which accident scenarios will be affected; (iii) how much risk will be introduced by the change, and (iv) how much safety improvement is brought in by the change. Such questions can be addressed using the ASCOS Risk Model. An analyst has to estimate, first, how the change influences the risk model. In other words she/he has to find out which scenario(s), or ESD(s), of the model are affected by the change. Second, the analyst should modify the (local) probability values of the affected parts of the current risk model to represent the implemented change. This will allow her/him to calculate a new (global) probability value for the concerned ESD end state, and compare this with the probability value that was valid prior to the introduction of the change (i.e. the probability value reflecting the baseline situation). Regarding this scenario, these steps were applied to the ESD element *ASC-32: Runway Incursion*. In particular, it was shown:

- The reasoning that led to the identification of the relevant pivotal element of this ESD—*ATC does not resolve the conflict (ASC 32 b1)*— that had to be modified;
- How to access the corresponding Fault Tree, i.e. the fault tree that has ASC 32 b1 as a top event;
- Which elements of this FT undergo a positive change (i.e. reduction of failure probability), and which ones undergo a negative change (i.e. increase of failure probability);
- Finally it was shown how these changes alter the probability value of the outcome event of the ESD in question compared to the initial probability value (of the same ESD).

Scenario 3.2: Tutorial with the Tool for risk assessment

This validation scenario introduced the Tool for risk assessment. The scenario provided participants with a basic understanding of the software interface and its functionalities or capabilities, i.e. how to add/modify ESD, FT, and barriers and how to compute overall risk figure. Also this case made use of the runway excursion at take-off scenario to illustrate the working details of the system. During this case participating experts were invited to play with the software and explore its functionalities. In order to speed up familiarization with the software tool, participants were provided with user manual and a link to access the software tool prior to the meeting.

Scenario 3.4: Introduction of Upset Prevention and Recovery Training (User Trial)

This scenario consisted of a guided group-exercise in which the participants had to modify—by using the ASCOS Risk Assessment Tool— the probabilities values associated to the introduction of a change: the "Upset Recovery and Prevention Training. This scenario was helpful to demonstrate the application of the ASCOS Tool for risk assessment in the context of emerging risk assessment in the Total Aviation System. In particular, it was helpful to demonstrate how the risk model elements element can be modified to represent the impact of a safety improvement (i.e. UPRT). The chosen scenario, UPRT, addressed the EASp action item AER4.16: "Develop regulations which ensure that initial and recurrent pilot training and checking is adequate to provide a pilot with the knowledge, skills and attitude to be competent in preventing and, if necessary, recovering from a loss-of-control in flight situation."

3.8 Materials

This section describes the materials that were used to gather the participants' feedback.

3.8.1 Questionnaires

A questionnaire was developed for each of the ASCOS results under evaluation, for a total of five questionnaire. Each questionnaire covered the areas of performance that were relevant for the particular result. These have been presented in section 3.4. For each area the participants were requested to express a rating on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaires are available in the appendix.

3.8.2 SUS questionnaire (Exercise 3)

Exercise 3 only made use of a standardized questionnaire: The System Usability Scale or SUS [23]. This is a ten-item questionnaire with five response options that yields a single score representing a composite measure of the overall usability of the system under evaluation. In the exercise, SUS was used for measuring the perceived ease of use, i.e. usability, of ASCOS Tool for risk assessment. Participants were asked to fill it shortly after having completed their interactive session with the tool (see "Tutorial with the software tool" under § 3.7.3).

3.9 Data analysis

The data collected during each exercise included: researchers' notes, questionnaire ratings, and audio recordings of the structured focus group discussions. Data was analysed separately for each exercise, and consisted of basic statistical analysis combined with Qualitative Content Analysis (QCA) [24]. Descriptive statistics was used to analyse the questionnaire ratings, including the SUS ratings collected in Exercise 3. The QCA was used to analyse the rationale behind these ratings. This was in fact the most important source of data for this exercise: The questionnaire items, indeed, played mostly the role of useful triggers for reflection and discussions—but the real and important data come from the qualitative description of the reasons behind the experts' ratings.

The QCA entailed three phases: A preparatory phase followed by two cycles of coding. The preparatory phase consisted of collecting and aggregating all researchers' notes, on listening the audio files, and on transcribing them. This material was reviewed extensively and provided the basis for the next two data analysis phases. The first cycle of coding involved searching the transcripts of the focus group discussions in order to look for passages that pointed at problems, attitude, description of experts' professional experiences which were relevant for the purpose of the study—i.e. understanding the value of the proposed ASCOS solutions. Whenever such passages were found they were marked and assigned a code. This phase allowed to develop a list of initial, first level codes. In the second coding cycle, the initial codes were searched in order to find similarities and differences between them: The aim was to double check these codes, discard the irrelevant ones, and cluster the relevant ones under higher order categories, i.e. categories that capture relevant key findings.

An intermediate result of this data analysis process was the preparation of the minutes of the meetings. These were instrumental in offering an opportunity to the participating experts to comment and amend what the research team understood from the discussions, prior the data was analysed. In short the minutes were used for the purpose of corroboration.

3.10 Limitations

The main limitation encountered during the planning of the exercises regarded the accessibility to experts. The number of ASCOS UG members interested on the exercises was insufficient to ensure the adequate level of participation (estimated between 6 and 12 for focus group research) across all of the three exercises. This difficulty was mitigated by recruiting additional experts external to the ASCOS UG. Eventually this solution ensured an adequate participation level for Exercise 1 and 3, but not for Exercise 2. This exercise was attended by a total of five experts, some of whom had to leave the meeting at the end of the morning session for reasons external to ASCOS. This affected, inevitably, the depth and length of Exercise 2 group discussions, as a result not all of the discussion topics (the KPAs) were addressed as planned.

None of the exercises benefited from EASA participation. This means that the range of potential benefits of the proposed ASCOS results was not explored in its entirety. Future EASA participation in validation exercises would provide additional benefits and e.g. gain further insight into:

- How the ASCOS solutions compare with current type certification practices. The EASA is the European certifying authority responsible for type certificates, and having its view represented directly in the familiarisation/validation exercises would have provided a further understanding of the potential that ASCOS may deliver in this certification domain;
- How the proposed solution compare with the different certification practices of different domains and countries. Because ofs its unique institutional role, the EASA enjoys a unique centralized view over the different perspectives of different national CAA, OEM, and maintenance organizations of different countries and certification domains. A better grasp of these perspectives, as “channelled” by relevant EASA experts directly in the exercises, would have allowed to gather more information about the potential for standardization and cross-domain integration of the ASCOS solutions.

However, it should be noted that EASA has thoroughly reviewed the initial proposed ASCOS certification approach [12, 38, 39] and the framework for Safety Performance Indicators [3], both subject of the validation exercises. The input and feedback of EASA has therefore been incorporated during the development phase of these products. EASA has provided significant inputs and feedback to ASCOS (in particular during the first 18 months of the project). EASA has hosted an ASCOS – EASA Workshop on the 19th of April 2013 (with about 25 participants from EASA), and has participated in the 1st User Group Workshop and Final Dissemination Forum.

Therefore, overall, the support from EASA over the whole duration of the ASCOS research project exceeds by far that of a typical European Commission safety research project. The outcome of the validation exercises may be biased negatively because it was not possible to take EASA's inputs into account during the exercises.

Finally, participation of experts was not constant across the three exercises. In practice only one expert managed to attend all of the exercises, while two other experts attended Exercise 1 and 3. This meant that the basic familiarization session with the ASCOS project and the certification approach had to be repeated for all the three exercises. Inevitably this limited the time that could be spent on further training and evaluation of the evaluated ASCOS results.

4 Results Exercise 1: The ASCOS certification approach

This section reports the result of ASCOS Validation Exercise 1. This exercise evaluated the proposed ASCOS certification approach. The section is organized as follow:

- Section 4.1 provides an overview of the group of experts involved in the exercise;
- Sections 4.2 to 4.8 reports on the results specific for the ASCOS certification approach. In particular each section reports the feedback collected for each key performance area (KPA) in which ASCOS is supposed to deliver an improvement;
- Section 4.9 summarizes the validation results of Exercise 1.

4.1 Expert group profile

Exercise 1 collected the views and feedback of a total of nine experts, which were recruited from the ASCOS User Group and among the ASCOS research team's contacts. The majority of the experts (n=6) had direct experience in aviation certification. Their mean experience in certification was 25 years (*stdev*=12.49) (see the table below). The non-certification experts had a background in safety management, commercial aviation piloting, and Air Traffic Management. The graph in Figure 2 shows the types of organizations from which the experts in the two came from: National Civil Aviation Authorities (CAA), Original Equipment Manufacturer (OEM), Air Operators, and Air National Service Providers (ANSP). Overall, the group composition and representation seemed adequate to the goal of the exercise.

Table 7. EXERCISE 1: Profile of the expert group

Size (n)	9
Average experience of participants in certification (years)	24
Min	7
Max	40
St Deviation	12.49

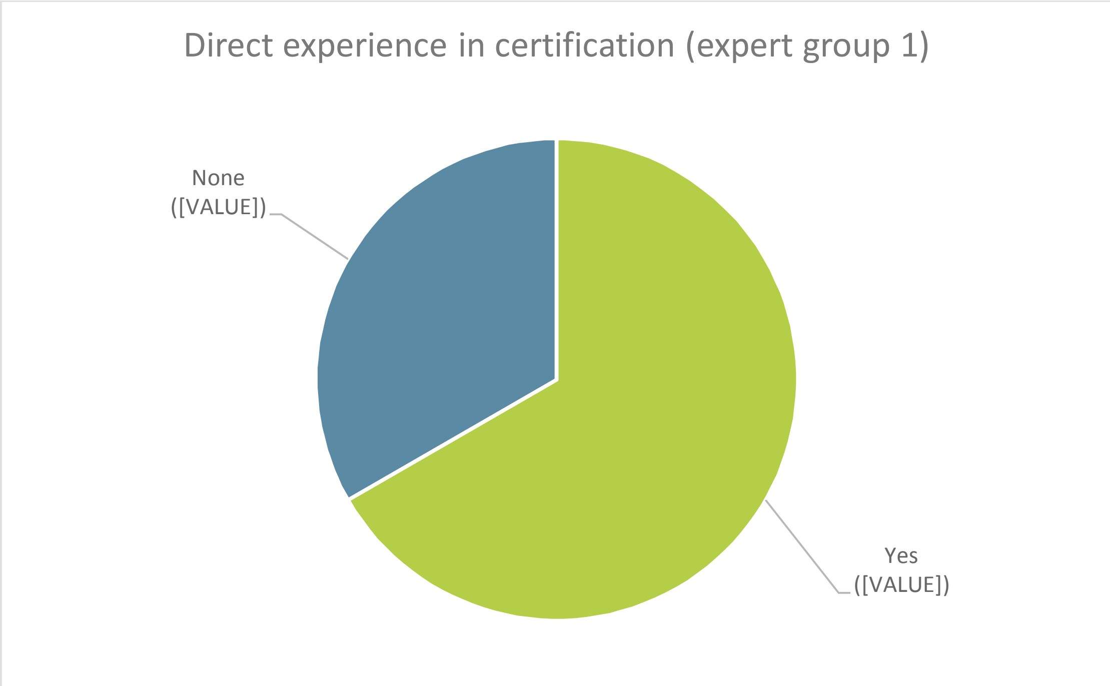


Figure 2. EXERCISE 1: Experience of the participating experts in certification.

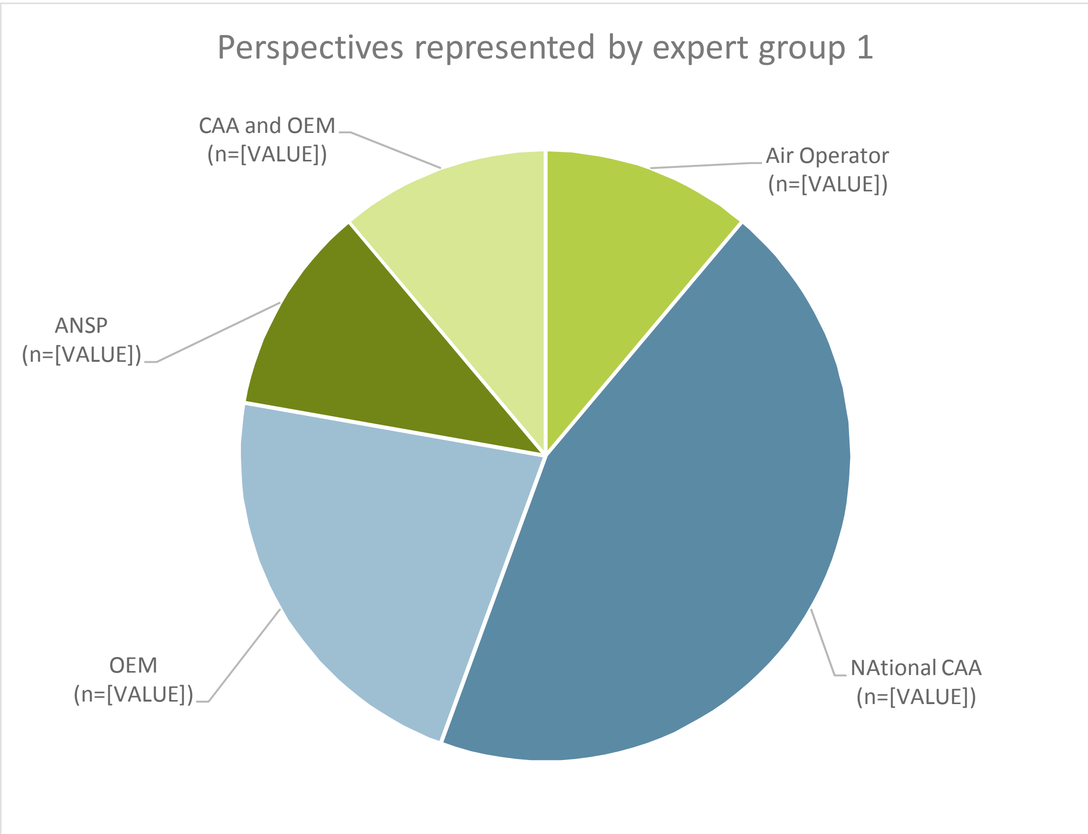


Figure 3. EXERCISE 1: Perspectives represented by the expert group 1.

4.2 Efficiency

This area of performance concerns the ability of ASCOS to make more cost-effective the initial phase of certification—i.e., the phase leading to the initial certification approval by an authority. In particular the impact of ASCOS on the (i) time, (ii) effort, and (iii) training required of both the certifying authority and the applicant to get the initial certificate was evaluated. The graph shows that none of the participating experts believe that the ASCOS approach could improve efficiency.

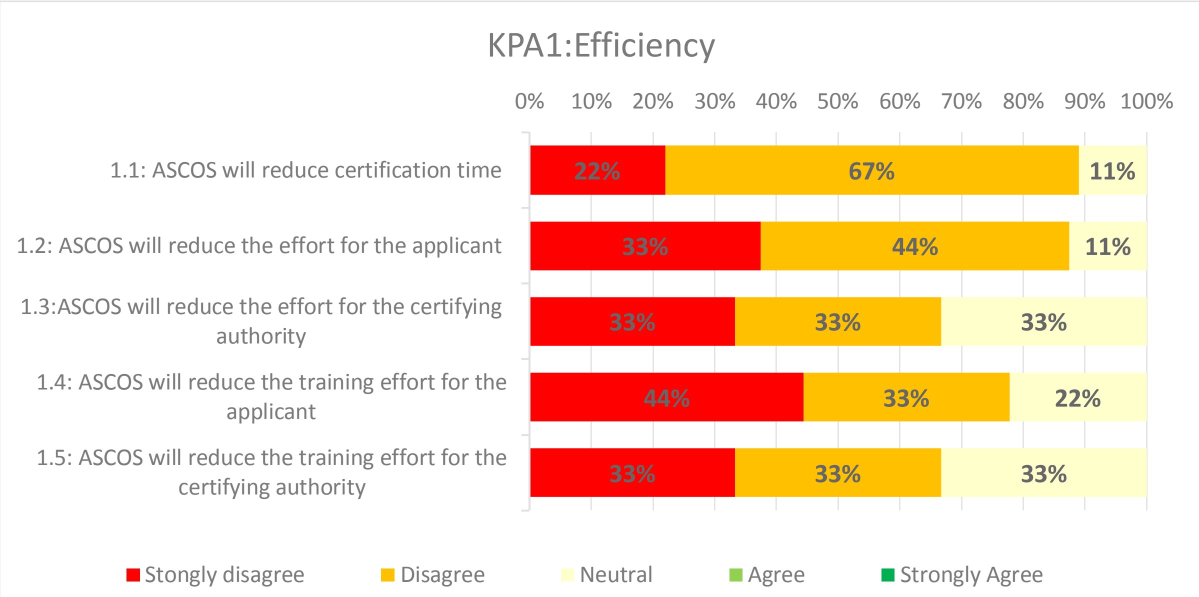


Figure 4. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA1-Efficiency.

Regarding time and effort, it was noted that considering the Total Aviation System (TAS) since the very early certification stages implies, as a minimum, bringing an additional step to the certification process as compared to today. Therefore, it is reasonable to expect that more specialists will be involved in the certification process. In particular, it will be necessary to have specialists able to consider the impact of the concerned change at the TAS level—i.e. beyond any single domain. In practice one can envisage that in order to follow the ASCOS approach existing specialists in each organization will be retained, but new ones—those with TAS safety assessment knowledge and skills—will also be required. Overall, the increased number of specialists will translate into an increased coordination effort—i.e., the effort needed to integrate the work of different specialists from different domains into a single argument structure.

In addition to this, one of the experts with experience in the certification of operators highlighted that the possibility to go beyond the compliance-based approach (as proposed by ASCOS) appears as a reasonable solution in some cases, but will by definition extend the time needed to get a certificate: "*When there is room for alternative means of compliance to prove to the authority that you will finally meet the requirements (means of compliance different from those indicated in the available regulations and guidance material), this approach could help […]. But the simple fact of deciding to follow a different path will inevitably require a* *longer period of time. Because both the authority and the applicant need time to think about the acceptability of these alternative means of compliance."*

Finally another expert argued that a reason for presuming that a longer time will be required when applying the proposed approach is the tendency of today's society to become more 'legalistic' compared to the time in which current certification processes were designed. A process which requires to work also across the borders of existing organizations is likely to trigger an increased need for legal considerations in certification activities and possible defensive attitudes by both the applicants and authorities, with an inevitable impact on both the effort and time needed to engage in the certification process.

Because of the above reasons the general perception among the experts was that ASCOS is unlikely to reduce the time and effort needed in earlier certification phases.

Partially mitigating this point was the observation that ASCOS may however generate benefits and savings during the subsequent stages of the product/system lifecycle—system operations, changes, and modifications. The approach may reduce the effort needed to keep the change safe over the whole lifecycle of a product or system. For instance, the early identification of cross-domain hazards (see also the following section) should in principle reduce the risk of hazards being identified too late in the certification process and thus requiring major rework, with consequent delays in the achievement of the certificate. However, it is unclear among the experts whether this positive effect on efficiency has the potential to actually counterbalance the effects mentioned above (expected increase in time and effort).

Some experts also highlighted that a decrease in the absolute cost of certification should be considered as undesirable: While making certification practices more cost-effective is an important and worthwhile objective, it is reasonable to expect that the growing complexity of novel products and systems will requires more and more sophisticated approaches and methods to be used in certification. Hence, focusing too much on absolute cost reduction alone may result in a negative side effect: streamlined certification approaches that would then be poorly fitted to the (increasing) complexity of current and future systems.

The ASCOS approach was perceived as requiring an additional effort also in the area of training—on both the applicant and certifying authority side. The main reason is that the ASCOS focus on the Total Aviation System requires the applicant and the certifier to equip their personnel with a larger competence or skill base. Their personnel will need to have knowledge of several domains, rather than just one; they will need to know the interactions between these domains; and finally they will need to have the ability to assess how risk can propagate across domains. Overall, this knowledge adds to, and does not replace, today's domain specific knowledge. The latter will be still needed in fact. Therefore, it is difficult to envisage how the training effort can be reduced compared to today.

One expert added that the increased training effort may be expected more at the level of project managers, rather than individual specialists. This position was motivated by the consideration that, most likely, it will be the role of team leaders and project managers to interact the most with equivalent roles in other domains. This view was however not further elaborated by the experts. At the same time it is worth noticing that it points at the need to further clarify what the roles and responsibilities will become for the stakeholders engaged in managing and coordinating the various parts of the TAS safety argument.

4.3 Soundness

This area of performance investigated the potential of the ASCOS certification approach to increase the soundness, or robustness, of safety assurance in certification. Experts agreed that, compared to today, the approach can increase the consideration of cross-domain hazards and safety requirements. First, such hazards and requirements would be considered right at the start of the certification process. In this phase the approach emphasizes, in fact, the explicit and early focus on the interfaces spanning across multiple domains, the related hazards which may arise across these interfaces, and the related safety requirements. Second, this process is supported by an explicit safety argument notation, which should guarantee increased visibility and traceability of cross-domain hazards and safety requirements.

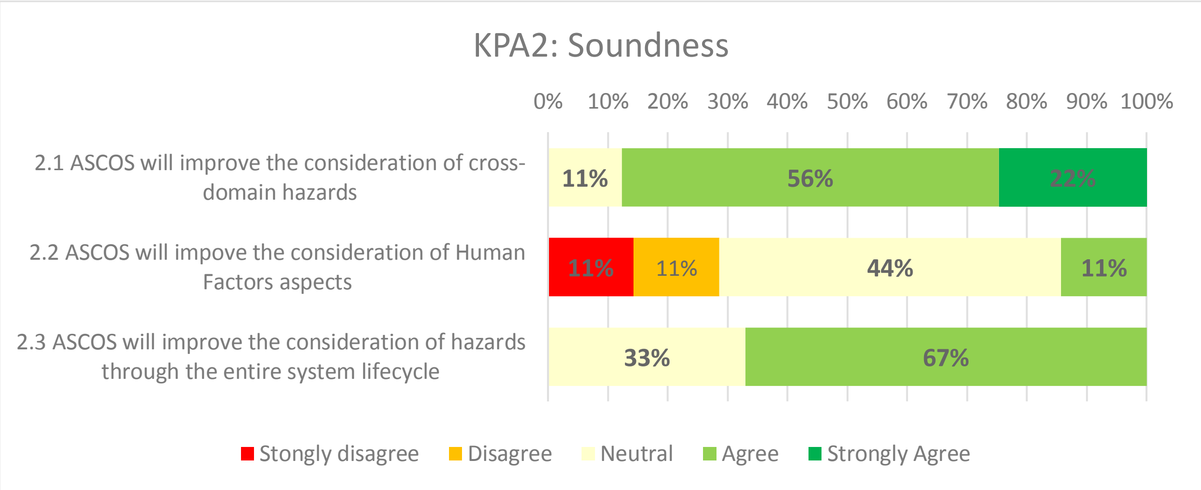


Figure 5. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA2-Soundness.

Partially contradicting these considerations was the observation that the timely identification of cross-domain hazards depends, indeed, on the expertise of the people involved, rather than solely on the certification process itself. In other words, the consideration of TAS level hazards is also a problem of getting the right group of experts of different domains working together, to identify how one's "own" hazards impact/propagate on other domains. Somehow, the experts' impression was that the way this has to happen needs further definition in ASCOS.

Regarding the ability of the approach to consider human factor related hazards, the majority of respondents was neutral or did not agree. This reflected the fact that from the training material—the WP4 use cases—it was not immediately clear how the approach could deliver benefits in this area. As a matter of fact, the proper identification and analysis of human factors related hazards seems to be more related to the integration of specific human factors methods, able to capture for example human performance and HMI design issues. Many of these methods are available in the literature, and most of them could be incorporated, presumably, into the generic and higher level ASCOS process.

4.4 Cross-domain Integration

This area of performance investigated the potential of ASCOS to improve the cooperation and exchange of safety risk information—such as hazards, safety objectives, requirements, target levels etc.—across the different domains of the Total Aviation System.

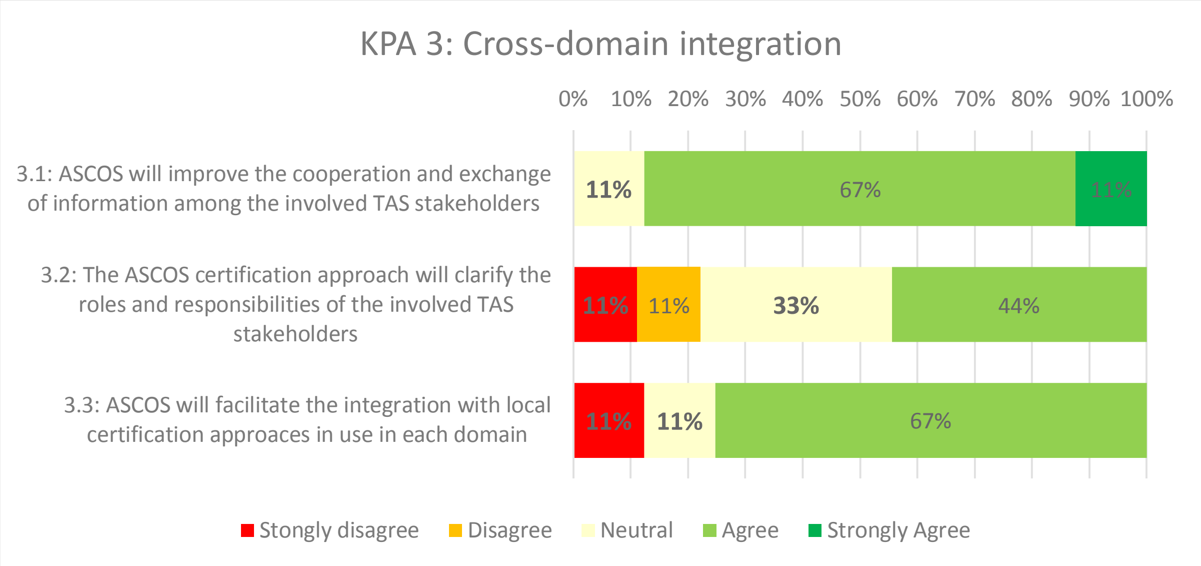


Figure 6. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA3: Cross-domain integration.

The majority of the experts (76%) agree that the ASCOS approach could increase the cooperation and exchange of information in the certification process of aeronautical products. They also noted, however, that this improvement is subject to the availability of a regulatory or legal requirement that makes it mandatory for the involved TAS domain partners to share risk information. Otherwise, it is very unrealistic to expect legal departments of applicant organizations to let safety department share proprietary information. This would in fact disclose weaknesses of systems and products to the public, thus potentially compromising OEMs reputation (they could be perceived as vulnerable). One expert with a background in aircraft certification stressed that already today it is difficult for inter-organizational working groups, established for safety improvement purposes, to have access to design assumptions and safety cases. Therefore, a legal requirement is necessary to make ASCOS feasible. This should specify not only how cooperation should take place but also what happens when the involved parties do not cooperate.

Experts expressed a very moderate rating about the ASCOS potential to clarify roles and responsibilities. In the experts' view it was not fully understood how roles and responsibilities would change compared to today. They noted that the evaluated version of the logical argument framework requires the definition of the required roles and responsibilities across the organizations involved in a change. In particular, the experts felt they lacked a clear definition of who would be the architect of the safety argument. This could be either a single person or an organization—but it needs to be specified to fully appreciate the potential of the approach.

ASCOS was perceived as being compatible with local certification approaches. The experts did not see any specific inconsistency with existing local approaches available across different domains. Quite the contrary. The use of a modular safety argument and the GSN notation can make the approach flexible enough to integrate the input coming from different approaches, independently of whether they are compliance or performance based. In this respect the approach seems promising. Partially countering this position was the observation that the added complexity resulting from accommodating and integrating different local approaches can lead to side effects that are unknown at the moment. These may consist of bottlenecks arising at the level of coordination between the actors of different organizations and domains. During the exercise no example was provided of such side effects; they reflect however an expert concern that may require consideration during future refinement and evaluation of the ASCOS approach.

4.5 Harmonization

This area of performance investigated the potential of ASCOS to be harmonized with the various approaches used by different stakeholders of the TAS.

The approach was perceived to be promising by the participating experts because it looks compatible—and not in conflict—with existing approaches. However, the following remarks were made:

- Appreciating the full potential for harmonization would require considering ASCOS harmonization with FAA certification practices;
- ASCOS may generate some cross-boundary side-effect that are difficult to anticipate at the current stage;
- ASCOS will require a different way of working; and this requires a lot of commitment from both applicant and certifying authorities.

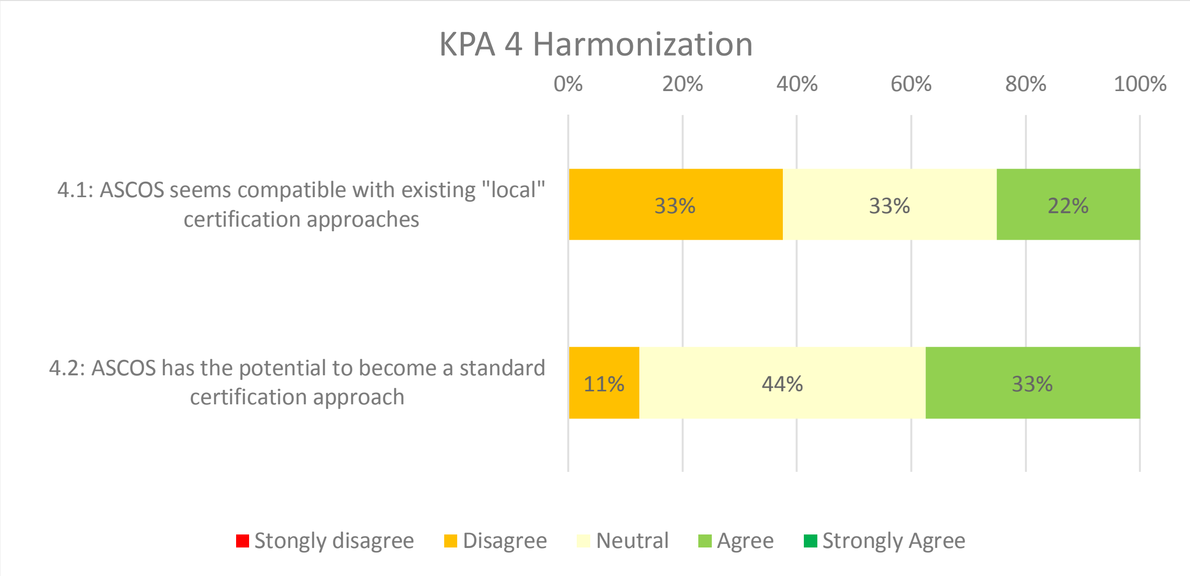


Figure 7. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA4-Harmonization.

Participants were supportive of the idea that ASCOS could become a standard approach for certification. However, it was noted this potential will depend on training and commitment by the various stakeholders. Also it may be a long process because it requires that different relevant stakeholders to be convinced of the added value.

4.6 Accommodation of innovation

This area of performance evaluated the potential of ASCOS to ease the certification of innovative products and services—i.e. products and services for which no standard exist. Some expert believed ASCOS can bring an improvement in this area, considering that the approach accommodates performance based certification. The majority of the ratings falling under the neutral category reflects the idea that while ASCOS may make certification possible for innovative products, this does not mean that ASCOS will necessarily simplify things for such products. ASCOS has the potential, in fact, to add clarity and structure to the certification of novel products, whose route to certification is notably characterized by higher uncertainty.

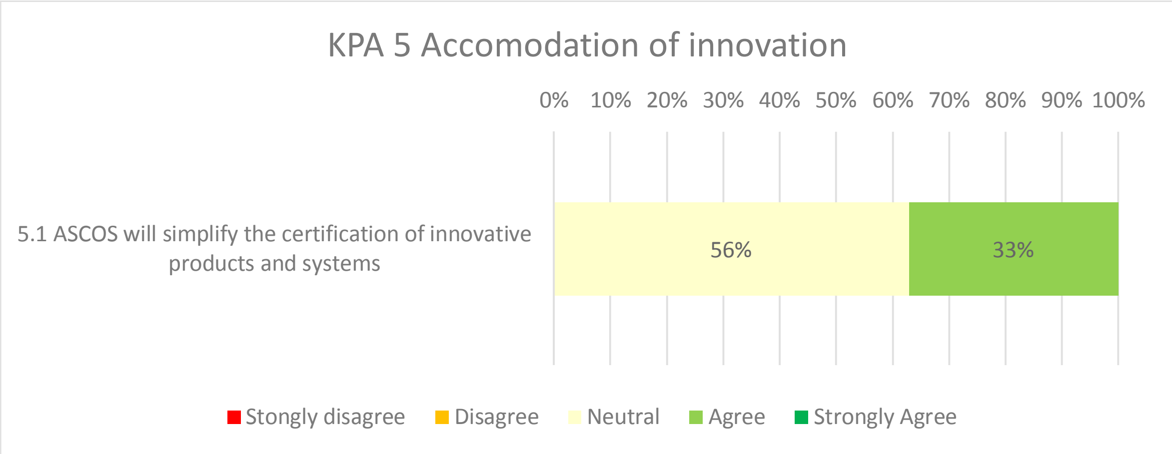


Figure 8. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA5-Accomodation of Innovation.

4.7 Acceptability

This area of performance looked at the acceptability of the ASCOS certification approach from both the applicant's and the certifier authority's perspective. Experts noted the approach looks promising, although its acceptability may vary depending on the certification domain. In particular, certifying authorities may see the approach more acceptable for airport and ATM, rather than for airworthiness and organizations. This is due to the fact that these latter certification domains have well established rules and standards in place. On the other hand, certification rules and standards for airport and ATM do not enjoy the same level of maturity: Rules for these domains are currently under development by EASA. Therefore, ASCOS can be perceived as more useful for airports and ATM as here it would help filling an existing gap.

Applicant organizations will have to do more as compared to today: ASCOS was perceived to demand applicants' significant effort, time and resources to be spent in the upstream phase of certification. Hence, acceptability requires that the approach will bring clear benefit for them "upstream" or "downstream". These benefits could include: increasing support in areas in which the applicant struggles with existing regulations; increased flexibility in showing compliance, especially when developing innovative products. Once the proposed certification approach will gain more maturity, a thorough assessment of how it delivers in areas such as these would allow a better appreciation of its acceptability.

Finally, the experts elaborated on which key stakeholders should be convinced first in order to promote wide spread adoption of the approach. They remarked, first, that the approach should be acceptable for national authorities, not only EASA. Most importantly, they noted the approach should be acceptable especially for leading OEMs: If large OEM organizations like Airbus adopts this approach—because it has some value for them—, then adoption by EASA and national certifier authority will follow suit.

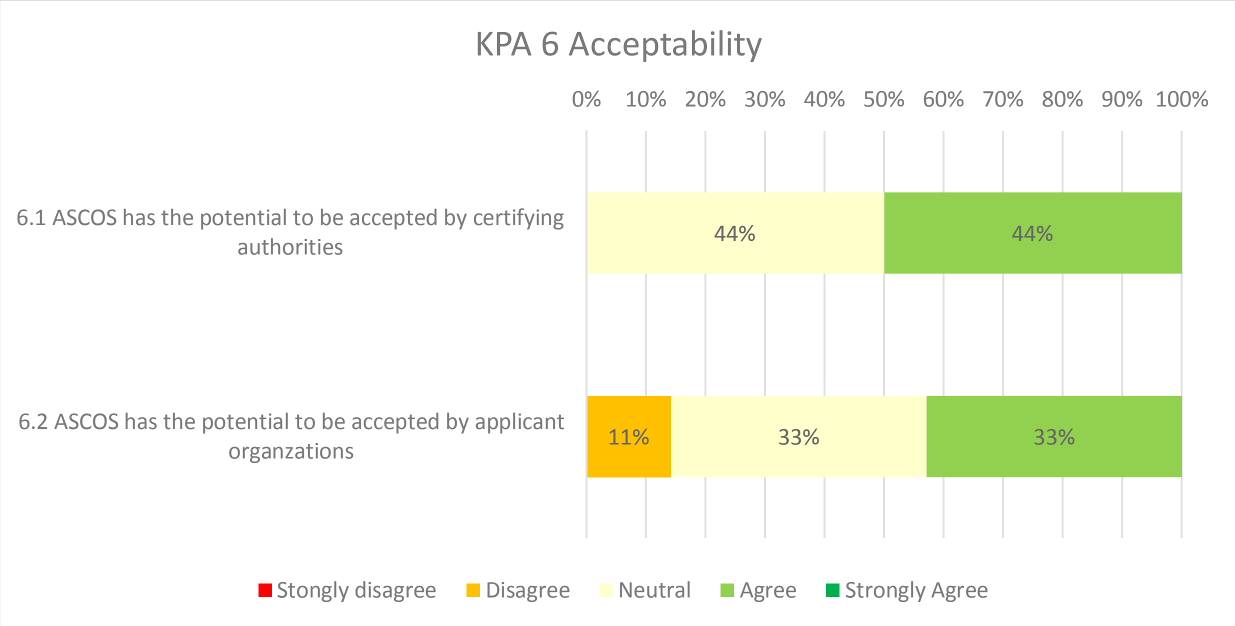


Figure 9. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA6-Acceptability.

4.8 Flexibility

Flexibility investigated the potential of ASCOS to be applicable to a wide range of products. Experts agreed that the approach seems very promising for innovative products, i.e. products for which there are no available standards and regulations. On the other hand, the approach did not look beneficial for the certification of derivative changes, i.e. a major change to a previously approved design type[2]. This is the case because these changes have as a certification basis the type certification issued for the original—or grandfather—design type.

[2] For example, the Douglas DC 9-30 is one among several derivative aircraft designs based on the original DC-9 type design [25, p. 16]

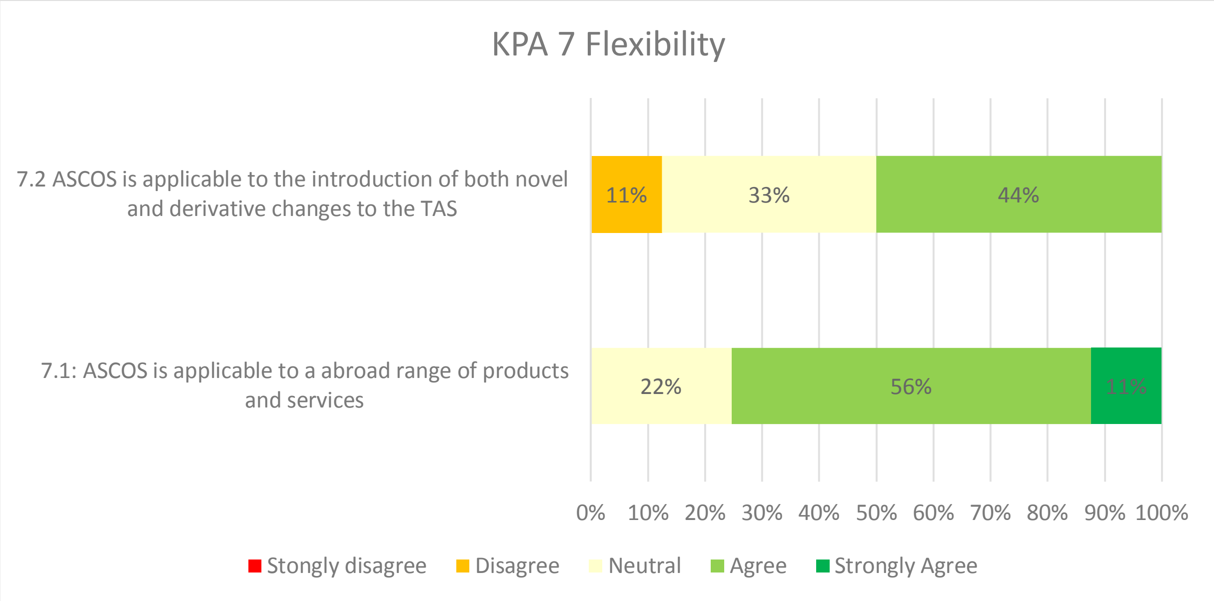


Figure 10. EXERCISE 1: Participants' response ratings to questionnaire items related to KPA7-Flexibility.

Finally, it was not clear how the approach would help the certification of an innovative service such as an independent de-icing service. First, it was noted that it is difficult to appreciate the potential of the approach in this area because today there is no specific certificate for this kind of services/organizations. Second, the appreciation of the potential of the approach in this area depends also on the adequacy of the risk model informing the ASCOS approach: This type of change requires a risk model able to capture organizational contributors to safety—such as adequate staffing, adequate resources, adequate definition of roles and responsibilities, etc. At the same time these organizational contributors to safety are not usually covered by "classic" safety risk models, i.e. risk models that tend to focus on the identification of sequences of events at the system sharp end. Therefore, to express an informed rating about the potential of the ASCOS approach to certify new organziations and services, the experts felt they should know more about the potential of the underlying ASCOS risk model for capturing organizational contributors to safety.

4.9 Summary of Exercise 1 validation results

The previous sections have reported the validation results specific to the ASCOS certification approach that have emerged during Exercise 1. The key results of the exercise are summarized in the table below.

Table 8. EXERCISE 1: Summary of the validation results for the proposed ASCOS certification approach.

KPA	Key Results
1. Efficiency	**Time & Effort** - More specialists—in particular, specialists knowledgeable on cross-domain risk assessment—are expected to be involved since the early certification stage in addition to existing domain specialists. This is likely to increase the coordination efforts needed to gain the initial certificate; - The expected increase in interactions between different organizations, as introduced by the approach, is likely to increase the effort required for considering legal aspects in certification; - Possible reductions in the overall certification effort are deemed possible when considering the whole system lifecycle, and not only the initial certification approval phase - The approach has the potential to reduce the risk of delays and major reworks later on in the process caused by relevant hazards becoming evident (too) late during the certification process; **Training** - ASCOS requires TAS expertise, therefore both the applicant and the competent Authorities will need to train their personnel accordingly.
2. Soundness	- ASCOS strong focus on cross-domain hazards and safety requirements identification (by means of a formal notation and since the initial certification phase) may reduce the likelihood of relevant hazards and requirements being missed or identified too late; - The effectiveness of the process heavily depends on the expertise of the people involved, rather than solely on the formal steps of process. Roles, processes, process owner, require further specification; - It is not immediately evident how the approach could improve the consideration of human factors aspects in certification.
3. Cross-Domain Integration	- ASCOS has the potential for cross-domain integration, however this depends on the availability of supporting regulation(s) mandating the sharing of safety risk information across the TAS stakeholders involved in a change; - The roles and responsibilities of the TAS architect(s) should be further specified in order to appreciate the potential for improved coordination;
4. Harmonization	- The approach looks compatible with local approaches—either performance or compliance based—used across different domains.

5. Accommodation of Innovation	- It is expected that the ASCOS approach will increase the rate of success regarding the certification of innovative concepts (concepts for which there are no reference standards available); - ASCOS can add clarity and structure to the certification of innovative products;
6. Acceptability	- Acceptability by both applicants and the certifying authorities may be challenged by the perceived significant effort required to adopt the approach; - There is an expectation that national CAAs will see the approach as more helpful (and therefore more acceptable) for airports and ATM, since these domains are more performance based compared to aircraft system certification (which is compliance based); - The wide spread adoption of the approach will be promoted by the acceptance of leading OEMs.
7. Flexibility	- The approach seems more promising for innovative products, while it looks potentially less useful for derivatives changes, i.e. changes for which an initial (pre-existing) certification base applies; - It was difficult to estimate the potential for certifying novel services—e.g. independent de-icing operations—because this type of change is not covered by a dedicated certificate nowadays.

5 Results Exercise 2: ASCOS Safety Performance Indicator Framework and ASCOS Tool for Continuous Safety Monitoring

This section reports on the feedback collected during ASCOS Validation Exercise 2, which focused on evaluating the main WP2 results: The ASCOS Safety Performance Indicator Framework and the ASCOS Tool for Continuous Safety Monitoring (ATCSM). The section is organized as such:

- Section 5.1 provides an overview of the group of experts involved in the exercise;
- Section 5.2 reports on the results specific for the ASCOS SPI framework;
- Section 5.3 reports on the results specific for the ATCSM;
- Section 5.4 summarizes the validation results of Exercise 2.

Note for this exercise a choice has been made to omit the results of the quantitative questionnaires. The limited group size (see § 3.10) implied that such questionnaires were not representative. The exercise however gathered also useful qualitative feedback, and this is reported in the next sections. Sections 5.2 reports the result related to the ASCOS Safety Performance Framework; Section 5.3, the results related to the ASCOS Tool for Continuous Safety Monitoring.

5.1 Expert group profile

Exercise 2 collected the views and perspectives of a small group of safety management experts composed by a total of five people. These were recruited from the ASCOS User Group and among the ASCOS research team's contacts. Experts' mean experience in safety management was 9.6 years (*stdev*=9.2). As shown by Figure 11, the panel included experts coming from two National Civil Aviation Authorities (CAA) (n=4) and one Air Operator (n=1). Overall, this suggests that the group composition and representation were adequate to the goal of the exercise.

Table 9. EXERCISE 2: Profile of the expert group.

Size	5
Average experience of participants in safety management (years)	9.6
Min	3
Max	25
Sdev	9.2

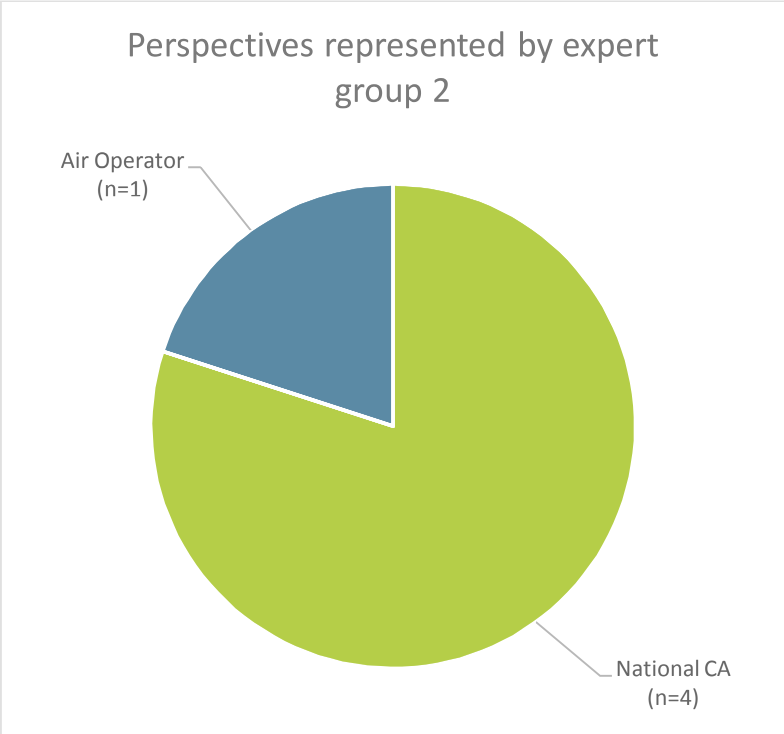


Figure 11. EXERCISE 2: Perspectives represented by the expert group.

5.2 The ASCOS Safety Performance Indicator Framework

The present section reports the results collected in relation to the ASCOS Safety Performance Indicator Framework.

5.2.1 Soundness

This area of performance investigated the extent to which the proposed safety performance indicator framework contributes to improving continuous safety monitoring practices in the context of certification. The following points were highlighted during the group discussions:

- ***Link between ASCOS SPI framework and local, change-specific, SPIs.*** The ASCOS continuous safety monitoring concept proposes to adopt a standard SPI framework. While acknowledging this intent, the experts expressed the concern that the framework may not acknowledge how the SPIs are to be defined in the context of the local change within a specific operator. The point was that the monitoring of the safety performance of a change requires, in fact, the definition of a change-specific set of SPIs—that is, the set of SPIs specifically defined for monitoring the safety performance of the specific change. The experts' concern was that the specific SPIs may not be included in the list of ASCOS defined SPIs. In other words, while ASCOS provides some requirements for good SPIs, the ASCOS defined SPIs cannot be used for each and every certification case, but its use requires extension/adaptation. For instance, during the meeting it was noted that one of the ASCOS organization SPIs considers "the number of safety management meeting"; and that such kind of SPI per se is a good SPI to consider, however, as stated is too generic: To be useful, such a SPI needs to be expressed in a way that directly connects with the change at hand. It could become for instance something like "number of safety meeting addressing the hazards associated with the introduction of linear holding procedures[3]" (in situations in which linear holding procedure is the change in question). These considerations do not invalidate the proposed ASCOS SPI framework. They point out, however, the need to further elaborate the way in which the framework and its SPIs can be enhanced, adapted, and expanded by the relevant stakeholder(s) in the context of the specific case at hand.

- ***Framework biased towards the consideration of lagging indicators mostly.*** The proposed ASCOS SPI framework is based mostly on lagging indicators, i.e. mostly incidents, accidents, and deviations from normal operations. However, it was noted that nowadays safety monitoring is not limited to the measurement of negative occurrences only. What is important is to track also good leading indicators—i.e., indicators that are good predictors of successful safety performance. Such indicators do not only tell what does not work, but also what works well. It was noted that this idea is rooted on the Safety-II view of safety [26]. This view captures recent developments in safety thinking, and its foundational ideas have been recently detailed by EUROCONTROL on a dedicated withe paper [26].

[3] Linear holding consists of arrival procedure that allow "ATC to delay, sequence, and integrate aircraft arrivals by giving routings along predefined variable legs to specific points, instead of providing radar heading. This procedure can reduce significantly the need for holding stacks." The procedure was briefly mentioned but not discussed during the meeting.

One of the main idea in Safety-II is that a focus on errors, malfunctions, and failures—the classic indicators under Safety-I—essentially lacks the consideration of the main positive activities that promote safety. These positive activities, if identified and reinforced, can further increase safety performance. In short, the Safety-II view of safety implies a move from ensuring that '*as few things as possible go wrong*' towards ensuring that '*as many things as possible go right*' [27]. The implications for ASCOS, is that of considering the possibility of enhancing the ASCOS SPIs framework with indicators of successful safety performance.

- **Including non-quantifiable elements in the scope of ASCOS Safety Monitoring**. As a general remark, it was noted that the proposed solutions—the ASCOS SPIs framework and the accompanying software—focus essentially on reported/reportable quantitative safety data, which can be used to quantify performance indicators. This scope was considered somehow narrow if compared against the set of safety issues that could be associated to the introduction of a change and which require monitoring. It was noted that some of these issues will not be measurable, and therefore may receive little benefit from occurrence reporting only. On this basis, it was commented that innovative safety monitoring solutions should not be limited to occurrence reporting only, but should contain provisions for of monitoring of other data sources that report on non-measurable events. This position appears to be based upon the definition of safety performance monitoring provided by ICAO in its Safety Management Manual [28]. Besides occurrence reporting, this definition includes (in the scope of safety monitoring) also other safety data sources: safety studies, safety reviews, safety surveys, audits, and internal investigations. And these sources are not limited to reportable occurrences. On elaborating on the implications of these considerations for ASCOS, it can be noted that they do not invalidate the SPI framework or the tool. They suggest that ASCOS could clarify the treatment of non-measurable events in the context of the ASCOS Process for Safety Performance Monitoring (see D2.3 [29]).

- ***Integration with WP1 and WP3.*** As a general remark, the experts noted that the SPIs framework (and the ATCSM tool) could be useful also in a general Safety Management System, as opposed to being specifically useful in certification. This consideration arose from the observation that the linkage between the SPI framework and the ASCOS proposed certification process (as well as the risk model (WP3) was not entirely clear to the experts. Therefore, this link should be further elaborated in order to understand more in detail the added value that the framework (and the accompanying tool) can bring to certification.

5.2.2 Completeness

This item addresses the extent to which the framework covers different domains of the Total Aviation System. It was noted that the SPI framework could be extended to cover also the domains of ground handling, maintenance, and airport operations. Essentially, the experts expected these domains to be covered, considering the TAS orientation of ASCOS. This remark was partly mitigated by the observation that the development of the ASCOS SPI framework was scoped around the EASA key operational issues (as defined in the European Aviation Safety Plan [22]), so its scope was somehow set by this initial requirement.

5.2.3 Standardization

Experts noted that it should be clarified how the ASCOS SPIs framework connects with the work done by the EASA Network of Analysts (NoA) subgroup in the area of safety performance indicators for State Safety Programmes (SSPs). This consideration was made because this latter work would be normally viewed as the primary reference for national authorities and the EASA. Therefore, the way in which the ASCOS's framework connects with it requires further clarification, else the former risks to be perceived as a redundant with the latter. In this context, note that EASA experts participating in the NoA sub-group for SPIs have in fact reviewed the framework for Safety Performance Indicators proposed by ASCOS [3] (comments processed satisfactorily).

5.3 The ASCOS Tool for Continuous Safety Monitoring

This section reports the experts' feedback related to the ASCOS tool for Continuous Safety Monitoring (ATCSM). This feedback has been collected during Exercise 2. It has to be noted that the potential of the tool has been demonstrated by means of (i) a presentation and (ii) an interactive demo only. In other words the participants did not interact directly with the tool for the reasons already explained under section 3.5.2.

5.3.1 Usefulness

Experts concurred that the ATCSM is a useful enhancement of ECCAIRS: The tool could allow both the competent authority and the applicant to generate graphs of SPIs trends from ECCAIRS data that are useful in the context of certification. For instance, the ATCSM (and ECCAIRS) data can be useful in the final phases of the type design approval—the phases ranging from the initial type certificate application request, put forward by the applicant, to the release of the (initial) type certificate by the authority. In this phase, the certifying authority has the need to access occurrence data of comparable designs to check what hazards could actually be envisaged for a new design type. This process can be viewed as a sort of "due diligence"—a sort of check done by the authority to ensure with all possible means that a new design is acceptably safe. Similarly, in the same phase the applicant has the need to access occurrence data to develop the safety case that has to be reviewed and approved by the competent authority. In both cases, an occurrence reporting system like ECCAIRS, supplemented by the ATCSM can be a useful support. These example should not limit the value of the ATCSM and ECCAIRS to type certificate. These systems can be useful also for the release and maintenance of the Certificate of Airworthiness—i.e. when an applicant, for instance, requires the approval (i) for introducing a new type of aircraft into its fleet, (ii) for replacing a piece of equipment of an existing aircraft, or (iii) when opening a new route. Overall, these remarks indicates that the ATCSM (and ECCAIRS) could be a helpful support for both the applicant and the certifying authority in the context of certification.

Other remarks were collected that points at the use of the ATCSM and ECCAIRS in the context of the local Safety Management System. In this context it was noted that the ATCSM, essentially, does not change how things are done today already in a SMS context. This should be expected, at least, for organziations with a mature SMS in place. The tool could be more useful, on the other hand, for organizations new to SMS and with no occurrence reporting systems in place. Such organizations would benefit from ECCAIRS and the ATCSM as a means to progress the maturity of their own SMS.

Limitations. One limitation on the use of the ATCSM has been observed. The link with ECCAIRS was considered limiting because organizations that are not using ECCAIRS have to code their occurrences according to ECCAIRS and they have to use a similar physical data structure as ECCAIRS 5 in order to be able to use the tool. This observation requires, however, consideration of the fact that ECCAIRS compatibility is not optional but mandatory according to EC regulation 376/2014 (see art. 7, paragraph 4). ASCOS is consistent, indeed, with this regulation. Therefore, this limitation, as formulated, may no longer be present once industry has managed to comply with the regulation in question.

5.4 Summary of Exercise 2 validation results

The previous sections have reported the validation results specific to the proposed ASCOS SPI framework and the ATCSM. These results have been collected during Exercise 2, and they are summarized in Table 10 and Table 11 respectively.

Table 10. EXERCISE 2: Summary of the validation results for the ASCOS SPI framework.

KPA	ASCOS Safety Performance Indicator Framework
1. Soundness	- It needs to be further defined how the SPI framework can be adapted/modified/extended for a specific certification case; - As the ICAO view of safety monitoring includes also non-quantitative data sources, it should be made explicit that ASCOS continuous safety monitoring focuses solely on quantifiable safety performance indicators based on reported occurrences (else safety experts would wonder why one should not consider non-quantifiable and non-reportable events); - The SPI framework seems to be oriented mainly towards the consideration of lagging indicators. The framework could be improved by including (leading) indicators that represent safety enhancing activities.
2. Completeness	- To address the variety of Total Aviation System, the SPI framework should include other domains, such as ground handling, maintenance, and airport operations. As an alternative the actual scope of the intended TAS should be better clarified.
3. Standardization	- The link with EASA SPI framework requires further clarification, otherwise there is the risk the ASCOS SPI framework be perceived as a duplication of the former.

Table 11. EXERCISE 2: Summary of the validation results for the ASCOS Tool for Continuous Safety Monitoring.

KPA	ASCOS Tool for continuous safety monitoring
1. Usefulness	- The ATCSM is viewed as a valuable ECCAIRS enhancement, useful for comparative safety assessment to be done by the authority prior to the releasing of the type certificate, and the applicant as part of building a safety case; - ATCSM has been defined as an ECCAIRS enhancement. While this complies with point 4(a) and 4(b) of Art. 7 of EASA regulation 376/2014, it should not be neglected that organizations not having ECCAIRS may have to (re) code their occurrences using ECCAIRS to be able to benefit from the tool.

6 Results Exercise 3: ASCOS Risk Model and the Tool for risk assessment

This section reports on the expert feedback collected during ASCOS Validation Exercise 3, which focused on the validation of the main WP3 results: The ASCOS Risk Model and the accompanying Tool for risk assessment. The section is organized as follows:

- Section 6.1 provides an overview of the group of experts involved in the exercise;
- Section 6.2 reports the results specific for the ASCOS Risk Model;
- Section 6.3 reports feedback relevant for both the ASCOS Risk Model and the ASCOS Tool for risk assessment;
- Section 6.4 reports the feedback specific to the Tool for risk assessment;
- Section 6.5 summarizes the validation results of Exercise 3.

Note that the ASCOS Risk Model is an updated and improved version of the original CATS model [33]. At the end of 2014, one additional validation exercise was performed on request of the Dutch Ministry of Infrastructure and Environment. This exercise concerned the original CATS Model and accompanying Tool Set. The question was to address several potential user needs and to see whether the Model and Tool Set could handle these user needs well. The results of this validation exercise have been discussed with the Ministry as well as with EASA, who also expressed interest in the results. Because of the close relationship between the CATS and the ASCOS Risk Models, the results of the CATS validation exercise (in particular the conclusions regarding the CATS Model) could also be applicable to the ASCOS Risk Model. Therefore, the details of the CATS validation exercise are documented in this report, see Appendix K.

6.1 Expert group profile

Exercise 3 collected the views and feedback from a group of seven safety and risk management experts, which were recruited from the ASCOS User Group and among the ASCOS research team's contacts. Experts' mean experience in safety and risk management was 26 years (*stdev*=15.6). The panels included experts from three different National Civil Aviation Authorities (CAA) (n=5), and one Original Equipment Manufacturer (n=1). Overall, the group composition and representation seemed adequate to the goal of the exercise.

Table 12. EXERCISE 3: Profile of the expert group.

Size	7
Average experience of participants in safety management (years)	26
Min	3
Max	45
Sdev	15.63878

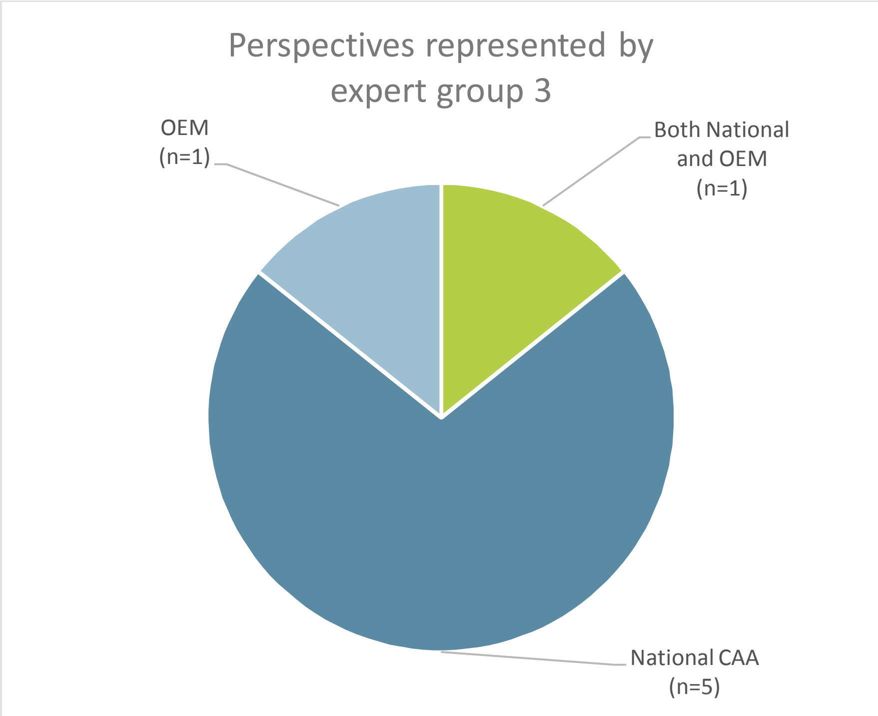


Figure 12. EXERCISE 3: Perspectives represented by the expert group.

6.2 The ASCOS Risk Model

The majority of the experts' ratings tend to aggregate in the centre with a moderate prevalence of neutral (n=10) and disagree (n=9) ratings. No questionnaire item received an overly positive (strongly agree) or overly negative (strongly disagree) rating. Overall, these ratings reflect the fact that the experts were somehow hesitant to express a definite rating about the actual value of the risk model and the supporting tool: The general impression was that the model has some potential value.

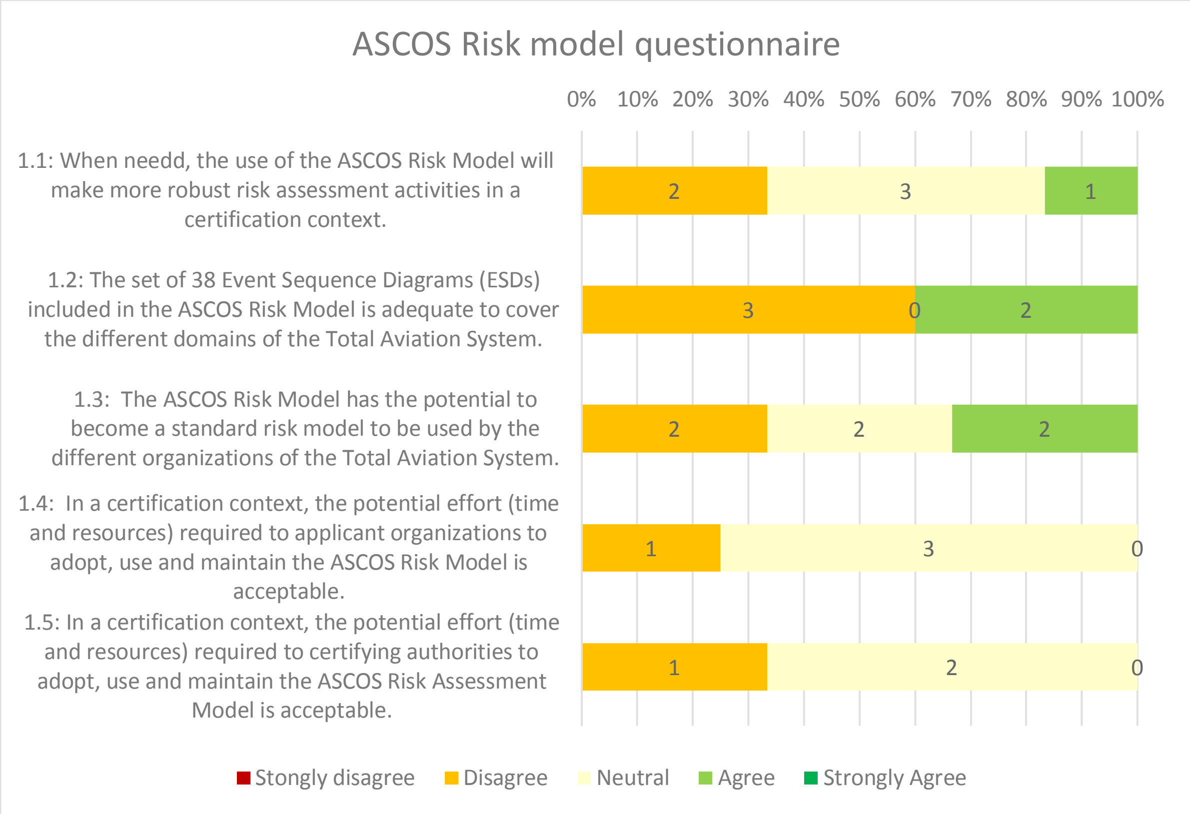


Figure 13. EXERCISE 3: ASCOS Risk model questionnaire.

6.2.1 Soundness.

This KPA refers the extent to which the ASCOS Risk model could improve risk assessment activities in a certification context. On the positive side it was noted that the model has potential provided that the level of detail of the Fault Tree events is sufficiently specific. The neutral ratings reflect the fact that it was not entirely clear how the model can deliver its value in a real life certification context. In particular:

1. Experts questioned the actual value of the ASCOS Risk model compared to available in-house risk models. Today individual organizations—OEMs, Air Navigation Service Providers (ANSPs), and authorities—have their own local models of risk, and these are more specific and detailed compared to the ASCOS Event Sequence Diagrams (ESDs) and Fault Trees (FTs). Experts said that this certainly is true for large operators as they have appropriate resources, skills, and manpower for developing and maintaining such more specific and detailed risk models, e.g. regarding the aviation domain one expert noted that Airbus is very likely to have their in-house risk models, which are more specific than the evaluated model. Moreover, it was noted that the value of the ASCOS Risk Model remains uncertain also for smaller operators. These smaller operators, not having the same resources as large operators, may benefit in principle from a ready-made risk model, but in practice their scale of operations may be so small that the model may be of limited support to them.

2. To the experts it was unclear how the model would work in practice—in particular how the model can improve cross-domains risk assessment in a certification context. For instance, it was unclear (i) how the model actually changes risk assessment; (ii) which organization/s is/are responsible for maintaining the model and the tool; (iii) how the model can be used by multiple stakeholders; and, in particular, (iv) how one stakeholder can use the model to assess the propagation of risk into another domain or set of domains. Note that these issues brought up by the experts do not reflect problems with the model itself—i.e. problems about the structure of the model's ESDs and FTs. These issues point, instead, at the practical management and governance of the model: the way in which the model is to be used by TAS stakeholders to identify cross-domain risks. These aspects should be further specified by WP1.5.

3. Experts' general concern was that the ASCOS model does not support the identification of new unknown risks, especially those accompanying the deployment of novel technologies that have no previous operational history. As one expert put it, *"With this new approach the "unknowns" remain unknown*." This concern highlights an important expectation of the experts—the expectation that the newly proposed approach embeds the capability to identify future and emerging risks across the multiple domains of the aviation system. To understand the implications of this expectation, it is necessary to consider that in ASCOS, this capability is not part of the risk model. The model is essentially a risk quantifier, and its use assumes, indeed, that hazards have been already identified by means of other safety risk assessment methods. The specific ASCOS element providing this capability is the Future Aviation Safety Team (FAST) approach—a consolidated international approach aimed at identifying unidentified hazards based on the analysis of on-going and future changes affecting the

aviation system[5], [30]. Thus, as ASCOS has considered by means of FAST the identification of novel risks, it can be concluded that the experts' concern in question, essentially, captures the importance of further specifying how the FAST approach can be used to inform and maintain the ASCOS Risk Model for future/emerging hazard identification and risk assessment.

6.2.2 Completeness.

This KPA assessed whether the ASCOS Risk model covered adequately the different domains of the Total Aviation System. For this question it is possible to observe a predominance of moderately negative (n=3) over moderately positive ratings (n=2). These ratings seems justified by the following:

1. It was noted that the ASCOS model elements seemed very flight operational oriented. In other words, the ESDs and FTs seems to consider aircraft systems and operation related failure mostly. On the other hand the ESD and FT models related to the ATM domain appeared to be somehow incomplete, e.g. "*There are many more relevant factors than those covered in the ASCOS FTs that should be included to fully cover ATM*". The relatively limited coverage of ATM can be appreciated especially by comparing the risk model against existing ATM specific models, such as the CAA Significant Seven Bow ties [31]. This kind of models can be taken as a reference for enhancing the ASCOS risk model with additional, and more ATM specific, model elements.

2. Also, the completeness of the model was questioned because the model is essentially based on historical data, and a model based mostly on hazards identified in past accidents was seen by definition incomplete. Two reasons were elaborated during the group discussions to support this point. First, such a model does not cover (dangerous) events that could occur, but which have not occurred yet. In other words, focusing only on past experiences limits the appreciation of the broader range of possibilities in which a system may fail—not everything that could possibly happen has already happened. Arguably, more analytical (imaginative/predictive) assessments are needed to identify potential hazards beyond those which have become manifest in past accidents. Second, a model based solely on past accidents covers only events related to the systems and environments that were in place at the time and place the accidents occurred. In other words, the model would not keep up with system evolution, as it does not fully reflect the range of situations in which novel systems in modern environments may fail. This expert concerns, points at the importance of defining guidance for ensuring that the risk model is regularly updated, both in the quantification of events and the scenarios themselves, to account for recent evolutions and development.

3. Also, the fact that the model is based on safety data from incidents and accidents in Europe limits its validity outside Europe. Experts noted that to be a true TAS model, however, the ASCOS Risk assessment model should ideally account for other areas of the world as well—third world countries and regions included. In these regions and countries aviation culture, working habits, and skills vary

considerably compared to Europe—and a truly TAS approach should strive to account for these differences.

4. It was observed that, besides quantification, the risk model could be useful in the context of risk assessment to verify that a basic set of relevant hazards across the different parts of the TAS has been considered.

6.2.3 Standardization.

This KPA investigated the potential of the model to become a standard risk model adopted by the various organizations of the TAS domain. The absence of a clear trend in the scores related to this item suggests that experts were unable to express an informative judgment about this KPA. As also noted earlier, they remarked that it was not fully clear what immediate benefit the current version of the ASCOS Risk Model would bring compared to existing more specific local models. Besides, the following remarks were made:

1. It was noted that the risk terminology used in ASCOS should be consistent with existing standard risk terminologies. In particular, it was observed that the definition of emerging risk was not immediately clear: ASCOS D3.1 defines it as a “familiar risk that is increasing or new risk that become apparent in new or unfamiliar conditions” [5]. This definition appeared not immediately clear to the experts: To them, emerging risks denotes new and unknown risks. This latter interpretation of the term seems consistent with the usage of the term “emerging” in ICAO Doc 9859 Safety Management Manual [28], which uses the term as opposed to known risks.

2. Upon reflecting on the value of the model, it was suggested that the model would make sense if seen as a cross-domain model—and not as domain specific. Due to the relatively higher specificity and granularity of local models, the ASCOS approach is unlikely to replace them. However, this does not mean that the approach has no potential for risk assessment. The value of the model could be that of increasing cross-domain communication between different domains. It was suggested by an expert that the model could be used as a trigger of cross-domain interactions: Depending on the change at hand, the model could be helpful to direct the initiator of a change towards the identification of the other parts of the TAS that may be affected, so that the relevant stakeholders, the stakeholder “owning” these parts, could be approached. In this perspective the tool would function as a sort of “router”: It would trigger cross-boundaries interactions amongst the network of the relevant stakeholders. It was suggested that this use of the model would be more valuable because it would complement—and not replace—existing local models. These latter models would be complemented with a TAS capability that they currently do not have. It was remarked that this use of the risk model would justify the complementary use of a less single-domain specific model.

3. Finally, it was remarked that the model may become a standard model provided it is clear how it connects to existing Safety Management Systems and State Safety Programmes. These will have to be implemented by operators and authorities respectively; therefore, the integration of a novel risk model under SMSs and State Safety Programmes (SSM) will make the widespread adoption of the model it more likely.

6.2.4 Acceptability

Questionnaire items 1.4 and 1.5 requested participants to score the acceptability of the effort needed to implement the model, this effort considered from both an applicant's (1.4) and a certifying authority's (1.5) perspective respectively. The majority of the respondents felt they did not have enough information to answer this question: In their view it was not obvious what changes would have to be made in applicant and certifying organizations—in terms of staff, procedures, rules, etc.—in order to adopt the model. One expert noted that a substantial effort is likely to be required to put the risk model in use, although this point was not further elaborated.

6.3 The ASCOS Risk Model and the Tool for risk assessment

This section reports the experts' feedback gathered for both the Risk Model and the Tool for risk assessment. This feedback was collected during the second group discussion of ASCOS Exercise 3. It can be noted that this discussion (and the accompanying questionnaire) was initially scoped around the risk assessment tool. However, it generated feedback also about the underlying ASCOS Risk Model. This was the case because, from the experts' perspective, it was difficult to make a clear distinction between the model and the tool. Conceptually, these two solutions were considered as strictly interwoven. In particular, the tool for risk assessment was perceived essentially as an editor of the ASCOS Risk Model. As a result two types of feedback emerged: (i) joint feedback about both the Risk Model and the Tool for risk assessment, and (ii) specific feedback about the tool only. For the sake of clarity a choice has been made to report the former in this section and the latter under section 6.4.2.

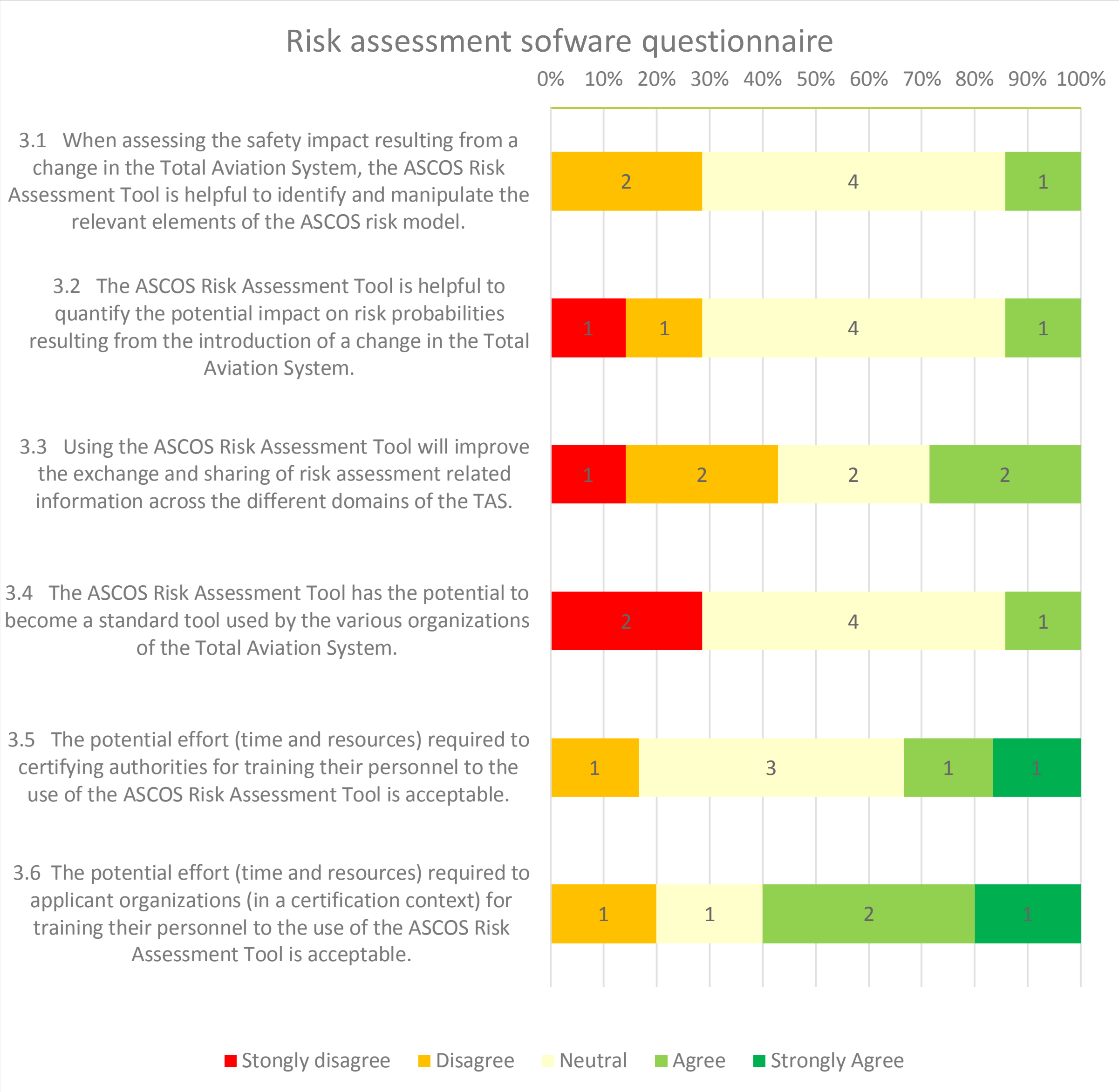


Figure 14. EXERCISE 3: Tool for risk assessment questionnaire

6.3.1 Manipulability

This KPA investigated the extent to which the tool supports the manipulation[4] of the elements of the ASCOS Risk Model. Experts viewed the tool as a useful means to interact with, edit, and modify the FT and ESD elements of the ASCOS Risk Model and change their probabilities values. They added that the quality of interaction with user interfaces could be improved, and suggested some areas for improvement (these are described under § 6.4.2).

Upon reflecting on the usefulness of the tool, experts' understanding was that its main function is probability calculation. Therefore, in their view the tool should be regarded more as an editor of the ASCOS Risk model, rather than a risk assessment tool. This latter definition suggests the tool supports the understanding of the impact of a change and is capable of assessing the severity level of (cross-domain) hazards as well. In fact the identification and severity assessment of (cross-domain) hazards has to be done using other safety methods than the tool. The background of this remark, already explained in the discussion on the risk model (see § 6.2.1), was the fact that some support was expected in the area of hazard identification. This expectation calls for a clearer definition of how the ASCOS Risk Model integrates with the FAST approach.

6.3.2 Quantification capability

This KPA investigated the adequateness of the Tool for risk assessment to quantify changes in risk levels following the introduction of a change. During the discussion of the quantification capabilities of the tool, the following points were made:

1. The tool (and the model) should account for severity levels, rather than probability only, to be a true risk calculator (risk model). The same hazard—e.g. ice on wing, inadvertent thrust reversal deployment in flight—may have very different consequences (i.e. severity levels) depending on the type of aircraft. At the moment the tool does not take differences in severity levels into account. Severity levels are not clearly present as an outcome of the risk model.

2. The point above captures also the fact that some users would like to apply the tool (and the model) to specific aircraft types. The current version is not suitable as it is a generic model, so every time a user wants to apply the model to a specific aircraft type, operation, and so on, s/he has to review and adapt the model according to her or his use case. The model and the tool are flexible and therefore, in principle, can support this adaptation; in practice, it will require expertise, time, and money.

3. Some participants mentioned that the tool (and the model) should provide a means to handle different probability units. Depending on the domain these can be expressed in different forms, e.g.,

[4] This KPA originates from the concept of *direct manipulation* [32], a Human Computer Interaction style that, as a minimum, involves the continuous graphical representation of the object of interest, which the user can manipulate directly by means of icons and commands (instead of using a complex syntax). As the ASCOS Tool for risk assessment supports a direct manipulation style, manipulability essentially "measured" the extent to which this style has been built into the software.

"per flight", "per hour of flight", etc. The current version of the model, however, cannot account for these differences, because the model's elements use the same unit to make quantification possible.

4. It was stressed that an important element limiting the usefulness of the numbers produced by the tool is the lack of confidence data. For instance, one should know more about the underlying (quality assurance) process by which data is provided rather than numbers alone. This is necessary for the tool to generate meaningful figures about the actual impact of a change. To explain this point, one authority representative from the ATM domain said that, when reviewing safety cases, he does not just look at the numerical evidence. This is not per se decisive. What really counts, besides numbers alone, is the rigour—i.e. the breath and the depth—of the underlying assessment process. This is really what a safety expert, at least in the ATM domain, needs to know to express a sound judgment on the quality of a safety case, rather than numerical values only. These comments are counterbalanced by experience in the aircraft system certification domain, in which quantitative evidence seems to be more decisive.

5. On the other hand, some of the participants felt that overemphasizing quantification can create an (undesirable) incentive for superficial risk assessment. The tool could convey, indeed, the erroneous message that it can simplify safety assessment. As noted by an expert, it seems that "*all is needed is to change some probability values, press some buttons, and calculate new risk levels*". The experts felt this way of working may promote the use of "quick and dirty" risk estimates, while limiting detailed thinking about the validity of these estimates. Overall, this situation could create a new and interesting problem for regulators, who should ensure that the risk assessment tool is used responsibly by the relevant stakeholders.

6.3.3 Cross-domain integration

This KPA inquired into the potential of the risk assessment tool to promote the exchange and sharing of risk assessment related information across the different domains of the TAS.

1. On the positive side, it was acknowledged that the tool, together with its underlying model, could promote the development of a standardized language—a common language shared across the multiple stakeholders of the TAS to talk about risks and hazards.

2. But experts agreed that, in order to implement a cross-domain approach, access to local proprietary safety data is essential. The problem is that OEMs would not normally share this data, and this would make it difficult to develop detailed, quantified risk models. To overcome this problem, experts concurred that the ASCOS approach requires a legal requirement to make the sharing of proprietary safety risk data mandatory. In this way they mirrored a consideration that was already expressed in relation to the certification process (see §4.4).

3. When discussing the contribution of the tool to cross-domain integration, the extent to which the tool (and the model) could be integrated with ECCAIRS was considered. ECCAIRS integration was justified by the observation that, in principle, a portion of the base elements of the ASCOS Risk Model's FTs could be quantified with ECCAIRS data. Doing so would allow to define of the model Fault Tree base events according to a standard taxonomy—ECCAIRS—with which many experts are already familiar. It was noted this could be one way to establish a clear connection with the products of ASCOS WP2. The idea of linking the model with the ECCAIRS taxonomy and data was countered by the observation that ECCAIRS output is not visible to applicants, but to authorities only.

6.3.4 Standardization

This KPA investigated the potential of the Tool for risk assessment to become a standard software tool to be used across the TAS domains. The majority of the ratings fell into the neutral score: The expert noted they did not have enough information to rate the potential for standardization of the tool. They suggested that information useful for expressing a more certain judgment about the potential for standardization would come from a comparative benchmark—a comparative benchmark with the functionalities of other comparable software, i.e. FT editors, available on the market. The reason for this suggestion comes from the fact that the ASCOS Risk model could be implemented, in principle, also into other comparable FT editing software. Thus, assessing the potential for standardization requires an understanding of the relative advantages the Tool for risk assessment offers in comparison to these available editors.

6.3.5 Acceptability from an Applicant and Certifier perspective

This KPA investigated how acceptable the risk assessment tool is from the perspective of a certifying authority and an applicant respectively. Two positions emerged regarding this question: First, one expert noted that the tool seems acceptable provided experienced personnel is available—whereby experienced means specialist with at least 5 to 10 years of experience in aviation. However, this point was not elaborated further during the discussions. Second, it was noted that a thorough appreciation of acceptability requires an understanding of the infrastructure of people and technology that need to be established by an organization in order to adopt the tool. Moreover, acceptability depends on a clear definition of the role of the tool. If the tool is perceived as redundant with existing (local) tools, the advantage of adopting it would not be clear for an organization. However, if the tool is perceived as useful for TAS risk assessment specifically, then it would have a much clearer purpose and added value to existing "local" tools.

6.4 ASCOS Tool for risk assessment

This section reports the results specific for the ASCOS Tool for risk assessment. These come from (i) the System Usability Questionnaire (SUS), which was administered to the experts immediately after their trial session, and (ii) from the second group discussion of Exercise 3.

6.4.1 Usability

The System Usability Scale (SUS) [23] yields a single score representing a composite measure of the overall usability of the system under evaluation. The ASCOS Risk assessment tool's SUS score is 57 (out of 100), a value that indicates that the tool requires usability improvement[5]. To note that the SUS score alone has no diagnostic power: It does not indicate what the problematic HMI areas and the required improvements are. For this reason, feedback has been collected during the exercise from the experts with regard to the specific aspects of the tool to be improved.

6.4.2 Suggested areas for improvement

The second group discussion of ASCOS Exercise 3 highlighted usability aspects of the tool that could be improved. These are listed below:

1. **Providing suggestions for identifying which ESDs/FTs could be affected by a change**. Currently, the use of the tool demands the user, a safety analyst, to be thoroughly familiar with the risk model—all the knowledge of the model has to be "in the head" of the safety analyst. Therefore, it was suggested to have a tool functionality that, depending on the change at hand, suggests which FTs of other ESDs could be affected. The tool should not change automatically the related FTs in other ESDs that should be considered. Rather it should only highlight them. This functionality would be particularly helpful in cross-domain risk assessment: While an analyst is usually expert of her/his own domain, she/he may be less familiar with the FTs of other domains, and therefore it could be useful if the tool highlights hazards (or FTs) to consider.

2. **Providing support for the graphical exploration of the model elements.** Suggestions were made to ease the search and exploration functions of the tool. Currently, access to the model elements is mostly based on key words. Therefore, it was suggested to adopt a more graphical format to ease model element exploration. Also, there should be a function to facilitate the identification of the relevant model elements that should not be limited to the name of ESD and FT events only, but should also include the definitions of model elements.

3. **Displaying hover box next to the model's element.** It was suggested to implement definitions of the model elements, and make these available to the user when needed by means of hover boxes. A hover box is a box displayed when the mouse moves hover a pre-set trigger area. Specifically to the ASCOS tool, when the user hovers over any of the model elements with the mouse pointer the system could display a corresponding hover box. This would provide details about the selected element.

4. **Supporting compatibility with other FT software.** Compatibility with other FTs software is important to consider. It was suggested that the tool should be able to export the results of an analysis to other

[5] SUS score can be interpreted based on the following: <50= not acceptable; 50-70= acceptance is marginal, some improvements are needed; >70: acceptable.

FT software programs. At the moment, the tool can export data to MS Excel, so compatibility depends on the ability of other FT software programs to read Excel files. Also, it was noted that the tool could have some sort of "hooks" for detailed FTs developed by service providers in other software packages.

5. **Enhancing user input of probability values.** The input of updated probability values could be enhanced in order to make it easier for the user while reducing the risk of data entry mistakes. When manually entering a new probability value, a simple typo, in fact, could result in a very significant—and mistaken—variation in the order of magnitude of the value. Such data input mistakes could be avoided by enhancing the tool's HMI with a safe guard that constraints the range of user inputs that can be accepted by the tool. Having such a protection was considered quite important by the experts to prevent erroneous entries of unrealistic probability values. In addition to that, it was suggested to further ease the data input method. In particular, the tool HMI (i) could suggest upper and lower threshold probability values, (ii) could provide pre-set modification factors of these values (e.g. 0.5, 1.5, 2, 2.5), (iii) could offer direct support for the calculation of an updated probability value depending on the chosen modification factor (currently, the user has to recur to calculators external to the tool, either paper or computer based).

6. **Integrating an audit trail capability**. It was remarked that the tool should include an "audit trail" capability, i.e. the capability to maintain a detailed chronological records of the changes made to probability values. Such records should contain information such as the person responsible for making the change, detailed description of the change, reasons, evidence and underlying assumptions for the change. This capability was considered important because it provides an important means to establish the confidence of the risk assessment outcomes produced by the tool. The tool already offer a functionality to keep track of changes; therefore, this could be further evolved to cover the need for an audit trail.

6.5 Summary of Exercise 3 validation results

The previous sections have reported the validation results specific to the proposed ASCOS Risk Model and Tool for risk assessment. These results have been collected during Exercise 3, and they are summarized in Table 13 and Table 14 respectively.

Table 13. EXERCISE 3: Summary of the validation results for the ASCOS Risk Model.

KPA	ASCOS Risk Model
1. Soundness	- The risk model appears too generic if compared against local risk models developed by individual stakeholders. Since it is likely that more specific ESDs and FTs will be available at local level, it should be further defined how to connect such models with the generic ASCOS risk model. Such a connection could be obtained by establishing proper hooks between the top even of (local) detailed FT to the relevant ESD or high level FT structure of the ASCOS risk model.

	- It should be clarified *how* the ASCOS risk model can be used and *by whom* in a TAS risk assessment. Roles and responsibilities associated with the use of the risk model should be further clarified; - The Risk Model creates an expectation that it will serve as support for the identification of novel and unknown risks, while this is not the case. In ASCOS this capability comes from FAST, and therefore it should be clarified how FAST can be used to inform and maintain the Risk Model. - To be a true risk model, the model should have the capability to assess severity levels and consider different probability units.
2. Completeness	- The models appear biased towards the coverage of aircraft operations mainly. ATM operations should have the same level of coverage; - Limiting the model to historic data on the one hand allows quantification, on the other makes the model incomplete with regard to the full range of risks that may occur; - The validity of the Risk Model could be enhanced by feeding it with worldwide safety data, so to cover also third world areas, which notably have working practices (and sources of hazards) very different compared to Europe.
3. Standardization	- The use of the term "emerging risks" associated to the ASCOS Risk Model was criticized for its potential ambiguity and limited compatibility with existing safety management standards.
4. Acceptability	- More information about the staff, expertise, and technological infrastructure needed to use the model would be needed to collect feedback on the acceptability of the risk model.

Table 14. EXERCISE 3: Summary of the validation results for the ASCOS Tool for risk assessment

KPA	ASCOS Risk Tool
2.1 Manipulability	- Perceived as a useful support tool to interact with/modify the elements of the risk model; - Perceived mainly as an editor of the ASCOS Risk Model, rather than a proper risk assessment tool, due to the lack of a severity assessment output/functionality.
2.2 Quantification capability	- It should have the capability to assess severity levels; - It should have the capability to handle different probability units of model elements; - The current lack of confidence data limits the meaningfulness of the probability estimates generated by the model/tool.
2.3 Cross-domain integration	- The wide spread adoption of the tool may help promoting the development of a standard risk language across the various TAS domains; - ECCAIRS data could be used to quantify a portion of the FT model and ESD events;

	- The use of the tool requires a regulatory requirement to make it obligatory for TAS stakeholders to share risk related information, so that quantitative risk assessment can be integrated (same point made for the certification approach).
2.4 Standardization	- The ASCOS risk model could be, in principle, implemented on different software. Thus, a comparative benchmark with other comparable FT editor is desirable to understand the true potential of the Tool in term of standardization and compatibility
2.5-2.6 Acceptability	- The tool seems acceptable to the user; however, a thorough assessment of acceptability requires an appreciation of the organizational and IT arrangements needed to operate the tool.
Usability	- Overall SUS score was 57, meaning that the tool usability falls in the marginal acceptance range, i.e. some improvements are needed. Identified areas of improvement include: - Providing suggestions for identifying which ESDs/FTs could be affected by a change; - Providing further support for the graphical exploration of the model elements; - Displaying hover boxes next to the selected model's element with information about the element; - Supporting compatibility with other FT software packages (available on the market); - Enhancing the user manual input of probability values; - Integrating an audit trail capability.

7 Conclusions

This deliverable has reported the results of the validation exercises carried out in the context of ASCOS Work Package 5.3. The exercises evaluated (i) the result of WP1, i.e. the Proposed ASCOS certification approach; (ii) the main results of WP2, i.e. the ASCOS SPI Framework and the ASCOS tool or Continuous Safety Monitoring; and (iii) the main results of WP3, i.e. the ASCOS Risk Model, and the ASCOS Tool for risk assessment. These three sets of results have been addressed each by a dedicated validation exercise—each exercise consisting of a one-day workshop, in which data has been collected by means of structured focus group discussions and individual questionnaires.

For each ASCOS solution, the exercises have identified bottlenecks, shortcomings, and, most importantly, areas of performances in which these solutions can improve current practices. The tables presented under sections 4.9, 5.4, and 6.5 present these results in a summarized version. These results require further elaboration in order to define consolidated recommendations for improvement. This will occur in the ASCOS WP5.4, which will take into account not only the results presented in this report, but also the results of the cases developed by the WP4. In turn, this material will provide the basis for the definition of recommendations for improving the proposed solutions. These recommendations will feed into ASCOS WP1.5 [1], which will develop a more refined version of the ASCOS certification approach.

In concluding this report, it is necessary to remind the informed reader of at least three points that justify the approach of the present study.

First, the chosen methodological format, a one-day workshop with selected experts, was justified by the initial estimated maturity level of the products (this value was initially estimated to fall in the range V1—V2 at the beginning of the WP5, i.e. at the time the validation strategy has been defined and the planning of the present work has begun). To note that this approach should not be equivocated with a simple familiarization: Familiarization was an important phase of the workshops, but it served mainly the purposes of preparing the experts to the main workshop phase, i.e. the phase which collected the experts views by means of focus group discussions and questionnaires. It is possible that the actual maturity level of specific products at the end of the project differs from the value estimated at the beginning. E.g. the tool for continuous safety monitoring is now anticipated to be put in the public domain by JRC for stakeholders use without need for further research.

Second, consistently with EUROCONTROL E-OCVM, the validation perspective adopted by the WP5 has focused on understanding the extent to which the proposed certification solutions meet their intended purpose. In other words the results that emerged in this study are results about validation, not verification. It was not an issue of investigating how well the ASCOS solutions meet their initial specifications (verification objective), but an issue of investigating how the prosed process may fit in the target domain (certification). Therefore, data collection has not been limited to a “ticking-the-box” approach, i.e. checking that the ASCOS solutions meet their initial specifications. It was more than that. It was, indeed, an issue of gaining insights into the views of experts, and then understand the potential value of ASCOS from this view.

Thirdly, each exercise investigated the value that the proposed ASCOS results may bring to certification. And this value was considered from the perspective of the target users of the ASCOS products: experts knowledgeable in the areas of certification, safety monitoring and risk assessment experts. The WP5 strove, to recruit experts knowledgeable in these areas. By virtue of their direct professional experience and independence from the project, they have brought insightful and fresh perspectives about the value of the ASCOS product in a way that was not possible to achieve by relying solely on the experts internal to project. Indeed, it was anticipated that failing to integrate these perspectives would have resulted in a very self-referential validation, significantly biased towards an overoptimistic view of the proposed solutions.

References

[1] ASCOS, “Description of Work.” Jul-2012.
[2] EUROCONTROL, “E-OCVM, Version 3.0, Volume I,” Feb. 2010.
[3] A. L. C. Roelen, J. Verstraeten, Save, L., and N. Aghdassi, “D2.1: Framework Safety Performance Indicators v1.5,” ASCOS Technical report D2.1, Jan. 2014.
[4] R. Menzel and W. Post, “D2.4: Tools for continuous safety monitoring,” ASCOS deliverable D2.4, v. 01, 7/24/2014, 2014.
[5] V. Bonvino, J.-F. Delaigue, and Heckman, J-P., “D3.2.3: Total Aviation System Safety Assessment Methodology. Representation of future risks models,” ASCOS Deliverable D3.2.3, version 1.1, 2013.
[6] H. Udluft, P. C. Roling, and R. Curran, “D3.3: Tool for risk assessment: User Manual,” ASCOS Deliverable D3.3, 2014.
[7] ACARE, *European Aeronautics Vision for 2020: Meeting society’s need and winning global leadership. Report of the group of personalities*. Belgium, 01.
[8] European Commission, “Aeronautics and Air Transport: Beyond Vision 2020 (Towards 2050): A background document from ACARE (The Advisory Council for Aeronautics in Europe).” .
[9] European Commision, “Flightpath 2050: Europe’s Vision for Aviation. Report of the High Level Working Group on Aviation Research.”
[10] B. Pauly, T. Longhurst, A. Iwaniuk, M. Idzikowski, and B. Dziugiel, “D1.1: Analysis of existing regulations and certification processes v1.3,” ASCOS Technical report D1.1, Aug. 2013.
[11] U. Dees, P. van der Geest, S. Bull, T. Longhurst, A. Eaton, G. Temme, and B. Pauly, “D1.2: Definition and evaluation o innovative certification approaches v1.3,” ASCOS Technical Report, Aug. 2013.
[12] A. Simpson, S. Bull, and T. Longhurst, “D1.3: Outline Proposed Certification Approach,” ASCOS deliverable D1.3, version 1.2, Dec. 2013.
[13] R. Wever, L. Save, S. Rozzi, and T. Longhurst, “D5.1: Validation Strategy,” ASCOS deliverable D5.1, 1.2, 2014.
[14] Rozzi, S., Save, L., S. Pozzi, M. Torelli, R. Wever, and H. Udluft, “D5.2: Validation Plan and Scenarios, version 1.1,” D5.2 (Report prepared for ASCOS: Aviation Safety and Certification of New Operations and Systems, 7th Framework Programme, Grant agreement n. 314299), November 20014.
[15] *A Guide to the Project Management Body of Knowledge (PMBOK Guide) 5th Edition*. Project Management Institute, 2013.
[16] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. Cambridge, Mass: The MIT Press, 2012.
[17] G. Boy, *Orchestrating Human-Centered Design*. London: Springer, 2012.
[18] P. Checkland, “Systems thinking, systems practice: includes a 30-year retrospective,” 1999.
[19] “D6.4: Exploitation Plan,” Technical report ASCOS D6.2.
[20] “ASCOS Website,” 2014. [Online]. Available: www.ascos-project.eu.
[21] FAA, “CPS Commercial Airplane Certification Process Study. An evaluation of selected aircraft certification, operations, and maintenance processes.” 2002.
[22] EASA, “European Aviation Safety Plan 2014-2017,” European Aviation Safety Agency.
[23] J. Brooke, “SUS-A quick and dirty usability scale,” *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.
[24] M. Schreier, *Qualitative Content Analysis in Practice*. Thousand Oaks, California: SAGE Publications Ltd, 2012.
[25] *Improving the Continued Airworthiness of Civil Aircraft: A Strategy for the FAA’s Aircraft Certification Service*. National Academies, 1998.
[26] EUROCONTROL, “From Safety-1 to Safety-II: A White Paper,” Bruxells, Whiote Paper, 2013.
[27] EUROCONTROL, “Safety - Focus on what goes right,” 2013. [Online]. Available: https://www.eurocontrol.int/news/safety-focus-what-goes-right. [Accessed: 03-Jun-2015].
[28] ICAO, “Doc 9859 Safety Management Manual (SMM) 3rd Ed.” International Civil Aviation Organization, 2013.

[29] A. Iwaniuk, Michalak, P., van Es, G, B. Dziugiel, W. Miksa, Mączka, M., N. Aghdassi, Menzel, R., and Save, L., "D2.3: Process for Safety Performance Monitoring," Technical report ASCOS deliverable D2.3 vo.1, Mar. 2014.

[30] "FAST Overview." [Online]. Available: http://www.nlr-atsi.nl/fast/overview/index.html. [Accessed: 04-Mar-2015].

[31] CAA UK, "Bowtie Document Library | Access the Significant Seven Bowtie Templates | Operations and Safety." [Online]. Available: http://www.caa.co.uk/default.aspx?catid=2792&pagetype=86&sglid=28. [Accessed: 04-Mar-2015].

[32] B. Shneiderman, *Designing the user interface: strategies for effective human-computer interaction*, vol. 2. Addison-Wesley Reading, MA, 1992.

[33] B. Ale et al, Causal Model for Air Transport Safety, Final Report, March 2009.

[34] DNV, Spreadsheet with CATS model. Appendix to the CATS Final Report, 2009.

[35] DEGAS, Advice on the Causal model of Air Transport Safety, DEGAS Advice 2008-068, October 2008.

[36] R. Wever, Working document in preparation of ASCOS validation exercises, 22 October, 2014

[37] DGLM, Beleidsagenda Luchtvaartveiligheid 2011-2015, 2011 (in Dutch).

[38] W. Kleine Beek, Review of principles for the outline proposed certification approach, ASCOS D1.3, version 1.0, EASA/WKB/E(2)2013(D)55896, 9 December 2013

[39] S. Bull, A. Simpson, T. Longhurst; Response to EASA/WKB/E(2)2013(D)55896, version 1.0, 28-01-2014.

Appendix A Definitions

Applicant. In a certification context, the organization that formally asks for a certificate to the certifier or certification authority.

Benefit Mechanism. It consists of a cause-effect description of the improvement proposed by a project. It shows how the change proposed by a project leads the intended benefit. It can be presented in both textual and diagrammatic form [27].

Certification. The process and set of activities aiming at the satisfaction of an authority that a "deliverable" (e.g. aircraft, aviation product, service, or organization) complies with a set of regulations in order to ensure its proper operation and to ensure continued performance of these items during their operational life.

Certification Safety Argument. A documented body of evidence that provides a convincing and valid argument that a system is acceptably safe for a given application in a given environment [31]. A safety case is composed by a set of explicit claims about the system, a body of supporting evidence, by a set of inference rules that link the claims to the evidence, an explicit set of fundamental assumptions and judgments in which the argument is valid. While this definition applies to safety arguments in general, we refer here to certification safety argument as the safety argument to be used in a certification context.

Certification Safety Assurance Documentation. The documentation produced by the applicant to demonstrate to the certifying authority that the change in question is acceptably safe. This documentation contains the evidence, the data, and the assumption that support the overall claim that the change is acceptably safe. After being produced by the applicant(s), this documentation is then reviewed by the certifiery body which will have to approve or reject such documentation. In countries like the UK the Certification Safety Assurance Documentation corresponds to the certification safety argument. However, while this latter is perhaps the most sophisticated methodology to demonstrate the acceptable safety of a given change, safety arguments are not adopted uniformly across all states and stakeholders, and some states might use other approaches.

Certifier. In a certification context, the authority that provides certification.

Change. In this context, the replacement or the introduction of a new procedure, service, operation, hardware or software system.

Cross acceptance. A situation where equipment in-service accepted by a particular authority is accepted for use by a different authority.

Emerging risk. This is defined as an existing hazard with a change in risk level, i.e. a change in severity level, a change in the likelihood, or both.

Future risk. This is defined as a risk associated with the future introduction of a novelty (e.g. new design, new procedure, and new organization).

Key Performance Area (KPA). "Key performance areas are broad categories that describe different areas of performance of an ATM system." [22]. "Key Performance areas are a way of categorising performance subjects related to high-level ambitions and expectations. ICAO has defined 11 KPAs: Safety, security, environmental impact, cost effectiveness, capacity, flight efficiency, flexibility, predictability, access and equity, participation and collaboration, interoperability." [28].

Key Performance Indicator (KPI). Key performance indicators "measure performance in key performance areas and are identified once key performance areas are known. A key performance indicator is a measure of some aspect of a concept or concept element, for example 'the total number of runway incursions per year', the 'mean arrival delay per week at airport X'." [22]. "Current/past performance, expected future performance (estimated as part of forecasting and performance modelling), as well as actual progress in achieving performance objectives is quantitatively expressed by means of indicators (sometimes called Key Performance Indicators, or KPIs). To be relevant, indicators need to correctly express the intention of the associated performance objective. Since indicators support objectives, they should not be defined without having a specific performance objective in mind. KPIs are not often directly measured. They are calculated from supporting metrics according to clearly defined formulas, e.g. cost-per-flight-indicator = Sum (cost) / Sum(flights). Performance measurement is therefore done through the collection of data for the supporting metrics." [28].

Metrics. "Supporting metrics are used to calculate the values of performance indicators. For example cost-per-flight indicator = Sum(cost) / Sum(flights). Performance measurement is done through the collection of data for the supporting metrics (e.g. this leads to a requirements for cost data collection)" [28].

Modularization. It is the process of breaking up complex or large arguments into manageable modules.

Performance Framework. A performance framework is "used to document and establish the framework for performance assessment. It typically consists of Key Performance Areas (KPAs), key performance indicators (KPIs), performance targets, metrics and measurement-related assumptions which are used to validate a concept. The performance framework may be enhanced to support the understanding of how benefit is produced and delivered and for the examination of performance trade off" [22]. A performance framework needs "to be in place at the very early stage to ensure that it is taken into account in the planning of the validation programme and exercises" [22].

Performance Target (PT). "Performance targets are closely associated with [key] performance indicators: they represent the values of performance indicators that need to be reached or exceeded to consider a performance objective as being fully achieved." [28].

Use Case

Validation Scenario. "A validation scenario is a specific scenario developed for the purposes of undertaking validation activities and to gather evidence relevant to the validation objectives. It is used to analyse the performance and interactions described or expected in the operational concept scenarios." (E-OCVM)

Validation. The process by which the fitness-for-purpose of a new system or operational concept being developed is established”. The objective of the validation of the ASCOS results is to demonstrate that they are suitable for their intended purpose or use and brings the expected benefits for the user

Verification. Verification is the set of activities aimed at testing or demonstrating that the product (e.g. a tool) meets the technical specifications. The verification aims to assess the technical quality and performance of the products.

Appendix B EXERCISE 1: Agenda

Agenda V4.1

ASCOS User Group Meeting 3

9 & 10 October 2014

Amsterdam, The Netherlands

<u>Agenda Friday 10 October</u>

Goal of the meeting: The second day of the User Group Meeting is organized in the form of a workshop in which the participants are asked to take part in a number of validation sessions. These aim at testing and collecting feedback on the new certification approach proposed by ASCOS based on the Key Performance Areas that were identified.

8.30– 9.00	**1. Introduction into the validation activities for the new certification approach – Luca Save** *Form: short presentation by WP5 leader* *Goal: To clarify the goals of the second day of the User Group Meeting and to explain the difference with the discussions on the first day of the User Group Meeting*
9.00 – 10.00	**2. Validation Session 1 – Efficiency and Soundness of ASCOS certification approach** *Form: The validation exercise will start with a plenary presentation about the key performance area 'Soundness and Efficiency', this is followed by completing a questionnaire and discussions in subgroups.* *Goal: to collect feedback from the user group members on the soundness and efficiency of the ASCOS certification approach.*
10.00 – 10.15	**Break**

10.15 – 11.00	**3. Continuation Validation Session 1 – Soundness and Efficiency of ASCOS certification approach**
11.00 – 12.30	**4. Validation Session 2 – Cross-domain integration and harmonization of ASCOS certification approach** *Form: The validation exercise will start with a plenary presentation about the key performance area 'Cross-domain integration and harmonization', this is followed by completing a questionnaire and discussions in subgroups.* *Goal: to collect feedback from the user group members on the cross-domain integration and harmonization of the ASCOS certification approach.*
12.30 – 13.00	**5. Validation Session 3 – Accommodation of Innovation, Operability and Flexibility of ASCOS certification approach** *Form: The validation exercise will start with a plenary presentation about the key performance area 'Accommodation of Innovation, Operability and Flexibility', this is followed by completing a questionnaire and discussions in subgroups.* *Goal: to collect feedback from the user group members on the Accommodation of Innovation, Operability and Flexibility of the ASCOS certification approach.*
13.00 -14.00	**Lunch**
14.00 – 15.00	**6. Continuation Validation Session 3 – Accommodation of Innovation, Operability and Flexibility of ASCOS certification approach**
15.00 – 15.30	**7. Debriefing – Luca Save & Simone Rozzi** Form: plenary presentation Goal: to inform all participants about the results of the day and the resulting next steps to be taken by the ASCOS team
15.30	**End of meeting – Gerard Temme**

Appendix C EXERCISE 1: Questionnaire

BIOGRAPHICAL DATA

NAME:	*IDENTIFICATION CODE*
ORGANIZATION:	
CURRENT POSITION IN THE ORGANIZATION:	
YEARS OF EXPERIENCE IN CERTIFICATION:	
INVOLVED IN CERTIFICATION ACTIVITIES IN THE ROLE OF: ☐ CERTIFIER ☐ APPLICANT ☐ BOTH CERTIFIER AND APPLICANT ☐ NONE OF THE ABOVE	
EXAMPLES OF PRODUCTS OR SERVICES OR YOU HAVE CERTIFIED: - - -	

1-EFFICIENCY OF THE CERTIFICATION PROCESS (5)

TIME

1.1: Compared to current certification practices, the ASCOS certification approach will reduce the overall time needed to complete a certification process from the initial application to the final approval by the authority.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

EFFORT

1.2: Compared to current certification practices, the ASCOS certification approach will reduce the effort (time and resources) needed by the certifying authority to follow the overall certification process and to review the final documentation produced by the applicant.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

1.3: Compared to current certification practices, the ASCOS certification approach will reduce the effort (time and resources) needed by the applicant to complete the overall certification process, including the production of the required final documentation.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

TRAINING

1.4: Compared to current certification practices, the ASCOS certification approach will reduce the effort required to the certifying authorities to train their personnel to fulfil the certifier role.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

1.5: Compared to current certification practices, the ASCOS certification approach will reduce the effort needed by the applicants to train their personnel to carry out certification activities.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

2 SOUNDNESS OF THE CERTIFICATION SAFETY ASSURANCE DOCUMENTATION (3)

2.1: Compared to current certification practices, the ASCOS certification approach will improve the consideration of cross-domain hazards in the safety assurance documentation.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

2.2: Compared to current certification practices, the ASCOS certification approach will improve the consideration of human factors related hazards in the safety assurance documentation.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

2.3: Compared to current certification practices, the ASCOS certification approach will promote the consideration of hazards through all the lifecycle phases (including 'transition into operations') when developing the safety assurance documentation.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

3 CROSS-DOMAIN INTEGRATION (3)

3.1 The ASCOS certification approach will promote and support the cooperation and exchange of information among stakeholders from different aviation domains during the development of a certification process.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

3.2 The ASCOS certification approach will help to clarify the roles and responsibilities of different stakeholders across different aviation domains starting from the early stages of a certification process.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

3.3 The ASCOS certification approach will facilitate the integration among local certification approaches in use in each domain, regardless of whether they are performance or compliance-based.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

4 HARMONIZATION (2)

4.1: The ASCOS certification approach is sufficiently compatible with existing certification practices across Europe (such as different domains, geographical areas and industries) to prevent disruption or waste of existing experience and competences.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

4.2: The ASCOS certification approach has the potential to become a standard reference approach in use across different domains and stakeholders.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

5 ACCOMMODATION OF INNOVATION (1)

5.1: Compared to current certification practices, the ASCOS certification approach will simplify the certification of innovative product and services.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

6 ACCEPTABILITY (2)

6.1: The introduction of the ASCOS certification approach has the potential to be accepted by the existing certifying authorities

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

6.2: The introduction of the ASCOS certification approach has the potential to be accepted by the applicant organizations.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

7 FLEXIBILITY (2)

7.1: The ASCOS certification approach is sufficiently flexible to be applicable to a wide range of products and services of different size and complexity.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

7.2: The ASCOS certification approach is fit for certifying both novel and derivative products and services.

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree
6. not sure/not applicable

Appendix D EXERCISE 2: Agenda

ASCOS Validation Workshop 2

27 November 2014

ENAC

Meeting room: Room SO10, Floor -1, Via Gaeta 3, Rome

Time	Session
9:00 – 9:20	**1. Welcome and Introduction – Luca Save & Simone Rozzi (DBL)**
9:20 – 10:00	**2. Overview of the ASCOS project and the proposed certification approach – Rombout Wever (NLR)**
10:00 – 10:45	**3. Safety monitoring in ASCOS – Nuno Aghdassi (Avanssa)**
10:45 – 11.00	*Coffee break*
11:00 – 12.00	**4. ASCOS tool for continuous safety monitoring: Introduction + Demo – Reinhard Menzel (JRC)**
12.00 – 13.00	**5. Group Discussion 1: Feedback on ASCOS SPIs and on ASCOS tool for continuous safety monitoring**
13.00 – 14.00	*Lunch*
14.00 - 14.45	**6. Group Discussion 2: ASCOS Safety Monitoring and Certification**
14.45 - 15.00	*Coffee break*
15:00 – 15:30	**7. Group Discussion 2 (*continuation*)**

Appendix E EXERCISE 2: QUESTIONNAIRE 1

NAME:	
ORGANIZATION:	
CURRENT POSITION IN THE ORGANIZATION:	
YEARS OF EXPERIENCE IN SAFETY MANAGEMENT:	
CURRENT OR PAST ROLE IN SAFETY MANAGEMENT ACTIVITIES:	

1.1: The **ASCOS Safety Performance Indicators (SPIs)** cover adequately the different domains of the Total Aviation System.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

1.2: The **ASCOS Safety Performance Indicators (SPIs)** have the potential to become **standard reference indicators** to be used across the different domains of the Total Aviation System.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

1.3: The **ASCOS Tool for Continuous Safety Monitoring** (ATCSM) makes the monitoring of SPI trends in a safety monitoring context easier.

Strongly disagree	1	2	3	4	5	Strongly Agree		Not sure/not applicable	

Please insert here any comment you may have

1.4: The **ASCOS Tool for Continuous Safety Monitoring** (ATCSM) has the potential to become a standard tool used by the various organizations of the Total Aviation System.

Strongly disagree	1	2	3	4	5	Strongly Agree		Not sure/not applicable	

Please insert here any comment you may have

1.5: The potential **effort** (time and resources) required to **certifying authorities** for **training their personnel** to the use of the ATCSM tool is acceptable

Strongly disagree	1	2	3	4	5	Strongly Agree		Not sure/not applicable	

Please insert here any comment you may have

1.6: The potential **effort** (time and resources) required to **applicant organizations for training their personnel** to the use of the ATCSM tool is acceptable.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**		**Not sure/not applicable**	

Please insert here any comment you may have

Appendix F EXERCISE 2: Questionnaire 2

NAME:	

3.1 The ASCOS Safety Monitoring Concept will **make more robust** the existing monitoring and maintenance of certification.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.2 The ASCOS Safety Monitoring Concept will **make more cost-effective** the existing monitoring and maintenance of certification.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.3 The ASCOS Safety Monitoring Concept will increase the **rate of success for the certification of innovative systems**—i.e. systems for which no standards and regulations exist and therefore a fully compliance-based approach does not apply.

Strongly disagree	**1**	**2**	**3**		**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

	Please insert here any comment you may have

3.4 The ASCOS Safety Monitoring Concept will be **accepted** by the certifying authority and the applicant.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

Appendix G EXERCISE 3: Agenda

Time	Session	
9:00 – 9:20	1	**Welcome and Introduction – Deep Blue**
9:20 – 9.30	2	**Overview of the ASCOS project – NLR**
9.30 – 10:00	3	**ASCOS certification approach – EBENI**
10:00 – 11:15	4	**Results from ASCOS WP3: The ASCOS Risk Model and the supporting ASCOS Risk Assessment Tool – NLR/TUD**
10:45 – 11.00		*Coffee break*
11:00 – 11.30	5	**Demonstration of the ASCOS Risk Model – NLR**
11.30 – 12.30	6	**GROUP DISCUSSION 1: The ASCOS Risk Model**
12.30 – 13.30		*Lunch*
13.30 – 15.00	7	**ASCOS Risk Assessment Tool: (a) Introductory interactive session and (b) risk assessment exercise with participants – NLR/TUD**
15.00 - 15.15		*Coffee break*
15:15 – 16:15	8	**GROUP DISCUSSION 2: The ASCOS Risk Assessment Tool**

Appendix H EXERCISE 3: Risk model questionnaire

NAME:	
ORGANIZATION:	
CURRENT POSITION IN THE ORGANIZATION:	
YEARS OF EXPERIENCE IN RISK/SAFETY >MANAGEMENT:	
CURRENT OR PAST ROLE(S) IN RISK/SAFETY MANAGEMENT ACTIVITIES:	

1.1: When needed, the use of the ASCOS Risk Model will **make more robust** risk assessment activities in a certification context.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

1.2: The set of 38 Event Sequence Diagrams (ESDs) included in the ASCOS Risk Model **is adequate** to cover the different domains of the Total Aviation System.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

1.3: The ASCOS Risk Model has the potential to become a **standard risk model** to be used by the different organizations of the Total Aviation System.

Strongly disagree	1	2	3	4	5	Strongly Agree	Not sure/not applicable	

Please insert here any comment you may have

1.4: In a certification context, the potential **effort** (time and resources) required to **applicant organizations** to **adopt, use and maintain** the ASCOS Risk Model is acceptable.

Strongly disagree	1	2	3	4	5	Strongly Agree	Not sure/not applicable	

Please insert here any comment you may have

1.5: In a certification context, the potential **effort** (time and resources) required to **certifying authorities** to **adopt, use and maintain** the ASCOS Risk Assessment Model is acceptable.

Strongly disagree	1	2	3	4	5	Strongly Agree	Not sure/not applicable	

Please insert here any comment you may have

Appendix I EXERCISE 3: System Usability Scale (SUS) questionnaire

NAME:	

	Strongly disagree				Strongly agree
2.1. I think that I would like to use this system frequently.	**1**	**2**	**3**	**4**	**5**
2.2. I found the system unnecessarily complex.	**1**	**2**	**3**	**4**	**5**
2.3. I thought the system was easy to use.	**1**	**2**	**3**	**4**	**5**
2.4. I think that I would need the support of a technical person to be able to use this system.	**1**	**2**	**3**	**4**	**5**
2.5. I found that the various functions in this system were well integrated.	**1**	**2**	**3**	**4**	**5**
2.6. I thought there was too much inconsistency in this system.	**1**	**2**	**3**	**4**	**5**
2.7. I would imagine that most people would learn to use this system very quickly.	**1**	**2**	**3**	**4**	**5**
2.8. I found the system very cumbersome to use.	**1**	**2**	**3**	**4**	**5**
2.9. I felt very confident using the system.	**1**	**2**	**3**	**4**	**5**
2.10. I need to learn a lot of things before I could get going with this system.	**1**	**2**	**3**	**4**	**5**

Appendix J EXERCISE 3: Risk Assessment Tool Questionnaire

NAME:	

3.5 When assessing the safety impact resulting from a change in the Total Aviation System, the **ASCOS Risk Assessment Tool** is helpful to **identify and manipulate** the relevant elements of the ASCOS risk model—i.e. the elements that need to be updated/modified because of the change.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.6 The ASCOS Risk Assessment Tool is helpful to **quantify the potential impact on risk probabilities** resulting from the introduction of a change in the Total Aviation System.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.7 Using the ASCOS Risk Assessment Tool **will improve the exchange and sharing of risk assessment related information** across the different domains of the TAS.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.8 The ASCOS Risk Assessment Tool has the potential to become a **standard tool** used by the various organizations of the Total Aviation System.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.9 The potential **effort** (time and resources) required to **certifying authorities for training** their personnel to the use of the ASCOS Risk Assessment Tool is acceptable.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

3.10The potential **effort** (time and resources) required to **applicant organizations** (in a certification context) **for training** their personnel to the use of the ASCOS Risk Assessment Tool is acceptable.

Strongly disagree	**1**	**2**	**3**	**4**	**5**	**Strongly Agree**	**Not sure/not applicable**	

Please insert here any comment you may have

Appendix K Causal model for Air Transport Safety (CATS) case studies

Appendix A.1 Introduction

Appendix A.1.1 Background

In ASCOS WP3, a total aviation systems safety assessment method and supporting tools have been developed in support of safety based design of new systems, products and/or operations. This includes the development of a risk based model based on accident scenarios. This risk based model is an improvement of the original Causal model for Air Transport Safety (CATS) accident scenarios [33]. Validation of the CATS model and toolset is part of the objectives of the Ministry of Infrastructure and the Environment (Directorate-General for Mobility and Transport (DGB)) of 2014. Obviously, the two projects could benefit from each other's validation activities. Moreover, the European Aviation Safety Agency (EASA) is interested in using CATS. They add a relevant user perspective on the use of the CATS and ASCOS models and toolsets.

Appendix A.1.2 Objective and approach

Before the objective of the validation of the CATS software tool is defined, there is a need to clarify what is meant in ASCOS and CATS when the term "validation" is used. The term validation often is used differently and as such could cause misunderstanding of activities and expected results.

In the ASCOS validation plan [14] the difference between validation and verification is described as follows: "Validation is an issue of establishing if a product satisfies the originally intended purpose for which it was designed and if it delivers the expected benefits to the user. In short, validation answers the question "have we built the right system?". In this way it differs from verification: this latter activity focuses in fact on answering the question "have we built the system right, i.e. according to the initial specification?"".

In the CATS Final Report [33], the validation that has been performed in the project is discussed. The CATS final report states that some of these activities were performed by the consortium (verification, calibration (partly), some face validity examples) but others were postponed to a later project (a.o. case validity, convergent validity).

Note that the validation activity in CATS is defined differently than in ASCOS.

Knowing that the validation of CATS is a task which should not be underestimated [35], and given the available budget and time, the objective for the activities that are documented in this report is to perform case studies with the CATS 2009 model and toolset to show that CATS is able to address strategic user needs and to provide the user feedback on the achieved results.

It is noted that we have to distinguish between the CATS model (Fault Trees, Event Sequence Diagrams, the Bayesian Belief Net), the quantification of the events in the model, and the supporting toolset. The case studies can provide insight on all three elements of CATS. Also, it is explicitly stated that it is not the intention to modify the elements of CATS, neither qualitatively nor quantitatively.

The approach to achieve the validation objective consists of the following steps:

- Identify a set of case studies, taking into account information from regulatory users DGB and EASA and from the ASCOS project.
- Define a baseline risk level which will be used for comparison
- Run the CATS tool set; and
- Report results and findings.

Appendix A.1.3 Structure of the appendix

In Appendix A.2, case studies are identified. Appendix A.3 briefly describes the CATS model and the CATS toolset, as presented in the CATS Final Report [33]. In Appendix A.4, the case studies are elaborated and results are presented. Also, some initial findings are given. These findings are collected and discussed in Appendix A.5.

Appendix A.2 Identification of case studies

Appendix A.2.1 DGB

Obviously, DGB is an important user of CATS. Until now, user needs have been identified during the CATS project. In the CATS Final Report [33], the requirements of potential users and the wishes of the client are summarised. These are compared with what could be achieved in the CATS project.

In 2014, DGB formulated a number of specific cases that CATS should be able to address:

1. According to the model, what are the most important, most frequent accidents and what are the probabilities?
2. What are the most important precursors to air transport accidents?
3. To how many accident types do these most important precursors relate? (This could be relevant to assess the importance of an integrated model)?
4. Do the results correspond to the accents that have been found in the Netherlands and the EU SSP?
5. Sensitivity analysis: When frequencies of the precursors are changed, how much do the accident frequencies change?
6. Which mitigating measures can be identified to reduce the frequencies of the precursors?
7. What is the effect of mitigating measures to the accident probability?
8. How important and how large is the risk of bird strikes. Is this different in the Netherlands when compared to the model?

9. Is this theme (bird strikes) important when compared to other themes?

Interesting case studies can also be found in the Dutch Policy Agenda for aviation safety 2011-2015" [37]. This policy agenda presents the areas in which the Dutch aviation safety system needs to be improved. Agenda items are:

- 3.1: International spearheads aviation safety
- 3.2: Implementation of international regulations in the Netherlands
- 3.3: Integral approach inside the kingdom
- 3.4: Technical assistance to third countries
- 3.5: Factors that influence an optimal air transport operation
- 4.1: Implementation and oversight safety management systems
- 4.2: Development concrete and measurable safety objectives
- 4.3: Improve data collection and analysis
- 4.4: Establish conditions for a mature safety culture
- 5.1: Airport safety – improve safe layout and use
- 5.2: Safety outside the airport terrain and external safety: presence of obstacles, lasers and vulnerable objects in the area
- 5.3: Reduce the risk of bird strikes
- 6.1: The aircraft
    - Common standard for UAS base on approach of ICAO/EASA
    - Approach to volcanic ash
- 6.2: Optimise training methods and training levels of aviation personnel
- 6.3: Prevent abuse of medication, alcohol, and use of drugs as part of SMS
- 6.4: Introduction of Fatigue Risk Management System (FRMS)
- 6.5: Optimisation of Air Navigation Services

Note that agenda item 5.3 also has been mentioned by DGB. Also, a forthcoming update of the policy agenda is expected to cover future agenda items.

Appendix A.2.2 EASA

EASA is a potential user of CATS. EASA publishes information on existing safety actions, main risk areas, and new actions to mitigate risk. EASA's safety policy is transformed into a safety plan (EASp) which is updated yearly. The latest version covers the years 2014-2017 [22]. The EASp covers three broad areas: systemic, operational and emerging issues. The risks identified in these areas are mitigated by safety actions. The EASp reports on the status of 88 standing actions and proposes 18 new actions. These 18 new actions are the following:

- SYS3.16: EASA should assess further, together with Member States, the benefits of FDM-based indicators for addressing national safety priorities

- SYS3.17: EASA should produce, together with Member States, best practice on the oversight of FDM programmes
- SYS5.8: EASA to support Competent Authorities a: in defining the right competences needed to properly discharge their safety oversight responsibilities; and b: in providing training to their staff
- SYS5.9: Promote the concept of ‘pooling’ available expertise among NAAs in order to make subject matter experts available in a cost effective way, to those States that need resources
- SYS5.10: EASA Standardisation to monitor the availability of staff at the NAAs
- SYS5.11: A thematic workshop, with the involvement of the NAA and the industry is to be organized to promote the issues and orientations published in the Training Implementation Policy
- AER1.11: Mandating existing technology to be installed on large aeroplanes – newly designed or newly produced
- AER4.16: Develop regulations which ensure that initial and recurrent pilot training and checking is adequate to provide a pilot with the knowledge, skills and attitude to be competent in preventing and, if necessary, recovering from a loss of control in flight situation.
- AER6.1. EASA to evaluate new opportunities to mitigate the risk of on-board fires
- AER6.2: Safety Issue shall be addressed by the MS on their SSPs. This will include as a minimum agreeing a set of actions and measuring their effectiveness
- AER6.3: Develop industry best practice to outline mitigations to the risks associated with the carriage of Lithium batteries
- HE1.5: EASA to make a proposal to arrange the helicopter section of the EASp and seek an agreement with the Helicopter community
- GA1.7: EASA to make a proposal to arrange the general aviation section of the EASp and seek an agreement with the General Aviation community
- EME3.5: Provide advice to stakeholders on best practice for the management of mandatory requirements
- EME3.6: Conduct Continued Airworthiness Industry seminars and meetings to promote the applicable rules and standards
- EME3.7: Monitor achievement through oversight
- EME3.8: National Authorities to encourage compliance with ADs during meetings with industry on a regular basis and monitor level of responsiveness
- EME3.9: EASA will study possibilities to use the risk picture provided by the EASp to support the transition to a more risk-based oversight approach

Appendix A.2.3 ASCOS

In ASCOS, validation activities with the ASCOS risk model and toolset are performed. The ASCOS risk model is derived from CATS and to manage this risk model, a new software tool has been developed. For information on the ASCOS validation activities, see for example the Validation Strategy [13] and the Validation Plan and Scenarios [14]. In the latter document, validation of the safety risk assessment model and methodology and the supporting software tool is covered by validation exercise 3. In this exercise, two specific validation scenarios (case studies) are defined that could be of interest to our validation activities:

- Scenario 3.3: Introduction of Remote Virtual Tower (RVT) operations;
- Scenario 3.4: Introduction of Upset Prevention and Recovery Training (UPRT).

Appendix A.2.4 Case studies

With the information in the previous appendices, the following potential case studies are identified:

Case study	User
What are the most important, most frequent accidents and what are the probabilities	DGB
What are the main precursors to accidents	DGB
How do we deal with bird strikes	DGB
Reduce the number of overrun events during landing	EASA
Initial and recurrent training in relation to loss of control in flight	EASA
Mitigate risk of on-board fires	EASA
Implementation of regulations related to smoke and fire	EASA
Use Case 5: Remote Virtual Tower	ASCOS
Use Case 6: Upset Prevention and Recovery training	ASCOS

Based on the information in the previous appendices, the following case studies are chosen for elaboration in this report:

- Case Study 1: Most frequent accidents and their probabilities – This is a basic functionality for CATS. It is of interest both for DGB as for EASA.
- Case Study 2: Main precursors to accidents – It is of interest both for DGB as for EASA.
- Case Study 3: How do we deal with bird strikes – This case study is mentioned several times by DGB as an interesting one. It is also discussed in the DGB's Policy Agenda.
- Case Study 4: Remote Virtual Tower (RVT) – This one is easy to implement in CATS and provides an opportunity to compare the CATS and ASCOS risk models and toolsets.

These case studies are performed in Appendix A.4. Before that, the CATS model and CATS toolset that are used by these case studies are briefly described.

Appendix A.3 CATS model and toolset

The CATS case studies will be performed with the CATS model and toolset from 2009. For a detailed description of CATS, the reader is referred to [33].

CATS model

The CATS model consists of 33 Event Sequence Diagrams:

Table 15 CATS (2009) ESDs

ESD	Initiating event	Flight phase
1	Aircraft system failure	Take off
2	ATC event	Take off
3	Aircraft handling by flight crew inappropriate	Take off
4	Aircraft directional control related systems failure	Take off
5	Operation of aircraft systems by flight crew inappropriate	Take off
6	Aircraft takes off with contaminated wing	Take off
7	Aircraft weight and balance outside limits	Take off
8	Aircraft encounters windshear after rotation	Take off
9	Single engine failure	Take off
10	Pitch control problem	Take off
11	Fire on board aircraft	En route
12	Flight crew member spatially disorientated	En route
13	Flight control system failure	En route
14	Flight crew incapacitation	En route
15	Anti-ice system not operating	En route
16	Flight instrument failure	En route
17	Aircraft encounters adverse weather	En route
18	Single engine failure	En route
19	Unstable approach	Landing
21	Aircraft weight and balance outside limits	Landing
23	Aircraft encounters windshear during approach/landing	Landing
25	Aircraft handling by flight crew during flare inappropriate	Landing

26	Aircraft handling by flight crew during roll inappropriate	Landing
27	Aircraft direction control related systems failure	Landing
28	Single engine failure	Landing
29	Thrust reverser failure	Landing
30	Aircraft encounters unexpected wind	Landing
31	Aircraft are positioned on collision course	En route
32	Incorrect presence of aircraft/vehicle on runway in use	Take off, Landing
33	Cracks in aircraft pressure cabin	En route
35	Flight crew decision error/operation of equipment error	Landing
36	Ground collision imminent	Take off, Landing
37	Wake vortex encounter	En route

Note that for ESD 32 and ESD 36 two versions of the ESDs have been developed: one for the flight phase take-off and one for the flight phase landing. The two versions have a similar structure on ESD level but could have a different quantification of events.

Quantification of events

The baseline version of CATS represents the world-wide air transport system for commercial air transport for aircraft with MTOW above 5700 kg and western-built aircraft only. The 33 accident scenarios including the quantified events are presented in DNV's spreadsheet containing all relevant information of the ESDs and FTs. The Final version of the spreadsheet is [34].

CATS toolset

The CATS toolset consists of the following elements (see *Figure 15*):

- maxiCATS – This is the interface for the user. Through maxiCATS, the user can set up modifications to the model, perform calculations and see the results. It also allows the user to start the other four tools (see the next four bullets).
- CATSPAWS - CATSPAWS is supplied as a separate program by which the displays of BBN node characteristics can be directly manipulated. The connection between the BBN and the CATSPAWS database is between the coded identifiers of the nodes. All information in CATS is held in a single database. Detailed descriptors can be supplied and edited through CATSPAWS. That is also the way to adapt values when new information is obtained by analysis of data. CATSPAWS will identify inconsistencies that may develop in the coding of nodes and of missing data.
- Uninet - The user determines the number of samples to be taken after which UNINET begins its calculation. It takes the user BBN parameter settings and starts calculating and feeds the data back to

maxiCATS for the user to see. Once the calculations are done the user gets a report on them, including any adaptations to the parameters that were needed to retain consistency of the model.

- Unisens – With this tool, the user can process data after finishing the calculations, and importance measures can be derived. These are used to determine what factors are mainly important for changes in the probabilities of accidents. Unisens can also be used to evaluate correlations.
- Unigraph – The importance measures as determined by Unisens can be presented in various graphical forms.

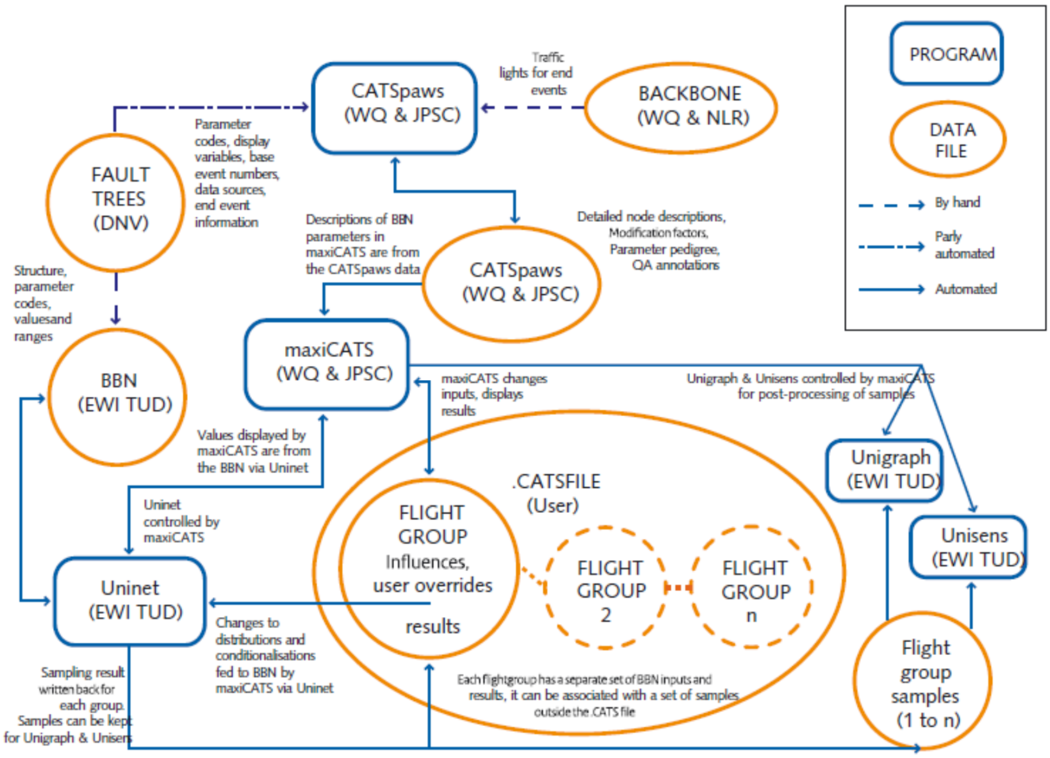


Figure 15 CATS toolset [33]

The baseline version that is used for the case studies uses the database CATSpaws94.mdb and the BBN version OVeryGBBN_09_05_06.bbn and is based on the DNV spreadsheet [34] (see the previous paragraph on quantification of events).

Appendix A.4 Case study results

Appendix A.4.1 Case study 1: Most frequent accidents and their probabilities

Question

What are the most important, most frequent accidents and what are the probabilities of these accidents? What is the level of risk in the total aviation system and how is it divided over the various accident scenarios?

Approach

To answer the question(s), the Flight Group "Baseline" as included in the CATS toolset of 2009 is used. It represents the event quantification as briefly described in Appendix A.3. The maxiCATS tool can be used to determine accident risks for the baseline situation.

Results

The following figures show some possibilities of generating accident risk related output using maxiCATS:

- Highest ESD end state accident frequency (*Figure 16*)
- Highest ESD end state fatal accident frequency (*Figure 17*)
- Highest accident frequency per ESD end event type (*Figure 18*)
- Highest frequency of accident types (*Figure 19*)

Highest ESD end state accident frequency

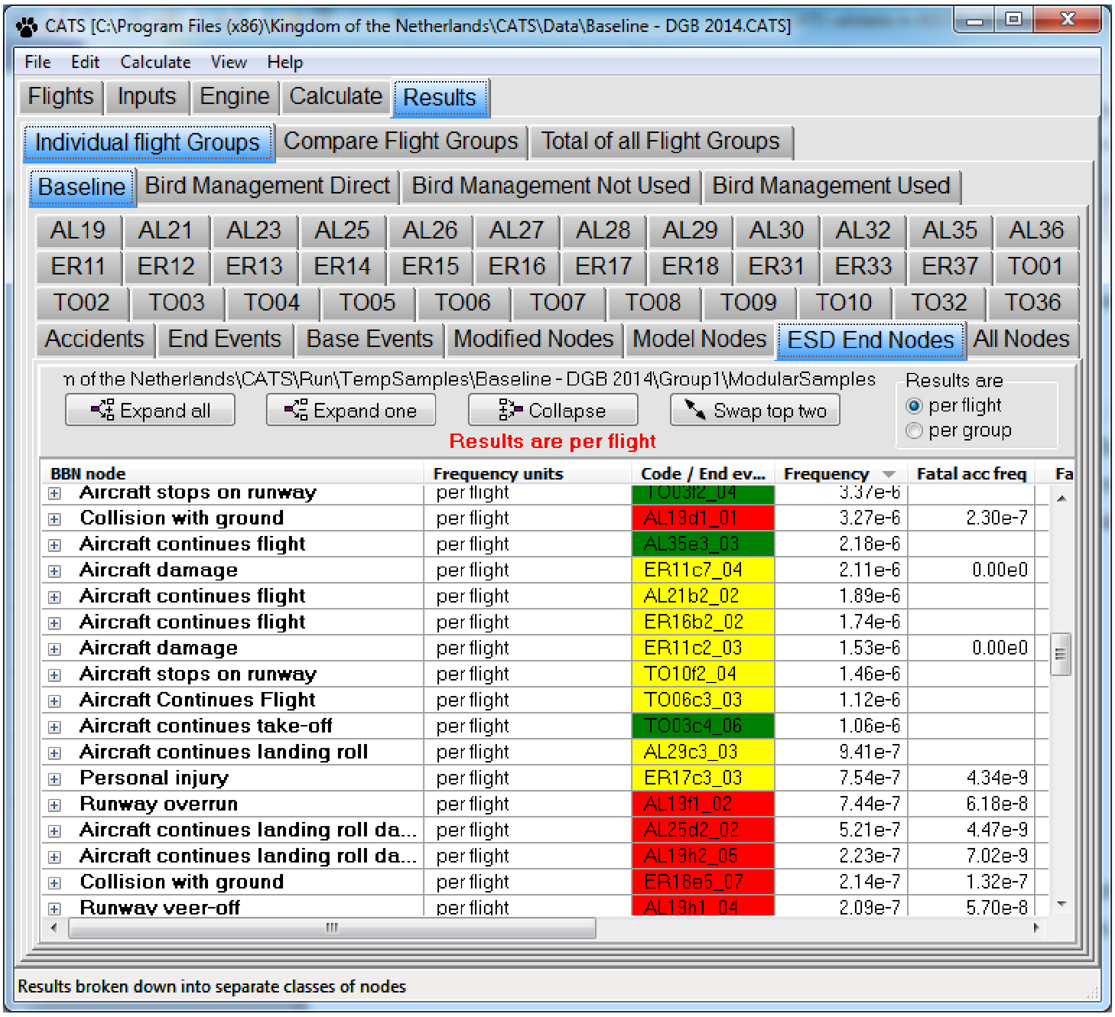


Figure 16 Highest accident frequency per ESD end node

The figure shows a subset of the results. In the list, all end states are included, not only accidents. In maxiCATS events are coloured green, yellow and red corresponding to the GOOD, SORT OF GOOD and BAD end states. maxiCATS provides the possibility to order the results (highest frequency to lowest frequency and vice versa) but does not provide the possibility to select only the accidents. Exporting these results to Excel and selecting only those ESD end states that represent an accident (or continuation of the flight with damage, the yellow ones), we have the following top 10 with respect to accident frequency:

BBN node	Frequency units	End event	Frequency
Aircraft damaged due to ground collision	per flight	TO36e1_01	2.86E-04
Aircraft damaged due to ground collision	per flight	AL36e1_01	1.24E-04
Aircraft continues flight damaged	per flight	ER33b1_02	2.06E-05
Collision with ground	per flight	AL19d1_01	3.27E-06
Aircraft damage	per flight	ER11c7_04	2.11E-06

Aircraft damage	per flight	ER11c2_03	1.53E-06
Personal injury	per flight	ER17c3_03	7.54E-07
Runway overrun	per flight	AL19f1_02	7.44E-07
Aircraft continues landing roll damaged	per flight	AL25d2_02	5.21E-07
Aircraft continues landing roll damaged	per flight	AL19h2_05	2.23E-07

Highest ESD end state fatal accident frequency

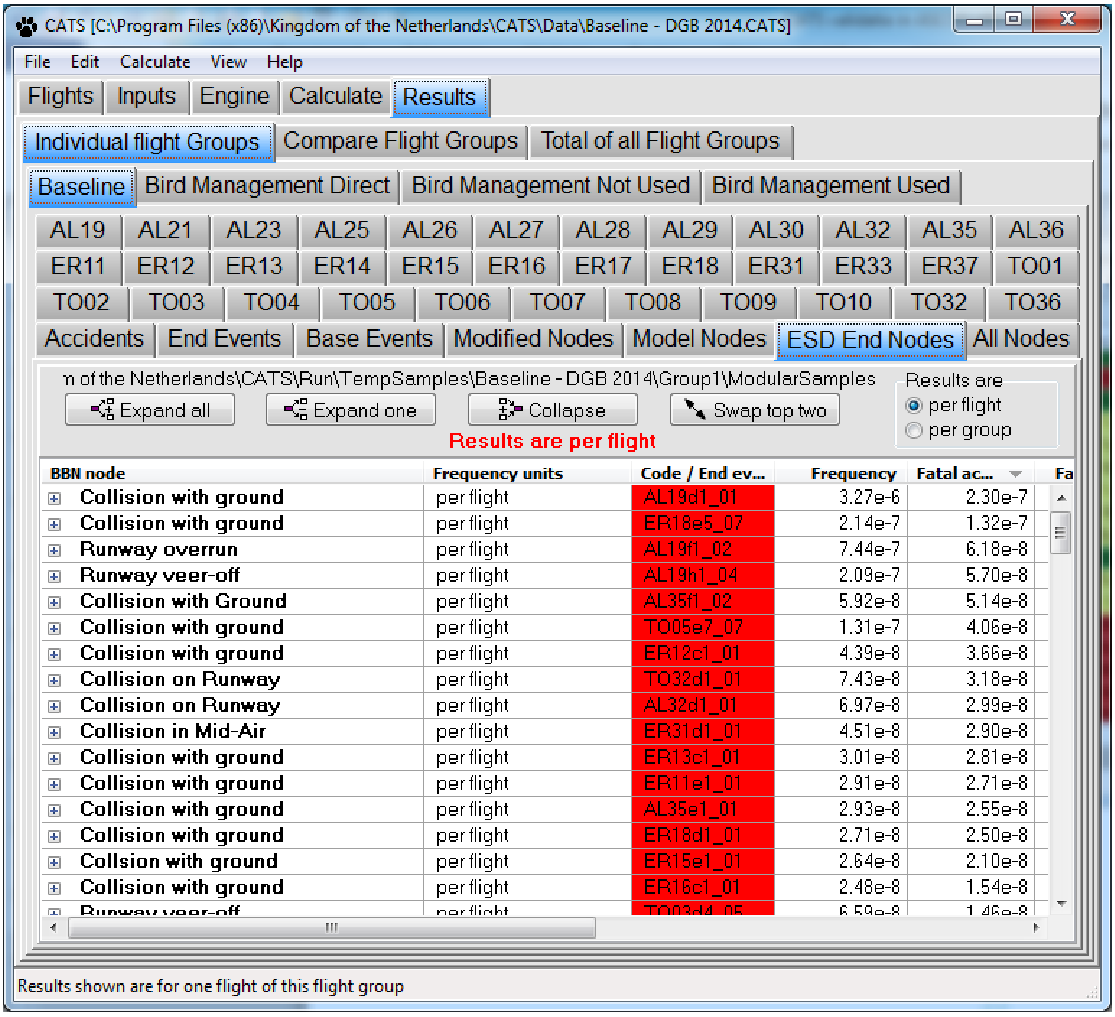
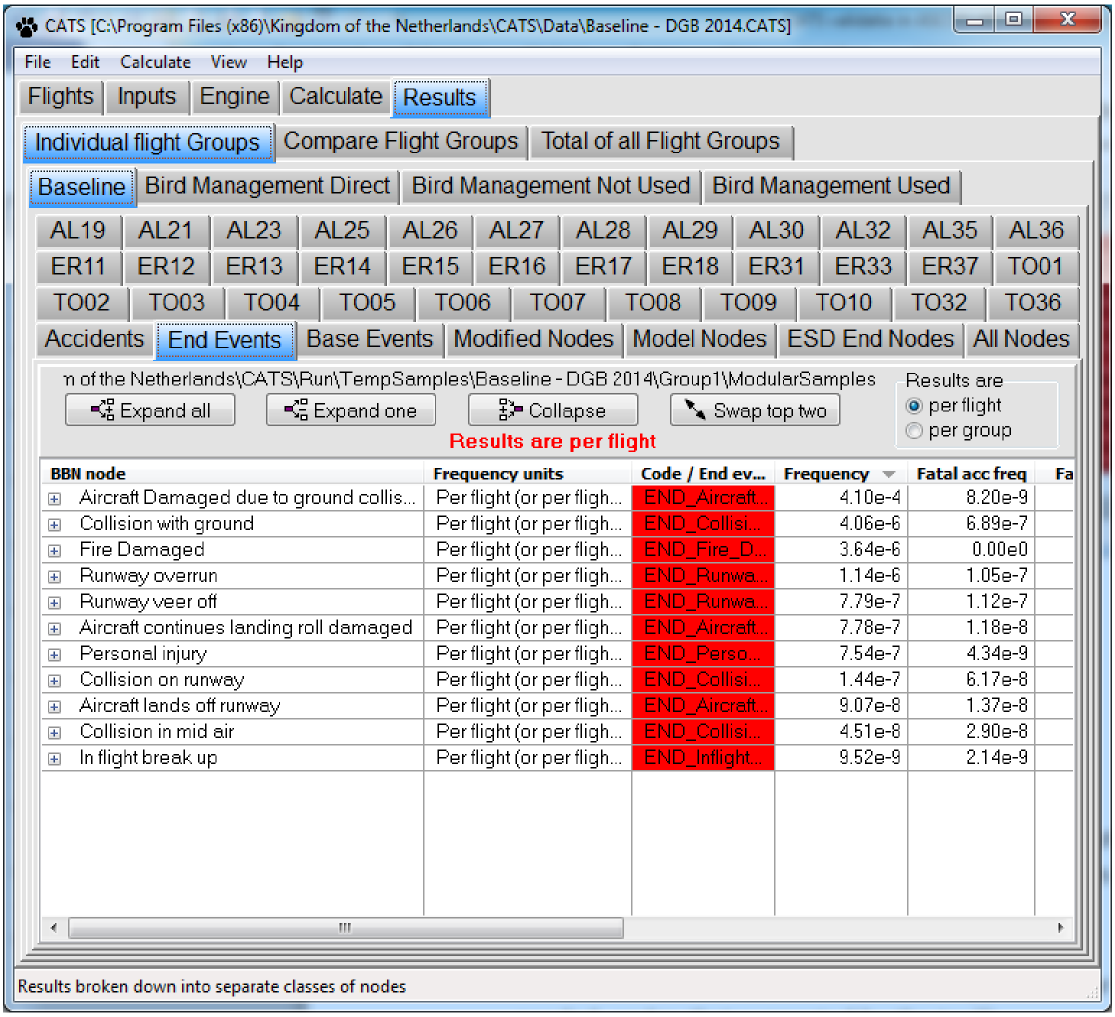


Figure 17 Highest fatal accident frequency per ESD end node

Because we look at fatal accident frequency, sorting the results for the fatal accident frequency from high to low is sufficient to give the required information. The end state with the highest fatal accident frequency is AL19d1_01, after an unstable approach.

Highest accident frequency per ESD end event type



BBN node	Frequency units	Code / End ev...	Frequency	Fatal acc freq	Fa
Aircraft Damaged due to ground collis...	Per flight (or per fligh...	END_Aircraft...	4.10e-4	8.20e-9	
Collision with ground	Per flight (or per fligh...	END_Collisi...	4.06e-6	6.89e-7	
Fire Damaged	Per flight (or per fligh...	END_Fire_D...	3.64e-6	0.00e0	
Runway overrun	Per flight (or per fligh...	END_Runwa...	1.14e-6	1.05e-7	
Runway veer off	Per flight (or per fligh...	END_Runwa...	7.79e-7	1.12e-7	
Aircraft continues landing roll damaged	Per flight (or per fligh...	END_Aircraft...	7.78e-7	1.18e-8	
Personal injury	Per flight (or per fligh...	END_Perso...	7.54e-7	4.34e-9	
Collision on runway	Per flight (or per fligh...	END_Collisi...	1.44e-7	6.17e-8	
Aircraft lands off runway	Per flight (or per fligh...	END_Aircraft...	9.07e-8	1.37e-8	
Collision in mid air	Per flight (or per fligh...	END_Collisi...	4.51e-8	2.90e-8	
In flight break up	Per flight (or per fligh...	END_Inflight...	9.52e-9	2.14e-9	

Figure 18 Highest accident frequency per end event type

Figure 18 shows the ranking by frequency of occurrence of all possible end event types as defined in CATS (Aircraft damaged due to ground collision, collision with ground, fire damaged, etc.). maxiCATS provides an easy way to sort these event types on frequency, fatal frequency etc.

Highest frequency of accident types

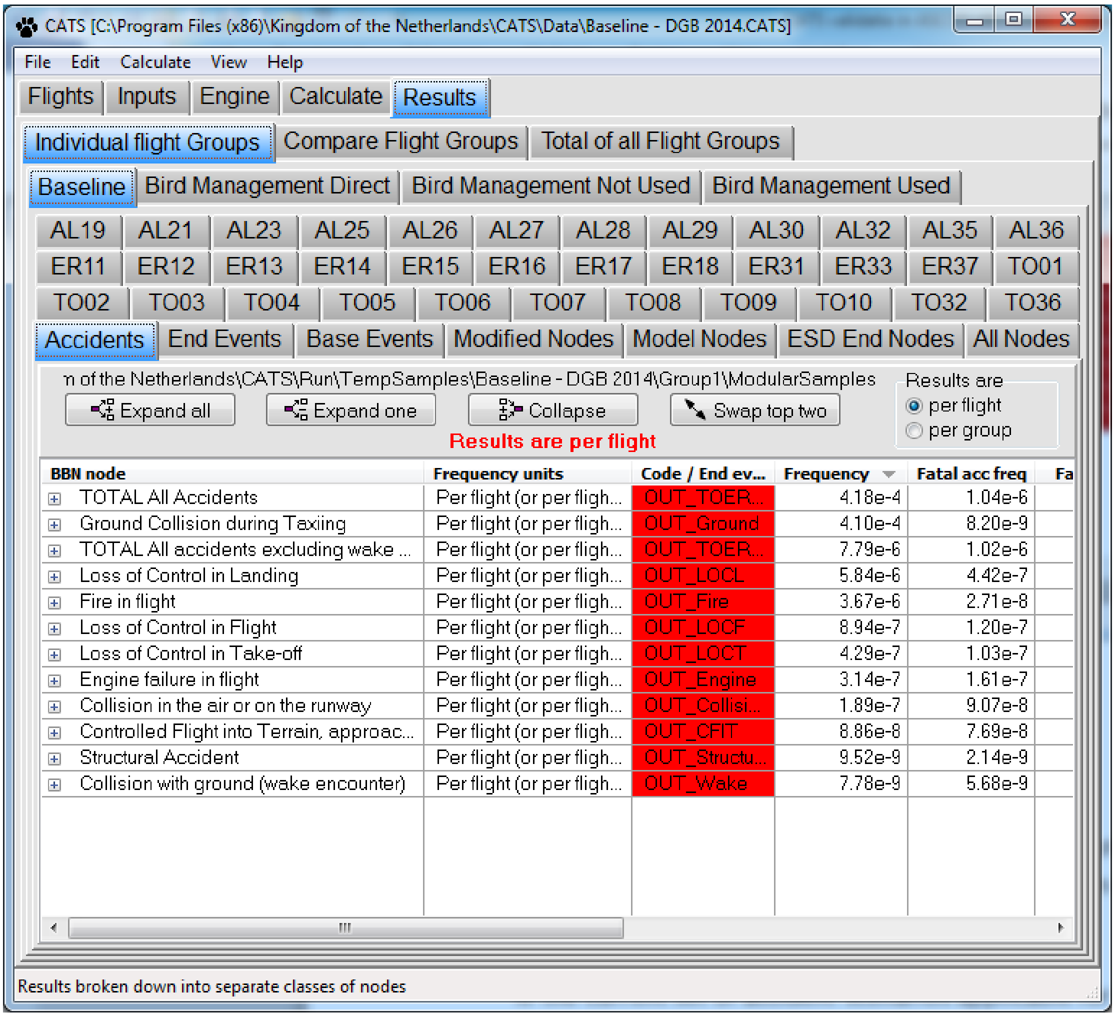


Figure 19 Highest frequency of accident types

Figure 19 shows the accident types (ground collision during taxiing, loss of control in landing, fire in flight etc.) as defined in CATS. These types are slightly different way of grouping accidents from the end event grouping of Figure 18. It also shows the total accident frequency: 4.18e-04 per flight. Of this frequency, a large part is caused by accidents on the ground during the taxi flight phase: 4.10e-04 per flight.

Findings

- It is easy to present the most important, most frequent accidents and the probabilities of these accidents with maxiCATS. maxiCATS provides various options of generation accident risk results and also provides a simple way of sorting the risk results in an increasing or decreasing order. This can be done for each of the output columns.
- One could imagine also other types of grouping of results, such as accident risk per ESD (summing the accident end states of one ESD), or the frequencies of accident end states only, that could be of

interest to the user. These are not available yet in maxiCATS, but could easily be generated by means of post processing with for example Microsoft Excel.

Appendix A.4.2 Case study 2: Identify precursors to accidents

Question

The question is to identify precursors to accidents, i.e. those events that are most important in generating the largest accident risk in the total aviation system.

Approach

The approach is as follows:

- Identify the ESDs with the largest accident frequencies;
- Find the base events with the largest contribution to the accident frequencies.

Results

From the results of the previous appendix, we can identify the ESDs with the highest accident frequencies (Figure 16). It appears that ESD 36, ESD 19 and ESD 33 are the top 3 ESDs (when looking at individual end state probabilities). When looking at the fatal accident frequencies (Figure 17), it appears that ESD 19 is at the top of the list. Therefore, for this case study we focus on the precursors to ESD 19.

The next step is to find the base event to ESD 19. For this we can use two "Tabs" in the Results of maxiCATS: the Tab "ESD19" (*Figure 20Figure 20 Events (nodes) of ESD 19 (AL19)*) and the Tab "Base events" (*Figure 21*).

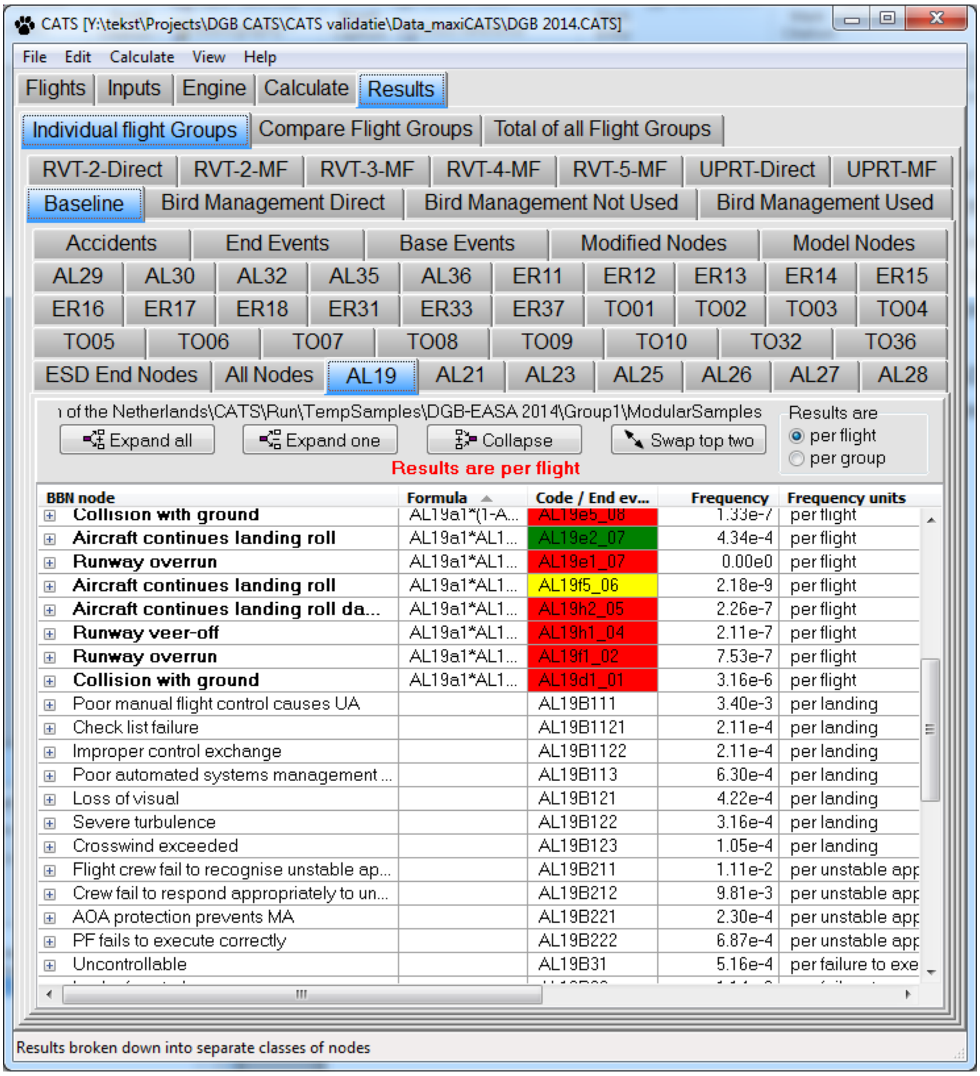


Figure 20 Events (nodes) of ESD 19 (AL19)

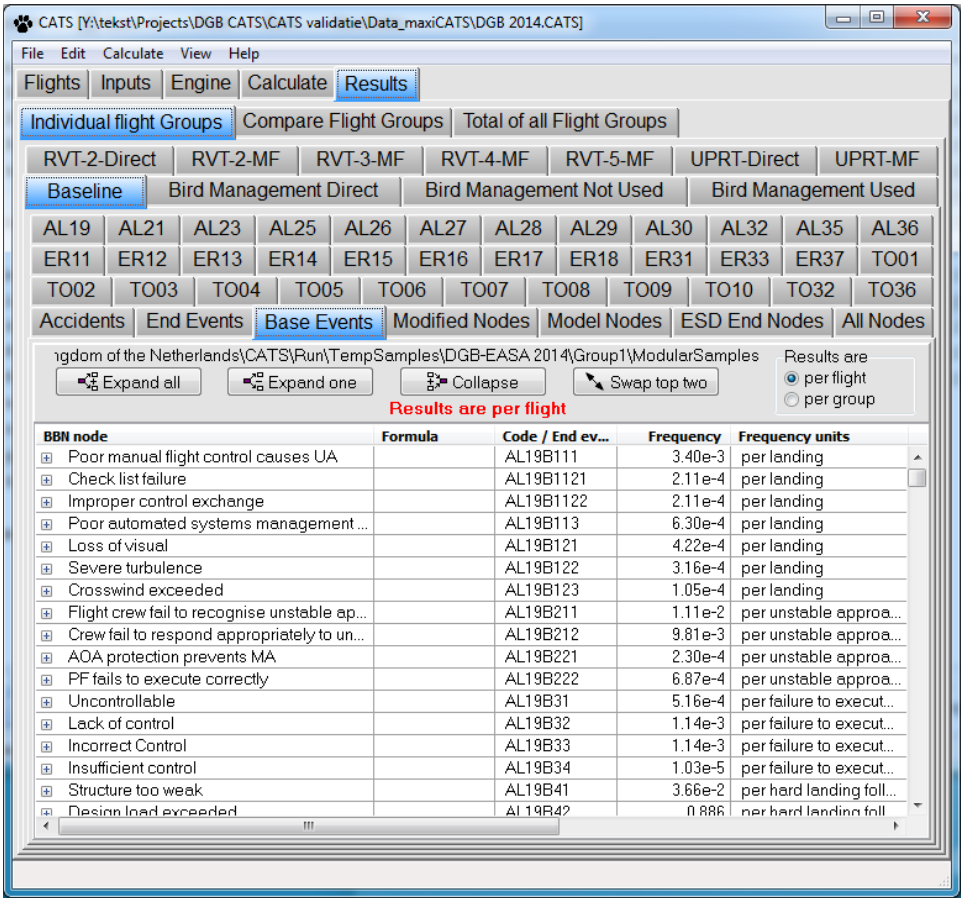


Figure 21 Base events of ESD 19 (selection)

Both Tabs do not give an easy overview of the required information: they only list base events, sorted by frequency. Moreover, the results do not present the contribution of the base event frequency to the accident risk of ESD 19, as is presented in the DNV spreadsheet [34]. This information could even be more relevant for identification of precursors than their individual frequency of occurrence. A high base event frequency of occurrence does not necessarily mean that its contribution to the accident probability is high because it might be multiplied for example by another very small frequency in an AND gate. The following tables show the base events, sorted by "contribution" (Table 16) and by "frequency" (Table 17). These tables are generated by exporting maxiCATS results to Microsoft Excel. The contributions in Table 16 are taken from the DNV spreadsheet.

Table 16 Top ten ESD 19 base events, sorted on their contribution to ESD 19 accident risk

BBN node	Frequency units	End event	Contribution	Frequency
Poor manual flight control causes UA	per landing	AL19B111	0.64	3.39E-03
Flight crew fail to recognise	per unstable approach	AL19B211	0.469	1.11E-02

unstable approach				
Crew fail to respond appropriately to unstable approach	per unstable approach	AL19B212	0.408	9.80E-03
Lack of control	per failure to execute missed approach	AL19B32	0.192	1.14E-03
Incorrect Control	per failure to execute missed approach	AL19B33	0.192	1.14E-03
Design load exceeded	per hard landing following unstable approach	AL19B42	0.145	0.886
Poor automated systems management causes UA	per landing	AL19B113	0.12	6.35E-04
PF fails to execute correctly	per unstable approach	AL19B222	0.093	6.90E-04
Uncontrollable	per failure to execute missed approach	AL19B31	0.087	5.16E-04
Loss of visual	per landing	AL19B121	0.08	4.22E-04

Table 17 Top ten ESD 19 base events, sorted on their frequency

BBN node	Frequency units	End event	Contribution	Frequency
Design load exceeded	per hard landing following unstable approach	AL19B42	0.145	0.886
Uncontrollable	per structural failure after hard landing	AL19B51	0.053	0.356
Structure too weak	per hard landing following unstable approach	AL19B41	0.006	3.58E-02
Flight crew fail to recognize unstable approach	per unstable approach	AL19B211	0.469	1.11E-02
Flight crew fail to notify ATC of inadequate reserves	per Inadequate fuel reserves	AL19B811	0.00111	1.02E-02
Crew fail to respond appropriately to unstable approach	per unstable approach	AL19B212	0.408	9.80E-03

Lack of control	per structural failure after hard landing	AL19B52	0.00075	5.17E-03
Incorrect Control	per structural failure after hard landing	AL19B53	0.000754	5.06E-03
Insufficient control	per structural failure after hard landing	AL19B54	0.000758	4.95E-03
Poor manual flight control causes UA	per landing	AL19B111	0.64	3.39E-03

Note that we focused in this case on precursors for one ESD (ESD 19) only. One could also consider looking for events that have a major impact on more than one ESD, therefore becoming more important to identify. This options has not been performed in this study.

Findings

- maxiCATS provides the possibility to identify precursors to accidents by presenting overviews and rankings of results of all events of individual ESDs and of all base events in the model;
- The information on the contribution of an event to the accident risk, as presented in the DNV fault trees [34] is not available in maxiCATS;
- Depending on the way a “precursor” is defined, some post processing (e.g. by means of Microsoft Excel) might be required to derive the required results.

Appendix A.4.3 Case study 3: Bird strikes

The Policy Agenda [37] provides information on the issue of bird strikes in the Netherlands. The number of geese has increased in recent years and the risk of bird strikes with fatal consequences is considered high. It is the ambition for 2011-2015 is to reduce the number of bird strikes, especially those that have the highest risk. Actions to achieve this ambition are e.g. the continuation of the deterrence and hunt for geese, no new wetlands in the airport area, introduction of radar. Indicators (and associated targets) for 2011-2015 are:

- % in airport area (25% less)
- % foraging ranges in airport area (50% less)
- % population reduction (50% reduction)
- X number of changes to air traffic after radar warnings (50x)

The quite generic question “How do we deal with bird strikes” is made more specific by DGB. They asked the following questions that CATS should be able to address:

- What is the risk of bird strikes?
- Is the risk of bird strikes different in the Netherlands than in Europe/World?
- Are bird strikes important when compared to other themes?

- What is the effect in terms of changed accident risk of the proposed mitigating measures?

To be able to answer these questions, we first have to answer the following questions:

- Which ESD(s) deal with bird strikes?
- Which base events are related to bird strikes and what is their (estimated) frequency?

These questions are elaborated next.

Identify relevant ESDs and events

Question

Which event sequence diagram(s) and events are related to bird strikes?

Approach

Search for "bird" in the software tool(s)

Results

There is no unique ESD that deals with bird strikes. In CATS, a bird strike only is considered when it leads to problems, e.g. an engine failure or a landing gear failure. Therefore, the CATS model has to be searched for relevant ESDs

Searching maxiCATS for the word "bird" (using the help function) leads to a hit in the topic "List of inputs". It appears that maxiCATS contains a modifying factor, an influence that the user case use to change the frequency of base events. The influence is called "Bird Management" (EM_Bird_mgt) and contains the cases Bird Management Used and Not Used. It is not clarified what these cases mean in practice.

Searching CATSPAWS for the word "bird" is not possible. Searching the DNV spreadsheet for the word "bird" gives the following results:

- Event TO09B14 Foreign Object Damage (Engine ingests objects such as debris left on the runway by other aircraft or it suffers a bird strike)
- Event TO10B1314 Foreign Object Damage (A foreign object strikes one of the control surfaces rendering it ineffective. Such objects include birds and runway debris)
- Event ER14F41232 Fuselage failure due to bird strike (Fuselage fails as a result of bird strike)
- Event ER18B131 Foreign Object Damage (The engine mount is struck by a foreign object. This can include but is not exclusive to birds and detachment of objects from the aircraft)
- Event AL27B113 Foreign object damage to landing gear (A foreign object strikes and damages the landing gear, including debris on runway and birds)
- Event ER33B112111 Bird strike (Aircraft suffers damage to pressure boundary due to bird strike)

This means that the following ESDs appear to be relevant:

- ESD09: Single engine failure during take-off

- ESD10: Pitch control problem during take-off
- ESD14: Flight crew incapacitation
- ESD18: Single engine failure during flight
- ESD27: Aircraft direction control related systems failure
- ESD33: Cracks in aircraft pressure cabin

Of these ESDs, only ESD14 and ESD33 deal explicitly with bird strikes. In the other ESDs, birds are considered as Foreign Object Damage and the model does not show the magnitude of the bird contribution to FOD.

Findings

- In maxiCATS and CATSPAWS, a function is missing for searching events on (part of) a certain name;
- How do we know if we have found all the relevant events and ESDs?
- Sometimes, the name of an event does not sufficiently clarify the situations that are actually covered by the event. For example, birds are considered as FOD and somehow the user must know this, otherwise events could be missed. The exact definition of an event is given in CATSPAWS. The user must be aware of this and should be supported by the toolset to find the right base events.

Risk of bird strikes

Questions

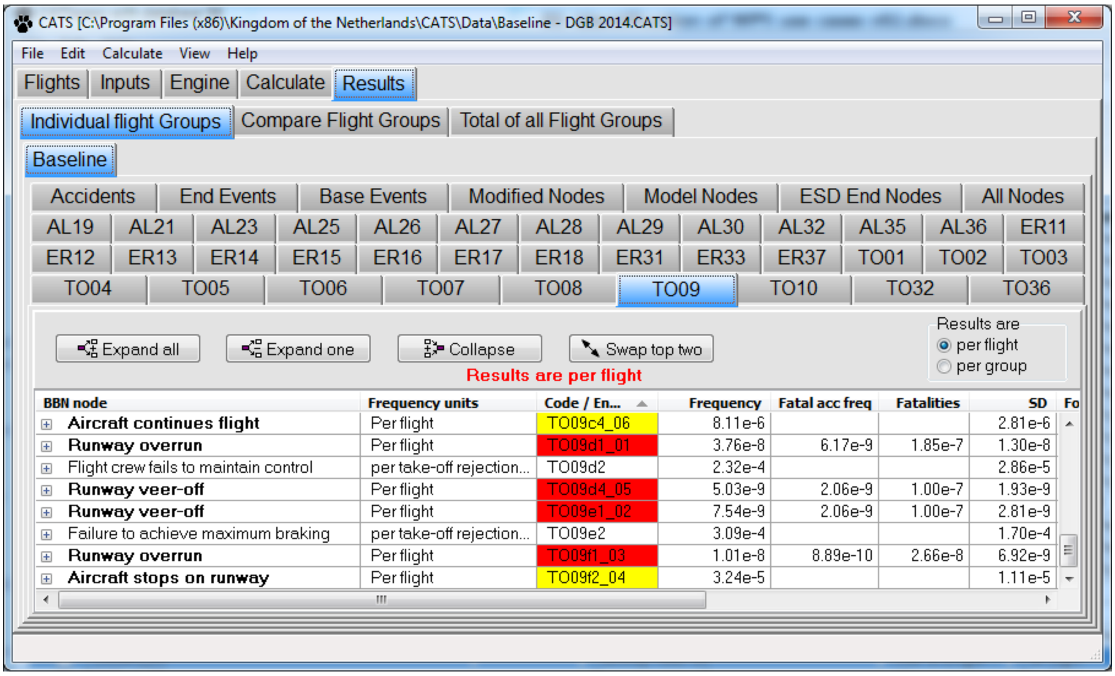
- What is the risk of bird strikes?
- Is the risk of bird strikes different in the Netherlands than in Europe/World?

Approach

Note that CATS only considers bird strikes when this actually leads to an initiating event of one of the identified ESDs. Bird strikes that result in a lower severity, that do not have an impact on the progress of the flight are not accounted for. So if we focus on accident risk, the approach is to list the accident probabilities of the above-mentioned ESDs and the probabilities of the related base events. These probabilities represent a world-wide average. For the comparison Netherlands – Europe/World we need additional data, which is outside the scope of this project.

Results

After calculating the baseline scenario with maxiCATS, the results for all events are available. The following figure illustrates the end state probabilities for ESD09 (Single engine failure during take-off):



The results for all relevant events are:

Event	Title	Frequency per flight	End state	maxiCATS results	DNV results
TO09B14	FOD	1.35E-05	TO09d1_01 TO09e1_02 TO09f1_03 TO09d4_05	3.76E-08 7.54E-09 1.01E-08 5.03e-09	3.76E-08 7.52E-09 1.00E-08 5.01e-09
TO10B1314	FOD	1.79E-06	TO10d1_01 TO10f1_03 TO10d4_05	1.13E-08 6.48E-09 3.03E-09	1.11E-08 4.42E-09 2.21E-09
ER14F41232	Fuselage failure due to bird strike	2.0E-07	ER14c1_01	1.04E-08	9.33E-09
ER18B131	FOD	2.07E-09	ER18d1_01 ER18e1_02 ER18e5_07	2.71E-08 7.28E-08 2.14E-07	2.01E-08 6.77E-08 1.73E-07

AL27B113	FOD to landing gear	2.70E-07	AL27c1_01	1.11E-07	1.10E-07
ER33B112111	Bird strike	1.00E-05	ER33c1_01	9.52E-09	9.13E-09

From these figures we should be able to determine the risk of a bird strike. Summing all end state probabilities gives an accident frequency that includes, but is not limited to bird strikes. Determining the portion of accidents that is caused by a bird strike can be achieved by looking at the base events in relation to base events that are not related to bird strikes. Also, data analysis would be necessary because it is not clear what part of the FOD-related base events relates to birds. This however is outside the scope of this project.

Findings

- maxiCATS does not provide a functionality that is able to determine the risk of a bird strike, without (substantial) post processing.

Bird strikes compared to other themes

Question

Are bird strikes important when compared to other themes?

Approach

First define "other themes". At least we could list the risks of all ESDs, which in a way also are "themes". Per ESD we have to sum the probabilities of accident end states. For example, the risk of an Unstable Approach (ESD 19) is:

End state	**Probability**
AL19d1_01	3.27E-06
AL19f1_02	7.44E-07
AL19h1_04	2.09E-07
AL19h2_05	2.23E-07
AL19e1_07	0.00
AL19e5_08	1.26E-07
AL19f4_09	1.78E-07
total	*4.69E-06*

This is the average, baseline number. Risk of unstable approaches can also be determined for Netherlands, Europe, or for a certain Airport/Runway. In these cases, flight groups have to be defined that represent these situations, which could be a non-trivial task.

Results

In the previous sub-appendix we have seen that we could not (as part of this project) determine the risk of a bird strike. Therefore, it is also not possible to make the comparison with other themes.

Findings

- maxiCATS does not provide a functionality to determine the risk of an ESD, e.g. the risk of an Unstable Approach. This functionality should add all probabilities of ESD end states that are accidents. This means that some post processing has to be done by the user.

Effect of mitigating measures

Question

What will be the effect in terms of accident/incident risk in case targets for the indicators are met?

Approach

We can use two approaches:

- Use the available Influence “Bird Management” to see the effect of the measures (two options: “used” and “not used”);
- Directly change the base event frequencies (user override).

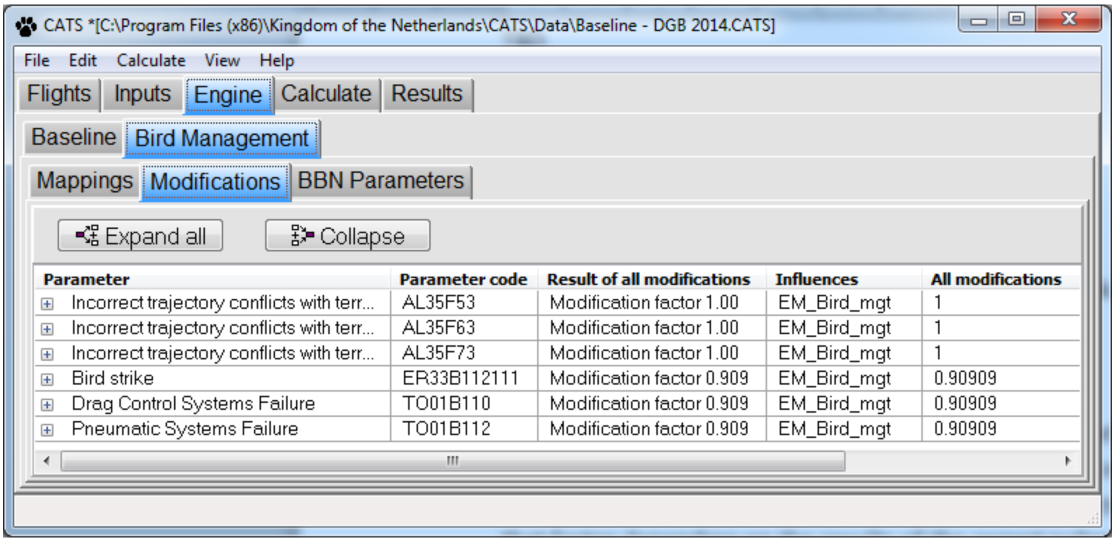
This can be achieved by defining three new flight groups in maxiCATS, for each of which modifications or influences have to be defined:

- Bird Management Used
- Bird Management Not Used
- Bird Management Direct

Note that these modifications, which link to a number of base events, should reflect the mitigating measures and actions that are proposed by DGB in their Policy Agenda [37]., The indicators (and associated targets) are:

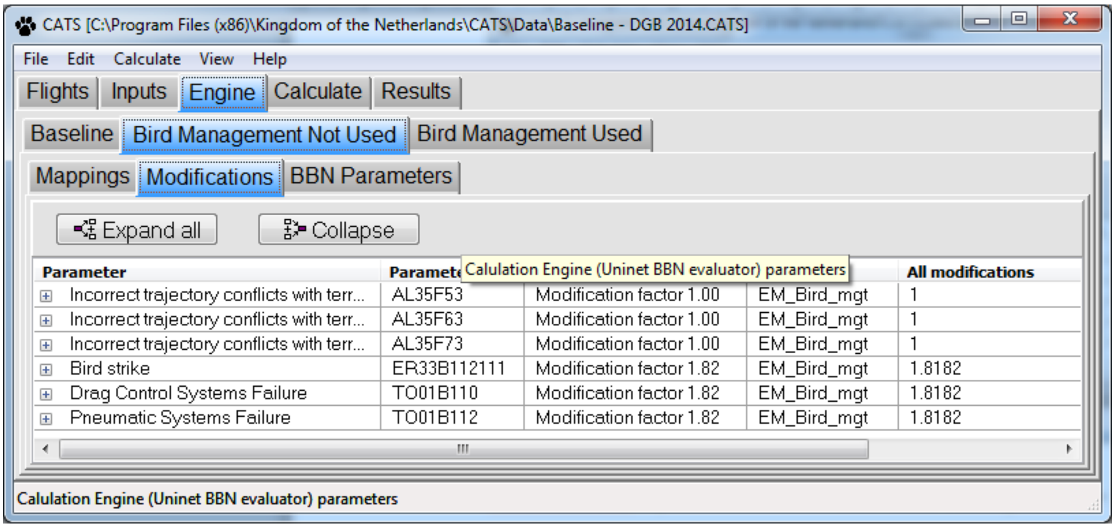
- % in airport area (25% less)
- % foraging ranges in airport area (50% less)
- % population reduction (50% reduction)
- X number of changes to air traffic after radar warnings (50x)

Translating these targets into modification factors could be a difficult task. For this exercise, modification factors are assumed. These factors represent a reduction in the frequencies of relevant base events, The baseline version of maxiCATS shows the following modifications in case of the Influence “Bird Management Used”:



Parameter	Parameter code	Result of all modifications	Influences	All modifications
Incorrect trajectory conflicts with terr...	AL35F53	Modification factor 1.00	EM_Bird_mgt	1
Incorrect trajectory conflicts with terr...	AL35F63	Modification factor 1.00	EM_Bird_mgt	1
Incorrect trajectory conflicts with terr...	AL35F73	Modification factor 1.00	EM_Bird_mgt	1
Bird strike	ER33B112111	Modification factor 0.909	EM_Bird_mgt	0.90909
Drag Control Systems Failure	TO01B110	Modification factor 0.909	EM_Bird_mgt	0.90909
Pneumatic Systems Failure	TO01B112	Modification factor 0.909	EM_Bird_mgt	0.90909

Modifications in case of the Influence “Bird Management – Not Used” are:

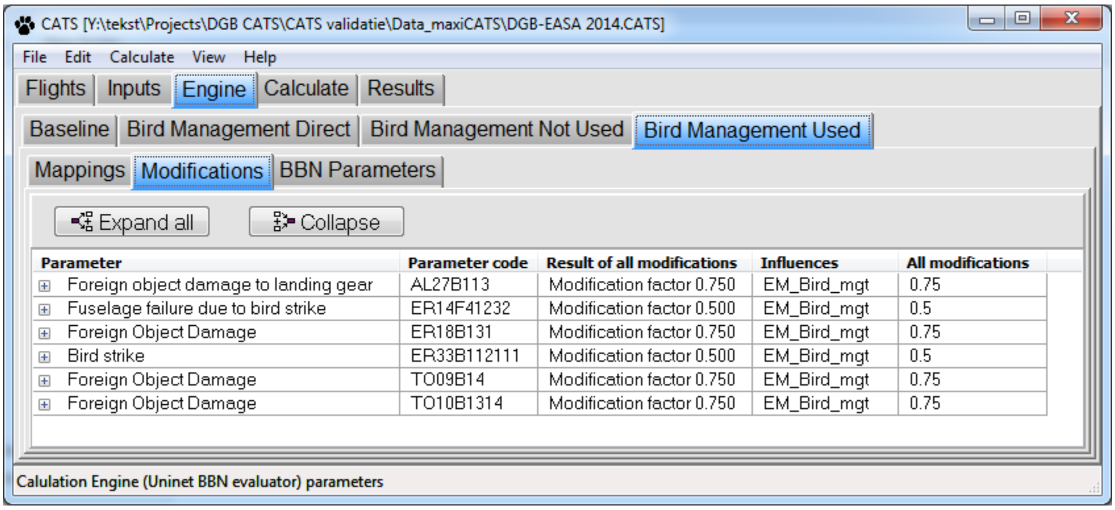


Parameter	Paramet			All modifications
Incorrect trajectory conflicts with terr...	AL35F53	Modification factor 1.00	EM_Bird_mgt	1
Incorrect trajectory conflicts with terr...	AL35F63	Modification factor 1.00	EM_Bird_mgt	1
Incorrect trajectory conflicts with terr...	AL35F73	Modification factor 1.00	EM_Bird_mgt	1
Bird strike	ER33B112111	Modification factor 1.82	EM_Bird_mgt	1.8182
Drag Control Systems Failure	TO01B110	Modification factor 1.82	EM_Bird_mgt	1.8182
Pneumatic Systems Failure	TO01B112	Modification factor 1.82	EM_Bird_mgt	1.8182

As one can see, the modifying factors are linked to ESD01, ESD33 and ESD35. This is different from those identified previously by searching for the word “bird” in base events, viz. ESD09, ESD10, ESD14, ESD18, ESD 27 and ESD33. It is not clear why maxiCATS proposes a different set.

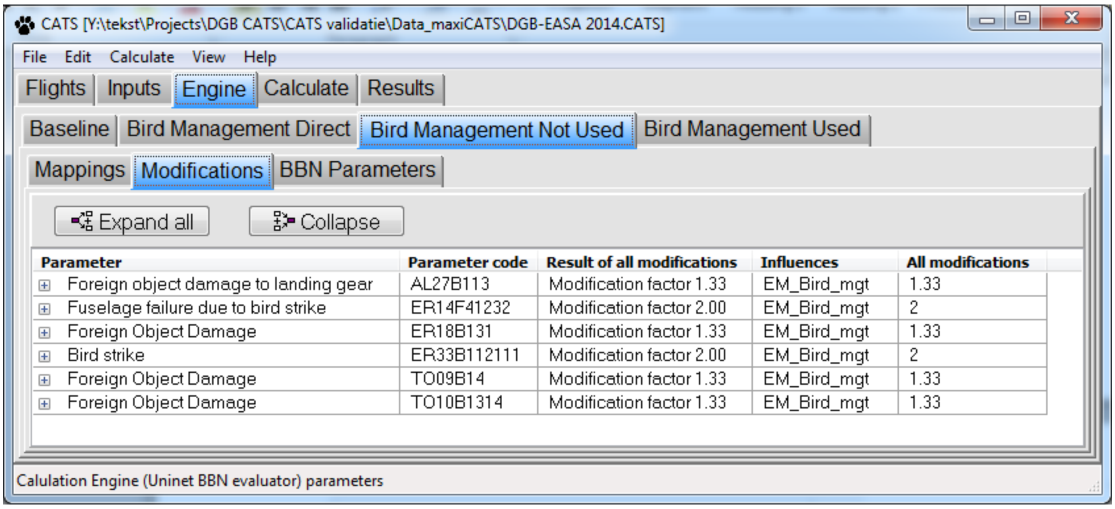
Therefore, the influence as defined in the baseline toolset was redefined to reflect the expectations. For this, the CATSPAWS database needed to be changed by means of Microsoft Access. This resulted in the following modifications (note that the magnitudes of the modifications are not expert based, they are just for illustration purposes):

Modifications in case of the Influence “Bird Management – Used”:



Parameter	Parameter code	Result of all modifications	Influences	All modifications
Foreign object damage to landing gear	AL27B113	Modification factor 0.750	EM_Bird_mgt	0.75
Fuselage failure due to bird strike	ER14F41232	Modification factor 0.500	EM_Bird_mgt	0.5
Foreign Object Damage	ER18B131	Modification factor 0.750	EM_Bird_mgt	0.75
Bird strike	ER33B112111	Modification factor 0.500	EM_Bird_mgt	0.5
Foreign Object Damage	TO09B14	Modification factor 0.750	EM_Bird_mgt	0.75
Foreign Object Damage	TO10B1314	Modification factor 0.750	EM_Bird_mgt	0.75

Modifications in case of the Influence “Bird Management – Not Used”:



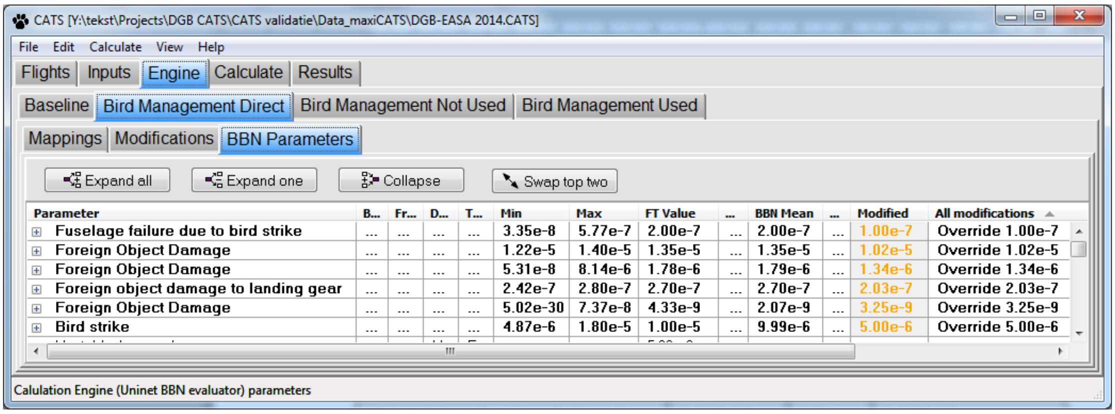
Parameter	Parameter code	Result of all modifications	Influences	All modifications
Foreign object damage to landing gear	AL27B113	Modification factor 1.33	EM_Bird_mgt	1.33
Fuselage failure due to bird strike	ER14F41232	Modification factor 2.00	EM_Bird_mgt	2
Foreign Object Damage	ER18B131	Modification factor 1.33	EM_Bird_mgt	1.33
Bird strike	ER33B112111	Modification factor 2.00	EM_Bird_mgt	2
Foreign Object Damage	TO09B14	Modification factor 1.33	EM_Bird_mgt	1.33
Foreign Object Damage	TO10B1314	Modification factor 1.33	EM_Bird_mgt	1.33

The user can also override base event frequencies directly. Using the same figures as the ones for the modifications of Bird Management Used, we have:

Event	Baseline frequency	Factor	Assumed frequency after mitigation
TO09B14 FOD	1.35E-05	0.75	1.01E-05
TO10B1314 FOD	1.78E-06	0.75	1.34E-06
ER14F41232 Fuselage failure due to bird strike	2.00E-07	0.5	1.00E-07
ER18B131 FOD	4.33E-09	0.75	3.25E-09
AL27B113 FOD to landing gear	2.70E-07	0.75	2.03E-07

ER33B112111 Bird strike	1.00E-05	0.5	5.00E-06

which is shown in maxiCATS as



Note that some of the modified values do not fall between the shown Min and Max values for the base event. In that case, it appears that CATS uses the Min or the Max value which is closest to the modified value! These Min and Max values are the Lower and Upper end of the distributions for these parameters which are used by the BBN. These distributions can be changed by means of the Uninet tool, resulting in different Min and Max settings.

Results

Results after running the tool (using a setting of 32500 samples) for the four flight groups (the baseline and the three new ones) are presented in the following figures, showing not the individual absolute results but the results relative to the baseline. Only a subset of results is shown.

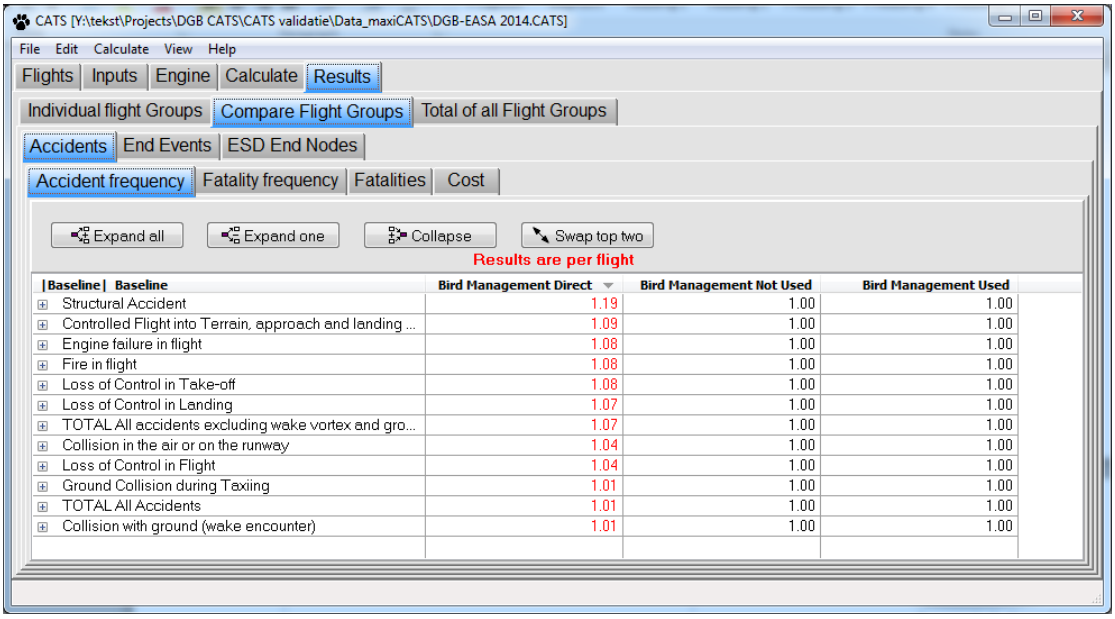
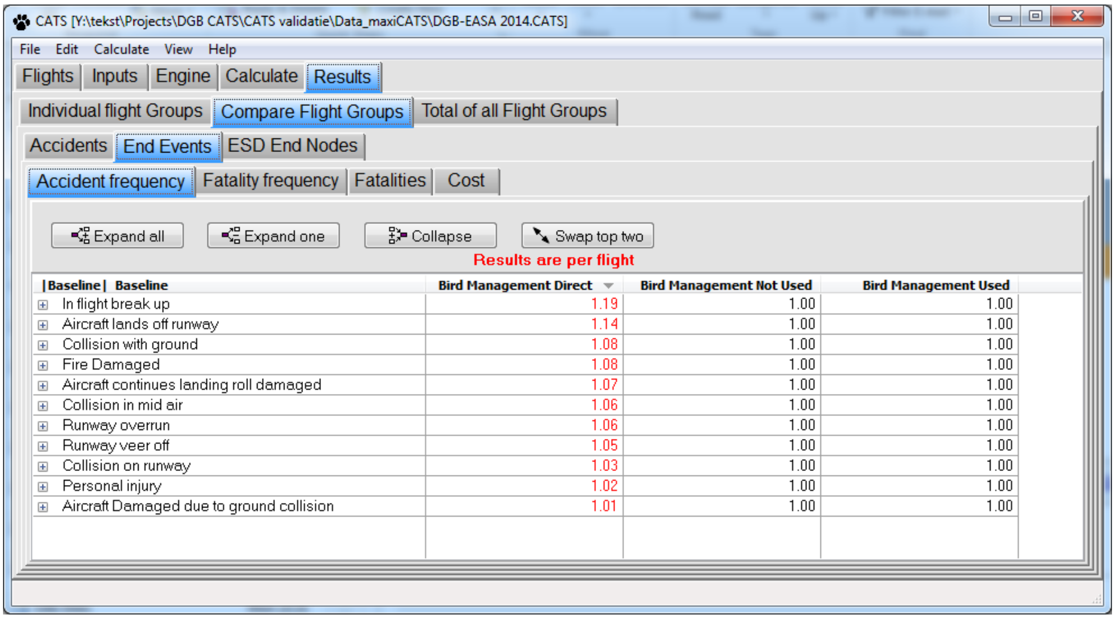
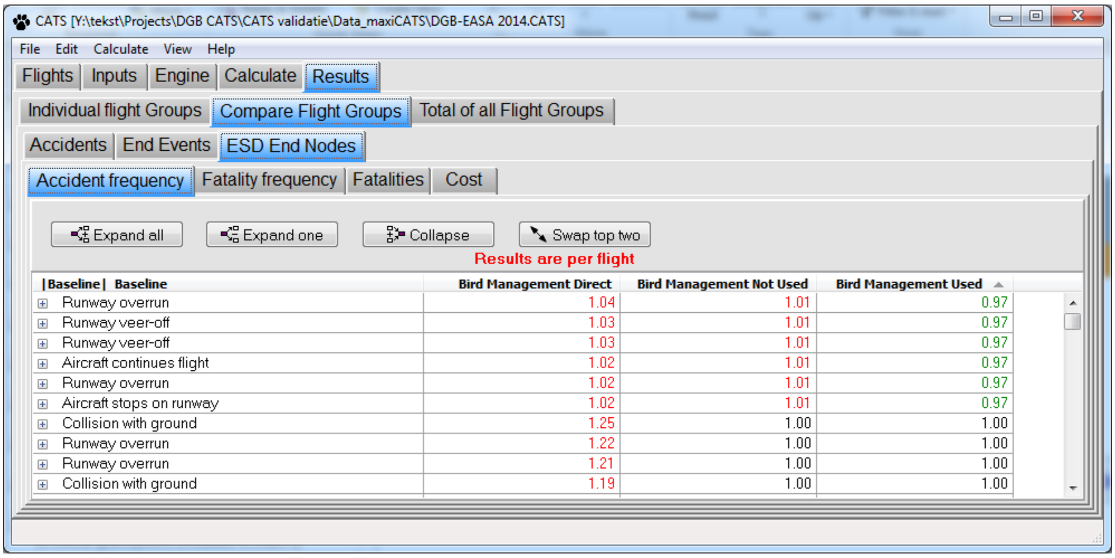


Figure 22 Relative changes in accident frequency for accident types wrt the baseline



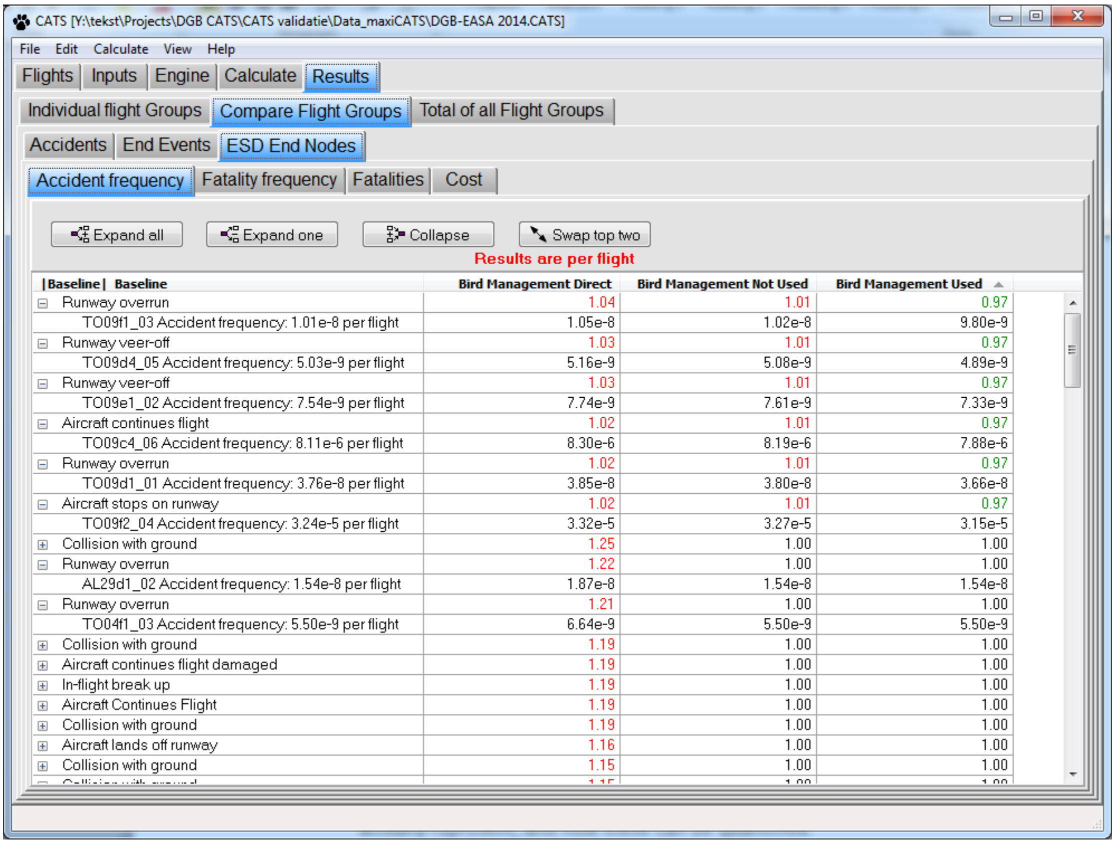
|Baseline| Baseline	Bird Management Direct	Bird Management Not Used	Bird Management Used
In flight break up	1.19	1.00	1.00
Aircraft lands off runway	1.14	1.00	1.00
Collision with ground	1.08	1.00	1.00
Fire Damaged	1.08	1.00	1.00
Aircraft continues landing roll damaged	1.07	1.00	1.00
Collision in mid air	1.06	1.00	1.00
Runway overrun	1.06	1.00	1.00
Runway veer off	1.05	1.00	1.00
Collision on runway	1.03	1.00	1.00
Personal injury	1.02	1.00	1.00
Aircraft Damaged due to ground collision	1.01	1.00	1.00

Figure 23 Relative changes in accident frequency for end events wrt the baseline



|Baseline| Baseline	Bird Management Direct	Bird Management Not Used	Bird Management Used
Runway overrun	1.04	1.01	0.97
Runway veer-off	1.03	1.01	0.97
Runway veer-off	1.03	1.01	0.97
Aircraft continues flight	1.02	1.01	0.97
Runway overrun	1.02	1.01	0.97
Aircraft stops on runway	1.02	1.01	0.97
Collision with ground	1.25	1.00	1.00
Runway overrun	1.22	1.00	1.00
Runway overrun	1.21	1.00	1.00
Collision with ground	1.19	1.00	1.00

Figure 24 Relative changes in accident frequency for ESD end nodes wrt the baseline



|Baseline| Baseline	Bird Management Direct	Bird Management Not Used	Bird Management Used
Runway overrun	1.04	1.01	0.97
TO09f1_03 Accident frequency: 1.01e-8 per flight	1.05e-8	1.02e-8	9.80e-9
Runway veer-off	1.03	1.01	0.97
TO09d4_05 Accident frequency: 5.03e-9 per flight	5.16e-9	5.08e-9	4.89e-9
Runway veer-off	1.03	1.01	0.97
TO09e1_02 Accident frequency: 7.54e-9 per flight	7.74e-9	7.61e-9	7.33e-9
Aircraft continues flight	1.02	1.01	0.97
TO09c4_06 Accident frequency: 8.11e-6 per flight	8.30e-6	8.19e-6	7.88e-6
Runway overrun	1.02	1.01	0.97
TO09d1_01 Accident frequency: 3.76e-8 per flight	3.85e-8	3.80e-8	3.66e-8
Aircraft stops on runway	1.02	1.01	0.97
TO09f2_04 Accident frequency: 3.24e-5 per flight	3.32e-5	3.27e-5	3.15e-5
Collision with ground	1.25	1.00	1.00
Runway overrun	1.22	1.00	1.00
AL29d1_02 Accident frequency: 1.54e-8 per flight	1.87e-8	1.54e-8	1.54e-8
Runway overrun	1.21	1.00	1.00
TO04f1_03 Accident frequency: 5.50e-9 per flight	6.64e-9	5.50e-9	5.50e-9
Collision with ground	1.19	1.00	1.00
Aircraft continues flight damaged	1.19	1.00	1.00
In-flight break up	1.19	1.00	1.00
Aircraft Continues Flight	1.19	1.00	1.00
Collision with ground	1.19	1.00	1.00
Aircraft lands off runway	1.16	1.00	1.00
Collision with ground	1.15	1.00	1.00

Figure 25 Detailed relative changes in accident frequency for ESD end nodes wrt the baseline

Discussion of results:

- The results in *Figure 22* and *Figure 23* show that the modifications of “Bird Management Used” and “Bird Management Not Used” show no changes on the aggregate level of accident types and end event types. This is well possible because these results are sums of various end states.
- Bird Management Used reduces accident frequencies for ESD end nodes and Bird Management Not Used increases accident frequencies of ESD end nodes as can be expected. Changes are for ESD09 only (see *Figure 25*). The end states of ESD10, ESD 14, ESD18, ESD27 and ESD33 are not affected. This can be explained by the fact that for these latter 5 ESDs, the contributions of the changed base events appear to be very small such that the change does not have any effect on the end state probability. The changed base event of ESD09 does have a large contribution to end state frequencies. Therefore, this result corresponds to expectations.
- The changes to the base events of “Bird Management Direct” and “Bird Management Used” are identical, but the results are apparently different. Also, where one would expect the results to improve, this is only the case for “Bird Management Used” but not for “Bird Management Direct”. In maxiCATS, there are three possibilities for defining influences: modification of parameters, conditionalizing parameters and conditionalizing by weight. The loggings of the simulations show that “Bird Management Used” uses a modification of parameters and “Bird Management Direct” uses the

conditionalization of parameters. This could induce different results. At this point, the meaning and impact of this difference is not clear.
- Modifications of "Bird Management Direct" also affect ESDs of which base event frequencies are not modified, e.g. ESD04, ESD29.

Findings

- It is possible to change pre-defined Influences in maxiCATS (although this had to be done directly in the CATSPAWS database due to a non-functioning button in maxiCATS). This is essential because users can define mitigating measures or operational changes themselves and define the potential influences on base event frequencies, including an estimated magnitude.
- The functionality of comparing flight groups is very useful. The user can easily check the changed end states after modification. Also, it is easy to change the order in which the results are shown (lowest to highest values, or highest to lowest values, for every output column);
- Using a "user override" function, base event frequencies can be set by the user. It appears that the new values should lie in a predefined interval (Min and Max values). This means that the user is not completely free to make choices for base event frequencies (e.g. extreme values of 0 or 1), without making changes to the probability density functions of the base events.
- The results achieved by the "user override" function are different from the results by using the predefined Influences. More explanation by the model and software developers is required to understand this issue.
- The baseline version of the CATS tool has defined the influence Bird Management differently from what could be expected when searching the model for events related to birds.

Appendix A.4.4 Case study 4: Remote Virtual Tower

This case study is described in [36] and is one of the validation exercises in the ASCOS project. It assesses the safety effect of the introduction of a remote tower operation in the total aviation system. A remote tower is the provision of Air Traffic Services at one or more airports from a control facility that is not located in the local ATS tower.

To ease the comparison of the results of two different versions of CATS, the same setup of the validation exercise is chosen. This means that the scope of the safety assessment is limited to ESD 32 in line with the ASCOS validation scenario even though more than one ESD may be affected by the change. Note that the validation in the context of this project determines the fitness for purpose of the CATS toolset. It is not the purpose to compare the numerical outcomes of the case study in the CATS and ASCOS tools. These two tools are not the same, so different results can be expected. Three basic differences are the a) quantification baseline (worldwide data is used in CATS for the quantification of scenarios as opposed to EU data in ASCOS), b) BBN including ESD dependencies and human performance models (CATS) versus ESDs and FTs (ASCOS) and c) the adaptations in risk model in ASCOS compared to CATS.)

Wever [36] describes 6 runs, the first 5 of which are:

- Run 1: Baseline risk picture
- Run 2: Baseline risk picture given the occurrence of a RIMCAS failure
- Run 3: Future risk picture in Failure Case (the ATCO does not see a visible conflict in time)
- Run 4: Future risk picture in Success Case (low visibility or darkness prevents conflict detection)
- Run 5: Future risk picture in Failure and Success Case (combination of Run 3 and Run 4)

Question

Determine the risk pictures of Run 2 – Run 5, as described above.

Approach

As for the bird strike case study, two approaches can be used to implement modification factors:

- Implement modifications in the CATSPAWS database
- User override of base event frequencies

The latter method only is applied to Run 2.

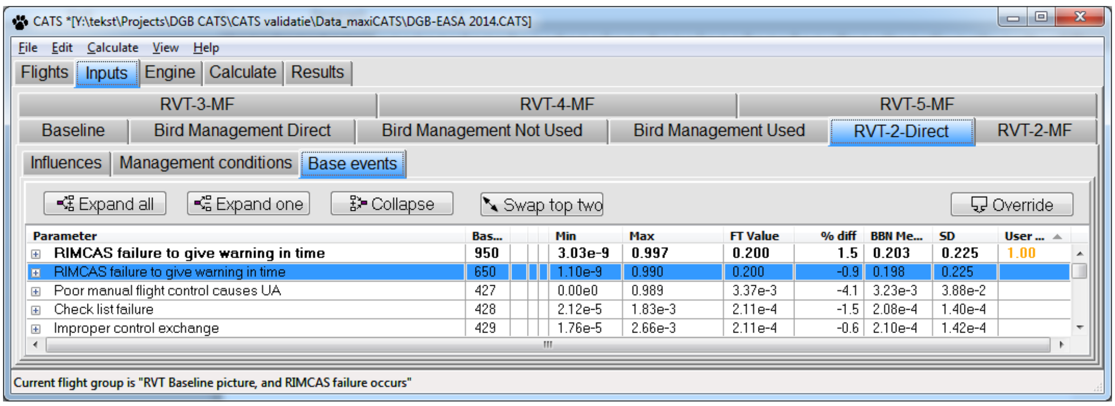
The following flight groups are defined:

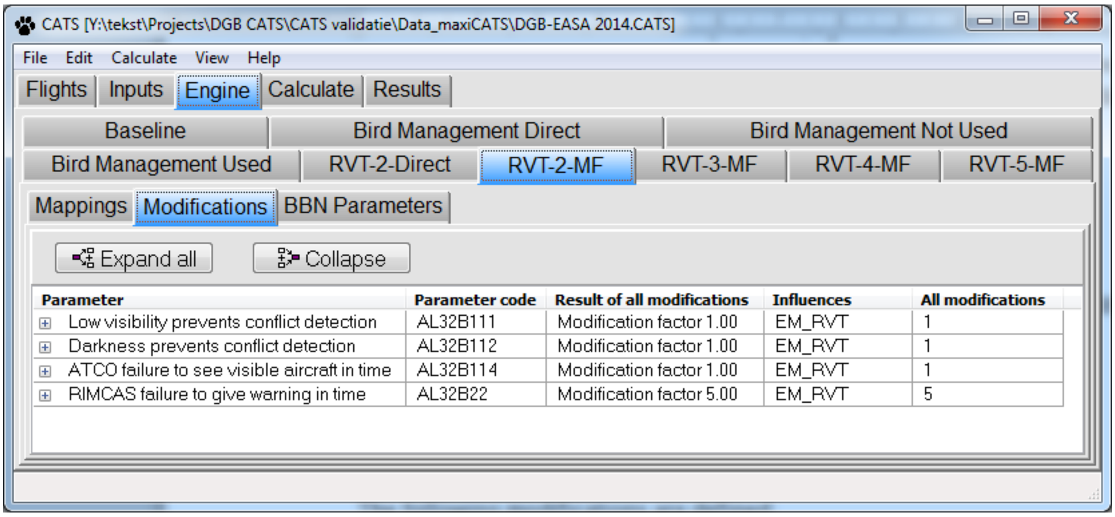
- RVT-2-MF – Modifications for Run 2 (RIMCAS failure)
- RVT-2-Direct – User override of base event frequencies for Run 2 (RIMCAS failure)
- RVT-3-MF – Modifications for Run 3 (Failure Case - the ATCO does not see a visible conflict in time)
- RVT-4-MF – Modifications for Run 4 (Success Case - low visibility or darkness prevents conflict detection)
- RVT-5-MF – Modifications for Run 5 (Failure Case and Success case)

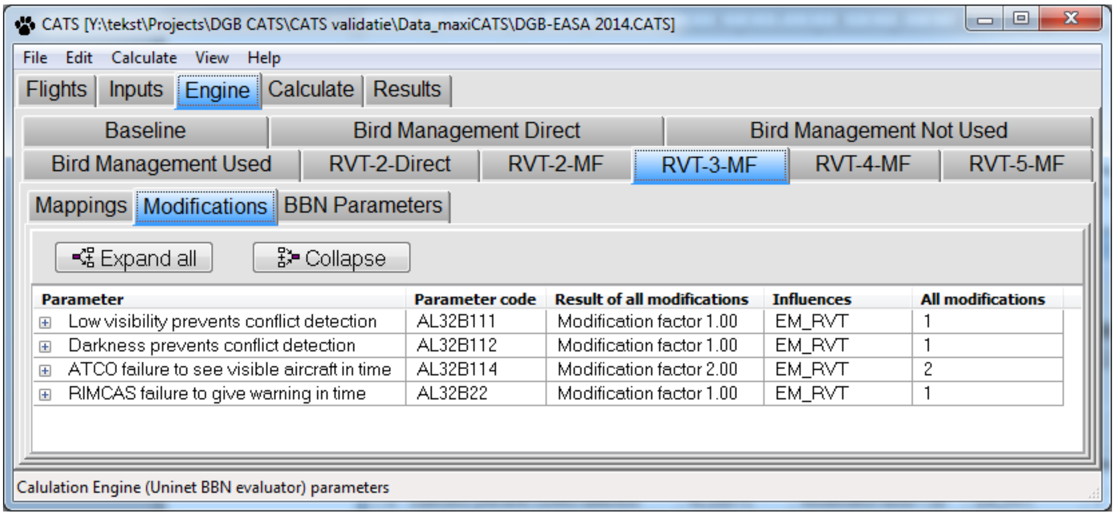
Following [36], the following changes are required:

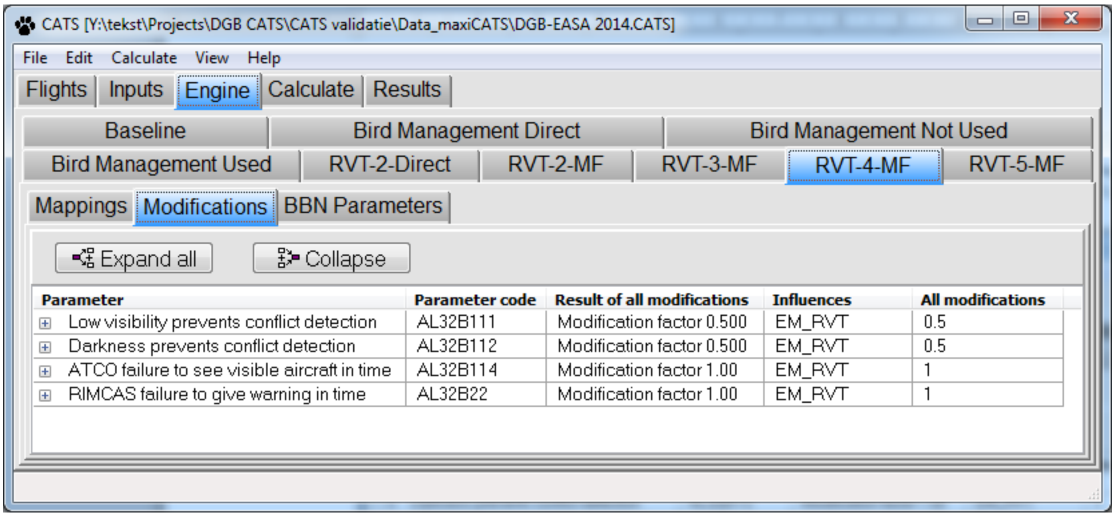
- Event ASC32b112 (AL32B22 in CATS 2008) frequency becomes 1 (Run 2);
- Event ASC32b123 (AL32B114 in CATS 2008) frequency becomes 0.2 (Run 3 and Run 5);
- Event ASC32b121 (AL32B111 and AL32B112 in CATS 2008) frequency becomes 0.4 (Run 4 and Run 5);

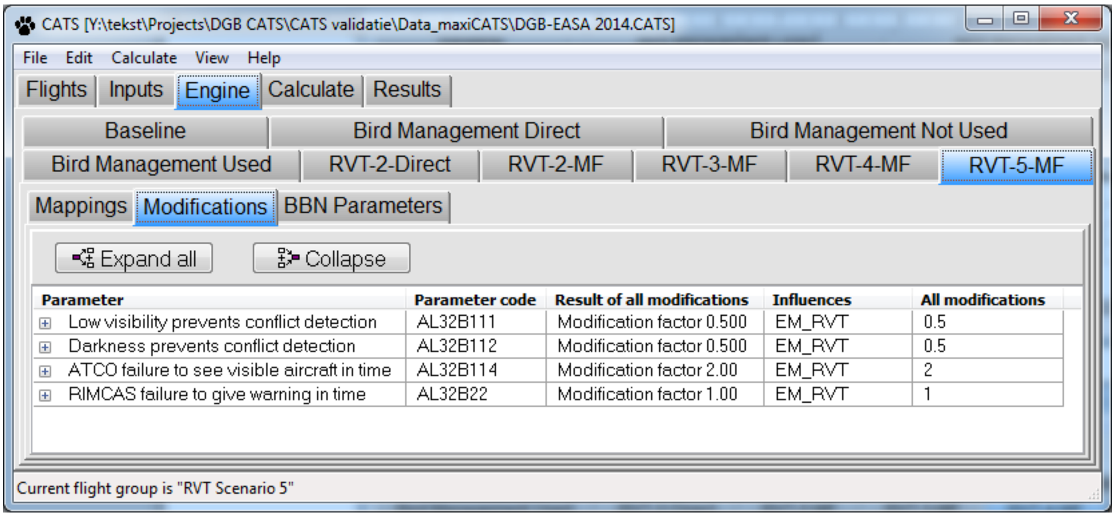
This leads to the following probabilities (RVT-2-Direct) and modification factors per Run (RVT-2-MF etc.):




CATS [Y:\tekst\Projects\DGB CATS\CATS validatie\Data_maxiCATS\DGB-EASA 2014.CATS]
File Edit Calculate View Help
Flights Inputs Engine Calculate Results
Baseline Bird Management Direct Bird Management Not Used
Bird Management Used RVT-2-Direct RVT-2-MF RVT-3-MF RVT-4-MF RVT-5-MF
Mappings Modifications BBN Parameters
Expand all Collapse
Parameter | Parameter code | Result of all modifications | Influences | All modifications
Low visibility prevents conflict detection | AL32B111 | Modification factor 1.00 | EM_RVT | 1
Darkness prevents conflict detection | AL32B112 | Modification factor 1.00 | EM_RVT | 1
ATCO failure to see visible aircraft in time | AL32B114 | Modification factor 1.00 | EM_RVT | 1
RIMCAS failure to give warning in time | AL32B22 | Modification factor 5.00 | EM_RVT | 5


CATS [Y:\tekst\Projects\DGB CATS\CATS validatie\Data_maxiCATS\DGB-EASA 2014.CATS]
File Edit Calculate View Help
Flights Inputs Engine Calculate Results
Baseline Bird Management Direct Bird Management Not Used
Bird Management Used RVT-2-Direct RVT-2-MF RVT-3-MF RVT-4-MF RVT-5-MF
Mappings Modifications BBN Parameters
Expand all Collapse
Parameter | Parameter code | Result of all modifications | Influences | All modifications
Low visibility prevents conflict detection | AL32B111 | Modification factor 1.00 | EM_RVT | 1
Darkness prevents conflict detection | AL32B112 | Modification factor 1.00 | EM_RVT | 1
ATCO failure to see visible aircraft in time | AL32B114 | Modification factor 2.00 | EM_RVT | 2
RIMCAS failure to give warning in time | AL32B22 | Modification factor 1.00 | EM_RVT | 1
Calulation Engine (Uninet BBN evaluator) parameters


CATS [Y:\tekst\Projects\DGB CATS\CATS validatie\Data_maxiCATS\DGB-EASA 2014.CATS]
File Edit Calculate View Help
Flights Inputs Engine Calculate Results
Baseline Bird Management Direct Bird Management Not Used
Bird Management Used RVT-2-Direct RVT-2-MF RVT-3-MF RVT-4-MF RVT-5-MF
Mappings Modifications BBN Parameters
Expand all Collapse
Parameter | Parameter code | Result of all modifications | Influences | All modifications
Low visibility prevents conflict detection | AL32B111 | Modification factor 0.500 | EM_RVT | 0.5
Darkness prevents conflict detection | AL32B112 | Modification factor 0.500 | EM_RVT | 0.5
ATCO failure to see visible aircraft in time | AL32B114 | Modification factor 1.00 | EM_RVT | 1
RIMCAS failure to give warning in time | AL32B22 | Modification factor 1.00 | EM_RVT | 1



Results

Results that were expected:

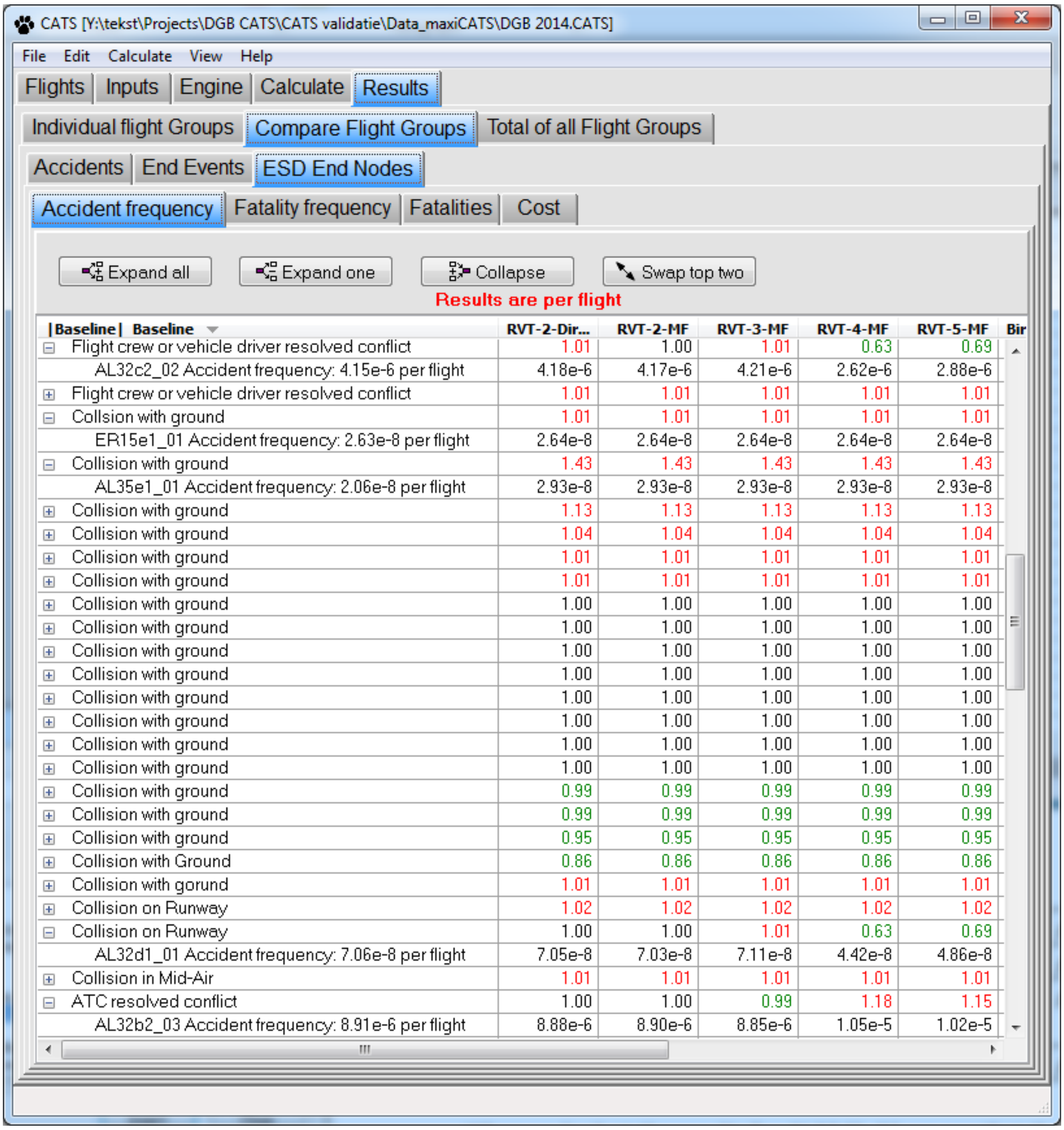
- RVT-2-Direct and RVT-2-MF should have the same results as the two approaches modify the same event (in two different ways).
- If RIMCAS fails, then the risk of ESD 32 increases (i.e. the accident end state probability increases).
- Accident frequencies of RVT-3-MF are expected to increase based on engineering judgement;
- Accident frequencies of RVT-4-MF are expected to decrease based on engineering judgement;;
- No expectations regarding accident frequencies of RVT-5-MF because the effect of the combination cannot be estimated beforehand based on engineering judgement;.

Results achieved by the ASCOS model and tool for the 5 runs:

		Baseline	Run 2		Run 3	
Event title	code	probability	probability	change	probability	change
Runway incursion	ASC32a1	1.24E-05	1.24E-05	1.00	1.24E-05	1.00
ATC does not resolve the conflict	ASC32b1	3.53E-01	3.58E-01	1.01	0.361485	1.02
Flight crew or vehicle driver does not resolve the conflict	ASC32c1	1.46E-02	1.46E-02	1.00	0.014622	1.00
Aircraft continues flight	ASC32c2	8.04E-06	7.99E-06	0.99	7.94E-06	0.99
Collision on runway	ASC32d1	6.43E-08	6.51E-08	1.01	6.57E-08	1.02
Aircraft continues flight	ASC32d2	4.33E-06	4.38E-06	1.01	4.43E-06	1.02

		Baseline	Run 4		Run 5	
Event title	**code**	**probability**	**probability**	**change**	**probability**	**change**
Runway incursion	ASC32a1	1.24E-05	1.24E-05	1.00	1.24E-05	1.00
ATC does not resolve the conflict	ASC32b1	3.53E-01	2.07E-01	0.59	0.231481	0.65
Flight crew or vehicle driver does not resolve the conflict	ASC32c1	1.46E-02	1.46E-02	1.00	0.014622	1.00
Aircraft continues flight	ASC32c2	8.04E-06	9.86E-06	1.23	9.56E-06	1.19
Collision on runway	ASC32d1	6.43E-08	3.77E-08	0.59	4.21E-08	0.65
Aircraft continues flight	ASC32d2	4.33E-06	2.54E-06	0.59	2.84E-06	0.65

CATS results:


CATS [Y:\tekst\Projects\DGB CATS\CATS validatie\Data_maxiCATS\DGB 2014.CATS]

File Edit Calculate View Help

Flights | Inputs | Engine | Calculate | Results

Individual flight Groups | Compare Flight Groups | Total of all Flight Groups

Accidents | End Events | ESD End Nodes

Accident frequency | Fatality frequency | Fatalities | Cost

Expand all | Expand one | Collapse | Swap top two

Results are per flight

|Baseline| Baseline	RVT-2-Dir...	RVT-2-MF	RVT-3-MF	RVT-4-MF	RVT-5-MF	Bir
⊟ Flight crew or vehicle driver resolved conflict	1.01	1.00	1.01	0.63	0.69	
AL32c2_02 Accident frequency: 4.15e-6 per flight	4.18e-6	4.17e-6	4.21e-6	2.62e-6	2.88e-6	
⊞ Flight crew or vehicle driver resolved conflict	1.01	1.01	1.01	1.01	1.01	
⊟ Collsion with ground	1.01	1.01	1.01	1.01	1.01	
ER15e1_01 Accident frequency: 2.63e-8 per flight	2.64e-8	2.64e-8	2.64e-8	2.64e-8	2.64e-8	
⊟ Collision with ground	1.43	1.43	1.43	1.43	1.43	
AL35e1_01 Accident frequency: 2.06e-8 per flight	2.93e-8	2.93e-8	2.93e-8	2.93e-8	2.93e-8	
⊞ Collision with ground	1.13	1.13	1.13	1.13	1.13	
⊞ Collision with ground	1.04	1.04	1.04	1.04	1.04	
⊞ Collision with ground	1.01	1.01	1.01	1.01	1.01	
⊞ Collision with ground	1.01	1.01	1.01	1.01	1.01	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	1.00	1.00	1.00	1.00	1.00	
⊞ Collision with ground	0.99	0.99	0.99	0.99	0.99	
⊞ Collision with ground	0.99	0.99	0.99	0.99	0.99	
⊞ Collision with ground	0.95	0.95	0.95	0.95	0.95	
⊞ Collision with Ground	0.86	0.86	0.86	0.86	0.86	
⊞ Collision with gorund	1.01	1.01	1.01	1.01	1.01	
⊞ Collision on Runway	1.02	1.02	1.02	1.02	1.02	
⊟ Collision on Runway	1.00	1.00	1.01	0.63	0.69	
AL32d1_01 Accident frequency: 7.06e-8 per flight	7.05e-8	7.03e-8	7.11e-8	4.42e-8	4.86e-8	
⊞ Collision in Mid-Air	1.01	1.01	1.01	1.01	1.01	
⊟ ATC resolved conflict	1.00	1.00	0.99	1.18	1.15	
AL32b2_03 Accident frequency: 8.91e-6 per flight	8.88e-6	8.90e-6	8.85e-6	1.05e-5	1.02e-5	

Figure 26 Accident frequencies as a result of case study 4

Comparing accident probabilities of CATS and ASCOS can be done by means of the following mapping (end states with the same meaning in CATS and ASCOS):

	CATS	ASCOS
Collision on runway	AL32d1_01	ASC32d1
Flight crew or vehicle driver resolved conflict	AL32c2_02	ASC32d2
ATC resolved conflict	AL32b2_03	ASC32c2

An important result of this case study is that the modification factors, on a qualitative level, have a similar positive or negative effect in both the CATS as the ASCOS runs. On a quantitative level, the CATS and ASCOS accident frequencies cannot be compared due to the fact that CATS and ASCOS are different models, based on different datasets, as explained in the beginning of this appendix.

Looking only at the CATS results of this case study (*Figure 26*), we notice that not only accident frequencies of end nodes of ESD 32 change but also accident frequencies of other ESDs, e.g. ESD15 (ER15e1_01) and ESD35 (AL35e1_01). For example, changing a RIMCAS probability (Run2) apparently also has an effect on the accident frequency after a non-operating anti-ice system en route (ESD 15) and the CFIT related accident frequency (ESD 35). This could be explained from the fact that CATS includes ATC, flight crew and maintenance human performance models that influence base events of various ESDs. Another aspect of the CATS model is that dependencies between ESDs are modelled such that changes in ESDs in the take-off flight phase could affect frequencies of ESDs in later flight phases. However, it is not exactly clear and should therefore be studied.

Findings

- Qualitative changes, i.e. negative or positive change in probability, to ESD accident frequencies are similar in ASCOS and CATS.
- The changes to accident frequencies in CATS are as expected qualitatively. No conclusions can be drawn regarding the quantitative results.
- Accident frequencies of ESDs other than ESD 32 are also affected by the modifications, such as ESD 15 and ESD 35. The mechanisms behind that should be well understood and clarified before being able to draw solid conclusions on the results.
- The outcomes of the two approaches used in Run 2 (modification factors versus user override of frequencies) are not exactly the same. It has to be clarified why this is not the case.

Appendix A.5 Overall findings

In this document the findings and results are described of some case studies that have been performed with the Causal Model of Air Transport (CATS) as described in the Final Report [33]. CATS consists of a model, a baseline quantification of the events in the model and a supporting toolset. Validation of CATS has been done on a certain level during the CATS project. To extend the experience with CATS, some additional case studies have been performed with the CATS model and toolset of 2009. Input for the identification of the case studies has been derived from DGB, EASA and the ASCOS project. Here, main findings are presented.

Risk picture

Risk pictures for a specific situation such as the baseline of the model are easy to generate with maxiCATS. The user can change the sorting order in each of the columns (increasing and decreasing results) and can also change the order of columns (from left to right). Risk pictures can be made on the level of accident types, ESD end event types and for all ESD end states. If the user wants to create a risk picture which is not available through maxiCATS (e.g. accident risk per ESD or the frequencies of accident end states only), he/she can do this by post processing of maxiCATS results, for example by exporting data to Microsoft Excel.

Precursors

MaxiCATS provides the possibility to identify precursors to risk by presenting overviews of events of individual ESDs and of all base events in the model. The user can rank base events by frequency of occurrence easily. The original DNV spreadsheet on which the toolset is based also contains information about the contribution of a base event to the accident probability of an ESD. This could be relevant for the identification of precursors. Contribution information is not included in maxiCATS. This means that some post processing by the user is required to derive the requested results.

Modifications to base event frequencies

Two case studies focused on the effects of making changes to frequencies of base events of the model. The aim was to check if the changes to the accident frequencies of end states would change according to expectation. If the end states frequencies do not change as expected, it should be possible to explain the difference by means of the qualitative model. The two case studies were a) application of bird management by an Airport and b) the introduction of a remote virtual tower operation in the air transport system. The main findings of these case studies are:

- Most often, the accident frequencies changed as expected (at least qualitatively) after making modifications to input parameters.
- There are two different approaches to making changes to base event frequencies: by using pre-defined modification factors and by using a user override function in which the user can directly change frequencies of specific events. Where one would expect similar results for these two approaches if in both approaches the changes to the same set of events was identical, this was not the case. Apparently, there is a different effect when the user uses a modification factor, or when the user conditionalizes the frequency of the event by a certain weight or a factor. This distinction should be well understood before continuing using the toolset.
- Using a “user override” function, base event frequencies can be set by the user. It appears that the new values should lie in a predefined interval (Min and Max values). This means that the user is not completely free to make choices for base event frequencies (e.g. extreme values of 0 or 1), without making changes to the probability density functions of the base events.
- Some pre-defined modification factors for Bird Management (cases: Used and Not Used) have been implemented in the CATS toolset. These modification factors change base event frequencies of ESD01, ESD33 and ESD35. This is different from those identified by searching the DNV spreadsheet, viz. ESD09, ESD10, ESD14, ESD18, ESD 27 and ESD33. This difference needs to be clarified.

- Due to the fact that CATS is implemented as one big Bayesian Belief Net, including BBNs for human performance models that affect base events of more than one ESD, more accident frequencies might change than expected after changing base event frequencies. The mechanisms behind that should be well understood by the user before being able to draw solid conclusions on the results.

CATS in general

CATS 2009 is an advanced BBN model supported by a toolset with many possibilities. That in itself is very interesting and worthwhile. However, this also means that working with the model and the toolset requires the user to have a good level of knowledge of the structure and quantification of the risk model. By doing exercises and case studies as those performed in this document, CATS users can learn a lot about required and useful functionalities that the CATS model and toolset should be able to support. With hands-on experience, the user is able to identify additional interface and functional requirements for the supporting toolset which can be an input for further model and tool development, such as:

- Adding in maxiCATS and CATSPAWS a function for searching events by (part of) a certain name, instead of by event code only;
- Adding a function that supports the user in selecting appropriate base events, not only by using the name of the event, but also by including the definition of the event. For example: search for “bird” should also lead to finding the event “FOD”, which includes “birds” in its definition. Somehow the user must be assured that the selection of events is complete.

The exercises however also showed some unexpected results which could not be explained (yet) and as such did not improve the understanding of the working of the model and the toolset.

As described in the software documentation, some functionalities of the software tool set have not been implemented yet. For example, new modification factors have to be implemented in the CATSPAWS database directly, instead of through the CATSPAWS programme.